Clinical Linguistics & Phonetics, April–May 2010; 24(4–5): 271–282

informa
healthcare

The Word Complexity Measure: Description and application to developmental phonology and disorders

CAROL STOEL-GAMMON

University of Washington, Seattle, WA, USA

(Received 8 November 2009; Accepted 20 December 2009)

Abstract
Miccio's work included a number of articles on the assessment of phonology in children with phonological disorders, typically using measures of correct articulation, using the PCC, or analyses of errors, using the framework of phonological processes. This paper introduces an approach to assessing phonology by examining the phonetic complexity of children's productions. Unlike most current measures of phonological development and disorders, this approach does not assess accuracy, but focuses exclusively on the complexity of segments, syllables, and words. Using independent analyses, the Word Complexity Measure is based on the phonetic complexity of a child's productions, with later acquired sound classes, syllable shapes, and word patterns coded as more complex. The Word Complexity Measure allows researchers and clinicians to compare the phonetic and phonological aspects of productions across children and over time. The measure can also be used to compare the complexity of a child's productions with that of the target words.

Keywords: *phonological assessment measures, phonological development, phonological disorders, independent analyses*

Introduction

Adele Miccio's contributions to the field of speech-language pathology cover a broad array of topics, ranging from investigations of language and literacy in low-income African-American and Hispanic children, to the effects of otitis media on speech and language development, to conversational interactions of children with speech disorders, bilingual children's language abilities and early reading outcomes, and Puerto Rican mothers' beliefs and home literacy practices. One of her prime interests was the assessment of phonological disorders, with particular attention to the use of multiple measures including stimulability (Miccio, Elbert, and Forrest, 1999; Miccio, 2002). In order to obtain a complete picture of a child's phonological system, Miccio advocated gathering speech samples that include single-word articulation tests as well as conversational speech, and using both independent and relational analyses to characterize the child's phonology (Miccio, 2002; Tyler, Miccio, Hoffman, Norris, Hodson, Scherz, et al., 2002). This paper extends some of Adele's ideas regarding

Correspondence: Carol Stoel-Gammon, PhD, Department of Speech and Hearing Sciences, University of Washington, 1417 N.E. 42nd Street, Seattle, WA 98105, USA. E-mail: csg@u.washington.edu

ISSN 0269-9206 print/ISSN 1464-5076 online © 2010 Informa UK Ltd.
DOI: 10.3109/02699200903581059

assessment by proposing a measure that can be used for young children with typical development and may be applied to the assessment of older children with speech and/or language delays. This measure is based on the phonetic complexity of a child's production as well as the complexity of target forms. The paper provides the following: (1) a rationale for the measure; (2) a description of the measure; (3) applications of the measure; and (4) a preliminary evaluation of the measure and comparisons with other measures.

Phonological assessment: Qualitative and quantitative measures for young children and/or children with language delay

As noted above, Adele Miccio's approach to assessing phonology incorporated multiple measures and used a variety of analyses. Of note, she felt that one needed a data set that contained elicited single words, to ensure that all phonemes are tested, and a running speech sample that would provide information on intelligibility, prosody, and interactions between speech and language. Her approach included independent analyses, based only on the child's production, and relational analyses that compare the child's production with the target form to determine levels of accuracy and the nature of errors. Adele's assessments provide a detailed account of the child's phonological abilities. However, for children with limited vocabularies or highly unintelligible speech, it is difficult to obtain all the information needed for this type of assessment. This is often the case for young children in the early stages of lexical development, and for older 'late-talking' children with limited vocabularies. With these populations, one needs a different, simpler analysis that can provide the information for clinical assessments or research purposes.

The question for phonological assessment for any population is 'what do we need to know and when do we need to know it?' In my view, for children in the early stages of phonological development, our assessment should focus on the *breadth* of a child's system. Rather than analysing the accuracy of individual phonemes, as is the case with a single-word articulation test, or determining the types of errors, our attention should be directed toward the range of sound classes and syllable/word structures that occur in the child's speech. The measure proposed here, the *Word Complexity Measure* (WCM), provides both quantitative and qualitative information about productions, and is appropriate for children with small vocabularies.

First, the WCM yields a single number based on phonetic complexity of productions analysed; this number can be used to compare across children and/or to track changes in a child's phonology over time. Second, the WCM presents qualitative information about the word patterns, syllable structures, and sound classes the child uses. In addition, both the qualitative and quantitative measures can be performed on the target words to determine the complexity of these words. Comparing the complexity of the child's forms with the target words shows the degree to which the child's productions are simplified. Unlike relational measures of accuracy or errors, this comparison is not based on phoneme accuracy, but on independent analyses of two sets of data. As noted in the examples below, assessment of the target forms can also provide information on lexical selection and avoidance patterns that may be present in the child's sample.

The Word Complexity Measure (WCM): Description and examples

The WCM involves independent analyses of word productions of children acquiring English. Because the measure is designed for use with children who may have limited vocabularies and would be unable to provide a full set of responses to a single-word articulation test, the

descriptions and examples below are based on words from conversational samples. It is possible, however, to use the WCM to analyse productions gathered with a single-word test.

Each word in a sample is awarded a complexity 'score' based on an approach that assesses different levels within the phonological system: word patterns, syllable structures, and sound classes. Higher scores indicate the presence of more complex/later acquired phonological parameters, with complexity determined by findings of investigations of early phonological acquisition (e.g. Stoel-Gammon, 1985; Robb and Bleile, 1994) and developmental norms (e.g. Prather, Hedrick, and Kern, 1975; Smit, Hand, Freilinger, Bernthal and Bird, 1990). It is well documented that certain sound classes and syllable structures tend to be established early in the speech of children learning English, namely the sound classes of stops, nasal, and glides, and the CV and CVCV syllable/word structures (Stoel-Gammon, 1985; 1987; Grunwell, 1987; Dyson, 1988). In the WCM analysis, phonological parameters that are expected to occur in the productions of a 24-month-old with typical speech and language development do not receive complexity points. The description below lists the eight 'complex' parameters and provides examples of scoring.

Word patterns

(1) Productions with more than two syllables receive 1 point.
(2) Productions with stress on any syllable but the first receive 1 point.

Syllable structures

(1) Productions with a word–final consonant receive 1 point.
(2) Productions with a consonant cluster (defined as a sequence of two or more consonants within a syllable) receive 1 point for each cluster.

Sound classes

(1) Productions with a velar consonant receive 1 point for each velar.
(2) Productions with a liquid, a syllabic liquid, or a rhotic vowel receive 1 point for each liquid, syllabic liquid, and rhotic vowel.
(3) Productions with a fricative or affricate receive 1 point for each fricative and affricate.
(4) Productions with a voiced fricative or affricate receive 1 point for each voiced fricative and affricate (in addition to the point received for #3).

Scoring word complexity with the Word Complexity Measure

The examples below, taken from child speech, illustrate the scoring procedures for the WCM:

(1) [beɪbi] (target: *baby*) WCM = 0 points
(2) [dædi] (target: *daddy*): WCM = 0 points
(3) [bɛt] (target: *bed*): WCM = 1 point (final consonant)
(4) [tɪtɪn] (target: *chicken*): WCM =1 point (final consonant)
(5) [baɪ.ɪn] (target: *crying*): WCM =1 point (final consonant)
(6) [waʊn] (target: *around*): WCM =1 point (final consonant)
(7) [bʊk] (target: *book*): WCM = 2 points (final consonant + velar)
(8) [dɪs] (target: *this*): WCM = 2 points (final consonant + fricative)
(9) [bʌdəfwaɪ] (target: *butterfly*): WCM = 3 points (> two syllables + cluster + fricative)
(10) [gwin] (target: *green*): WCM = 3 points (cluster + final consonant + velar)

Total complexity points: 14

Complexity points for the same words produced by an adult are as follows:

(1) [beɪbi] (*baby*): WCM = 0 points
(2) [dædi] (*daddy*): WCM = 0 points
(3) [bɛd] (*bed*): WCM = 1 point (final consonant)
(4) [tʃɪkɪn] (*chicken*): WCM = 3 points (final consonant + affricate + velar)
(5) [kraɪ.ɪŋ] (*crying*): WCM = 5 points (final consonant + cluster + liquid + 2 velars)
(6) [ə'raʊnd] (*around*): WCM =4 points (non-initial stress + cluster + final consonant + liquid)
(7) [bʊk] (*book*): WCM = 2 points (final consonant + velar)
(8) [ðɪs] (*this*): WCM = 4 points (final consonant + 2 fricatives + voiced fricative)
(9) [bʌɾɚflaɪ] (*butterfly*): WCM = 5 points (> 2 syllables + cluster + fricative + rhotic vowel + liquid)
(10) [ɡrin] (*green*): WCM = 4 points (cluster + final consonant + velar + liquid)

Total complexity points: 28.

As shown, the sample of 10 words produced by the child was awarded 14 points, yielding an average of 1.4 points per word. For the target words, the point total is 28, yielding an average complexity of 2.8. The ratio of child to adult complexity is .5, indicating that the complexity of the child's productions is half the complexity of the target words. It should be noted that this small data set is presented to illustrate the WCM coding system; 10 words is too small to be a reliable indicator of overall word complexity. A larger sample appears in Table I the desired number of words in a sample is discussed later. Although all productions in the child's sample data set were attempts at identifiable words, unintelligible forms may be included in the data set and scored the same way. In this case, however, it would not be possible to compare their complexity with the target word.

Applying the WCM to speech samples

A child sample and target words

Table I presents a data set in which the WCM is applied to 36 words and illustrates how the measure can be used to obtain quantitative and qualitative assessments of a child's phonology. The words are selected from conversational sample gathered from 'K' age 22 months (data from Bleile, 1991). The number of complexity points for each word is shown in the last column. Across words, the complexity varies from 0–4 points and the average is 1.27. Nine words (25% of the sample) received zero points, indicating very simple productions in terms of sound classes and syllable types. The word with the highest complexity score was *glasses,* produced as [bæsɪz], a production that contains a final consonant and two fricatives, one of which is voiced.

The table also presents the extent of use of each parameter. As shown by the totals at the bottom of the columns, seven of the eight parameters occurred in at least one of K's productions, indicating that K's output included a wide range of sound classes and syllable/word structures. There was, however, a wide range of use across the parameters: the sample had no productions with non-initial stress, only one with a cluster, and one with a liquid; there were, in contrast, 13 occurrences of fricatives or affricates and 21 productions with a final consonant. A preliminary interpretation of these numbers is that K is quite proficient at producing closed syllables and fricatives/affricates but has difficulty with multisyllabic words, words with non-initial stress, and words with liquids.

Table I. WCM scoring of child productions.

Word	IPA	> 2 syl	Stress	Cluster	Final C	Velar	Fricative, affricate	Vcd fr, affric	Liquid, rhotic V	Total points
Airplane	ɛpeɪn				1					1
Apple	æpʊ									0
Baby	beɪbi									0
Ball	bɔ									0
Bear	bɛr				1				1	2
Bed	bɛd				1					1
Bird	bɔɪt				1					1
Blue	bu									0
Book	bʊk				1	1				2
Bread	bʌd				1					1
Butterfly	bʌɾəfaɪ	1					1			2
Chicken	tɪtɪn				1					1
Circle	tɪtu									0
Clown	taʊn				1					1
Coat	toʊt				1					1
Crying	baɪ.ɪn				1					1
Daddy	dædi									0
Duck	dʌk				1	1				2
Elephant	ɛfən				1		1			2
Eyes	aɪs				1		1			2
Flower	faʊwə						1			1
Foot	fʊt				1		1			2
Frog	fɔg				1	1	1			3
Glasses	bæsɪz				1		2	1		4
Green	gwin			1	1	1				3
Hurt	hʊ						1			1
Monkey	mʌŋki					2				2
Puzzle	pʌzu						1	1		2
Rabbit	wæbɪt				1					1
Sad	sæ						1			1
Sleeping	sipɪn				1		1			2
Slide	baɪ									0
Snack	neɪk				1	1				2
Spider	paɪju									0
This	dɪs				1		1			2
Turtle	tʊtʊ									0
	Totals	1	0	1	21	7	13	2	1	av: 1.27

Although the analyses in Table I provide qualitative and quantitative information on K's production patterns, they do not permit a full interpretation of the results. For example, K produced no words with non-initial stress and only one word with a consonant cluster; this could mean that she makes errors when attempting target words with these parameters; alternatively, it could mean that the sample included few words with non-initial stress and/or clusters. A full interpretation of the results should include analyses of the target words in the sample: by calculating average word complexity and determining the number of occurrences of each of the complexity parameters, the two samples can be compared. This analysis appears in Table II.

Analyses of the target words provide two kinds of information. First, they allow us to determine the complexity of the words attempted by the child. In the very early stages of

Table II. WCM scoring of target words.

Word	IPA	> 2 syl	Stress	Cluster	Final C	Velar	Fricative, affricate	Vcd fr, affric	Liquid, rhotic V	Total points
Airplane	ɛrpleɪn			1	1				2	4
Apple	æpl								1	1
Baby	beɪbi									0
Ball	bɔl				1				1	2
Bear	bɛr				1				1	2
Bed	bɛd				1					1
Bird	bɝd				1				1	2
Blue	blu			1					1	2
Book	bʊk				1	1				2
Bread	brɛd			1	1				1	3
Butterfly	bʌɾɚflaɪ	1		1			1		2	5
Chicken	tʃɪkɪn				1	1	1			3
Circle	sɝkl					1	1		2	4
Clown	klaʊn			1	1	1			1	4
Coat	koʊt				1	1				2
Crying	kraɪ.ɪŋ			1	1	2			1	5
Daddy	dædi									0
Duck	dʌk				1	1				2
Elephant	ɛləfɪnt	1		1	1		1		1	5
Eyes	aɪz				1		1	1		3
Flower	flaʊwɚ			1			1		2	4
Foot	fʊt				1		1			2
Frog	frɔg			1	1	1	1		1	5
Glasses	glæsɪz			1	1	1	2	1	1	7
Green	grin			1	1	1			1	4
Hurt	hɝt				1		1		1	3
Monkey	mʌŋki					2				2
Puzzle	pʌzl						1	1	1	3
Rabbit	ræbɪt				1				1	2
Sad	sæd					1	1			2
Sleeping	slipɪŋ			1	1	1	1		1	5
Slide	slaɪd			1	1		1		1	4
Snack	snæk			1	1	1	1			4
Spider	spaɪdɚ			1			1		1	3
This	ðɪs				1		2	1		4
Turtle	tɝdl								2	2
	Totals	2	0	15	25	15	19	4	28	av: 3.00

lexical acquisition, these words tend to be phonologically simple, predominated by forms like *mommy*, *daddy*, *baby*, *hi*, and *bye*. Of these target forms, only the word *hi* would receive a complexity point (for /h/). The table shows that K, a typically-developing toddler of 22 months, attempted words with varying complexities, ranging from simple forms receiving no points like *baby* and *daddy* to a word such *glasses* which received 7 points. The average complexity for the target words in this sample is 3.00. Table I shows that the average complexity of K's productions was 1.27, much lower than the target forms. A difference of this magnitude is not unexpected for a 22-month-old child who is acquiring a range of syllable structures and beginning to produce consonants from later-developing sound classes. Second, the analyses allow us to compare the number of occurrences of particular syllable/word patterns and sound classes in the two samples. As noted above, K's productions

Table III. Number of occurrences of each complexity parameter in K's sample, number of opportunities in the target words, and ratio of occurrences to opportunities.

Parameter	K's sample: # occurrences	Target words: # opportunities	Ratio
> 2 syllables	1	2	.50
non-initial stress	0	0	—
clusters	1	15	.07
final cons	21	25	.84
velars	7	15	.47
fric/affric	13	19	.68
vcd fric/affric	2	4	.50
liquid/rhotic V	1	28	.04

included at least one exemplar of each complexity parameter with the exception of non-initial stress. The data in Table II show that the sample did not include any target words with non-initial stress; hence we have no way of determining her ability to produce words with this stress pattern.

A comparison of the two data sets appears in Table III, showing the number of occurrences of each parameter in K's sample and number of opportunities for occurrence in the target sample; the table also provides a ratio comparing overall occurrences of each parameter in the two samples. It should be noted that the ratios are not based on word-by-word comparisons, but on the total numbers of occurrences of each parameter in the samples.

Table III shows that the complexity of K's productions lies primarily in the use of final consonants, a syllable structure parameter, and in the occurrence of fricatives and affricates, a sound-class parameter. Both parameters mirror aspects of the target words with ratios of occurrence of .84 and .68, respectively. K's productions differ from the targets dramatically in the use of clusters and liquids/rhotic vowels: the target words contained 15 clusters and 28 liquids/rhotic vowels; in spite of numerous opportunities, K produced only one cluster and one liquid. The use of velar consonants appears to be emerging in K's speech with a ratio of occurrence of .47. Given that the target data set included only two multisyllabic words, no words with non-initial stress, and a limited number of voiced fricatives and affricates, it is not possible to make definitive statements about K's abilities to produce these parameters.

Outcomes of the WCM applied to seven samples of child speech and target words

Table IV shows the application of the WCM to seven speech samples, five from children with typical development and two from children with phonological delay or disorder. The children range in age from 16–48 months and the sample sizes vary from 35–100 words. For the youngest children, aged 17 months, the productions represent their total vocabulary. Several of the data sources included variant pronunciations of some words; for example, 'Je' produced the word *juice* as [dʒu] and [dʒus], and the word *rock* as [wa] and [wati]; in these cases the most common variant appeared in the sample. (A comparison of three data sets with and without variant pronunciations indicated that differences in overall complexity were relatively minor; for example, the average word complexity of Je's sample including variant pronunciations was .67; without variants, the average was .63.)

Table IV presents the WCM analyses for the child productions and target words along with ratios of these two measures, the ranges of word complexity for the child and target, and the number of words with zero points in each data set. Except for 'F', age 48 months, the

Table IV. Application of the WCM to child productions and target words.

		WCM			WCM: range		% words with 0 points	
Source*	Sample # words	Child	Target	WCM ratio Child/Target	Child	Target	Child	Target
Typical development (child: age)								
S: 17 mos	35	.60	1.66	.36	0–3	0–5	60	29
D: 17 mos	35	1.02	1.80	.57	0–3	0–5	51	26
Je: 17 mos	40	.63	1.78	.36	0–4	0–5	55	15
Jo: 21 mos	45	.93	1.78	.52	0–3	0–5	40	11
K: 22 mos	65	1.49	2.89	.52	0–5	0–6	26	9
Phonological delay/disorder (child: age)								
Sm: 21 mos	100	.22	2.08	.11	0–3	0–6	83	12
F: 48 mos	43	.66	3.49	.19	0–4	0–8	51	2

* S, D (Stoel-Gammon and Dunn, 1985); Je (Ingram, 1981); Jo (Velten, 1943); K (Bleile, 1991); Sm (Ingram and Ingram, unpub ms.); F (Williams, 2003).

children's productions come from diary studies or conversational interactions. Among children with typical development, the analyses show that average word complexity ranges from .60–1.49 and tends to increase with age. The last column in the table compares the proportion of 'simple' words: the proportions for children with typical development vary from 9–29% for the targets and 26–60% for the children's productions; in both cases, the proportion of non-complex words declines with age.

The measures for 'D' show a high level of complexity relative to the other children of about the same age. A closer examination of his data set reveals two important factors regarding his WCM measures: (1) D's target words include a high number of velars, indicating a pattern of lexical selection; specifically, there were 13 occurrences of velar consonants in the target forms. In his productions of these words, D often used the process of Velar Assimilation producing the targets *duck*, *whack*, *frog* as [gak]. As a consequence, D's sample included 17 occurrences of velars, four more than occurred in the target words. Given that each velar consonant received one complexity point, D's average word complexity was increased by his use of consonant assimilation. This example serves as a reminder that multiple aspects of a child's sample must be considered in any analysis.

The measures for the two children with phonological delay/disorder present profiles that differ extensively from those of children with typical development. Notably, the WCM ratios, an indication of how closely the complexity of the child's productions mirrors the complexity of the targets, reveal a wide discrepancy between the two groups. In the case of 'Sm', a boy with phonological delay, average word complexity is .22, compared with .60–1.49 for younger children with typical development, while the average complexity of his target words is 2.08, similar to those of the children with typical development. Within Sm's sample, 83% of his productions received no points, meaning they were V or CV syllable/word structures, composed of vowels and stops, nasals or glides. The WCM ratio comparing the complexity of Sm's productions with that of his target words was .11, well below the ratios for children in the typically-developing group. In terms of parameters, Sm produced velar consonants (12 productions in 100 words), final consonants (five productions), and a few words with more than one syllable and/or with non-initial stress. Overall, however, Sm's profile is very different from those children with typical phonological development.

The measures for F, a 4-year-old with a phonological disorder (Williams, 2003), are based on a sub-set of words from an articulation test (Goldman and Fristoe, 2000) designed to assess the phonemes of English in various word positions. As with other articulation tests, the average complexity of the target words is quite high: 3.49. Here again, the WCM analyses show a large discrepancy between the average WCM of the child (.66) and the target (3.49), with a ratio of .19. Of note, 51% of F's productions received no complexity points, whereas only 2% (one word) of the target words received no points.

Preliminary evaluation of the WCM as an assessment tool

A new measure of phonological assessment should provide information that (1) is useful for describing/assessing phonologies, and (2) is not available from existing measures. As noted above, the WCM provides three types of information based on independent analyses of transcribed data. First, it measures the phonetic complexity of a word (or utterance, if the target cannot be identified) and assigns a 'complexity index' for each word; it then provides an average complexity of all words in the sample. This number can be used to compare complexity across children of the same age or language level, or to track changes in complexity within a child over time. Second, the WCM provides a measure of the use of eight complexity parameters across three domains: word patterns (number of syllables; stress placement); syllable structures (clusters; word-final consonants); and sound classes (velars; fricatives and affricates; liquids). These measures yield qualitative information regarding the sources of phonetic complexity in a child's productions. Third, it measures the complexity of target words in the sample and describes the extent of occurrence of each of the eight parameters. This information is helpful in interpreting the measures derived from analyses of child productions.

Are there other measures that provide the same information? There are many analyses that yield a single number as a measure of a child's phonology. The most common of these is a relational analysis, the Percentage of Consonants Correct (PCC) (Shriberg, Austin, Lewis, McSweeny, and Wilson, 1997). This measure compares the child's production and the target form on a phoneme-by-phoneme basis and provides a single number reflecting the proportion of consonants produced correctly. As noted in the introduction, relational analyses may not be the most appropriate approach to assessing the phonology for young children. In particular, measures such as the PCC are limited to analysis of a single domain, accuracy, and cannot be used with unintelligible speech.

Ingram (2002) proposed a 'whole-word' measure, the 'Phonological Mean Length of Utterance' (PMLU), in which each word receives points based on two factors: (1) the number of segments in a word, an independent analysis (each consonant and vowel receives one point), and (2) the number of consonants correct, a relational analysis (each correct consonant receives one point). The point total for each word is the sum of the two measures. This measure provides some information beyond traditional relational analyses and, as with MLU, yields a single number that may be useful in establishing developmental norms. It is not designed, however, to provide qualitative information about the nature of a child's phonological system.

Single number measures based on independent analyses have also been used to compare groups of children (e.g. Paul and Jennings, 1992; Rescorla and Ratner, 1996; Carson, Klee, Carson, and Hime, 2003). In most cases, these analyses present the number of different consonants and/or number of syllable shapes in a speech sample. Inventory-based approaches such as these have shown differences between the phonologies of young children with typical development from those with speech or language delay. In some cases, the single number

measure is accompanied by a list of phones that occur frequently in the samples of children in the two groups. For example, in their study comparing children with normal language with those with specific expressive language impairment, Rescorla and Ratner presented phonetic inventory counts (the number of different consonants and vowels) and listed the consonantal phones and the number of children in each group that produced each phone. This latter analysis provides some qualitative information on the children's phonological systems, but is not as broad as the WCM, which includes information on three 'levels' of the phonological system: word patterns, syllable shapes, and sound classes.

One concern about inventory measures is their temporal 'stability' (i.e. test–re-test reliability) across samples. In a recent study, Morris (2009) gathered two 20-minute samples within a 1-week period from 10 children aged 18–22 months; she then transcribed all meaningful productions, performed a series of independent analyses, and evaluated the test–re-test correlations of five phonetic measures: (1) number of different consonants in word-initial position; (2) number of different consonants in word-final position; (3) number of different word shapes; (4) syllable structure level (based on Stoel-Gammon, 1989; Paul and Jennings, 1992); and (5) the index of phonetic complexity (Jakielski, Maytasse, and Doyle, 2006). This latter measure is similar to the Word Complexity Index described in this paper and is discussed below.

Morris (2009) reported that the number of initial consonants was the least stable of the five measures, with a relatively low test–re-test correlation ($\rho = .226$); correlations for measures of final consonants and word shapes were moderate, and those for syllable structure level and phonetic complexity were high. The only correlation that reached statistical significance was syllable structure. The lack of test–re-test stability is likely to be a problem for all analyses based on spontaneous speech samples, including relational measures such as PCC, because of differences in the words across the two samples. This is an aspect of phonological assessment that should receive attention in future studies.

The WCM measure shares several features with the Index of Phonetic Complexity (IPC) proposed by Jakielski et al. (2006). To date, the only descriptions of the measure come from a conference poster (Jakielski et al., 2006) and the study by Morris (2009) cited above. Like the WCM, the IPC is based on a system in which words are awarded points based on the presence of complex parameters. Several parameters appear in both the ICP and the WCM. For example, both measures award points to words of more than three syllables and to final consonants and to fricatives and liquids. The two approaches also differ in several ways. For instance, the IPC gives a point to each word with place variegation for single consonants; thus, the word *doggie* produced as [dagi] would receive a point for a velar consonant in both measures, but the IPC would add one more point for place variegation (the occurrence of singleton consonants with two places of articulation). In a similar vein, both measures award points for consonant clusters, but the IPC awards an additional point for place variegation within a cluster, e.g. [st-] receives one point while [sp-] receives two. Finally, the WCM awards a point for non-initial stress; the IPC does not consider stress. The IPC, as presented in 2002, does not provide a comparison of the complexity of child sample and the target words or include information data on the types of complexity occurring in a child's output.

Questions and future directions

As with any new measure, there are several aspects of the WCM that deserve further research. First, are the eight parameters involved in the measure the most appropriate? The measures were selected based on previous investigations of early phonological development of English; they incorporate a variety of word, syllable, and sound class parameters that are expected to

change over time. One strength of the measure is that it is easy to modify—if a researcher or clinician were interested in different parameters, in vowels for example, the WCM could be structured so that points are awarded for particular vowel types. The same type of modification could be made if the focus were on place variegation as in the IPC described above. As long as both child and target samples are coded in the same way, comparisons across data sets can be made.

An unresolved issue for the WCM is the optimal sample size. At this point, it is not possible to specify the number of words/utterances that are needed. Table IV presents WCM measures based on data from seven samples, with the number of words per sample ranging from 35–100; for these samples, only one pronunciation of each target word was included. One way to increase sample size is to include variant productions; if such pronunciations are included, I would suggest a maximum of two variants for a given target. It is interesting to note that, in spite of the differences in number of words, the measures in Table IV show a general increase in complexity with age. One of the target populations for the WCM is late-talking children with limited vocabularies; the summaries presented in Table IV suggest that a relatively small sample will yield useful quantitative and qualitative information for this population. Of course, if a child's productions are restricted to non-complex forms, as defined by the WCM, the average complexity will be zero, regardless of the number of words in the sample.

In sum, this paper presents a new assessment tool based on independent analyses of word complexity. The tool yields information about the production patterns of young children with typical development and of children with speech and/or language delay. This tool should prove useful to both researchers and clinicians interested in assessing phonological development and disorders.

Declaration of interest: The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

References

Bleile, K. (1991). *Child Phonology: A Book of Exercises for Students*. San Diego: Singular Publishing Group, Inc.

Carson, C., Klee, T., Carson, D., & Hime, L. (2003). Phonological profiles of 2-year-olds with delayed language development: Predicting clinical outcomes at age 3. *American Journal of Speech-Language Pathology*, *12*, 28–39.

Dyson, A. (1988). Phonetic inventories of 2- and 3-year-old children. *Journal of Speech and Hearing Disorders*, *53*, 89–93.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2*. Circle Pines: American Guidance Service.

Grunwell, P. (1987). *Clinical Phonology*. 2nd Ed. London: Croom Helm.

Ingram, D. (1981). *Procedures for the Phonological Analysis of Children's Language*. Baltimore: University Park Press.

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

Ingram, D., & Ingram, K. Frame-content theory as an account of early phonological acquisition. Unpublished manuscript.

Jakielski, K., Maytasse, R., & Doyle, E. (2006). Acquisition of phonetic complexity in children 12–36 months of age. *Poster presented at the convention of the American Speech-Language-Hearing Association*, Miami, FL.

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

Miccio, A., Elbert, M., & Forrest, K. (1999). The relationship between stimulability and phonological acquisition in children with normally developing and disordered phonologies. *American Journal of Speech-Language Pathology*, *8*, 213–214.

Morris, S. (2009). Test-retest reliability of independent measures of phonology in the assessment of toddlers' speech. *Language, Speech, and Hearing in Schools*, *40*, 46–52.

Paul, R., & Jennings, P. (1992). Phonological behaviors in toddlers with slow expressive language development. *Journal of Speech and Hearing Research*, *35*, 99–107.

Prather, E., Hedrick, D., & Kern, C. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders*, *40*, 179–191.

Rescorla, L., & Ratner, N. B. (1996). Phonetic profiles of toddlers with specific language impairment (SLI-E). *Journal of Speech and Hearing Research*, *39*, 153–165.

Robb, M., & Bleile, K. (1994). Consonant inventories of young children from 8 to 25 months. *Clinical Linguistics and Phonetics*, *8*, 295–320.

Shriberg, L., Austin, D., Lewis, B., McSweeny, J., & Wilson, D. (1997). The Percentage of Consonants Correct (PCC) metric: extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, *40*, 708–722.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, *55*, 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 in Schools*, *40*, 46–52.

Stoel-Gammon, C. (1989). Pre-speech and early speech development of two late talkers. *First Language*, *9*, 207–223.

Stoel-Gammon, C., & Dunn, C. (1985). Normal and Disorderd Phonology in Children. Austin, TX: Pro-Ed.

Tyler, A., Miccio, A., Hoffman, P., Norris, J., Hodson, B., Scherz, J., et al. (2002). Five views of the elephant: perspectives on the assessment of articulation and phonology in preschoolers. *American Journal of Speech-Language Pathology*, *11*, 213–214.

Velten, H. (1943). The growth of phonemic and lexical patterns in infant language. *Language*, *19*, 281–292.

Williams, L. (2003). *Speech Disorders Resource Guide for Preschool Children*. San Diego: Singular Press.