Using Standardized Tests to Inventory Consonant and Vowel Production: A Comparison of 11 Tests of Articulation and Phonology

Sarita L. Eisenberg
Elaine R. Hitchcock
Montclair State University, Montclair, NJ

Inventory analysis is recommended as a standard part of a comprehensive phonological assessment in textbooks and reference guides on articulation and phonological disorders (e.g., Bauman-Waengler, 2004; Bleile, 2004; Creaghead, Newman, & Secord, 1989; Pena-Brooks & Hedge, 2007; Smit, 2004; Velleman, 2003; Williams, 2003). In a survey by Skahan, Watson, and Lof (2007), 80% of speech-language pathologists (SLP) reported using phonetic inventory analysis at least some of the time, with 31% reporting always including phonetic inventory analysis as part of their phonological assessments. The inventory analysis most typically focuses on consonants, either combining productions across word positions or looking separately at production by word position. Although less common, a vowel inventory analysis can also be done.

Establishing a phonetic inventory is important because this inventory can be used as the basis for determining the developmental appropriateness of a child's sound production (Bleile, 2004), identifying targets for stimulability testing (Powell & Miccio, 1996), and/or selecting goals (Davis, 2005; Gierut, 2005) or treatment targets within a phonological pattern (Hodson & Paden, 1991). SLPs who have already given a norm-referenced standardized test to identify a child with an articulation and phonological disorder (henceforth termed *speech sound disorder* and abbreviated as SSD) may want to use the already available data from that test to carry out a phonetic inventory analysis. Of the 11 tests considered in this report (see Appendix A for details), some include an inventory analysis of the sounds produced by the child (see Table 1). The current report considers the validity of using the single-word production data from these tests for the purpose of compiling a phonetic inventory. We will start by explaining what is meant by a phonetic inventory.

A phonetic inventory[1] is part of an independent analysis of a child's phonological system, revealing the sounds produced by the

[1]As noted by Shriberg and his colleagues (Shriberg & Kwiatkowski, 1982; Shriberg et al., 1997), the term *phonetic inventory* is not really an accurate one because this type of inventory is based on broad transcription and counts any recognizable production of a particular speech sound, including distortions, toward the inventory. They suggest, instead, merely categorizing sounds as *in the inventory* or *out of the inventory.* We have chosen to retain the term phonetic inventory in this report because it is commonly used (e.g., Elbert & Gierut, 1986) and serves to distinguish this inventory type from phonemic and other inventories.

ABSTRACT: **Purpose:** This report considered the validity of making conclusions about a child's phonetic inventory (the sounds a child can and cannot produce spontaneously without a prior model or other stimulation) based on the data from standardized single-word tests of articulation or phonology.
Method: We evaluated the opportunities for production of word-initial consonants, word-final consonants, and vowels within the words included on 11 tests. Only words that met specific phonetic criteria (termed phonetically controlled words) were counted as opportunities for each consonant or vowel.
Results: None of the tests provided sufficient coverage of consonants or vowels for establishing a phonetic inventory and making conclusions about the segments that a child can and cannot produce.
Conclusion: Use of the data from a single standardized test of articulation or phonology would not be sufficient for completely inventorying a child's consonant and vowel production and selecting targets for therapy. It is recommended that clinicians supplement test data by probing production in additional phonetically controlled words.

KEY WORDS: articulation assessment, phonological assessment, phonetic inventory, standardized tests

Table 1. Single-word tests of articulation and phonology.

Name of test	*Norm-referenced*	*Consonant positions tested*[a]	*Vowels tested*	*Sample size (# of words)*	*Inventory chart*
Arizona Articulation Proficiency Scale, 3rd ed. (AAPS–3; Fudala, 2000)	Yes	IF	Yes	46	No
Bankson-Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990)	Yes	IF	No	81	Yes
Clinical Assessment of Articulation and Phonology (CAAP; Secord & Donohoe, 2002)	Yes	IF	No	44	Yes
Diagnostic Evaluation of Articulation and Phonology: Articulation Assessment (DEAP; Dodd, Hua, Crosbie, Holm, & Ozanne, 2006)	Yes	IF	Yes	30	No
Fisher-Logemann Test of Articulation (FLTA; Fisher & Logemann, 1971)	No	IMF	Yes	104	No
Goldman-Fristoe Test of Articulation, 2nd ed. (GFTA–2; Goldman & Fristoe, 2000); Khan-Lewis Phonological Analysis, 2nd ed. (KLPA–2; Khan & Lewis, 2002)[b]	Yes	IMF	No	53	Yes[b]
Hodson Assessment of Phonological Patterns, 3rd ed. (HAPP–3; Hodson, 2004)	Yes	NA	Yes	50	Yes
Photo-Articulation Test, 3rd ed. (PAT–3; Lippke, Dickey, Selmar, & Soder, 1997)	Yes	IMF	Yes	77	No
Smit-Hand Articulation and Phonology Evaluation (SHAPE; Smit & Hand, 1992)	Yes	IF	No	81	Yes
Structured Photographic Articulation Test, 2nd ed. (SPAT–II; Dawson & Tattersall, 2001)	Yes	IMF	No	45	Yes
Templin-Darley Tests of Articulation (TDTA; Templin & Darley, 1968)	Yes	IMF	Yes	141	No

[a]I = initial, M = medial, F = final. [b]The KLPA–2 uses the words from the GFTA–2. These two tests are, therefore, considered together. The phonetic inventory chart is included on the KLPA–2 and not on the GFTA–2.

child regardless of correctness relative to the target (Elbert & Gierut, 1986; Powell & Miccio, 1996) and providing the set of sounds that are available to the child for forming words (Stokes, Klee, Carson, & Carson, 2005). Phonetic inventory analysis is used to identify both the sounds that a child produces and the sounds that a child does not produce. Only sounds that a child produces spontaneously are attributed to the child's phonetic inventory; sounds that are produced only in response to auditory and/or visual stimulation are not considered to be in the child's phonetic inventory (Powell & Miccio, 1996). A typical criteria for crediting a sound to a child's phonetic inventory is two productions of that sound (Powell & Miccio, 1996; Stoel-Gammon, 1985). Sounds produced only once are considered marginal (Grunwell, 1985). Sounds that are not produced are considered to be missing from the child's phonetic inventory (Powell & Miccio, 1996).

The phonetic inventory is distinct from two other independent analyses—a phonemic inventory and a stimulability inventory. *Phonemic inventory analysis* aims to establish not just the sounds that are produced by a child, but the sounds that are used contrastively (Elbert & Gierut, 1986) and that enable the child to make distinctions between words (Stokes et al., 2005). *Stimulability assessment* aims to determine the additional sounds not attributed to the phonetic inventory that the child can produce in a supportive condition, most typically in response to an auditory model (Glaspey & Stoel-Gammon, 2005; Lof, 1996; Powell & Miccio, 1996).

These independent analyses can be distinguished from relational analyses that compare a child's production to an adult target. Model/replica or contrastive analysis (Grunwell, 1985; Klein, 1984) and phonological process analysis (Grunwell, 1981; Ingram, 1989) are examples of relational analyses that show a child's pattern of errors relative to an adult target. Also part of a relational analysis is a determination of the consistency of production relative to the target (Ingram & Ingram, 2001). Some researchers include an accuracy criterion for attributing sounds to a child's inventory (Ingram & Ingram, 2001; Stokes et al., 2005). That is, they add a relational component to the inventory analysis (cf. Stoel-Gammon & Dunn, 1985). Ingram and Ingram (2001), for instance, suggested a criteria of two correct productions or a 50% accuracy rate for including a sound in the inventory. A more stringent criteria would be to require that sounds be consistently produced relative to the target in order to establish an inventory of mastered sounds (Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997).

The data needed for carrying out each of these independent and relational analyses are different. In this report, we consider the nature of the stimuli needed for establishing a phonetic inventory—the sounds produced by a child without stimulation and regardless of correctness relative to the target—and whether these data are provided on 11 standardized single-word tests.

Why Tests Are Used

The use of norm-referenced, standardized, single-word tests is a typical part of an evaluation for SSD (Skahan et al., 2007). Although single-word production testing may not provide as complete a picture of a child's speech sound production as connected speech sampling does (Andrews & Fey, 1986; Bernhardt & Holdgrafer, 2001a; Morrison & Shriberg, 1992), there are several disadvantages of connected speech sampling that might lead SLPs to rely on the data from single-word tests instead, not only for identifying children as having SSD, but also for analyzing production and determining goals (e.g., Hodson, Scherz, & Strattman, 2002; Khan, 2002; Tyler & Tolbert, 2002).

Transcription and analysis of a connected speech sample requires more time than may be available in most clinical settings (Bleile, 2002). Hence, in a recent clinical forum on phonological assessment in the *American Journal of Speech-Language Pathology,* several of the authors recommended transcribing only the single-word sample from a standardized test and not a conversational sample in the interest of making the assessment more time efficient (Hodson et al., 2002; Khan, 2002; Tyler & Tolbert, 2002). Some children may be reluctant to talk and so it may also be time consuming or difficult to obtain a sufficient sample for analysis (Smit, 2004; Wolk & Meisler, 1998). Transcription of conversational speech samples may not be possible for highly unintelligible children because the target words may not be identifiable (Grunwell, 1985; Paden & Moss, 1985; Shriberg et al., 1997; Stoel-Gammon & Dunn, 1985). In addition, because children choose the words they produce in their spontaneous speech and may avoid difficult words, the range of sounds and word shapes produced, or even attempted, may be restricted (Ingram, 1989; Stoel-Gammon & Dunn, 1985; Williams, 2003).

Single-word standardized tests have been designed to provide a time-efficient means to elicit a diverse speech sample that includes opportunities for production of all consonants and, in some tests, for production of vowels. An additional advantage of single-word tests is that the SLP knows what the child is trying to say. This can facilitate transcription and also enable the SLP to compare the child's productions to adult targets (Bernhardt & Holdgrafer, 2001a; Grunwell, 1985; Hodson et al., 2002; Velleman, 1998).

One of the most compelling reasons for SLPs to use single-word standardized tests is the availability of norms for identifying SSD (Skahan et al., 2007). This is important because quantitative test results may be required by school districts to qualify for services or may be required by third-party payers to receive coverage for SLP services (Klein, 1996; Tyler & Tolbert, 2002). However, norm-referenced standardized tests that have been designed for the purpose of identifying SSD might not be appropriate for inventorying production and selecting goals (cf. Plante, 1996). McCauley and Swisher (1984) noted that standardized test data can lead to faulty conclusions about both areas of deficit (e.g., missing sounds) and achieved skills (e.g., known sounds). This is because selection of test content—that is, the type and range of words—for a test designed to differentiate between children with and without SSD may result in a set of words that is unsuitable for determining which sounds a child is and is not capable of producing. In addition, tests designed only for assessing consonant production might not be appropriate for assessing vowel production.

Most single-word tests have been developed to provide a representative sample of English consonants. In traditional articulation tests, such as the Goldman-Fristoe Test of Articulation, 2nd edition (GFTA–2; Goldman & Fristoe, 2000) and the Photo-Articulation Test, 3rd edition (PAT–3, Lippke, Dickey, Selmar, & Soder, 1997), production of each sound is targeted and is recorded once for each position within a single word. However, just looking at a single targeted production may not be sufficient to reflect the variability of production that is characteristic of SSD (James, 2001). To maximize the information from the test, Klein (1984) recommended transcribing whole words rather than just the target sounds. Some older tests, such as the GFTA–2, have been revised to allow this option, and more recent phonological tests, such as the Bankson-Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990) and the Clinical Assessment of Articulation and Phonology (CAAP; Secord & Donohoe, 2002), record all consonants in every word and include all attempts at each sound. However, this does not ensure that there will be more than one opportunity for every sound.

Under the assumption that vowels are fully acquired by age 3;0 (years;months; Donegan, 2002), which is the lower age limit for most tests, single-word tests have focused mostly on consonants. This high degree of accuracy for vowel production was confirmed by Dodd, Holm, Hua, and Crosbie (2003), who reported > 97% accuracy in vowel production for 3-year-old children. However, some children with SSD do produce vowel errors (Pollock, 1991; Pollock & Berni, 2003; Pollock & Keiser, 1990; Stoel-Gammon & Herrington, 1990). In a study by Pollock and Berni (2003), for example, 11% of children with phonological disorder demonstrated vowel production accuracies of < 85%, and > 40% of the children with the most severe phonological impairment produced vowel errors. It is, therefore, important to have tests that can be used to assess vowel production. Tests designed to assess vowels typically target each vowel in one word (Pollock, 1991). SLPs can obtain information about vowel production by transcribing whole words rather than just the targeted consonants, but this does not ensure that all vowels will be sampled or that each vowel will be sampled more than once to capture any contextual variability in production.

Test Content

In this report, we consider whether the words used in standardized single-word tests are appropriate and sufficient for compiling a phonetic inventory of consonants and vowels. This is a content validity issue. Content validity delimits the trait being assessed and how that trait will be measured (Sabers, 1996). Content validity involves specification of a test's domain (referred to as content relevance) and the degree of representativeness with which a test samples that domain (referred to as content coverage; Lieberman & Michael, 1986). As for all aspects of validity, content relevance and coverage must be determined for each purpose. The trait we investigated in this report is a phonetic inventory—the set of consonants and vowels that a child is capable of producing without stimulation and regardless of correctness. For consonants, we focused on positional inventories for word-initial and word-final positions. For vowels, we focused only on the vowels that can occur in stressed syllables.[2]

We would like to be able make conclusions about which sounds a child can and cannot produce. In order to conclude with any certainty that a nonproduced sound is lacking from a child's inventory, the child must have sufficient opportunities to produce that sound, and the phonetic makeup of the words used to assess each sound must be controlled. Words with difficult phonetic contexts that might prevent a child from producing a consonant or vowel that the child is able to produce in simpler words would not be representative of the child's phonetic inventory. Words that scaffold production of a consonant or vowel that would not otherwise be produced also would not be representative of the child's phonetic inventory. Such contexts that affect the child's typical production of these segments represent measurement confounds (cf. Sabers, 1996).

Content validity must be evaluated for each type of use (Messick, 1989; Sabers, 1996). A test with appropriate content relevance and coverage for identifying SSD may not have appropriate content for

[2] The term *vowel* has been used throughout this report to refer collectively to monophthong vowels and diphthongs.

completing a phonetic inventory. We recognized that busy clinicians may want to use the same test data for identification of SSD and for completing a phonetic inventory analysis. Therefore, we investigated whether the speech sample data from these standardized single-word tests would provide a set of words that are phonetically controlled in a way that is appropriate for both establishing the consonants and vowels that are in a child's phonetic inventory and making conclusions about which sounds are lacking from the inventory.

Sample Size Considerations

For conversational speech samples, it has been suggested that a sample size of at least 100 words is necessary in order to reveal any variability in production as well as to inventory production of the majority of adult targets (Bernhardt & Holdgrafer, 2001a; Crystal, 1982; Grunwell, 1985). Crystal (1982) suggested that a much larger sample of 250 words would be preferable. Weston, Shriberg, and Miller (1989) suggested that a sample of either 90 different words or 225 total words, whichever occurs first, would be sufficient to provide a representative sample for phonological analysis. The sample sizes of norm-referenced single-word tests used for identifying SSD are mostly smaller than these recommended sample sizes for conversation, including fewer than 90 to 100 words (see Table 1). Researchers have questioned the sufficiency of these tests for other purposes, such as profiling a child's production and determining goals, and have recommended that larger samples be obtained (Bernhardt & Holdgrafer, 2001a; Miccio, 2002).

A number of authors have suggested the need to test sounds in more than one word because the effect of surrounding phonemes can influence a child's ability to produce the target sound (e.g., Elbert & Gierut, 1986; Grunwell, 1985; Ingram, 1989; Pollock, 1991; Weiner, 1979). Currently, we do not have research evidence on what constitutes a sufficient number of opportunities for concluding that a child does or does not produce a given sound. In looking at the content coverage of single-word tests, we wanted to balance comprehensiveness with the reality of the known time constraints of clinical practice.

Normative studies of single-word production have used sound production in two (Smit, Hand, Freilinger, Bernthal, & Bird, 1990) or three word positions (e.g., Templin, 1957) as evidence for crediting a child with a sound. However, this does not allow children who have acquired a sound but limit its occurrence across word positions to be credited with that sound. We considered, instead, the criteria used in studies of conversational speech of two exemplars in a given word position for attributing a sound to a child's inventory (Ingram & Ingram, 2001; Stoel-Gammon, 1987).

Based on these criteria from conversational speech studies, we recommend that sounds be attributed to a child's phonetic inventory only if they are produced at least twice. For consonants, the criteria would be for two productions in a single word position. In order to achieve this criterion, a child would need to be given at least two opportunities to produce each sound. However, consonant and vowel inventories are based on the sounds that are produced by a child and will, therefore, be affected by the words and sounds that are attempted by that child. In the following sections, we review variables that could affect the sounds produced by a child and, based on this review, we make suggestions for phonetically controlling the characteristics of the stimulus words used for establishing a phonetic inventory.

Considerations for Inventorying Consonants

Shriberg and Kwiatkowski (1980) noted inconsistencies in the word shapes in which consonants were tested on single-word tests. Specifically, these authors observed that word-initial fricatives, affricates, and liquids were less often tested in simple word shapes than were nasals and stops. Such variability in phonetic context is important because differences in phonological complexity can affect sound production (Branigan, 1976; Ingram, Christensen, Veach, & Webster, 1980; Klein, 1982). In this section, we review aspects of phonetic context that could affect consonant production.

Syllable stress. Ingram et al. (1980) reported that word-initial consonants were more prone to error when they occurred in unstressed syllables than in stressed syllables. Similarly, Klein (1982) observed fewer errors on medial /r/ when it started a stressed syllable (as in the word *giRAFFE*) than when it started an unstressed syllable (as in the word *CArrot*). Bernhardt and Stemberger (1998, cited in Bernhardt & Holdgrafer, 2001a) reported the same pattern for final consonants. For instance, these authors observed children who correctly produced the final consonant in the word *hat* but omitted the same final sound in the word *BUcket*. In addition, children may omit initial unstressed syllables (Ingram, 1989), thus reducing opportunities for attempts at word-initial consonants in words with initial unstressed syllables. Therefore, unstressed syllables would not be good opportunities for testing consonant production. This would rule out words such as *pajamas* or *balloons* for assessing production of word-initial consonants and words such as *wagon* or *carrot* for inventorying word-final production.[3]

Number of syllables. Ingram et al. (1980) observed a later acquisition age for /v/ reported by Templin (1957), who tested production of /v/ in the words *vacuum cleaner* and *valentine,* than reported by Wellman, Case, Mengert, and Bradbury (1931), who tested production of /v/ in the words *vase* and *vest.* Ingram et al. investigated the effect of number of syllables on the production of word-initial fricatives and affricates and confirmed that there was a higher accuracy rate for production of those sounds in monosyllabic words than in words of two or more syllables. Word-final consonants have also been reported to be produced earlier and more accurately in monosyllabic words than in longer words (Branigan, 1976). Multisyllabic words, such as *butterfly* or *television,* would therefore not be good opportunities for inventorying word-initial consonant production even though the initial syllable is stressed. Vance, Stackhouse, and Wells (2005), however, found less of a difference between one- and two-syllable words than between two-syllable words and longer words of three or four syllables. Bisyllabic words would, therefore, be more acceptable contexts for inventorying production as long as the target sound occurred in the stressed syllable. This includes words such as *candy* and *zipper* for word-initial consonants and *balloon* and *garage* for word-final consonants.

Cluster contexts. Children are likely to omit the more marked consonant (characterized by Velleman, 1998, as consonants that are less common and/or that involve less common characteristics) in clusters than when that consonant occurs as a singleton (Smit, 1993). Omissions from clusters may reflect difficulty with producing the consonant sequence rather than a problem with the consonant itself. Alternatively, clusters can serve as facilitating contexts for sounds

[3]These words and the words used throughout this report to illustrate phonetic characteristics have been drawn from the tests of articulation and phonology included in the current report.

that are not produced as singletons, with some children producing a particular consonant only within a highly co-articulated cluster context. This has been reported for /s/ (Hodson, 1997; Kent, 1982), for /r/ (Kent, 1982), and for velars (Bleile, 2004). Cluster contexts could, therefore, lead to underestimation or overestimation of consonant production and so would not be appropriate contexts for determining whether or not a consonant is in a child's phonetic inventory.

Morpheme structure. Morpheme structure can also affect production because SSD often co-occurs with language impairment (Shriberg & Kwiatkowski, 1994). For children with morphosyntactic deficits, omission of a final consonant in plural nouns, such as *keys,* or past-tense verbs, such as *locked,* may relate to morpheme omissions rather than to omission of the sound per se. In addition, the present participle *–ing,* in words such as *swimming,* is often produced colloquially as [ɪn] and may not represent production in monomorphemic contexts such as *ring*. Production in bimorphemic contexts could, therefore, lead to underestimation of consonant production and so would not be appropriate contexts for inventorying consonant production.

Harmonic and assimilation contexts. Final consonants may first develop in harmonic contexts (Velleman, 2003). Assessing final consonants in words with the same first and last consonant might lead to a higher rate of final consonant production. Therefore, it is important that final consonants be tested in words with different initial consonants. Although we found no reports of a similar influence on word-initial consonant production, we suggest that words used for inventorying word-initial as well as word-final consonant production be nonharmonic. Words such as *baby* or *cake* would, therefore, not be appropriate for inventorying consonant production.

Assimilation is commonly seen in children with SSD (Hodson & Paden, 1981), and this may result in production of sounds that children might not otherwise produce. Assimilation can affect place or manner. For instance, assimilatory place changes can be either alveolar-to-velar or velar-to-alveolar, and assimilatory manner changes can be either liquid-to-glide or glide-to-liquid. These assimilations may result in harmonic productions such as producing *dog* as either [gɔg] or [dɔd] or *yellow* as either [jɛjou] or [lɛlou]. Given the highly individual nature of assimilation and the large range of potential assimilation patterns across children, it is not possible to eliminate all words that might lead to assimilation. Rather, the occurrence of assimilation can be taken into account in interpreting inventory results.

Vowel–consonant associations. Vowels and consonants can influence each other in production (Pollock, 1994). A number of researchers have reported associations between particular vowels and consonants in early speech (Camarata & Gandour, 1984; Grunwell, 1981; Leonard, Devescovi & Ossella, 1987; Stoel-Gammon & Dunn, 1985; Wolfe & Blocker, 1990). Stoel-Gammon and Dunn (1985), for instance, reported a co-occurrence of alveolar consonants with high front vowels. For typically developing children around age 18 to 24 months, as the child's vocabulary increases, the child's phonology undergoes a radical change from a holistic representation of words to a phoneme-based system (Vihman, 1996; Walley, 1993). An essential production characteristic that demonstrates this change is that consonants and vowels combine freely. It is, therefore, important to sample each consonant with more than one vowel in order to establish that consonant as an independent segment in a child's phonetic inventory.

Considerations for Inventorying Vowels

Vowels can also be influenced by phonetic context. Because of this, Pollock (1991) suggested that vowels be tested with more than one postvocalic consonant. To eliminate the effect of a following consonant, tense vowels could also be tested in open syllables (Pollock, 1991). Pollock and Berni (2003) eliminated postvocalic liquid contexts, such as *chair* and *school,* for testing vowels because these consonants result in changed vowel quality. Because the primary context effect is from the following consonant (Pollock, 1994), it does not seem to be necessary to control for the prevocalic consonant or to test vowels in word-initial position (that is, without a preceding consonant), as was done by Pollock and Keiser (1990) when inventorying vowel production.

Syllable stress. Syllable stress has been shown to affect vowel production. Vowels in unstressed syllables tend to be neutralized, and so unstressed syllables would not provide good opportunities to assess vowel production. Unstressed syllables preceding a strong syllable may be deleted (Ingram, 1989) and so would reduce opportunities to produce vowels in initial unstressed syllables. This would rule out unstressed syllables in words such as *behind* or *crayons* for assessing vowels.

Number of syllables. James, van Doorn, and McLeod (2001) investigated the accuracy of vowel production as a function of number of syllables. Accuracy of vowel production in three-syllable words was significantly less than accuracy of production in monosyllabic and bisyllabic words. Multisyllabic words, such as *television* or *potatoes,* would, therefore, not be the best contexts for determining the vowels that a child is able to produce.

Suggested Characteristics for Words Used to Establish a Phonetic Inventory

In this section, we make suggestions concerning the stimuli that would be appropriate for eliciting consonant or vowel production for the purpose of determining a phonetic inventory. These suggestions take into account both content relevance and content coverage. For content relevance, we suggest controlling word characteristics to avoid both scaffolding production of sounds a child could not otherwise produce and preventing production of sounds the child is capable of producing. For content coverage, we suggest providing more than one opportunity for production and providing opportunities in more than one context so that there is sufficient evidence that the child has achieved production of the sound as an independent segment.

For consonants, we concluded that phonetically controlled words should provide a nonharmonic singleton context. We further concluded that the words should either be monosyllabic or include the target sound in the stressed syllable of a bisyllabic word. Because our interest was in the adequacy of tests for determining the consonants that a child can produce, we asked whether each consonant was included in at least two such words in both word-initial and word-final positions. To establish a consonant as a separate phoneme and to rule out holistic productions with a single vowel, we further considered whether word-initial consonants were followed by two different vowels and whether word-final consonants were preceded by two different vowels.

For word-initial consonants, we accepted CVCC words (that is, words with final clusters), and for word-final consonants, we accepted CCVC words (that is, words with initial clusters) because we did not find any reports that the presence of a cluster elsewhere in a word decreases the accuracy rate for singleton consonants. This decision allows a wider range of words to be used as stimuli. However, for some children, the added complexity of CVCC and CCVC word shapes might negatively impact their production of individual segments within a word.

For assessing vowels, we concluded that phonetically controlled words should be a monosyllabic word in which the vowel either occurred in an open syllable or was followed by a consonant other than a liquid or occurred in the stressed syllable of a bisyllabic word. Because our interest was in the adequacy of tests for determining the vowels that a child can produce, we asked whether each vowel was included in at least two such words. We accepted words with initial and final consonant clusters because we did not find any reports that the presence of a consonant cluster decreases the accuracy rate for vowels.

Purpose

In this report, we evaluated the opportunities for production of the word-initial and word-final singleton consonants and vowels of General American English provided by 11 standardized tests of articulation and phonology. Because the purpose was to determine if test content was sufficient for determining a child's phonetic inventory, we looked for opportunities in phonetically controlled contexts that would neither confound nor scaffold the child's ability to produce each sound. We evaluated all occurrences of each consonant and vowel rather than looking only at the specific target words designated by the test for each segment. We checked for content coverage, first whether or not each test included all consonants and/or vowels at least once and then for inclusion in two words. We also checked for content relevance, that is, the inclusion of consonants and vowels in phonetically controlled contexts. To check for coverage, we asked the following specific questions:

- Do the tests include one phonetically controlled word for testing each word-initial consonant and each word-final consonant?
- Do the tests include at least two phonetically controlled words with different neighboring vowels for testing all word-initial and word-final consonants?
- Do the tests include one phonetically controlled word for testing each vowel?
- Do the tests include at least two phonetically controlled words with different neighboring consonants for testing all vowels?

We then considered whether the data from these tests were sufficient and valid for making conclusions about a child's phonetic inventory.

METHOD

Eleven tests, listed in Table 1, were included in this report. We included 10 tests that provide norms for identifying children with SSD because we wanted to know whether such tests also provide sufficient data for inventorying a child's phonetic inventory. We included an additional test, the Fisher-Logemann Test of Articulation (FLTA; Fisher & Logemann, 1971), although this test does not provide norm-referenced data, because we wanted to include an additional test that had been developed to assess vowels. The FLTA and the Templin-Darley Tests of Articulation (TDTA; Templin & Darley, 1968) are older tests that have not been recently revised. We included these tests because they were both reported to be used by SLPs in the Skahan et al. (2007) survey, and they both include a larger number of words than the other tests.

There were some other tests in the Skahan et al. (2007) survey that we did not include in this report. For example, we did not include the Assessment Link Between Phonology and Language (ALPHA; Lowe, 1986, cited in Skahan et al., 2007) because the responses on this test are imitative and we were looking for words that the child can produce on his or her own. We included only the Hodson Assessment of Phonological Patterns, 3rd edition (HAPP–3; Hodson, 2004). Consistent with the view that only the most recent version of a test should be used (Jakubowitz & Schill, 2008), we did not include an earlier version of this test, the Assessment of Phonological Processes—Revised (APP–R, Hodson, 1986, cited in Skahan et al., 2007). The Khan-Lewis Phonological Analysis, 2nd edition (KLPA–2, Khan & Lewis, 2002) was combined with the GFTA–2 because it uses the same words as that test.

We also did not include two other recently published tests in this report. Although it is the most widely used test in the United Kingdom (Joffe & Pring, 2008), the South Tyneside Assessment of Phonology (STAP; Armstrong & Ainsley, 1988, cited in Joffe & Pring, 2008) was not included in this report because it is not readily available to SLPs in the United States and does not have norms for an American population. The Computerized Articulation and Phonology Evaluation System (CAPES; Masterson & Bernhardt, 2001, cited in Masterson, Bernhardt, & Hofheinz, 2005) also was not included in this report because the content of the test varies across children, depending on their performance on the first part of the test.

Some of the 11 tests that we did include in this report include a battery of tasks to assess various aspects of articulation and phonology. The current report focuses only on the subtest designed to elicit single-word productions and assess consonant and/or vowel production. Five of the tests (FLTA, GFTA–2, PAT–3, Structured Photographic Articulation Test, 2nd edition [SPAT–II; Dawson & Tattersall, 2001], TDTA) were designed to assess word-initial, word-medial, and word-final consonant production, and five were designed to assess consonants only in word-initial and word-final positions (Arizona Articulation Proficiency Scale, 3rd edition [AAPS–3; Fudala, 2000], BBTOP, CAAP, Diagnostic Evaluation of Articulation and Phonology [DEAP; Dodd, Hua, Crosbie, Holm, & Ozanne, 2006], SHAPE). Six of the tests (AAPS–3, DEAP, FLTA, HAPP–3, PAT–3, TDTA) were designed to assess vowels. Five of the tests (BBTOP, CAAP, HAPP–3, SHAPE, SPAT–II) include a phonetic inventory chart on the record form; there is an inventory analysis for the words on the GFTA–2 included on the KLPA–2. Additional information about each test is provided in Appendix A.

Conventions for Transcribing Test Words

Because the conventions for transcription are not the same for all 11 tests, some decisions had to be made regarding the transcription of some segments to allow for consistent comparisons across the tests. Postvocalic liquids were included with the consonants (cf. Bernhardt, & Holdgrafer, 2001b; Shriberg & Kwiatkowski, 1982). Words such as *chair* and *door* were, therefore, transcribed as including a postvocalic consonant /r/ (cf. Shriberg & Kent, 2003; Small, 2005), and these words were classified as monosyllabic. Unstressed syllables with sonorants, such as *pencil* or *station,* were transcribed as including a [əC] rather than a syllabic consonant and were, therefore, considered to be occurrences of those consonants in an unstressed syllable. Unstressed vowel contexts, such as *wagon* and *carrot,* in which /ə/ and /ɪ/ can occur in free variation, were transcribed as /ə/ and were not, therefore, counted as opportunities for /ɪ/.

We used broad transcription rather than phonetically transcribing allophonic variations. Transcription followed the conventions for General American English (Garn-Nunn & Lynn, 2004; Shriberg & Kent, 2003; Small, 2005). This particularly affected decisions about vowel transcription. The vowel in the words *dog* and *frog* was transcribed as /ɔ/, although some dialects would produce this as /ɑ/. Conversely, the vowel in the words *watch* and *water* was transcribed as /ɑ/, although some dialects would produce it as /ɔ/. The vowels /ɪ/ and /ɛ/ were distinguished before /n/ in words such as *pin* and *pen,* although this contrast is neutralized in some dialects. Likewise, the tense-lax vowel pairs /i, ɪ/, /ɛ, e/, and /u, ʊ/ were distinguished before /l/ in word pairs such as *pool* and *pull,* although some dialects neutralize these vowel contrasts.

Data

On each test, we looked for opportunities to produce all of the consonants of American English, including 22 consonants that occur in word-initial position and 21 that occur in word-final position. We also looked for opportunities to produce each of the 15 vowels of American English that can occur in stressed syllables (hereafter referred to as "stressed vowels"). Because we were looking at content coverage for inventorying production, we looked at all of the sounds in all of the words on each test (as suggested, for example, by Klein, 1984) rather than limiting our analyses to the specific sounds targeted in each word.

We limited our analyses of consonant coverage to word-initial and word-final positions for several reasons. Not all of the tests include medial position (see Table 1), and the phonetic contexts included as medial varied considerably within and across tests. Rvachew and Andrews (2002) reported that production in at least one type of medial position, intervocalic context, is likely to be the same as production in final position. Other researchers have suggested that medial position is best viewed as either syllable-initial or syllable-final within word (e.g., Grunwell, 1985). In addition, information about production in medial position may be less crucial for identifying beginning treatment targets as clinicians typically target sound production in word-initial and word-final positions before working on a sound in medial position (e.g., Bernthal & Bankson, 2004).

Separate charts were constructed for word-initial and word-final consonants and for vowels that enabled us to separate words that met our phonetic criteria from words that did not. The criteria for judging words to be phonetically controlled are listed in Table 2.

RESULTS

The current report considered whether the data from 11 standardized single-word tests of articulation and phonology were sufficient for making conclusions about which sounds were and were not in a child's phonetic inventory. To address this question, we examined the words on 11 tests of articulation and phonology for the opportunities they provided to produce all of the consonants and vowels of General American English. Because phonetic context can affect a child's production of consonants and vowels, we set criteria for the phonetic makeup of the words and only counted as opportunities words that met these criteria (termed phonetically controlled words). In order to establish the independence of each sound, we added a criterion of two different vowel contexts for each word-initial consonant and each word-final consonant, as well as a criteria of two different postvocalic contexts, an open syllable and/or different consonants, for each vowel.

Table 2. Criteria for phonetically controlled words.

Segment type	*Criteria*
Word-initial consonants	• Singleton • Monosyllabic or bisyllabic word • For bisyllabic words, target consonant in the initial stressed syllable • Nonharmonic (initial target consonant different from the final and/or medial consonant)
Word-final consonants	• Singleton • Monosyllabic or bisyllabic word • For bisyllabic words, target consonant in the second stressed syllable • Nonharmonic (final target consonant different from the initial and/or medial consonant) • Monomorphemic (final target consonant not part of a separate morpheme)
Vowels	• Monosyllabic or bisyllabic word • For bisyllabic words, target vowel in the stressed syllable • Not followed by a liquid consonant

Consonant Coverage

The consonants included on each test are shown in Table 3.

Production in one word. We first asked whether each test provided one phonetically controlled word for each word-initial and word-final consonant. The mean coverage rate for word-initial consonants was 96% (232/242; M = 21.1, SD = .09, Range = 20–22). Four of the tests (BBTOP, DEAP, FLTA, SPAT–II) included at least one phonetically controlled word for all 22 consonants in word-initial position, with no test providing a phonetically controlled word for less than 20 of the word-initial consonants.

In contrast, none of the tests included a phonetically controlled word for all 21 consonants in word-final position. The mean coverage rate for word-final consonants was 81% (188/231; M = 16.9, SD = 1.9, Range = 14–20). One of the tests (FLTA) included a phonetically controlled word for 20 of the word-final consonants, and two other tests (CAAP and TDTA) included a phonetically controlled word for 19 of the word-final consonants. Three of the tests—GFTA–2, HAPP–3, and SPAT–II—included a phonetically controlled word for only 14 or 15 of the word-final consonants.

The most frequently missing word-final consonants were /ð/ and /ʒ/. Because these are the least frequent sounds in English (Crystal, 1995; Fletcher, 1953), particularly in the vocabularies of young children, this might not represent as significant a gap in test coverage as other consonants would. However, there were other consonants, such as /g/and /ʤ/, that were missing in one word position on some of the tests.

Table 3. Consonant coverage.

Test	Word position	In at least one phonetically controlled word	Consonants[a] Number	%	In two phonetically controlled words with different vowels	Consonants[a] Number	%
AAPS	Initial	p b t d k m n w j h f θ ð s z ʃ ʧ ʤ l r	20	91	b t k m n w f h r	9	41
	Final	p t d g m n ŋ f v θ s z ʃ ʧ l r	16	76	d g m n s r	6	29
BBTOP	Initial	p b t d k g m n w j h f v θ ð s z ʃ ʧ ʤ l r	22	100	p b t d k g m n w j f v s z ʃ ʧ ʤ l r	19	84
	Final	p b t d k g m n f v θ s z ʃ ʧ ʤ l r	18	86	p b t k g m n f v θ s ʧ ʤ r	14	67
CAAP	Initial	p b t d k g m n w h f v θ s z ʃ ʧ ʤ l r	20	91	k w h r	4	18
	Final	p b t d k g m n ŋ f v θ s z ʃ ʧ ʤ l r	19	90	g n f v ʤ l	6	29
DEAP	Initial	p b t d k g m n w j h f v θ ð s z ʃ ʧ ʤ l r	22	100	b f	2	9
	Final	p b t d k g m n ŋ f v θ s ʃ ʧ l r	17	81	k m n s l r	6	29
FLTA	Initial	p b t d k g m n w j h f v θ ð s z ʃ ʧ ʤ l r	22	100	p b t d k m n w j h f s ʃ l r	15	71
	Final	p t d k g m n ŋ f v θ ð s z ʃ ʒ ʧ ʤ l r	20	95	p t d k g n f s l r	10	48
GFTA–2	Initial	p b d k g m n w j h f v θ ð s z ʃ ʧ ʤ l r	21	95	b k w f r	5	23
	Final	p d k g m n ŋ f v θ s ʃ ʧ l r	15	71	k n s l r	5	23
HAPP–3	Initial	p b t k g m n w j h f v θ s z ʃ ʧ ʤ l r	20	91	b m h f	4	18
	Final	p d k m n ŋ f v θ s z ʃ ʧ ʤ r	15	71	p k n r	4	19
PAT–3	Initial	p b t d k g m n w h j f v θ ð s z ʃ ʧ ʤ l	21	95	p b t k m n w h f s	11	50
	Final	p t d k g m n ŋ f v θ ð s ʃ ʧ l r	17	81	d g m n s ʃ r	7	33
SHAPE	Initial	b t d k g m n w j h f v θ ð s z ʃ ʧ ʤ l r	21	95	b d k g w h s ʃ ʧ l r	11	50
	Final	p t d k g m n ŋ f v θ s z ʃ ʧ ʤ l r	18	86	p t d k g m n s ʃ ʤ l r	12	57
SPAT–II	Initial	p b t d k g m n w j h f v θ ð s z ʃ ʧ ʤ l r	22	100	p t w f ʃ l	6	27
	Final	p d k g m n f v θ s z ʃ ʧ r	14	67	k s	2	10
TDTA	Initial	p b t d k g m n w j h f θ ð s z ʃ ʧ ʤ l r	21	95	p b t d k g m n w h f s ʃ l r	15	71
	Final	b t d k g m n ŋ f v θ ð s z ʃ ʧ ʤ l r	19	90	t d k m n f s ʃ ʧ ʤ l r	12	57

[a]Out of 22 word-initial consonants and 21 word-final consonants.

Production in two words with different vowels. Requiring production of a consonant with more than one vowel before putting that consonant in a child's inventory guards against overcrediting the child with consonants that were produced as holistic productions with a single vowel. Therefore, we checked whether each test included a minimum of two phonetically controlled words with differing vowels for a child to produce each consonant. The tests were not constructed to do this, and few did.

The mean rate for providing two such words for word-initial consonants was 42% (101/242; M = 8.6, SD = 5.4, Range = 2–19). For word-final consonants, the mean rate was 36% (84/231; M = 7.2, SD = 3.7, Range = 2–14). The BBTOP had the most complete coverage in both word positions, providing two words meeting the phonetic criteria for 19 of the 22 word-initial consonants and 14 of the 21 word-final consonants. None of the other tests provided two phonetically controlled words for more than 15 word-initial consonants or more than 12 word-final consonants.

Notably, the BBTOP was one of the least frequently used tests reported by clinicians on the Skahan et al. (2007) survey. The tests with the next best coverage were three other infrequently used tests (Skahan et al., 2007): the FLTA, TDTA, and Smit-Hand Articulation and Phonology Evaluation (SHAPE; Smit & Hand, 1992). The FLTA and TDTA are the two oldest tests. Both of these tests provided two phonetically controlled words for 15 word-initial consonants and two phonetically controlled words for 10 or 12 word-final consonants. The SHAPE, a more recent phonological test, included two phonetically controlled words for 11 word-initial consonants and for 12 word-final consonants. Notably, these were the four longest tests, ranging from 81 to 141 total words, which may be one reason for their limited popularity, as Huang, Hopkins, and Nippold (1997) found that the most commonly used norm-referenced tests are those that are quick to administer.

None of the other tests included two phonetically controlled words for more than 15 word-initial consonants or 10 word-final consonants. This included the most commonly used test based on the Skahan et al. (2007) survey, the GFTA–2, which provided two phonetically controlled words for only five word-initial consonants and five word-final consonants. This test was also one of the tests showing the least phonetic control for consonants in word-final position, not providing even one phonetically controlled word for six of the word-final consonants. The GFTA–2, however, has several characteristics that make it desirable for testing young children. Although not having the fewest number of words, it is relatively short, having 53 words, and is quick to administer. It also includes colorful pictures that are engaging to young children.

Vowel Coverage

The vowels included on each test are shown in Table 4.

Production in at least one word. We first asked whether each test provided one phonetically controlled word for each vowel. The mean coverage rate was 85% (141/165; M = 12.6, SD = 1.6, Range = 11–15), with two tests including a word for every vowel. The most commonly non-included vowels were /ɝ/ by six tests and /ɔɪ/ by eight tests. Two other vowels, /ɔ/ and /ʊ/, were not included on four of the tests.

Only six of the tests were designed to test for vowels, and only two of these tests (FLTA, TDTA) included a phonetically controlled word for all 15 vowels. Two other of the tests designed to test vowels (AAPS–3, PAT–3), as well as one test not designed to assess

Table 4. Vowel coverage.

Test	In one phonetically controlled word	Vowels[a]		In two phonetically controlled words with different consonants	Vowels[a]	
		Number	%		Number	%
AAPS	i ɪ eɪ ɛ æ ɝ ʌ u ʊ oʊ ɔ ɑ aɪ aʊ	14	93	i ɪ eɪ ɛ æ ʌ oʊ aɪ aʊ	9	60
BBTOP	i ɪ eɪ ɛ æ ɝɝ ʌ u ʊ oʊ ɔ ɑ aɪ aʊ	14	93	i ɪ eɪ ɛ æ ʌ u oʊ aɪ aʊ	10	67
CAAP	i ɪ eɪ ɛ æ ʌ oʊ ɔ ɑ aɪ aʊ	11	73	i ɪ eɪ ɛ æ ʌ aɪ aʊ	8	53
DEAP	i ɪ eɪ ɛ æ ɝ ʌ u ʊ ɑ aɪ ɔɪ aʊ	13	87	i ɪ ɛ æ ɑ aɪ	6	40
FLTA	i ɪ eɪ ɛ æ ɝ ʌ u ʊ oʊ ɔ ɑ aɪ ɔɪ aʊ	15	100	i ɪ eɪ ɛ æ ʌ u ʊ oʊ ɑ aɪ aʊ	12	80
GFTA–2	i ɪ eɪ ɛ æ ʌ u ɔ ɑ aɪ aʊ	11	73	i ɪ ɛ æ ʌ u aɪ aʊ	8	53
HAPP–3	i ɪ eɪ ɛ æ ʌ u oʊ ɑ aɪ aʊ	11	73	i ɪ eɪ æ ʌ u oʊ ɑ aɪ aʊ	10	67
PAT–3	i ɪ eɪ ɛ æ ʌ u ʊ oʊ ɔ ɑ aɪ ɔɪ aʊ	14	93	i ɪ eɪ ɛ æ ʌ u ɔ aɪ ɔɪ aʊ	11	73
SHAPE	i ɪ eɪ ɛ æ ʌ u oʊ ɔ ɑ aɪ aʊ	12	80	i ɪ eɪ ɛ æ ʌ u oʊ ɔ ɑ aɪ aʊ	12	80
SPAT–II	i ɪ eɪ ɛ æ ʌ u ʊ oʊ ɑ aʊ	11	73	i ɪ eɪ ɛ æ ʌ u oʊ ɑ	9	60
TDTA	i ɪ eɪ ɛ æ ʌ ɝ u ʊ oʊ ɔ ɑ aɪ ɔɪ aʊ	15	100	i ɪ eɪ ɛ æ ʌ ɝ u oʊ ɑ aɪ aʊ	12	80

[a]Out of 15 vowels.

vowels (BBTOP), included a phonetically controlled word for 14 of the vowels. The most commonly missing vowels were /ɔɪ/, /ɝ/, /ʊ/, and /ɔ/. Because these are the vowels reported to be problematic for children with SSD (Shriberg & Kwiatkowski, 1994), this represents an important limitation for using tests to check a child's vowel inventory.

Production in two words. Requiring more than one production of a vowel in different consonant contexts guards against overcrediting a child with particular vowels that may be holistic productions with a single consonant. Therefore, we checked whether each test included a minimum of two phonetically controlled words with differing consonants for each vowel. The tests were not constructed to do this, and few did. The mean rate for providing two such words for each vowel was 65% (107/165; $M = 9.5$, $SD = 1.9$, Range = 6–12).

The tests with the most coverage were the FLTA, SHAPE, and TDTA, all three of which included two words for 12 of the vowels. None of the other tests provided two words for more than 10 of the vowels. The most common reason for this was the inclusion of words with a liquid following the vowel. None of the tests provided two words for /ɝ/ and /ɔɪ/, and only one test provided two words for /ʊ/.

Of the tests designed to assess vowels, the one reported to be used most frequently was the PAT–3 (Skahan et al., 2007), which provided two phonetically controlled words for 11 of the vowels. The AAPS–3, the next most popular of the tests designed to assess vowels (Skahan et al., 2007), included two phonetically controlled words for only nine of the vowels.

DISCUSSION

In this report, we focused on one aspect of validity—content validity of standardized tests—for the purpose of establishing a child's phonetic inventory. We proposed phonetic criteria for the stimulus words used for compiling this inventory type. These criteria were set to guard against measurement confounds from phonetic contexts that might scaffold or inhibit correct production of a sound. We further proposed that conclusions that a sound is within a child's phonetic inventory must be based on a minimum of two opportunities for producing the sound in words that meet these phonetic criteria. In addition, as evidence that each consonant or vowel functions as an independent segment rather than as part of a holistic production, we suggested that each consonant be included in two different vowel contexts and that each vowel be included in two different consonant contexts.

None of the tests met these conditions for all consonants or all vowels. We therefore concluded that none of the tests are valid as the sole basis for generating a phonetic inventory for consonants or vowels. Importantly, this was true for tests that included a phonetic inventory analysis as part of the test as well as for tests that did not include an inventory analysis.

One reason this might be has to do with the characteristics of the words that are included on the tests. James, van Doorn, and McLeod (2008), for instance, found a higher rate of errors to occur on polysyllabic words than on words with one or two syllables. These authors noted that tests that do not include more complicated words may underdiagnose the occurrence and severity of SSD. In contrast, the appropriate words for determining a child's phonetic inventory involve simpler word shapes. The words selected for each of these purposes need to be different so that a test that is designed to identify SSD may not include the types of words suited for determining the sounds that a child can and cannot produce.

In a 1984 article, McCauley and Swisher discussed the misuse of norm-referenced standardized tests. These authors noted several issues that make it invalid to make conclusions about specific deficit areas and to select therapy targets. For one thing, test construction for identifying a disorder leaves gaps in the skills covered by the test. In our review of tests, we found that there were consonants and vowels that were not included on specific tests. As noted by McCauley and Swisher, norm-referenced tests do not provide sufficient items even for the skills that are included on tests. In our review, we considered all of the sounds in every word although some of the tests used a single word for each target. Despite this, we still found only one word for many of the consonants and vowels that were included on the tests.

McCauley and Swisher (1984) further noted that both correct and incorrect performance on a single item could relate to factors other than a child's true ability on that skill. In the current review, we found that tests did not include phonetically controlled words for all consonants and vowels. Most of the words that did not meet the phonetic criteria included contexts that could reduce a child's ability to produce the sound, but there were also some words that involved contexts that could scaffold correct production of a sound that the child would not otherwise produce.

Clinical Implications

Although useful for identifying children with SSD, currently available standardized tests do not provide the requisite number of opportunities in the right types of words for generating a complete phonetic inventory of either the consonants or the vowels of English. The test words must, therefore, be supplemented with probes of phonetically controlled words in order to establish the sounds that are and are not in a child's phonetic inventory. Optimally, the words for this should be monosyllabic and should include only singleton consonants, that is, CV or CVC words for word-initial position consonants and vowels and VC or CVC words for word-final position consonants. In this report, we accepted words that had somewhat more complicated word shapes, including clusters in the nontarget word position and bisyllabic words with the target sound in the stressed syllable. However, it is possible that even this added complexity might confound a child's ability to produce certain sounds, and that sounds not produced in these harder contexts should be re-probed in simpler word shapes.

In order to make conclusions about the sounds in a child's phonetic inventory, we have to ensure that the child has opportunities to attempt those sounds. Nonproduction of a sound on a test cannot always be taken as evidence that a child lacks a particular sound. An example of this is BK, a 6-year-old boy who was given the PAT–3. BK did not produce final /s/ in *house,* the target word for that sound, but produced final /s/ in two other words on the tests—correctly in the word *yes* and as a substitute for /ʃ/ in the word *fish*. BK's nonproduction of final /s/ in the single target word was, therefore, not an indication of his ability to produce /s/ in final position. Before concluding that a child is missing a sound, SLPs should check that the sound was, in fact, included on the test, and that it was included in words that were not "too difficult" phonetically. Sounds that were not tested at all or that were tested in phonetically difficult words should be probed in additional phonetically controlled words to confirm that the sound is really missing from the child's inventory. The best words to use for this would be monosyllabic words containing one or two singleton consonants (i.e., CV, VC, and CVC word shapes). For probing vowels, words with a liquid after the vowel should be avoided.

Alternatively, correct production of a sound on a test cannot always be taken as evidence that a child "knows" that sound. For example, CB, a 3½-year-old boy, successfully produced /k/ in word-initial position in the word cake (/kek/), but produced [tar] instead of /kar/ in the word *car*. His correct production of /k/ could have reflected assimilation from the velar phoneme rather than providing evidence that /k/ was an independent segment in word-initial position. Sounds produced in phonetically scaffolding contexts should, therefore, be probed in nonscaffolding contexts. Harmonic words with the target consonant elsewhere in the word should not, therefore, be used for probing consonants.

When making conclusions about the sounds in a child's inventory, clinicians also need to consider the impact of a child's own assimilation and harmony processes in addition to the characteristics of the target word. Most of the tests include words with both velar and alveolar stops. However, consider a child who is prone to velar assimilation. Such a child might produce the word *duck* as [gʌk] and the word *cat* as [kæk]. These assimilatory productions of /k/ and /g/ should not be taken as evidence that the child has these velar stops in his or her inventory in both word-initial and word-final positions. Rather, production of the velars should be probed in target words that do not include alveolars. Conversely, the nonproduction of /t/ and /d/ in these words should not be concluded to mean that the child lacks these sounds. Rather, the alveolar stops should be probed in target words that do not include velars. Sounds that are produced with harmony or showing place assimilation should be probed in words that do not have the influencing sound(s). This could mean probing velars in words without alveolars and alveolars in words without velars. Another common pattern to look out for is manner assimilation and harmony between glides and liquids, for instance, in words such as *yellow* and *lion*.

In addition, one production of a consonant or vowel in a single phonetic context is not sufficient for concluding that that consonant or vowel functions as a separate segment in a child's phonology. Consider JN, a 3-year-old boy. On the BBTOP, JN produced [nou] for the word *nose* and [wæʔə] for the word *wagon,* but did not produce these word-initial sounds in the other words on the test or in additional probe words. Crediting JN with /n/ and /w/ as separate segments based on his production in a single vowel context would, therefore, have overestimated his inventory; in fact, JN did not readily generalize production of these sounds to other vowel contexts in therapy. We suggest setting a minimum criterion of two productions in differing contexts before crediting a sound to a child's inventory. Word-initial or word-final consonants produced in only one vowel context and vowels produced in only one consonant context would be considered "marginal" (Grunwell, 1981). Therefore, we suggest providing two opportunities for attempting each segment and probing those sounds that are produced in only one context.

A list of the phonetically controlled words from all 11 tests is provided as Appendix B. These words can be used to supplement a test that includes less than two phonetically controlled words for any word-initial consonant, word-final consonant, or vowel. These words can also be used to probe production of marginal or missing segments. Additional words for probing consonants can be found in Gierut, Elbert, and Dinnsen (1987; Appendix A). Additional words for probing vowels can be found in Pollack (1991; Table 5).

Study Limitations

We have suggested a standard for attributing segments to a child's phonetic inventory. This standard was based on suggestions from previous research but, in fact, has not been validated empirically. The standard is a minimal one, just sufficient to establish some combinatorial use rather than being restricted to a single context. This standard would not be sufficient for allowing conclusions about whether these segments function as phonemes for a particular child. It would also not be sufficient for assessing consistency of production. Both of these areas would require additional data to be collected.

We have considered some phonetic aspects of words that must be controlled for: phonetic complexity, number of syllables, and stress pattern. There are additional factors that have not been addressed in this report that might also influence sound production. These include specific co-articulatory influences of various phonetic environments, word familiarity, and neighborhood density to name a few. In this report, we considered the possible confound of assessing sounds in longer and more difficult words. However, assessment of phones in phonetically simple words does not provide a complete picture of a child's articulation and phonology. After establishing the sounds that a child can produce in phonetically simpler words, it is also important to check the child's production in polysyllabic words, words with complex word shapes, words with unstressed syllables, and words with potential assimilation and harmony affects to check stability of sound production.

In addition, we have only considered individual consonants and vowels and have not looked at other aspects of a comprehensive phonological evaluation. These include independent and relational analyses for the word shapes, numbers of syllables, and stress patterns in the child's productions. These would also be important areas to assess when determining the aspects of phonology that need to be targeted for a particular child.

Conclusion

Clinical decision making involves making conclusions in the face of limited data. In this report, we have suggested a criteria of two productions in different contexts for attributing a sound to a child's phonetic inventory. This guards against overattributing a sound based on a single production while enabling inventory decisions to be made based on a finite amount of data that can be gathered within a reasonable time frame.

In this report, we have questioned the validity of using just the data from one test to establish a phonetic inventory of consonants or vowels for a child. We are not advocating the abandonment of norm-referenced standardized tests for identifying children with SSD. Nor are we concluding that any of the tests is inherently invalid because test validity must be evaluated for each type of use (Messick, 1989; Sabers, 1996). In our opinion, each test alone did not include a sufficiently complete set of phonetically controlled words for making conclusions about a child's phonetic inventory. In practical terms, this would necessitate supplementing a test with additional words to probe for production of each sound.

A phonetic inventory is one part of a comprehensive phonological assessment. Sounds that are missing from the phonetic inventory can be probed for stimulability. Once it has been established that a child produces a sound in a phonetically simple word, the SLP may want to check production in more complicated contexts, such as unstressed syllables and polysyllabic words. This becomes increasingly important for older children due to the expected increased length and complexity of their utterances.

ACKNOWLEDGMENTS

The authors wish to thank Jessica Morreale and Roberta Stromberg for their assistance with preparation of the data for this report. Portions of this report were presented at the 2004 annual convention of the American Speech-Language-Hearing Association in Philadelphia, PA, and the 2009 convention in New Orleans, LA.

REFERENCES

Andrews, N., & Fey, M. E. (1986). Analysis of the speech of phonologically impaired children in two sampling conditions. *Language, Speech, and Hearing Services in Schools, 17*(3), 187–198.

Bankson, N., & Bernthal, J. (1990). *Bankson-Bernthal Test of Phonology*. Chicago, IL: Riverside.

Bauman-Waengler, J. (2004). *Articulatory and phonological impairments: A clinical focus* (2nd ed.). Boston, MA: Pearson.

Bernhardt, B. H., & Holdgrafer, G. (2001a). Beyond the basics I: The need for strategic sampling for in-depth phonological analysis. *Language, Speech, and Hearing Services in Schools, 32,* 18–27.

Bernhardt, B. H., & Holdgrafer, G. (2001b). Beyond the basics II: Supplemental sampling for in-depth phonological analysis. *Language, Speech, and Hearing Services in Schools, 32,* 28–37.

Bernthal, J. E., & Bankson, N. W. (2004). *Articulation and phonological impairments* (5th ed.). Boston, MA: Allyn & Bacon.

Bleile, K. (2002). Evaluating articulation and phonological disorders when the clock is running. *American Journal of Speech-Language Pathology, 11*(3), 243–249.

Bleile, K. (2004). *Manual of articulation and phonological disorders: Infancy through adulthood* (2nd ed.). Clifton Park, NY: Delmar.

Branigan, G. (1976). Syllabic structure and the acquisition of consonants: The great conspiracy in word formation. *Journal of Psycholinguistic Research, 5*(2), 117–133.

Camarata, S., & Gandour, J. (1984). Describing idiosyncratic phonologic systems. *Journal of Speech and Hearing Disorders, 49,* 262–266.

Creaghead, N. A., Newman, P. W., & Secord, W. A. (1989). *Assessment and remediation of articulatory and phonological disorders* (2nd ed.). Columbus, OH: Merrill.

Crystal, D. (1982). *Grammatical analysis of language disability* (2nd ed.). San Diego, CA: Singular.

Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. Cambridge, UK: Cambridge University Press.

Davis, B. L. (2005). Goal and target selection for developmental speech disorders. In A. G. Kamhi & K. E. Pollock (Eds.), *Phonological disorders in children: Clinical decision making in assessment and intervention* (pp. 89–100). Baltimore, MD: Brookes.

Dawson, J. I., & Tattersall, P. J. (2001). *Structured Photographic Articulation Test—II*. DeKalb, IL: Janelle Publications.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics, 17*(8), 617–643.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology* (U.S. ed.). San Antonio, TX: The Psychological Corporation.

Donegan, P. (2002). Normal vowel development. In M. J. Ball & F. E. Gibbon (Eds.), *Vowel disorders* (pp. 1–35). Woburn, MA: Butterworth-Heinemann.

Elbert, M., & Gierut, J. A. (1986). *Handbook of clinical phonology: Approaches to assessment and treatment*. Boston, MA: College-Hill Press.

Fisher, H. B., & Logemann, J. A. (1971). *Fisher-Logemann Test of Articulation Competence*. Iowa City, IA: Houghton Mifflin.

Fletcher, H. (1953). *Speech and hearing in communication*. Princeton, NJ: Van Nostrand.

Fudala, J. B. (2000). *Arizona Articulation Proficiency Scale, 3rd revision*. Los Angeles, CA: Western Psychological Services.

Garn-Nunn, P. G., & Lynn, J. M. (2004). *Calvert's descriptive phonetics* (3rd ed.). New York, NY: Thieme.

Gierut, J. A. (2005). Phonological intervention: The how or the what? In A. G. Kamhi & K. E. Pollock (Eds.), *Phonological disorders in children: Clinical decision making in assessment and intervention* (pp. 201–210). Baltimore, MD: Brookes.

Gierut, J. A., Elbert, M., & Dinnsen, D. A. (1987). A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research, 30,* 462–479.

Glaspey, A. M., & Stoel-Gammon, C. (2005). Dynamic assessment in phonological disorders. *Topics in Language Disorders, 25,* 220–230.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation* (2nd ed.). Circle Pines, MN: AGS.

Grunwell, P. (1981). *The nature of phonological disability in children*. London, UK: Academic Press.

Grunwell, P. (1985). *Phonological assessment of child speech*. San Diego, CA: College-Hill.

Hodson, B., & Paden, E. (1991). *Targeting unintelligible speech: A phonological approach to remediation* (2nd ed.). Austin, TX: Pro-Ed.

Hodson, B. W. (1997). Disordered phonologies: What have we learned about assessment and treatment? In B. W. Hodson & M. L. Edwards (Eds.), *Perspectives in applied phonology* (pp. 197–224). Gaithersburg, MD: Aspen.

Hodson, B. W. (2004). *Hodson Assessment of Phonological Patterns* (3rd ed.). Austin, TX: Pro-Ed.

Hodson, B. W., & Paden, E. P. (1981). Phonological processes which characterize unintelligible and intelligible speech in early childhood. *Journal of Speech and Hearing Disorders, 46,* 369–373.

Hodson, B. W., Scherz, J. A., & Strattman, K. H. (2002). Evaluating communicative abilities of a highly unintelligible preschooler. *American Journal of Speech-Language Pathology, 11*(3), 236–242.

Huang, R., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28,* 12–29.

Ingram, D. (1989). *Phonological disability in children* (2nd ed.). London, UK: Cole & Whurr.

Ingram, D., Christensen, L., Veach, S., & Webster, B. (1980). The acquisition of word-initial fricatives and affricates in English between 2 and 6 years. In G. H. Yen-Komshian, J. F. Cavanaugh, & C. A. Ferguson (Eds.), *Child phonology: Volume 1 Production* (pp. 169–192). New York, NY: Academic Press.

Ingram, D., & Ingram, K. D. (2001). A whole-word approach to phonological analysis and intervention. *Language, Speech, and Hearing Services in Schools, 32,* 271–283.

Jakubowitz, M., & Schill, M. J. (2008). Ethical implications of using outdated standardized tests. *Perspective on School-Based Issues, 9,* 79–83.

James, D. G. H. (2001). An item analysis of Australian words for an articulation and phonological test for children aged 2 to 7 years. *Clinical Linguistics and Phonetics, 15*(6), 457–485.

James, D. G. H., van Doorn, J., & McLeod, S. (2001). Vowel production in mono-, di- and polysyllabic words in children aged 3;0 to 7;11 years. In L. Wilson & S. Hewat (Eds.), *Proceedings of the Speech Pathology Australia National Conference* (pp. 127–136). Melbourne, Australia: Speech Pathology Australia.

James, D. G. H., van Doorn, J., & McLeod, S. (2008). The contribution of polysyllabic words in clinical decision making about children's speech. *Clinical Linguistics and Phonetics, 22,* 345–353.

Joffe, V., & Pring, T. (2008). Children with phonological problems: A survey of clinical practice. *International Journal of Language & Communication Disorders, 43,* 154–164.

Kent, R. D. (1982). Contextual facilitation of correct sound production. *Language, Speech, and Hearing Services in Schools, 13,* 66–76.

Khan, L. M. (2002). The sixth view: Assessing preschoolers' articulation and phonology from the trenches. *American Journal of Speech-Language Pathology, 11*(3), 250–254.

Khan, L. M., & Lewis, N. (2002). *Khan-Lewis Phonological Analysis* (2nd ed.). Circle Pines, MN: AGS.

Klein, E. S. (1996). *Clinical phonology: Assessment and treatment of articulation disorders in children and adults*. San Diego, CA: Singular.

Klein, H. B. (1982). Hippopotomas is hard to say: The integration of lexical and polysyllabic lexical items. *Seminars in Speech, Language, and Hearing, 3,* 127–137.

Klein, H. B. (1984). Procedures for maximizing phonological information from single-word responses. *Language, Speech, and Hearing Services in Schools, 15*(4), 267–274.

Leonard, L., Devescovi, A., & Ossella, T. (1987). Context-sensitive phonological patterns in children with poor intelligibility. *Child Language Teaching and Therapy, 3,* 125–132.

Lieberman, R. J., & Michael, A. (1986). Content relevance and content coverage in tests of grammatical ability. *Journal of Speech and Hearing Disorders, 51,* 71–81.

Lippke, B. A., Dickey, S. E., Selmar, J. W., & Soder, A. L. (1997). *Photo-Articulation Test* (3rd ed.). Austin, TX: Pro-Ed.

Lof, G. L. (1996). Factors associated with speech-sound stimulability. *Journal of Communication Disorders, 29,* 255–278.

Masterson, J. J., Bernhardt, B. H., & Hofheinz, M. K. (2005). A comparison of single word and conversational speech in phonological evaluation. *American Journal of Speech-Language Pathology, 14,* 229–241.

McCauley, R. J., & Swisher, L. (1984). Use and misuse of norm-referenced test in clinical assessment: A hypothetical case. *Journal of Speech and Hearing Disorders, 49,* 338–348.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18,* 5–11.

Miccio, A. W. (2002). Clinical problem solving: Assessment of phonological disorders. *American Journal of Speech-Language Pathology, 11*(3), 221–229.

Morrison, J. A., & Shriberg, L. D. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research, 35*(2), 259–273.

Paden, E. P., & Moss, S. A. (1985). Comparison of three phonological analysis procedures. *Language, Speech, and Hearing Services in Schools, 16,* 103–109.

Pena-Brooks, A., & Hedge, M. N. (2007). *Assessment and treatment of articulation and phonological disorders in children* (2nd ed.). Austin, TX: Pro-Ed.

Plante, E. (1996). Observing and interpreting behaviors: An introduction to the clinical forum. *Language, Speech, and Hearing Services in Schools, 27,* 99–101.

Pollock, K. E. (1991). The identification of vowel errors using traditional articulation or phonological process test stimuli. *Language, Speech, and Hearing Services in Schools, 22*(2), 39–50.

Pollock, K. E. (1994). Assessment and remediation of vowel misarticulations. *Clinics in Communication Disorders, 4,* 23–37.

Pollock, K. E., & Berni, M. C. (2003). Incidence of non-rhotic vowel errors in children: Data from the Memphis vowel project. *Clinical Linguistics and Phonetics, 17*(4–5), 393–401.

Pollock, K. E., & Keiser, N. (1990). An examination of vowel errors in phonologically disordered children. *Clinical Linguistics and Phonetics, 4*(2), 161–178.

Powell, T. W., & Miccio, A. W. (1996). Stimulability: A useful clinical tool. *Journal of Communication Disorders, 29,* 237–253.

Rvachew, S., & Andrews, E. (2002). The influence of syllable position on children's production of consonants. *Clinical Linguistics and Phonetics, 16*(3), 183–198.

Sabers, D. L. (1996). By their tests we will know them. *Language, Speech, and Hearing Services in Schools, 27,* 102–109.

Secord, W. A., & Donohoe, J. S. (2002). *Clinical Assessment of Articulation and Phonology.* Greenville, SC: Super Duper Publications.

Shriberg, L., & Kwiatkowski, J. (1980). *Natural process analysis [NPA]: A procedure for phonological analysis of continuous speech samples*. New York, NY: Wiley.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40,* 708–722.

Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics* (3rd ed.). Boston, MA: Allyn & Bacon.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47*(3), 256–270.

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Research, 37*(5), 1100–1126.

Skahan, S. M., Watson, M., & Lof, G. L. (2007). Speech-language pathologists' assessment practices for children with suspected speech sound disorders: Results of a national survey. *American Journal of Speech-Language Pathology, 16,* 246–259.

Small, L. H. (2005). *Fundamentals of phonetics* (2nd ed.). Boston, MA: Allyn & Bacon.

Smit, A. B. (1993). Phonologic error distribution in the Iowa–Nebraska articulation norms project: Word initial consonant clusters. *Journal of Speech and Hearing Research, 36*(5), 931–947.

Smit, A. B. (2004). *Articulation and phonology resource guide for school-age children and adults*. Clifton Park, NY: Delmar Learning.

Smit, A. B., & Hand, L. (1992). *Smit-Hand Articulation and Phonology Evaluation*. Los Angeles, CA: Eastern Psychological Services.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779–798.

Stoel-Gammon, C. (1985). Phonetic inventories, 15–24 months: A longitudinal study. *Journal of Speech and Hearing Research, 28,* 505–512.

Stoel-Gammon, C. (1987). Phonological skills of 2-year-olds. *Language, Speech, and Hearing Services in Schools, 18,* 323–329.

Stoel-Gammon, C., & Dunn, C. (1985). *Normal and disordered phonology in children*. Austin, TX: Pro-Ed.

Stoel-Gammon, C., & Herrington, P. B. (1990). Vowel systems of normally developing and phonologically disordered children. *Clinical Linguistics and Phonetics, 4*(2), 145–160.

Stokes, S. F., Klee, T., Carson, C. P., & Carson, D. (2005). A phonemic implicational feature hierarchy of phonological contrasts for English-speaking children. *Journal of Speech, Language, and Hearing Research, 48,* 817–833.

Templin, M. (1957). *Certain language skills in children*. Minneapolis: University of Minnesota Press.

Templin, M. C., & Darley, F. L. (1968). *The Templin-Darley Tests of Articulation*. Iowa City: University of Iowa Press.

Tyler, A. A., & Tolbert, L. C. (2002). Speech-language assessment in the clinical setting. *American Journal of Speech-Language Pathology, 11*(3), 215–220.

Vance, M., Stackhouse, J., & Wells, B. (2005). Speech production skills in children aged 3–7 years. *International Journal of Language and Communication Disorders, 40*(1), 29–48.

Velleman, S. L. (1998). *Making phonology functional: What do I do first?* Woburn, MA: Butterworth-Heinemann.

Velleman, S. L. (2003). *Childhood apraxia of speech resource guide*. Clifton Park, NY: Delmar Learning.

Vihman, M. (1996). *Phonological development: The origins of language in the child*. Cambridge, MA: Blackwell.

Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*(3), 286–350.

Weiner, F. (1979). *Phonological process analysis*. Baltimore, MD: University Park Press.

Wellman, B. L., Case, I. M., Mengert, I. G., & Bradbury, D. E. (1931). Speech sounds of young children. *University of Iowa Studies in Child Welfare, 5*(2), 1–82.

Weston, A. D., Shriberg, L. D., & Miller, J. J. (1989). Analysis of speech samples with SALT and PEPPER. *Journal of Speech and Hearing Research, 32,* 755–766.

Williams, A. L. (2003). *Speech disorders resource guide for preschool children*. Boston, MA: Allyn & Bacon.

Wolfe, V. I., & Blocker, S. D. (1990). Consonant-vowel interaction in an unusual phonological system. *Journal of Speech and Hearing Disorders, 55,* 551–566.

Wolk, L., & Meisler, A. W. (1998). Phonological assessment: A systematic comparison of conversation and picture naming. *Journal of Communication Disorders, 31*(4), 291–313.

Received November 18, 2008
Revision received April 1, 2009
Accepted July 17, 2009
DOI: 10.1044/0161-1461(2009/08-0125)

Contact author: Sarita Eisenberg, Montclair State University, Communication Sciences and Disorders, 1 Normal Avenue, Upper Montclair, NJ 07473. E-mail: eisenbergs@mail.montclair.edu.

APPENDIX A. TESTS CONSIDERED IN THIS REPORT

Tests have been classified as articulation tests or phonological tests based on their designation in the test manuals and in reference books (e.g., Pena-Brooks & Hegde, 2007; Velleman, 1998; Williams, 2003). All of the phonological tests include phonological process analysis.

Articulation Tests

Arizona Articulation Proficiency Scale, 3rd edition (AAPS–3; Fudala, 2000). The AAPS–3 is a two-position test that was designed to cover all of the major speech sounds of English, including consonants and vowels. The test was designed to identify misarticulations and does not include an inventory chart on the record form.

Fisher-Logemann Test of Articulation (FLTA; Fisher & Logemann, 1971). The FLTA is a three-position test with a separate word for each consonant in each position and for each vowel. The test was designed to look for error patterns based on consonant and vowel features. Although the test manual does not explicitly talk about an inventory analysis, the record form is set up to follow the traditional place/manner/voicing consonant chart and the traditional vowel quadrilateral.

Goldman-Fristoe Test of Articulation, 2nd edition (GFTA–2; Goldman & Fristoe, 2000) and *Khan-Lewis Phonological Analysis,* 2nd edition (KLPA–2; Khan & Lewis, 2002). The GFTA–2 Sounds-in-Words subtest is a three-position test that was constructed to include all consonants except those that "have a low intervention priority compared with the other consonants [ʒ] " (Goldman & Fristoe, 2000, p. 11). The GFTA–2 does not include an inventory analysis, but the record form is set up to chart consonant production to allow the examiner to note which sounds were misarticulated. In addition, the GFTA–2 manual refers speech-language pathologists to the KLPA–2, which includes a phonetic inventory analysis on the record form. Although the GFTA–2 is not designed to assess vowels, the manual states that vowel deviations can be noted. The GFTA–2 also includes a subtest for eliciting a connected speech sample (Sounds-in-Sentences) and a subtest for assessing stimulability (Stimulability). The KLPA–2 provides a phonological process analysis based on the words from the GFTA–2. The GFTA–2 was the most commonly used articulation test on the Skahan et al. (2007) survey.

Photo-Articulation Test, 3rd edition (PAT–3; Lippke, Dickey, Selmar, & Soder, 1997). The PAT–3 is a three-position test. The manual states that the test is "an efficient inventory" (p. 1) and includes all consonants and vowels. The aim of the test is to identify articulation errors.

Structured Photographic Articulation Test, 2nd edition (SPAT–II; Dawson & Tattersall, 2001). The *SPAT–II* is a three-position test that was designed to assess all consonants except /ʒ/. The record form includes a positional consonant inventory chart. The manual suggests that clinicians record vowel errors but notes that vowels are not fully represented on the test. The manual also includes a probe for assessing imitation of each vowel in one word.

Templin-Darley Tests of Articulation (TDTA; Templin & Darley, 1968). The TDTA is a three-position test that was constructed to include all consonants and vowels. The test allows the determination of sounds produced incorrectly in one position and also of sounds that a child has failed to master (i.e., sounds not correctly produced in two word positions). In addition to the Diagnostic Test, which uses all of the test words, the TDTA has a shorter version Screening Test.

Phonological Tests

Bankson-Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990). The BBTOP was designed to include "most English consonants in initial and final position" (p. 1). The test includes a consonant inventory.

Clinical Assessment of Articulation and Phonology (CAAP; Secord & Donohoe, 2002). The CAAP Articulation Inventory subtest is a two-position test that was designed to include all consonants. The test provides a means to summarize the child's consonant inventory. In addition to a chart for recording production in each word, the record form includes a separate summary section for charting consonants in each position. The CAAP includes supplementary phonological process checklists.

Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd, Hua, Crosbie, Holm, & Ozanne, 2006). The DEAP includes three subtests. The Articulation Assessment subtest assesses production of individual consonants and vowels and enables clinicians "to establish the child's phoneme inventory" (p. 25). The Diagnostic Screen subtest and Phonology Assessment subtest assess stimulability.

Hodson Assessment of Phonological Patterns, 3rd edition (HAPP–3; Hodson, 2004). The HAPP–3 was designed to assess phonological error patterns rather than individual sounds. The test includes both a consonant and vowel inventory chart on the record form as part of the comprehensive phonological assessment section. The HAPP–3 also includes a preschool phonological screening and, for older children, a multisyllabic word screening.

Smit-Hand Articulation and Phonology Evaluation (SHAPE; Smit & Hand, 1992). The SHAPE is a two-position test that was designed to include most consonants in two or more words in at least one position. The test includes a phonetic inventory analysis of consonants grouped by manner, place, and word position.

APPENDIX B (P. 1 OF 2). PHONETICALLY CONTROLLED WORDS FROM 11 TESTS OF ARTICULATION AND PHONOLOGY

Word-Initial Consonants

	Monosyllabic words			Bisyllabic words		
	CV	*CVC*	*CVCC*	*CVCV*	*CVCVC*	*Other*
p	pie	page, patch, pen, pig, pin		paper	pages, pennies, pocket, puzzle	pencil, pencils
b	boy	bag, ball, bat, bath, bathe, bed, bell, belt, bird, boat, book, bus	beard, boats, books, box	bunny	bottle, buckle	basket
t	two	tail, teeth, ties, toes, top, tub		tiger, tv	table	toothache, toothbrush
d		deer, dish, dive, dog, door, duck		dinner		dishes, doctor
k	key	cage, can, car, cat, coat, comb, cone, cow, cup, keys, king	card, cold, corn		carrot	candle, candy, carrots
g		gate, girl, goat, gum, gun				
m		man, milk, moon, mouse, mouth	mask, milk	mother	matches, mitten	mittens, monkey
n		knife, nose	nails, nest			
w		watch, web, wing, witch, wolf		washer, water	wagon, watches, whistle	window
j	you	yes	yard, yawn	yellow		
h		hat, hive, hose, house	hand, heart, help, horse	hammer		hanger
f		feet, fire, fish, five, foot, phone	fence, fork	feather, finger	feathers, fishing	
v		van, vase	vest			vacuum
θ		thumb				thankyou
ð		that, there, this				
s	saw	seal, soap, sock, sun	socks		scissors	sandwich
z	zoo	zip		zipper		zebra
ʃ	shoe	sharp, sheep, shoes	shirt		shovel	
ʧ		chair, cheese, chip, church				
ʤ		jam, jar, judge	jars, jump	jello		jumping
l		leaf	lamp, large, legs, lift, locked	ladder, letter	lion	
r		rain, rake, red, ring, rock, rope		rubber	rabbit, rocket	

Word-Final Consonants

	Monosyllabic			Bisyllabic
	VC	*CVC*	*CCVC & CCCVC*	*CVCVC*
p		chip, cup, rope, sheep, soap, top, zip		
b		tub, web	crab	
t		bat, boat, cat, coat, feet, foot, gate, goat, hat, that	fruit, plate, skate	
d		bed, bird, red	bread, sled, slide	
k		book, duck, rake, rock, sock	black, block, quack, smoke, snake, truck	
g	egg	bag, dog, pig	flag, frog	
m		comb, gum, jam, thumb	broom, drum	
n		can, cone, gun, man, moon, pen, phone, pin, rain, sun, van, yawn	clown, green, plane, queen, screen, spoon, train	balloon
ŋ		king, ring, wing	string, swing	
f		knife, leaf		
v		dive, five, hive	glove, stove	
θ		bath, mouth, teeth	smooth	
ð		bathe	smooth	
s		bus, house, mouse, this, vase, yes	dress, glass, grass	
z		cheese, nose		
ʃ		dish, fish	brush, splash	
ʒ				garage
ʧ		church, patch, watch, witch	scratch	
ʤ		cage, page	bridge	
l		ball, bell, girl, seal, tail	school, snail	
r	ear	car, chair, deer, door, fire, jar, there	square, star	

APPENDIX B (P. 2 OF 2). PHONETICALLY CONTROLLED WORDS FROM 11 TESTS OF ARTICULATION AND PHONOLOGY

Vowels

	Monosyllabic			Bisyllabic		
	CV & CCV	*VC & CVC*	*Other*	*CVCV*	*CVCVC*	*Other*
i	key, three, tree	cheese, keys, leaf, seat, sheep, teeth	green, queen, screen, trees	meow, tv		zebra
ɪ		bib, chip, dish, fish, fit, king, pig, pin, ring, this, wing, witch, zip	bridge, lift, swing, string, twins	dinner, zipper	dishes, fishing, mitten, scissors, whistle	finger, mittens, sprinkler, swimming, window
eɪ	spray	bathe, cage, cake, gate, page, rain, vase	plane, plate, skate, skates, snake, train	baby, paper	table, pages	angels, crayon, crayons, station
ɛ		bed, red, egg, pen them, web, yes	bread, dress, fence, legs, nest, sled, steps, vest	feather, jelly, letter, yellow	feathers, pennies	dresses, pencil, pencils, present
æ		bag, bat, can, cat, hat, jam, man, patch, that, van	black, crab, flag, glass, grass, hand, lamp, mask, scratch, splash, stamps	hammer, ladder	apple, matches, rabbit, wagon	apples, basket, bathtub, candle, candy, cracker, crackers, glasses, hanger, planting. sandwich, thank-you, vacuum
ɝ		bird, church	shirt			
ʌ		bus, cup, duck, gum, gun, judge, sun, thumb, tub	brush, drum, glove, jump, truck, trunk	color, mother, rubber	buckle	buckle, brother monkey, jumping, thunder
u	blue, shoe, two, zoo	moon, shoes	broom, smooth, spoon		balloon, toothache	balloons, music, toothbrush
ʊ		book, foot	books			
oʊ		boat, coat, comb, cone, goat, hose, nose, phone, rope, soap, toes	boats, cold, smoke, stove	yoyo		snowman
ɔ	saw	dog, yawn	frog, straw	bottle, water	garage, rocket	
ɑ		rock, sock, top, watch	block, blocks, box, clock, locked, scarf, socks, stopped	washer	bottle, pocket	doctor
aɪ	cry, eye, pie	dive, five, hive, knife, nine, pipe, ties	slide	tiger	lion	ice-cubes, lightning, spider
ɔɪ	boy					
aʊ	cow	house, mouse, mouth	clouds, clown			flower, flowers