Published in final edited form as:

Augment Altern Commun. 2014 March ; 30(1): 71–82. doi:10.3109/07434618.2014.880190.

Validity of a Non-Speech Dynamic Assessment of Phonemic Awareness via the Alphabetic Principle

R. Michael Barker,

Schiefelbusch Institute for Life Span Studies, University of Kansas, United States

Mindy Sittner Bridges, and

Schiefelbusch Institute for Life Span Studies, University of Kansas, United States

Kathryn J. Saunders

Schiefelbusch Institute for Life Span Studies, University of Kansas, United States

Abstract

Most assessments of phonemic awareness require speech responses and cannot be used with individuals with severe speech impairments who may use augmentative and alternative communication (AAC). This study investigated the reliability and construct validity of the Dynamic Assessment of Phonemic Awareness via the Alphabetic Principle (DAPA-AP), which does not require speech. In all, 17 adults with mild to moderate intellectual disabilities completed the DAPA-AP, a letter-sound knowledge task, four measures of phonological awareness, and two reading assessments. Results indicated the DAPA-AP was both a reliable and valid assessment of phonemic awareness for this sample. Consequently, the DAPA-AP represents an important step in developing phonemic awareness assessments that have the potential to be suitable for use with a wide range of individuals, including those with SSI.

Keywords

Phonemic awareness; Reading; Assessment; Severe speech impairment; Alphabetic principle; Complex communication needs

Validity of a Non-Speech Dynamic Assessment of Phonemic Awareness via the Alphabetic Principle

Children with severe speech impairments (SSI) are likely to have considerable difficulty learning to read (Erickson, 2005; Foley, 1993; Foley & Pollatsek, 1999; Koppenhaver & Yoder, 1992), and these difficulties are likely to persist into adulthood. Furthermore, individuals with severe speech impairments often communicate using augmentative and alternative communication (AAC) systems, which can include selecting pictures or icons or spelling on a keyboard (Beukelman & Mirenda, 2005). These devices often include a

speech-generating component that will produce synthetic or recorded speech when the child uses the device. As adaptable as these systems are, they are inherently constrained by context and the vocabulary that is pre-programmed into the system. For example, a person using a system with preprogrammed vocabulary cannot order a specific item at a restaurant if that item is not already part of the communication device. This example highlights the importance of individuals having the skills to read and write in order to use devices that can generate speech from text input, as these devices are infinitely generative. As a result, adequate reading and writing skills can serve as an important alternative modality to communicate in lieu of speech, particularly when paired with a speech-generating device (Barker, Saunders, & Brady, 2012).

From the beginning of formal education, summative and formative assessments are used to measure attainment of important goals. Summative assessments evaluate whether students have achieved overarching curricular goals; formative assessments provide ongoing progress monitoring on the way to those goals (Yorke, 2003). Together, they tell teachers and clinicians about individuals' school readiness, whether they have met important educational and developmental milestones, and whether they are ready to proceed to the next grade. In recent decades, emphasis on the formative assessment of phonemic awareness and other subskills related to decoding, and not just decoding itself, has increased. Because such assessments can determine reading-instruction placement and subsequent progress monitoring, obtaining reliable and valid information is critical to a child receiving appropriate instruction, and achieving desired literacy goals. Consequently, there is a strong need for assessments that reliably and validly characterize students' skills. Unfortunately, most of these assessments rely on speech as a response mode and are thus not appropriate for individuals with severe speech impairments. This study represents a first step in the development of tests designed to assess the precursors to decoding skills in individuals with severe speech impairments. Its goal is to describe the reliability and construct validity of a new dynamic assessment of phonemic awareness that does not require speech responses and is appropriate for individuals with severe speech impairments.

Phonemic Awareness and the Alphabetic Principle

Although a wide variety of skills are necessary for skilled reading (Adams, 1990; Whitehurst & Lonigan, 1998), the acquisition of the alphabetic principle underpins the development of decoding skills (Byrne, 1998). The alphabetic principle has been defined as knowledge that "phonemes can be represented by letters, such that whenever a particular phoneme occurs in a word, and in whatever position, it can be represented by the same letter" (Byrne & Fielding-Barnsley, 1989, p. 313). The "in-a-word" part of the definition bears emphasis. The alphabetic principle involves relations between sounds that are embedded within whole spoken words (e.g., phonemes) and letters within printed words.

The development of the alphabetic principle rests on several subskills, including phonological awareness. Phonological awareness is a general term that refers to the ability to detect, manipulate, and/or analyze the sounds of speech (Adams, 1990; National Early Literacy Panel [NELP], 2008; National Institute of Child Health and Human Development [NICHD], 2000). Phonemic awareness, an aspect of phonological awareness, refers

specifically to the ability to detect sounds that are embedded within spoken syllables. As perhaps the most critical component of the alphabetic principle, phonemic awareness is one of the best predictors of reading success for all children (National Early Literacy Panel [NELP], 2008; NICHD, 2000; Snow, Burns, & Griffin, 1998; Wagner & Torgesen, 1987; Wagner, Torgesen, & Rashotte, 1994).

The discovery that phonemic awareness plays a key role in the acquisition of reading skills has had a profound effect on research and practice. Numerous assessment procedures have been designed and standardized. These typically involve tasks that require matching, deleting, moving, blending, or segmenting phonemes in spoken words. In a sound-matching (or categorization) task, participants match spoken words that have the same phoneme in the same position. For example, the participant might be asked, *Which word begins with the same sound as sat, sum or mum?* In an elision task, a person must delete a target phoneme from a word and then speak the new word. For example, a person might be asked to *Say tiger without saying /g/*. In a blending task, a person must combine individually presented phonemes and say the word. For example, a person might be asked, *What word do the sounds /d/ /o/ /g/ make?* In a segmenting task, a person must speak a target word one phoneme at a time. For example, a person might be asked to *Say cat one sound at a time*.

These examples are typical in that they require spoken responses. On the widely used Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999), all four subscales designed for ages 7 to 24 -- elision, blending words, blending nonwords, and segmenting nonwords -- require speech responses. Only the sound-matching subscale of the CTOPP does not require speech responses. Sound matching uses pictures to represent spoken words and requires children to choose a word that starts or ends with the same sound as a spoken example by either naming or pointing to the correct picture. This subscale is designed for young children with age-appropriate pictures, and it only assesses awareness of onsets (the first sound in a word) and codas (the last sound in a word). Consequently, there are very few options for comprehensive standardized assessments of phonological awareness that do not require speech responses and thus are appropriate for individuals who have severe speech impairments (Barker et al., 2012).

Assessment for Readers with Severe Speech Impairments

For individuals with severe speech impairments who communicate via augmented means, assessment of phonological awareness with standard assessments is difficult, if not impossible. Assessment items that require verbal responses must be modified so that non-verbal responses can be given, including pointing, scanning, or yes/no responses (Barker et al., 2012; Beukelman & Mirenda, 2005; Gillam, Fargo, Foley, & Olszewski, 2011). Doing so interferes with the psychometric properties of these standardized assessments and makes them difficult to interpret. Moreover, as described in the previous examples, the standard assessment methods frequently require complex verbal instructions that may not be understood by individuals with language delays.

Ideally, an assessment of phonemic awareness for individuals with severe speech impairments would consist of features that would directly address some of the problems associated with modifying standard assessments. First, it would be designed so that the

response mode would not require speech (i.e., no modifications would be needed to allow a person to respond without using speech). This could include touching-response options such as using one's finger or a mouth stick; there also should be support for providing responses via scanning. Second, it would require very limited verbal instructions to ensure task comprehension for young children or older children and adults who have language problems. Third, and related to the second point, it would be dynamic and thus provide feedback and teach the task as the individual moved through the assessment. And finally, it would be given on a computer (with oversight of an administrator) to fully automate the assessment process, increase procedural fidelity, and facilitate an objective measurement of responses.

There have been formal attempts to develop assessments with some of these characteristics for use with individuals with severe speech impairments (e.g., Iacono, 2004; Iacono & Cupples, 2002; Vandervelden & Siegel, 2001). Vandervelden and Siegel described four phonological awareness tasks that relied on recognition: initial, final, and complex phoneme recognition; and a visually adapted deletion/substitution task. In the initial and final phoneme recognition tasks, participants provided yes/no responses when judging whether a given spoken word contained a target phoneme in the initial or final position. For complex phoneme recognition, participants also indicated the position of the target phoneme by indicating first or last. In the visually adapted phoneme deletion/substitution task, participants chose pictures that represented a target word that was constructed by either removing the initial or final phoneme, or replacing the initial or final phoneme with another phoneme. Iacono and Cupples created an assessment of phonological awareness with computerized components for individuals with severe speech impairments, the Assessment of Phonological Awareness and Reading (APAR; available for download at <http://www.elr.com.au/apar/>). Subtests of the APAR are blending real words, blending nonwords, phoneme counting, and phoneme analysis. Blending real words, phoneme analysis, and counting phonemes are visually adapted such that participants provide answers by choosing pictures that represent words or numbers following the spoken test item. For blending nonwords, the administrator says a nonword one sound at a time, and then speaks either the same or a different non-word blended together; the participant provides a yes/no response indicating whether the two match.

Although both Iacono (2004) and Vandervelden and Siegel (2001) demonstrated some evidence of construct validity for their assessments, there are several aspects of these assessments that can be improved. Importantly, both of these examples represent attempts to modify existing static approaches to measuring phonological awareness so that they no longer require speech responses. In addition, both assessments required participants to comprehend complex verbal instructions in order to respond correctly. For example, instructions for a trial of the phoneme deletion/substitution task were, *Listen for /tin/. Take away the /t/. What word is left? Show me* (Vandervelden & Siegel, 2001, p. 43). People with limited language skills, such as those with intellectual disability, may have trouble understanding instructions such as these. Furthermore, as typical of static assessments, none provided feedback during the assessment to help participants learn the task, and, although the stimuli from the APAR can be administered via computer, neither of the assessments was fully computerized. Finally, some tests required significant pre-teaching, such as teaching

abstract symbols or pictures that were unknown prior to assessment, a process that can be time consuming.

Assessing Phonemic Awareness via the Alphabetic Principle

In the current study, a different approach was taken to eliminating the speech response, resulting in the creation of a task that simultaneously assesses the alphabetic principle and phonemic awareness. The procedures are modeled after a seminal series of studies on the development of the alphabetic principle in young children at the prereading stage of literacy development (summarized in Byrne, 1998). Following the logic that the alphabetic principle is necessary but not sufficient for decoding, Byrne's goal was to assess the alphabetic principle as cleanly as possible using procedures that did not require decoding.

To set the stage, we first describe the Byrne procedures, which required spoken responses. The initial studies focused on onsets (Byrne & Fielding-Barnsley, 1989). As noted, the children were nonreaders who knew few letter names and virtually no letter sounds. Five pairs of words using the onsets /m/ and /s/ (e.g., sat/mat, sum/mum, etc.) were included. Children were taught to read one of the five word pairs, to produce the sounds given the printed letters m and s, and to respond correctly to sound-categorization items incorporating all 10 words. In each trial of the forced-choice generalization test, the children were shown one of the 8 untaught printed words (e.g., sum) and given two spoken-word choices that differed only in the onset (e.g., *Does this say sum or mum?*). Saying the word that corresponded to the printed word was designated as correct.

Our version of the forced-choice task eliminates the speech requirement by reversing the roles of the spoken and printed words. Each trial presents a single spoken syllable (e.g., *sum*), and participants select between two printed-CVC-syllable choices that differed only in the target sound (e.g., sum and mum). The procedures are otherwise parallel to Byrne's. Responding correctly requires phonemic awareness, that is, in order to select the correct printed word, the participant must be able to isolate the target phoneme within a whole spoken syllable. An important caveat is that, in providing evidence of the alphabetic principle, the procedures are relatively pure; however, as evidence of phonemic awareness, the inclusion of letters clearly goes beyond pure phonemic awareness. Thus, although success on our task demonstrates phonemic awareness, failure does not indicate its absence; failure may be due to not understanding the alphabetic principle, which requires phonemic awareness and the knowledge that sounds map onto letters.

An additional reason for our choice of task is that it lends itself well to dynamic assessment. Most assessments of phonological awareness are static measures; the term refers to assessments of already learned abilities at one point in time (Lidz, 1991). In a static assessment, individuals answer a set of items with little or no feedback. As a result, failure could be due to a lack of understanding of the task, thus underestimating the individual's skills. Accordingly, educators and researchers have proposed dynamic assessment as an alternative assessment method. Dynamic assessment refers to procedures that embed instruction in the assessment process. In dynamic assessment, the examiner takes an active role by teaching a task or providing explicit prompts. Success is measured by a student's level of both independent and assisted performance (i.e., progress). Dynamic assessment

takes into account both the process and the product of learning; in other words, it considers growth in response to some sort of instruction (see Sternberg & Grigorenko, 1998, for a review). As a result, dynamic assessment can provide information about an individual's ability to respond to instruction that is not obtainable through more traditional assessment sources (Grigorenko & Sternberg, 1998; Lidz, 1991).

The Computerized Assessment of Phonemic Awareness

Our assessment, the Dynamic Assessment of Phonemic Awareness via the Alphabetic Principle (DAPA-AP), was designed to address many of the assessment issues faced by individuals with severe speech impairments. First, as stated previously, individuals respond to items on the DAPA-AP by touching a target on a computer screen instead of providing a verbal response. Second, the DAPA-AP requires limited verbal instructions, such that comprehension demands are intended to be very low. Third, the DAPA-AP has a dynamic component that helps test-takers understand the task by providing them with prompts, when needed, as they progress through the test items. In addition, the DAPA-AP incorporates the use of response items that consist of letters. The use of letters has been shown to help children demonstrate phonological awareness (Boyer & Ehri, 2011; Ehri et al., 2001), and it has been suggested that letters should be integrated into phonological awareness instruction for children with and without disabilities (Boyer & Ehri, 2011; Browder et al., 2009). Finally, the DAPA-AP is administered via a computer program, which can reduce administration time and administrator error. Reduced administration time is important for children who may fatigue easily.

The DAPA-AP uses pairs of consonant-vowel-consonant syllables -- mostly nonwords --to test awareness of target phonemes in a given location within a word. For example, one syllable pair in the onset subtest is mib/sib. For each trial, the computer plays an audio recording of one of the spoken syllables (e.g., "mib" and "sib") while also presenting the printed words mib and sib on the screen. To answer correctly, the participant must touch the printed word with the onset letter that matches the onset phoneme of the spoken word. Importantly, because the other phonemes in the word are the same (i.e., "ib"), the participant must differentiate the words based only on the onset (i.e., either /m/ or /s/).

Goals of Study

To establish that the DAPA-AP measured the construct of phonological awareness, we assessed its concurrent and convergent validity. Concurrent validity refers to an assessment's association with other well-established assessments of a construct (Coaley, 2010). Convergent validity refers to an assessment's association with other constructs that are known to be associated with the construct that an assessment purports to measure (Coaley, 2010). Consequently, we sought to establish the DAPA-AP's concurrent validity with well-established measures of phonological awareness, and its convergent validity with measures of reading. Achieving these goals required participants who could provide spoken responses yet had limited reading skills and sufficient attention skills for lengthy testing. Thus, we selected a group of adults with intellectual disabilities and limited reading skills. We established reliability by evaluating the internal consistency of the DAPA-AP and its subtests. We hypothesized that the DAPA-AP would demonstrate adequate reliability, as

evidenced by high values on measures of internal consistency. We evaluated concurrent validity by calculating correlation coefficients between DAPA-AP scores and subscales of the CTOPP. We hypothesized that the DAPA-AP would demonstrate high concurrent validity, as scores on the DAPA-AP would be highly correlated with other measures of phonological awareness. We evaluated convergent validity by calculating correlation coefficients between DAPA-AP scores and measures of real word and nonword reading. We hypothesized that the DAPA-AP would demonstrate high convergent validity, as scores on the DAPA-AP would be highly correlated with other measures of reading.

Method

Participants

Participants were 17 adults with mild to moderate intellectual disabilities who were enrolled in a residential facility that specializes in behavior problems. Level of intellectual disability was determined using records provided by the residential facility. Participants' mean age was 32 years (range = 16 to 58); 15 participants were male and 2 were female. All had sufficient speech skills to provide spoken responses to standard assessments, where required. Mean and median grade equivalent scores on the word identification and word attack subscales of the Woodcock Reading Mastery Test – Revised (WRMT-R; Woodcock, 1998) were 1.7 ($SD = 1.2$) and 1.6 (range = 0.0 to 5.0) and 1.3 ($SD = 2.0$) and 0.0 (range = 0.0 to 6.4), respectively. On average, participants identified 14 of 26 letter sounds ($SD = 9.74$; full description of the letter-sound task follows). We chose adult participants because a total of 11 assessments, counting all subtests, were included in the study, and the amount of time required to complete the DAPA-AP was unknown. Research was approved through the institutional review board at the University of Kansas and informed consent was obtained prior to participation. It is worth noting that we have a long-standing research relationship with the residential facility from which participants were recruited. Among the residents, there was a positive culture of participating in research and wide acceptance by residents, staff, and researchers of a well-established token economy. For their participation, participants were given credits to an on-site store where they could purchase sundries. They received the equivalent of \$2 USD for each session.

Procedure

A female research assistant, who was blinded to the purpose of the study, administered all assessments in a quiet, private room at a university research facility in the Midwestern United States. The DAPA-AP required one to four sessions, and was followed by one session for the standardized assessments. Sessions lasted approximately 30 min. For all except the first five participants (discussed later), each subtest of the DAPA-AP started with pretraining trials, where participants matched the printed syllable-pairs on the computer screen. After pretraining, participants received the following instructions from the research assistant: *The computer is going to say some words and I want you to touch the word that you hear.* If distracted, the researcher redirected participants and encouraged them to select an answer by pointing to the computer screen and using phrases like, *Which word?* and *What do you hear?* No other verbal instructions were given. Participants who performed well on the DAPA-AP typically completed it in a single session. Participants with lower accuracy

necessarily took longer to complete the DAPA-AP, because errors caused the DAPA-AP program to branch to additional prompted trials. The DAPA-AP was always administered first with the subtests in the following order: onset, rime, coda, and vowel. Following the completion of the DAPA-AP, in a separate session, participants completed a battery of standard assessments that required spoken responses. The first author scored the DAPA-AP results; the research assistant scored all other assessments.

DAPA-AP—The DAPA-AP was administered via the Match to Sample application (Dube, 1991) on a 12” (30.48 cm) laptop computer. The participants used a KTMT-1214TM¹ touch-screen overlay to indicate responses on the DAPA-AP. Printed syllables were displayed in black 24-point Geneva font on a white background. All auditory stimuli were digital recordings of a female speaking standard mid-western English.

The DAPA-AP consisted of four subtests: onset, rime, coda, and vowel. The syllable-pairs used in each subtest are presented in Table 1. The syllables were chosen to minimize the possibility of being recognized by sight. Few were real words, and with the exception of the first pair for the onset subtest (i.e., mat/sat), pairs never contained two real words. With the exception of mat and sat, the rimes within the rime, coda, and vowel subtests were not themselves real words. For example, syllable-pairs did not contain the rimes -it or -up. The onset, rime, and vowel subtests used six syllable-pairs; the coda subtest used nine syllable-pairs.

All four subtests were constructed according to the same logic. Each syllable-pair isolated the targeted segment by contrasting words that differed only in that segment, thus making that segment the only possible basis for a correct selection. The participant’s task was to listen to the recorded spoken stimulus, and choose the corresponding printed target. For example, in the onset subtest, the two printed choices had different onsets and the same rime (e.g., mib and sib). In order to answer correctly, the participant would have to recognize that the spoken “mib” started with /m/, and then associate that sound with the printed mib, which differed from sib only in the onset position. The pairs in the vowel and coda tests also differed by a single phoneme/letter. Note that, in the rime test, the incorrect choice differed in both the vowel and final consonant (e.g., “kog” as the sample, with kog and kib as choices). This is the only test with a target unit larger than a phoneme. (See Kirtley, Bryant, MacLean, & Bradley, 1989, for an example of similar logic in designing tests to isolate phonological awareness at the onset-rime versus single-phoneme level.) Also note that, in the coda subtest, we included two different vowel-consonant rimes for the word pairs (ot/og and ep/ek) to ensure we were testing discrimination between more than one vowel-consonant contrast.

Two types of trials were used in the DAPA-AP -- non-prompted and prompted (see Figure 1) -- arranged in non-prompted or prompted blocks of six trials. The two spoken syllables of the pair were presented in quasi-random order across trials, with the constraint that the same syllable was presented on no more than two consecutive trials. Each spoken syllable was presented 3 times. For each trial in a non-prompted block, the computer presented the

¹The KTMT-1214 Add-On Touch Screen for 12-14” Notebook is a product of Keytech, Inc., Garland, TX.

spoken target word, while displaying a small black box in the center of the screen. Touching the black box produced printed syllable-choice stimuli in two of the four corners of the screen, while continuing to present the spoken syllable every 2s (see the left side of Figure 1). Selecting the correct printed syllable produced a display of stars accompanied by a chime; incorrect selections produced a black display accompanied by a buzz. Prompted blocks differed in that the black box was replaced with a printed-syllable target (see the right side of Figure 1), which was then displayed along with the two printed choices. As in the non-prompted block, the spoken syllable was presented every 2s. The prompted blocks provided an opportunity for the participant to match printed syllables in order to learn the relationship between the spoken syllable and its printed form.

Each syllable pair was presented in either one non-prompted block (6 trials total) or a combination of three non-prompted and prompted blocks (18 trials total). Block 1 was always non-prompted. Participants who failed to meet criterion on Block 1 were presented with a second and third block. Block 2 was always prompted. Block 3 was non-prompted if criterion was met on block 2; Block 3 was prompted if criterion was not met on Block 2. Figure 2 depicts scoring as a function of each possible pathway through the blocks for the syllable pair, *mib/sib*. For 3 points, the participant must have met criterion (at least 5 of 6 correct; indicated by a + in Figure 2) on Block 1. For 2 points, the participant must have not met criterion on Block 1 (less than 5 of 6 correct; indicated by a – In Figure 2), met criterion on Block 2 (a prompted block), and then met criterion on Block 3 (a non-prompted block). For 1 point, the participant must have not met criterion on Block 1, met criterion on Block 2, and not met criterion on Block 3 (a non-prompted block); or *not* met criterion on Block 1, not met criterion on Block 2, but did meet criterion on Block 3 (a prompted block). For 0 points, the participant must have not met criterion on any of the three blocks.

Points for syllable-pairs were summed within each subtest and divided by the number of syllable pairs in that subtest. The range of possible scores for each subtest was 0 to 3. A participant with a subtest score close to 3 needed very few prompts. A score of approximately 2 meant that a participant answered correctly only after a prompted block for most items. A score of approximately 1 meant that a participant only met criterion on a prompted block, and not after the prompts were removed (did not learn from the prompts). A score close to 0 meant that the participant did not show evidence of identity matching (rarely met criterion on prompted blocks). DAPA-AP total score was calculated by summing across the subtests; the possible range of the DAPA-AP total score was 0 to 12.

Assessments requiring spoken responses—Participants completed five tests typically used to assess letter–sound knowledge, phonological awareness, and reading skills in individuals who can provide spoken responses, in the following order: a letter–sound knowledge task, four subscales from the Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999), and two subscales from the WRMT-R (Woodcock, 1998). The standardized tests were administered in accordance with the published testing manuals.

Letter–Sound Knowledge: The letter–sound knowledge (LSK) task was adapted from the Curriculum Based Measurement for Early Literacy (Speece, Case, & Molloy, 2011). Participants were shown 26, 3" × 5" (7.62 cm × 12.70 cm) flashcards, one for each lower-

case letter of the English alphabet printed in 48-point Arial font. The assessment administrator asked the participant, *What sound does this letter make?* when showing each card. If the participant named the letter, the administrator asked, *But what sound does it make?* Only short-vowel sounds were accepted as correct. Hard or soft sounds were accepted as correct for letters c and g (as in candy/cent and good/giant). For letter x, /ks/ as in box and /z/ as in xylophone were accepted.

CTOPP: The sound matching, blending words, elision, and rapid letter-naming subscales from the CTOPP were administered (Wagner et al., 1999) in the order described here. The sound matching subscale measured participants' ability to match a given word (and picture) to another based on the target word's onset or coda. Blending words measured participants' ability to combine phonemes that are presented one at a time into a real word and speak the word. Elision measured participants' ability to isolate a target phoneme from a spoken word, delete the phoneme, and speak the new word created by deleting the phoneme. Rapid letter naming measured the time it takes for a participant to orally name six letters (a, t, k, s, c, n), presented 9 times each, in an array with four rows and nine columns. Two arrays were presented and times were summed across the arrays. The alpha coefficients for sound matching, blending words, elision, and rapid letter naming were .93, .84, .89, and .82, respectively, per the CTOPP examiner's manual.

Woodcock Reading Mastery Test – Revised: The word identification and word attack subscales from the WRMT-R (Woodcock, 1998) were administered. Word identification measured participants' ability to read aloud a list of real words that increased in difficulty. Word attack measured participants' ability to decode and speak a list of non-words that increased in difficulty. The WRMT-R examiner's manual reports the split-half reliabilities for word identification and word attack to be .98 and .94, respectively.

Results

Because of the small sample size for this study, we present statistics with bootstrapped 95% confidence intervals where appropriate. Bootstrapping is a statistical technique where k samples of n size are drawn randomly with replacement from the collected data. These bootstrapped samples are used to create a confidence interval around the estimates derived from the sample data. We estimated $k = 1000$ bootstrapped samples of $n = 17$ for our results. The confidence intervals we report indicate that the estimate in question fell within that interval in 950 of the 1000 bootstrapped samples we created; thus, it is 95% likely that the true population parameter for the estimate falls within the bootstrapped 95% confidence interval.

We analyzed raw scores for the CTOPP and WRMT-R. Scores for all subscales except rapid letter naming represent the number of items answered correctly. Rapid letter naming indicated the number of seconds to speak the names of all items in the array, thus higher scores were indicative of poorer performance. We collected complete data on all but 2 of the 17 participants, who could not accurately respond to the practice items on rapid letter naming and, per the instructions, were not given the assessment. The DAPA-AP total score ranged from 0 to 12; DAPA-AP subtests ranged from 0 to 3, as described previously.

General Description of DAPA-AP Outcomes

Table 2 shows the DAPA-AP scores for each of the 17 participants (names are pseudonyms) broken down by subtest. The first 5 participants listed in the table did not receive identity-matching pretraining, as it had not yet been included as part of the DAPA-AP when they were tested. The first 3 participants, Kevin, Zac, and Ethan, earned perfect scores on the DAPA-AP, thus the lack of pretraining did not impact their performance. Clare and Josh, the other 2 participants who did not receive pretraining, demonstrated increasing scores from onset, to rime, to coda. This indicated that they may have been learning to identity match printed syllables as they progressed through the subtests. A closer look at their data revealed that both had low scores on the prompted trials for the first four syllable-pairs, but began scoring perfectly or almost perfectly on the prompted blocks at the min/sin pair of the onset subtest. This motivated us to add the pretraining component for all subsequent participants, to ensure that they were familiar with identity matching printed stimuli prior to beginning each subtest. Means, standard deviations, and medians for the CTOPP and the WRMT-R, and the DAPA-AP are presented in Table 3. Bootstrapped 95% confidence intervals are presented in brackets below each estimate.

Reliability

Internal consistency of the DAPA-AP was assessed using Cronbach's alpha. Results indicated excellent reliability for the items of the DAPA-AP. For all items, regardless of subtest, $\alpha = .98$. For onset, rime, coda, and vowel, α s = .93, .96, .97, and .93, respectively. In addition, the DAPA-AP total score was highly correlated with each subtest, r s = .94 to .98, p s < .01, bootstrapped 95% CIs .84 to 1.00, and each subtest was highly correlated with every other subtest, r s = .85 to .95, p s < .01, bootstrapped 95% CIs .55 to 1.00. The remaining analyses will use the DAPA-AP total score because of its high correlations with the other subtests of the DAPA-AP.

Characteristics of the Distributions

Floor effects—We also investigated the distributions of the CTOPP subscales in order to determine whether floor effects were present. The presence of a floor effect may indicate that a measure does not differentiate individuals' performance at the low end of the scale (Catts, Petscher, Schatschneider, Bridges, & Mendoza, 2009). Floor effects can be demonstrated by distributions that have positive skews, a high proportion of scores of zero, medians that are lower than the mean, and/or standard deviations that are as large or larger than the distance from lowest score to the mean. The simplest way to determine whether a distribution has a positive skew is the skew statistic (see Table 3). Of the assessments that required spoken responses, elision and blending words demonstrated the highest degree of positive skew, 1.95 and .70, respectively: 7 and 3 of 17 participants scored zero on elision and blending words, respectively. Both distributions had medians that were lower and standard deviations that were larger than their respective means. Word identification and word attack showed a similar pattern, with 2 and 9 of 17 participants scoring zero, respectively. Consequently, all four measures demonstrated floor effects.

Ceiling effects—The opposite was true of the subtests and overall scores for the DAPA-AP, as these measures showed evidence of ceiling effects. All four subtests demonstrated negative skew, a high proportion of maximum scores (3), medians that were higher than the mean, and standard deviations that were as large or larger than the distance from the mean to 3 (see Table 3). For the onset, rime, coda, and vowel subscales, 6, 4, 6, and 3 of 17 participants scored 3, respectively. Furthermore, 5 of 17 participants scored 12 (the maximum) on the DAPA-AP total score.

If the DAPA-AP is a measure of the phonemic awareness that uses the alphabetic principle, and the alphabetic principle constitutes a major component of decoding, one might expect individuals who showed some decoding skills on the Woodcock word-attack subtest to do very well on the DAPA-AP. We tested this prediction by calculating the odds that participants who could decode more than five items on word attack also scored at or near 12 (the maximum score) on the DAPA-AP. In all, 6 of 17 participants answered more than five items correctly on word attack; 8 of 17 participants scored greater than 11.5 on the DAPA-AP total score. Participants who decoded more than five nonwords were 13.33 times more likely to score greater than 11.5 on the DAPA-AP than participants who decoded five or fewer nonwords ($\chi^2(1, N = 17) = 4.90, p = .03$).

Construct Validity

Pearson correlation coefficients were calculated to establish the concurrent validity of the DAPA-AP total score with other known measures of phonological awareness. Likewise, correlations were calculated to establish the convergent validity with the reading measures (see Table 4). Results indicated that, in spite of the small sample size, the DAPA-AP total score was significantly and strongly correlated with letter–sound knowledge, sound matching, blending words, and word identification; and the correlation between DAPA-AP total score and elision approached conventional levels of statistical significance ($p = .06$). Moreover, the bootstrapped 95% confidence intervals indicated large effect sizes for letter–sound knowledge, sound matching, blending words, and word identification; and medium to large effect sizes for elision and word attack (Cohen, 1988). The smaller effect sizes for elision and word attack were likely due to the restricted distributions and positive skew for both of these variables (see Table 3).

We recalculated these correlations excluding the six participants who scored more than five items correct on word attack to ensure that those participants' scores were not responsible for the large correlations we observed. As seen in Table 4, the same general pattern held, even after excluding these participants; the correlations between the DAPA-AP total score and the other measures changed very little, with the exception of rapid letter naming. The bootstrapped 95% confidence intervals did increase in range, but the lowest values -- for elision and blending words -- still indicated at least small effects. All other estimates indicated at least medium to large effects.

Discussion

The goal of this study was to evaluate the construct validity of a new assessment of phonemic awareness that has features that make it appropriate for individuals with severe

speech impairments. These features include a non-speech response mode, limited verbal instructions, dynamic feedback, and computerized administration. Previous studies have demonstrated strong correlations between measures of phonological awareness and reading skills (National Early Literacy Panel [NELP], 2008; NICHD, 2000). Our strategy was to demonstrate correlations between the DAPA-AP and conventional tests for phonological awareness (concurrent validity) and individual-word reading (convergent validity) in adults with intellectual disabilities who could speak. Results provided evidence that the DAPA-AP may be a reliable and valid assessment tool for adults with intellectual disabilities.

Participants with decoding skills of grade level 1.5 or higher were very likely to score at or near the ceiling level of 3 on all subtests, meaning that they never required a prompted block. Given that many of the printed syllable-pairs were nonwords, participants who scored a 3 for a pair demonstrated both phonemic awareness and the alphabetic principle. This finding is to be expected for an assessment of phonemic awareness via alphabetic principle, and thus important at this stage in the development of our assessment. From a predictive-assessment perspective, there was insufficient variability to differentiate among individuals with decoding skills above this level. This finding is consistent with other research that showed phonological awareness is no longer a strong predictor of reading skills after the second grade (Hogan, Catts, & Little, 2005)

For individuals at this high performance level, a brief test that sampled syllable-pairs from all four subtests would likely have produced the same high accuracy. In practice, these participants' decoding skills would provide evidence of their phonological awareness; however, the present study's goal of assessing the validity of the test for individuals with severe speech impairments, made the inclusion of participants who scored at the ceiling level important.

At the opposite end of the distribution, one participant (Devon) did not consistently achieve a score of at least one, that is, he did not master the printed-word identity-matching task by the end of the assessment. This outcome is important to future development of the DAPA-AP for two interrelated reasons: First, the discrimination of printed words is necessary for accurate performance of the assessment task, which involved selecting one of two printed words that differed by a single letter. Second, the identity-matching task played a critical role in the dynamic component of the assessment: it served to prompt, and possibly teach, the correct response. Thus, administering this phonological awareness assessment to participants without identity-matching skills would yield difficult-to-interpret results. To make the test more efficient, and also more specific to individual skills, individuals at this level could receive a dynamic assessment of printed-word matching, which, only if passed, could lead into the phonemic awareness assessment. Worth noting is that, as a component of reading, the discrimination of printed words is important in and of itself. Consequently, this information should be helpful even though it is not a direct indicator of phonemic awareness.

The spoken-syllable-to-print tests administered here will be most informative with regard to progress towards the development of phonemic awareness and the alphabetic principle when scores are in between 3 (the ceiling) and 2. A score of 2 on a syllable-pair indicated that the participant demonstrated high accuracy on non-prompted trials, but only after receiving a

block of prompted trials. This shows that the participant rapidly learned the relationship between the pair of spoken syllables and their print counterparts. A score of 2 on all syllable-pairs in a subtest suggests that the participant may have been on the threshold of understanding the alphabetic principle. This information highlights the potential value of the dynamic aspect of the assessment: instruction for such a participant would likely be quite different from instruction for a participant who did not learn from the prompted trials.

No method of assessing precursors to decoding is without its drawbacks. In assessing phonemic awareness using the alphabetic principle, our assessment goes beyond available assessments for precursors of decoding. Achieving the maximum score required knowledge of the specific phoneme-grapheme relations included in the test. It also required the discrimination of three-letter printed words that differed by a single letter, which may be difficult for nonreaders (Yoo & Saunders, in press). In contrast, over the last few decades, the primary focus of assessments of phonological awareness that don't require speech has been on using pictures as response choices or tasks that require yes/no responses. These methods do not require printed-word discrimination. However, both involve relatively complex instructions (which, for pictures, includes the picture names) and require participants to process two or three spoken words presented in succession. Participants may know different names for the pictures prior to the assessment, and it is unclear how this might affect the outcome. In contrast, we used syllables that were represented by print and, as a result, the test cannot produce meaningful conclusions about individuals who do not discriminate printed words. However, in its structural simplicity, our assessment task is arguably more amenable to both computerization and the dynamic component of the DAPA-AP, which includes feedback and prompts to allow test-takers to learn the task. The comparability and relative merits of these differing methods, and the guidelines for selecting the best method for an individual's current level of knowledge, will be explored in future research.

Limitations and Future Directions

This pilot study had several limitations. First, the sample size was small. This was likely responsible for the handful of correlations that approached, but did not reach, conventional levels of statistical significance. With a larger sample size, these moderate-to-strong correlations would most certainly reach conventional levels of statistical significance.

Second, participants in this study were adults rather than children. We chose not to recruit children because it was unclear how long the DAPA-AP would take to administer. We assumed that adults were less likely to be fatigued by longer testing sessions. The choice to recruit adults, however, would not have affected the correlations between the DAPA-AP and subscales of the CTOPP, which were the focus of this study, particularly considering that participants were not strong readers. Nonetheless, generalizability of these findings to children with severe speech impairments has yet to be demonstrated and remains a goal for future research.

Third, for some participants, accuracy in responding to the prompts increased over the course of the assessment (see Table 2). These participants (Clare, Josh, Charles, Larry, and Kory) initially failed some prompted blocks (did not match identical printed words), which

accounts for their rising scores across the onset, rime, and coda subtests. Ideally, in a dynamic assessment, the prompt should reliably demonstrate the correct response from the outset of the assessment, making it possible to learn from the prompt. Future studies should include identity-matching pretraining trials to ensure accurate responses to the prompt prior to initiating the spoken-word-to-print portion of the assessment. This would ensure that any increases shown across the course of the assessment were due to learning from the prompts.

There are also practical limitations in this pilot version. Although computerization can potentially make testing more efficient, efficiency was not realized in this version of the test. Each of four subtests required a minimum of 36 trials, with up to 3 times that number if prompted trials were included. Clearly the test will require substantial modification and further testing to reduce the number of trials. Computerization would facilitate the application of decision rules that could be applied on a trial-to-trial basis to incorporate prompts or to skip to more- or less-advanced items. For example, each subtest could be preceded by a brief pretest of the included syllables. Results would determine whether the participant receives the dynamic assessment for that subtest, or moves to the next subtest. Alternatively, further research might determine a benchmark test that, if passed, reliably predicts performance on the other subtests.

Future research should focus on establishing the construct validity of the DAPA-AP in children, in individuals with severe speech impairments, and in other populations who may have speech difficulties. The DAPA-AP may prove to be an appropriate assessment for anyone who may have difficulty with these assessments, including individuals with severe speech impairments, individuals with other types of speech impairments (e.g., severe stuttering), and individuals with developmental disabilities, regardless of age. Therefore, future studies should seek to establish the concurrent validity of the DAPA-AP with other non-speech measures of phonological awareness, such as the APAR (Iacono & Cupples, 2002), as well as more widely used assessments such as sound matching from the CTOPP.

A goal of this work is to develop an assessment that can be easily presented using a wide range of computer technologies, including laptops and tablets. The proliferation of low-cost, high-power computers in the classroom makes computerized assessment a viable and important development. Moreover, computerization has many advantages. First, in addition to presenting the spoken and printed words for each test trial, the ultimate version of a computerized assessment would automatically score the student's response to each item and summarize the results for a report. These features help maintain procedural fidelity, which may be especially challenging for tests in which nonwords, or individual phonemes, spoken by the examiner are incorporated in the test items (Wagner et al., 1999). An additional advantage is that the computer allows items to be presented at a rate of 10 to 11 trials per minute. Finally, although the present study used a touch screen, in principle, any input device that can be interfaced with a computer would be possible. For example, an individual with quadriplegia could use a mouth stick or selection via scanning.

In conclusion, reading and writing skills afford children the ability to communicate in generative ways through speech-generating devices via text input. To this end, there is a strong and unmet need for assessments of phonological awareness that do not require speech

responses. This study represents an important first step in developing dynamic assessment of phonemic awareness appropriate for individuals with severe speech impairments.

Acknowledgments

This article was written with the support of NIH grants T32 HD057844, P30 HD002528, and R01 HD048528, awarded to the University of Kansas. We extend a special thank you to Hugh W. Catts for his assistance in critiquing this manuscript.

References

- Adams, MJ. Beginning to read: Thinking and learning about print. Cambridge, MA: The MIT Press; 1990.
- Barker RM, Saunders KJ, Brady NC. Reading instruction for children who use AAC: Considerations in the pursuit of generalizable results. *Augmentative and Alternative Communication*. 2012; 28:160–170. [PubMed: 22946991]
- Beukelman, DR, Mirenda, P. *Augmentative and alternative communication: Management of severe communication disorders in children and adults*. 3rd ed. Baltimore: Paul H. Brookes; 2005.
- Boyer N, Ehri LC. Contribution of phonemic segmentation instruction with letters and articulation pictures to word reading and spelling in beginners. *Scientific Studies of Reading*. 2011; 15:440–470.
- Browder D, Gibbs S, Ahlgrim-Delzell L, Courtade GR, Mraz M, Flowers C. Literacy for students with severe developmental disabilities: What should we teach and what should we hope to achieve? *Remedial and Special Education*. 2009; 30:269–282.
- Byrne, B. *The foundation of literacy: The child's acquisition of the alphabetic principle*. East Sussex: Psychology Press Ltd; 1998.
- Byrne B, Fielding-Barnsley R. Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*. 1989; 81:313–321.
- Catts HW, Petscher Y, Schatschneider C, Bridges MS, Mendoza K. Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities*. 2009; 42:163–176. [PubMed: 19098274]
- Coaley, K. *An introduction to psychological assessment and psychometrics*. Thousand Oaks, CA: Sage; 2010.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Dube WV. Computer software for stimulus control research with Macintosh computers. *Experimental Analysis of Human Behavior Bulliten*. 1991; 9:28–30.
- Ehri LC, Nunes SR, Willows DM, Schuster BV, Yaghoub-Zadeh Z, Shanahan T. Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*. 2001; 36:250–287.
- Erickson KA. Literacy and persons with developmental disabilities: Why and how? Paper commissioned for the EFA Global Monitoring Report 2006, Literacy for Life. 2005
- Foley BE. The development of literacy in individuals with severe congenital speech and motor impairments. *Topics in Language Disorders*. 1993; 13:16–32.
- Foley BE, Pollatsek A. Phonological processing and reading abilities in adolescents and adults with severe congenital speech impairments. *Augmentative and Alternative Communication*. 1999; 15:156–173.
- Gillam SL, Fargo J, Foley B, Olszewski A. A nonverbal phoneme deletion task administered in a dynamic assessment format. *Journal of Communication Disorders*. 2011; 44:236–245. [PubMed: 21168145]
- Grigorenko EL, Sternberg RJ. Dynamic testing. *Psychological Bulletin*. 1998; 124:75–111.

- Hogan TP, Catts HW, Little TD. The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools*. 2005; 36:285–293.
- Iacono TA. Accessible reading intervention: A work in progress. *Augmentative and Alternative Communication*. 2004; 20:179–190.
- Iacono TA, Cupples L. Assessment of phonological awareness and reading (Version 1.15) [Assessment]. 2002
- Kirtley C, Bryant P, MacLean M, Bradley L. Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology*. 1989; 48:224–245. [PubMed: 2794855]
- Koppenhaver, DA, Yoder, D. Literacy issues in persons with severe physical and speech impairments. In: Gaylord-Ros, R, editor. *Issues and research in special education*. Vol. 2. New York: Teachers College; 1992. 156–201.
- Lidz, CS. *Practitioner's guide to dynamic assessment*. New York: Guilford Press; 1991.
- National Early Literacy Panel. *Developing early literacy: Report of the National Early Literacy Panel*. 2008. Retrieved from <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- National Institute of Child Health and Human Development. *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office; 2000.
- Snow, CE, Burns, MS, Griffin, P, editors. *Preventing reading difficulties in young children*. Washington, D.C: National Academic Press; 1998.
- Speece, D; Case, LP; Molloy, DE. Curriculum based measurement for early literacy [Assessment]. 2011. Retrieved from <http://terpconnect.umd.edu/~dlspeece/cbmreading/index.html>
- Sternberg, RJ, Grigorenko, EL. *Dynamic assessment: The nature and measurement of learning potential*. Cambridge: Cambridge University Press; 1998.
- Vandervelden MC, Siegel LS. Phonological processing in written word learning: Assessment for children who use augmentative and alternative communication. *Augmentative and Alternative Communication*. 2001; 17:37–51.
- Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*. 1987; 101:192–212.
- Wagner RK, Torgesen JK, Rashotte CA. Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*. 1994; 30:73–87.
- Wagner, RK, Torgesen, JK, Rashotte, CA. *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed, Inc; 1999.
- Whitehurst GJ, Lonigan CJ. Child development and emergent literacy. *Child Development*. 1998; 69:848–872. [PubMed: 9680688]
- Woodcock, RW. *Woodcock reading mastery test-revised*. Circle Pines, MN: American Guidance Service; 1998.
- Yoo JH, Saunders KJ. The discrimination of printed words by prereading children. *European Journal of Behavior Analysis*.
- Yorke M. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*. 2003; 45:477–501.

Figure 1.
Comparison of non-prompted and prompted trials for the printed syllable-pair mib/sib.

Figure 2.

Example of scoring outcomes for the printed syllable-pair mib/sib. Blocks with a + indicate criterion was met; blocks with a - indicate that criterion was not met. Criterion was defined as at least 5 of 6 correct.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Syllable Word Pairs within each DAPA-AP Subtest

Onset	Rime	Coda	Vowel
mat/sat	kog/kib	mot/mog	kog/kag
mib/sib	sog/sib	sot/sog	sog/sag
mob/sob	nog/nib	bot/bog	nog/nag
mup/sup	tog/tib	rot/rog	tog/tag
min/sin	mog/mib	tep/tek	mog/mag
med/seed	pog/pib	rep/rek	pog/pag
—	—	nep/nek	—
—	—	fep/fek	—
—	—	jep/jek	—

Table 2

Individual Scores on DAPA-AP

Participant	Pretraining	Onset	Rime	Coda	Vowel	Total
Kevin	No	3.00	3.00	3.00	3.00	12.00
Zac	No	3.00	3.00	3.00	3.00	12.00
Ethan	No	3.00	3.00	3.00	3.00	12.00
Clare	No	1.00	1.17	2.11	0.50	4.78
Josh	No	1.33	2.67	2.89	1.83	8.72
Charles	Yes	2.67	3.00	3.00	3.00	11.67
Ed	Yes	0.83	0.50	0.22	0.17	1.72
James	Yes	1.67	1.67	1.67	1.17	6.17
Evan	Yes	3.00	3.00	3.00	2.17	11.17
Robert	Yes	3.00	3.00	2.89	3.00	11.89
Tucker	Yes	3.00	3.00	3.00	2.17	11.17
Larry	Yes	2.17	3.00	2.44	2.17	9.78
Roy	Yes	2.67	2.33	2.22	1.67	8.89
Matthew	Yes	3.00	3.00	3.00	3.00	12.00
Vicky	Yes	3.00	3.00	3.00	3.00	12.00
Kory	Yes	2.67	3.00	3.00	3.00	11.67
Devon	Yes	0.00	0.00	0.33	0.67	1.00

Note. Maximum score for each subtest was 3; minimum was 0. Maximum total score was 12, minimum was 0.

Table 3

Central Tendency for the CTOPP, Woodcock, and DAPA-AP

Assessment	Mean	SD	Median	Skew
Letter-sound knowledge	14.06 [9.65, 18.53]	9.74 [6.79, 11.25]	16.00 [5.01, 22.00]	-0.44 [-1.31, 0.40]
CTOPP				
Sound matching	10.18 [7.47, 13.00]	5.90 [4.27, 6.99]	11.00 [6.00, 15.00]	-0.14 [-0.94, 0.61]
Elision	2.53 [1.18, 4.39]	3.48 [1.52, 4.93]	1.00 [0, 3.00]	1.95 [.35, 2.96]
Blending words	6.47 [4.43, 8.94]	5.01 [3.04, 6.34]	6.00 [3.50, 8.00]	0.70 [-0.50, 1.43]
Rapid letter naming ^a	74.62 [61.50, 87.96]	28.52 [20.82, 33.60]	69.12 [51.50, 98.00]	0.23 [-0.62, 1.16]
Woodcock Johnson				
Word identification	24.76 [14.56, 36.62]	23.27 [15.50, 28.06]	20.00 [6.50, 47.00]	0.69 [-0.18, 1.69]
Word attack	6.12 [2.06, 11.76]	10.39 [3.89, 14.00]	0.00 [0, 7.00]	1.85 [0.61, 3.61]
DAPA-AP				
Onset	2.29 [1.85, 2.71]	0.96 [0.52, 1.21]	2.67 [1.67, 3.00]	-1.24 [-2.63, -0.24]
Rime	2.43 [1.96, 2.84]	0.98 [0.37, 1.27]	3.00 [2.33, 3.00]	-1.64 [-3.50, -0.50]
Coda	2.46 [2.00, 2.84]	0.91 [0.32, 1.23]	3.00 [2.22, 3.00]	-1.84 [-3.35, -0.64]
Vowel	2.15 [1.70, 2.59]	1.00 [0.61, 1.21]	2.17 [1.67, 3.00]	-0.83 [-1.79, 0.09]
DAPA-AP total	9.33 [7.57, 10.90]	3.70 [1.70, 4.66]	11.17 [8.72, 12.00]	-1.40 [-2.75, -0.45]

Note. Numbers in brackets are bootstrapped 95% confidence intervals.

^aCalculations based on $n = 15$ because of missing data.

Table 4

Pearson Correlations with DAPA-AP Total Score

Statistic	LSK	SM	Elision	BW	RLN ^a	WID	WA
Full sample (<i>n</i> = 17)							
<i>r</i>	.80	.79	.47	.63	-.46	.66	.44
<i>p</i>	.00	.00	.06	.01	.09	.00	.08
95% CI	[.64, .94]	[.56, .91]	[.31, .68]	[.45, .81]	[-.82, -.03]	[.51, .82]	[.30, .62]
> 5 on WA excluded (<i>n</i> = 11)							
<i>r</i>	.73	.74	.46	.57	-.27	.68	.47
<i>p</i>	.01	.01	.16	.07	.46	.02	.15
95% CI	[.31, .95]	[.38, .97]	[.13, .79]	[.12, .89]	[-.96, .43]	[.45, .90]	[.27, .82]

Note. LSK = letter sound knowledge; SM = sound matching; BW = blending words; RLN = rapid letter naming; WID = word identification; WA = word attack.

^aCalculations are based on *n* = 15 for the full sample and *n* = 9 for > 5 on WA excluded.