

Correlation of nonword repetition accuracy with vocabulary size and other measures of linguistic development in children with phonological disorders

Jan Edwards

Dept. of Speech and Hearing Science, Ohio State University

Mary E. Beckman

Dept. of Linguistics, Ohio State University

Benjamin Munson

Dept. of Communication Disorders, University of Minnesota

Please send inquiries to:

Jan Edwards

Dept. of Speech and Hearing Science

Ohio State University

Address until 7/15/03:

13 Al. Fleming St.

Thessaloniki 54642

Greece

Phone: 011 30 2310 839994

Email: [edwards.212@osu.edu](mailto:edwards.212@osu.edu)

Running title: Nonword repetition and phonological disorders

### Abstract

A growing body of research has documented effects of phonotactic probability on young children's phonological processing. This study extends this research in two ways. It compares nonword repetitions by 40 young children with phonological disorders with those by 40 age peers with typical phonological development and it also examines the relationship between the frequency effect in the nonword repetition task and other measures of phonological ability across a larger group of children. All children repeated low-frequency sequences less accurately than high-frequency sequences. The children with phonological disorders were less accurate overall, but showed no larger disadvantage for the low-frequency sequences. Across the larger group, the size of the frequency effect was correlated with vocabulary size, but it was independent of measures of speech perception and articulatory ability. These results support a view of phonological acquisition as the gradual development of multiple types of phonological representation, with robust higher-order symbolic categories (phonemes) emerging only later, as generalizations about sublexical sound patterns in the enlarging lexicon. Within this framework, phonological disorders seem to involve deficits at a more basic level of phonological representation.

Edwards, Beckman, and Munson (2003) examined the influence of phonotactic probability on nonword repetition accuracy in typically developing children of different ages and in adults. Their two major findings were that the number of real words in which a target phoneme sequence is attested influenced repetition accuracy and that there was an interaction between phoneme sequence frequency and vocabulary size. Low-frequency sequences were repeated less accurately, and the larger the vocabulary size, the smaller this effect of sequence frequency on accuracy. These results support a picture of the core of phonology as a set of emergent generalizations over the child's lexicon. If correct, this view of normal phonological acquisition has implications for our understanding of phonological disorders in children. Perhaps the age-inappropriate speech production of these children is the result of some difficulty in making generalizations about the sound structure of known lexical items. If so, this difficulty might involve the same kind of phonological knowledge that is involved in the interaction between vocabulary size and sequence frequency in Edwards et al. (2003). Alternatively, since children with phonological disorders do not characteristically have smaller vocabularies than their typically developing peers, the difficulty might involve a different kind of phonological generalization, one that is less affected by vocabulary growth.

In other words, to characterize phonological disorder within this framework, we need to know what *kind* of phonological knowledge is implicated in the frequency effects documented in Edwards et al. (2003). Acquiring the phonology of a spoken language clearly requires the child to establish highly detailed auditory representations of the sound patterns that make up the words and larger utterances of that language. It also requires the child to establish detailed articulatory representations that can be robustly generalized over all of the different segmental and prosodic contexts in which the child might need to produce any given sound pattern. The increase in repetition accuracy with vocabulary growth might simply reflect more robust generalizations in one or both of these two primary phonological domains. That is, the children with larger vocabularies might have been more accurate in repeating the nonword stimuli simply because their acoustic representations of the component sound patterns are better developed from being more robustly generalized over their experience of hearing these patterns in many different real words. Alternatively, their motor representations of the component constellations of articulatory gestures might be better developed, from more experience with producing them in many different real words. Children with phonological disorders, then, might simply have difficulty in one or both of these primary phonological domains and these difficulties might account for their phonological disorders. Such an account of phonological disorder would be consistent with views of phonological disorder as stemming from deficits in articulatory-motor and/or acoustic-auditory representations (e.g., Edwards, Fourakis, Beckman, & Fox, 1999; Forrest, Chin, Pisoni, & Barlow, 1995; Gibbon, 1999).

However, there is an alternative account of what kind of knowledge it is that develops with the growth of the lexicon in early childhood. It could be that, as the child acquires more and more words containing the same consonants and vowels in different novel combinations, a layer of symbolic units (consonant and vowel phonemes) emerges. This layer would intervene between the more primary acoustic and articulatory representations in the same way that symbolic units (words) stand between the phonological form and the semantic interpretation of a sentence (see Figure 1). It would supplement the "forward model" (Bailly, Laboissiere, & Schwartz, 1991; Jordan, 1990) that the infant develops in babbling, to allow for a "fast mapping" of the phonological structure of a new word that is comparable to the fast mapping of its morphology and syntactic class (Carey, 1978). Plaut and Kello (1999) have proposed such a

model in which higher-order phonological generalizations emerge as stable activation patterns in a cluster of hidden nodes at the center of a triangular recurrent network connecting auditory and articulatory representations of words to each other as well as to their associated semantic representations. These symbolic phonological units facilitate the inverse mapping between the auditory representation of a new word that is spoken to the child and a fluent articulatory representation for the child to use in repeating the word in a subsequent utterance. Developing these hidden nodes allows a new word to be parsed as a novel string of phonemes, which maps onto a set of familiar gestural constellations, which then can be flexibly re-assembled into a fluently coarticulated whole-word motor score that the child has never produced before.

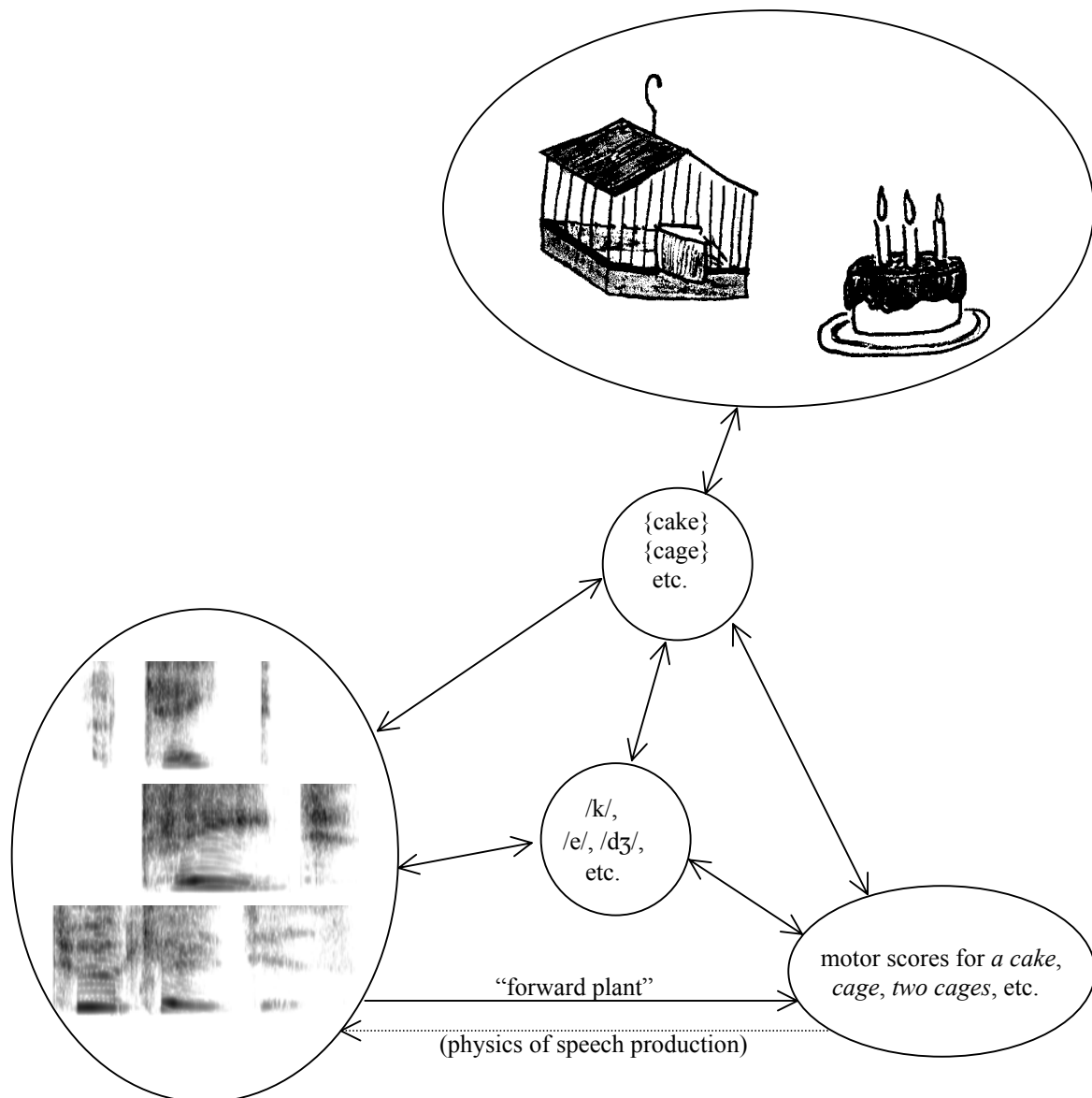


Figure 1. Relationships among three types of phonological representation that are posited to be associated with words in the lexicon.

Within this second account of the relationship between phonological development and vocabulary growth, children with phonological disorders might have difficulty with building this intervening layer of representations. While this view differs from accounts of phonological disorder as a deficit in higher-level phonological rules or constraints (e.g., Barlow & Dinnsen, 1998; Barlow & Gierut, 1999; Bernhardt, 1994; Bernhardt & Stoel-Gammon, 1996; Ingram, 1976; Shriberg & Kwiatowski, 1994), it is similar to these accounts in that phonological disorder is seen as a deficit in higher-level representations instead of (or in addition to) difficulty with the primary articulatory and auditory representations.

The focus of this paper is twofold. First, this paper seeks to distinguish between the two explanations of the interaction between vocabulary size and the effect of phonotactic probability that are available within the view of phonology described above. Second, within this framework, this paper examines the phonological knowledge deficits exhibited by children with phonological disorder through their performance on the same nonword repetition task used by Edwards et al. (2003).

For the first purpose of this paper, there are several sources of evidence that can be used to distinguish between the two explanations of the interaction between vocabulary size and the phonotactic probability effect. In general, if this interaction reflects developmental changes in the primary articulatory and acoustic representations — i.e., a finer and finer tuning of these representations for the patterns that occur in many words that the child knows — then we would expect the sequence frequency effect to be correlated with other measures of the development of robust articulatory and auditory representations for speech production and perception. On the other hand, if the interaction reflects the development of more robustly abstracted symbolic units as the child acquires more and more words, then we would expect the effect to be uncorrelated with any other differences between younger and older children that are not also related to vocabulary growth, such as maturational changes in motor control or other aspects of cognitive development. Therefore, in this paper, we examine the relationship between the effect of target sequence frequency on repetition accuracy and several independent measures of speech perception and articulatory ability.

Disentangling the effects of general articulatory ability is also important for another reason. It might be that children produce low-frequency sequences less accurately simply because they contain low-frequency phonemes that are inherently more difficult to produce (e.g., Ingram, 1988; Stemberger & Bernhardt, 1999; Winitz, 1969). Edwards et al. (2003) avoided very late-acquired sounds, such as /ɪ/ and /tʃ/ (Smit, Hand, Freilinger, Bernthal, & Bird, 1990) in the construction of the nonword stimuli. However, it was impossible to match the paired high-versus low-frequency target sequences for component phoneme frequency. By definition, a high-frequency sequence contains high-frequency phonemes and low-frequency sequences tend to contain low-frequency phonemes. If we compare each child's articulatory accuracy at producing the same phonemes in familiar real words and in the high-frequency and low-frequency target sequences, then we can determine whether phonemes in low-frequency sequences were produced less accurately simply because they were low-frequency, inherently difficult phonemes or because they were part of low-frequency sequences.

For the second purpose of examining phonological disorder within the conceptualization of typical phonological development proposed by Edwards et al. (2003), evidence can begin with looking at how children with phonological disorders perform on the nonword repetition task relative to age peers with typical phonological development. Other evidence regarding the characterization of phonological disorders within this framework can come from looking at

interactions between the two groups' performance on this task and on measures of phonological knowledge in the primary auditory and articulatory domains on which the groups might differ.

There is already considerable evidence that children with phonological disorder have subtle perceptual deficits and fine-grained motor problems that give rise to less robust phonological knowledge in these primary auditory and articulatory representational domains. For example, a number of studies have shown that children with phonological disorders have difficulty in discriminating contrasts that they cannot produce (e.g., Hoffman, Daniloff, Bengoa, & Schuckers, 1985; Ohde & Scharf, 1988; Rvachew & Jamieson, 1989). A few studies have also shown evidence of more general perceptual problems in children with phonological disorders (Edwards, Fox, & Rogers, 2002; Forrest et al., 1995). For example, Edwards et al. (2002) found that children with phonological disorders had more difficulty than typically developing age peers in discriminating between CVC words that differed only in the identity of the final consonant, in both gated and ungated conditions. Furthermore, their difficulties were unrelated to whether or not they were able to produce final consonants.

There are also a number of studies that suggest that children with phonological disorders have less robust articulatory representations as well. For example, there is both acoustic and articulatory evidence that children with phonological disorders who do not produce a contrast between alveolar and velar stop consonants do not simply substitute /t/ for /k/ or /k/ for /t/ (Edwards, Gibbon, & Fourakis, 1997; Gibbon, 1990, 1999; White, 2001). Rather, many of these children seem to produce "undifferentiated lingual gestures" (Gibbon, 1999) in which /k/ and /t/ are merged into a centralized lingual gesture that may be perceived as /k/ or /t/, depending on various factors such as the vowel context and the lingual contact pattern during stop release.

A number of researchers have suggested that children with phonological disorders have higher-level phonological deficits in addition to these difficulties with the two primary domain-specific representations (e.g., Barlow & Dinnsen, 1998; Barlow & Gierut, 1999; Bernhardt, 1994; Shriberg & Kwiatowski, 1994). The results of an earlier study by Beckman and Edwards (2000) may provide additional support for this claim, within the view of phonological acquisition described above. The third study reported in Beckman and Edwards (2000) found that children with phonological disorder, like the younger children in Edwards et al. (2003), had a somewhat larger effect of target sequence frequency on accuracy than their typically developing peers. This is what we would expect if the interaction of vocabulary size with the effect of target sequence frequency reflects the progressive development of robust symbolic units standing between these two primary phonological representations *and* if children with phonological disorders had difficulty with developing this intervening layer of representation. However, only six pairs of children participated in that study and the apparent interaction between participant group and target sequence frequency was not statistically significant. In this study, therefore, we examined a much larger group of children with phonological disorders in relation to typically developing age peers.

Thus, two questions were addressed in this study. First, what kind of phonological knowledge is reflected in the interaction between the effect of phonotactic probability and vocabulary size observed in Edwards et al. (2003)? Is it the development of more finely-differentiated primary auditory and articulatory representations, or is it the development of an intervening layer of symbolic units that maps between these two primary representations? Second, within this framework, how can we describe the phonological knowledge deficits exhibited by children with phonological disorders relative to their age peers?

To address these questions, we began by comparing the performance on the nonword repetition task of 40 children with phonological disorders relative to a matched group of 40 typically developing children. The results of this initial comparison did not confirm the hypothesis suggested by the third study in Beckman and Edwards (2000). Although children with phonological disorders were less accurate overall, there was no difference between the two groups in the size of the effect of sequence frequency. Moreover, the type of errors that predominated in the productions by the children with phonological disorders was characteristic of errors in productions by the youngest children in Edwards et al. (2003). To interpret these results further, we did several further analyses in which a larger group of typically developing children was included in order to examine relationships between nonword repetition accuracy and several other measures of phonological knowledge across a larger range of ages, vocabulary sizes, and articulatory abilities.

## Method

### Participants

The participants were 40 children with phonological disorders (PD), ranging in age from three to six years, and 104 children with typical phonological development (TD), ranging in age from three to eight years. All children were participating in a larger study on the etiology of phonological disorder. The 104 children with TD are the same children who participated in the study described in Edwards et al. (2003). All of the participants, both the children with PD and the children with TD, met the following criteria: 1) normal hearing, as evidenced by passing a hearing screening at 20 dB at 500, 1000, 2000, and 4000 Hz; 2) normal non-verbal intelligence, as evidenced by a standard score no more than one standard deviation below the mean on the *Columbia Mental Maturity Scale* (CMMS, Burgemeister, Blum, & Lorge, 1972); and 3) normal structure and function of the peripheral speech mechanism, as evidenced by a standard score no more than one standard deviation below the mean on the oral movement subtest of the *Kaufman Speech Praxis Test for Children* (KSPT, Kaufman, 1995). (Two children with phonological disorders had standard scores below 85 on the KSPT but were still included in the study because their low scores were not related to problems with neuromotor control. One child received a 78, with errors related to a short lingual frenum, and the other child received an 81, with errors related to his refusal to touch his tongue to his lips because of severely chapped lips). All of the children with PD scored at or below the 10<sup>th</sup> percentile on the *Goldman-Fristoe Test of Articulation* (GFTA, Goldman & Fristoe, 1986), while all of the children with TD scored above the 16<sup>th</sup> percentile (one standard deviation below the mean) on the GFTA.

We also selected 40 children from the larger group of children with TD to serve as age controls for the children with PD. All of these age controls scored at or above the 40<sup>th</sup> percentile on the GFTA and were individually matched to the children with PD on the basis of age (within 6 months), gender, and non-verbal IQ score on the CMMS (within 10 points – i.e., twice the standard error of measurement). Since these two groups are defined by their GFTA scores, they will be termed “GFTA groups” in figures and tables that report group means.

In this study, as in Edwards et al. (2003), we examined the relationship between the effect of sublexical sequence frequency on accuracy and vocabulary size. We used the *Peabody Picture Vocabulary Test-III* (PPVT-III, Dunn & Dunn, 1997) to estimate receptive vocabulary size, and we used the *Expressive Vocabulary Test* (EVT, Williams, 1997) to estimate expressive vocabulary size. These two tests are co-normed for participants aged 2 through 90.

Table 1 gives descriptive information for the 40 children in each of the two GFTA groups. One-way ANOVAs revealed significant group differences on GFTA percentile rankings ( $F[1, 78] = 441.35, p < .001, \eta^2 = .85$ ) and also on PPVT-III standard scores ( $F[1, 78] = 7.54, p = .007, \eta^2 = .08$ ). However, the group difference on the PPVT-III was much smaller than the group difference on the GFTA; furthermore, all children with PD scored within normal limits (no more than one standard deviation below the mean) on both measures of vocabulary size. There were no significant differences between GFTA groups on CMMS or EVT standard scores.

Descriptive information on the larger group of children with TD is given in Table 2 of Edwards et al. (2003). Since we report results for this larger group primarily as a reference for comparing the effects of GFTA grouping to the effects of age and age-correlated measures such as vocabulary size, relevant subgroups for analyses that report group means are the different “age groups” in that table — namely, three- to four-year-olds, five- to six- year-olds, and seven- to eight-year-olds.

Table 1. Sample size and the number of males in the two GFTA groups, along with mean age, age range, and test scores (with standard deviations in parentheses).

	Children with PD <sup>c</sup>	Children with TD <sup>d</sup>
Age in months	57 (9)	58 (10)
Age range in months	40 - 76	39 - 75
Gender	27 male	27 male
GFTA percentile ranking <sup>a, b</sup>	5 (3)	69 (19) *
CMMS standard score	107 (11)	109 (11)
EVT standard score <sup>b</sup>	105 (13)	110 (11)
PPVT-III standard score <sup>b</sup>	106 (12)	113 (11)*

<sup>a</sup>GFTA percentile rankings below the first percentile were assumed to be .5 for purposes of computing mean and standard deviation.

<sup>b</sup>\* indicates that scores were significantly different between the two groups,  $p < .05$ .

<sup>c</sup>PD is phonological disorder; <sup>d</sup>TD is typical phonological development.

### Nonword repetition

A detailed description of the stimuli for the nonword repetition task, the procedure for data collection, and the methods used for transcribing the data and for scoring segmental accuracy is provided in Edwards et al. (2003). Identical procedures were used to collect and transcribe the data of the children with PD. A single researcher transcribed all of the responses of the children with PD (she was the same transcriber who had transcribed the responses of the children with TD). A second transcriber independently transcribed 15 percent of the data from the children with PD (two three-year-olds, two four-year-olds, two five-year-olds). Phoneme-by-phoneme inter-rater reliability ranged from 87 to 92 percent for data from individual participants, with a mean of 89 percent across these six children with PD. Inter-rater reliability for the children with TD is given in Edwards et al. (2003).

Segmental accuracy was scored as in Edwards et al. (2003). Briefly, each target sequence could receive a maximum of six points, or three points per component phoneme. Consonants received one point each for correct place, correct manner, and correct voicing, while vowels received one point each for correct front-back position, correct vowel height, and correct “length” (i.e., tense or lax for a monophthong target and monophthong or diphthong for a

diphthong target). In addition to the segmental accuracy score, we also coded a “prosody score” for each response of each participant. For the prosody score, the response was coded either as correct or as incorrect. A response was considered to be prosodically correct when the participant correctly produced the target prosody, even if there were feature-changing substitutions for one or both of the phonemes in the target sequence (e.g., if the child produced /buskit/ for /bufkit/). A response was considered to be prosodically incorrect when the participant changed the prosodic position of one or both of the target phonemes (e.g., by deleting the consonant in a VC sequence to make the vowel the nucleus of an open syllable, such as /itən/ for /iptən/, or by inserting a vowel in a medial CC sequence to make the first consonant an onset rather than a coda, such as /bufəkɪt/ for /bufkɪt/).

Productions that were coded as not completely correct on the segmental accuracy score, but correct on the prosody score could then be identified as “feature substitution” errors. Producing a high proportion of prosody-changing errors would suggest that the child has less well-developed articulatory representations of the fine temporal coordination needed to produce the precisely coarticulated VC transitions of a closed syllable rime or the precisely overlapped constrictions of a word-internal consonant cluster. We were interested in whether there was any effect of GFTA group or age group on error type. That is, we wanted to determine whether children with phonological disorders relative to age peers (or younger typically developing children relative to older children) produced more errors that simplified the prosodic structure. Such a finding would suggest that they might not have robustly abstracted representations of individual consonant and vowel phonemes away from the preferred CV syllable structure. We were also interested to know whether this proportion was primarily related to the GFTA group (or age group) of the speaker or also differed between low- and high-frequency sequences.

#### Speech perception ability

All of the children had also participated in a speech perception task in which they were asked to choose between two CVC words which differed only in the identity of the final consonant in both ungated and gated conditions (Edwards et al., 2002). (Due to technical problems, data were missing for two children with TD and two children with PD.) We chose a single measure from this experiment, the d-prime values from the ungated condition for two different stimulus sets (averaged together), as an indicator of the participants’ ability to attend to fine phonetic detail in speech perception. Edwards et al. (2002) had selected this measure for use in an examination of the relationships among speech perception, age, vocabulary size, and articulatory ability because TD participants in all age groups and the children with PD performed above chance in the ungated conditions. They had found that this measure was correlated with both receptive vocabulary size, as measured by PPVT-III raw score, and with articulatory ability, as measured by GFTA raw score.

#### Articulatory ability

Two additional measures of phonological ability were calculated for each child. One was the raw score on the GFTA and the other was the number of times each target sound was produced correctly on a non-standardized Phonetic Inventory test (PI). The PI is a picture-naming task that we developed to provide a more detailed probe of articulation skills than the GFTA (Isermann, 2001). Both the GFTA and the PI use colored line drawings of familiar words to elicit single word productions of each consonant of English in initial, medial, and final word positions. The GFTA elicits each consonant in English one time in each word position and also elicits initial consonant clusters. The raw score on the GFTA is the number of errors produced. The PI provides a more in-depth assessment of a child’s articulation skills than the GFTA because it

elicits each target consonant in three different words for each target position and it also assesses vowel production. The GFTA was given to all participants, both the children with PD and the children with TD. The PI was given to all children with PD and to the 81 children with TD who were six and under, but it was not given to the 23 children who were seven or older. (We reasoned that the children with TD who were seven or older were likely to produce very few articulatory errors.) Because of recording problems, data from the PI were lost for seven of the children with TD and two children with PD. No data were lost from the GFTA.

The same researcher who had transcribed the nonwords also transcribed the words elicited with the GFTA, using the same methodology for both sets of words. The recordings were transferred from the DAT to a digital file on a PC and the productions were transcribed in the International Phonetic Alphabet at the level of a careful, broad phonemic transcription. The transcriber used a waveform editor, so that she could easily listen to each word as many times as necessary and could also examine the spectrogram in case of doubt. Another researcher transcribed the words elicited with the PI for all of the children, using this same method. The first researcher independently transcribed the words elicited with the PI for a subset of the children comprising two three-year olds, two four-year olds, and two five-year olds, with one child from each of the two GFTA groups for each of these three ages. The second researcher also independently transcribed the words elicited with the GFTA for these same six children. For the GFTA words, phoneme-by-phoneme inter-rater reliability for presence/absence of error ranged from 91% to 97%, with a mean of 95%. For the PI words, it ranged from 89% to 95%, with a mean of 92%.

#### Assessing accuracy of individual phoneme segments

The PI was originally designed as a supplement to the GFTA, which would allow us to assess whether incorrect production of any target phoneme on the GFTA was habitual. That is, one of our goals in eliciting the targets as many times as we did was to assess whether children were consistently incorrect (or consistently correct) across tokens of the same segment in different words (Isermann, 2001; Martin, 1989). Each production of a target segment was scored as correct or incorrect, and number of correct productions added to get an accuracy rate for that segment. Vowels were elicited in stressed syllables of three words and the consonants /h/, /j/, and /w/ were elicited in word-initial position in three words, yielding a maximum accuracy rate of 3. Nearly all of the other consonants were elicited in three different positions (initial, medial, and final) yielding a maximum accuracy rate of 9. (The only exceptions were /θ/, which was not elicited in word-medial position, and /ð/, which was not elicited in word-final position.)

In the current study, we used the productions elicited with the PI to assess whether differences in the accuracy scores for phonemes on the nonword repetition task were due to the phonotactic probabilities of the sequences in which they occurred in the nonword stimuli or to something intrinsic to the phonemes themselves. To make this assessment, we calculated total segmental accuracy scores for the segments in the high- and low-frequency target sequences on the nonword task, and we converted the accuracy rates for the same phonemes elicited as targets on the PI into weighted total scores. Since we expected vowels to be generally more accurate than consonants, we calculated separate totals for the consonants and for the vowels, as well as separate totals for the segments that were in the high-frequency target sequences and those that were in the low-frequency target sequences.

For the nonword task, these four total segmental accuracy scores were simply the sums of the accuracy scores for individual phonemes in each category. As described above, three points was the highest segmental accuracy score that a target phoneme could receive on the nonword task.

Therefore, since there were 14 stimulus pairs with CV or VC targets, the maximum total accuracy score possible for the vowels in the high- versus low-frequency sequences is 42 — i.e., 14 sequences containing one vowel each scoring a maximum of 3 points. Similarly, since there are 8 stimulus pairs with CC targets in addition to these 14 CV or VC targets, the maximum total accuracy possible for the consonants in the high- versus low-frequency sequences is 90 — i.e., the 16 consonants in the eight CC sequences plus the 14 consonants in the CV and VC sequences each scoring a maximum of 3 points.

To make comparable totals for the same phonemes on the PI, we needed to weight the accuracy rates for the target phonemes elicited in the real words by the number of times each sound occurred in the target sequences in the nonword stimuli. To do this, we treated the number of correct productions of the target phoneme on the PI as an analog to the three-point segmental accuracy score on the nonword repetition task, and multiplied this “accuracy score” by the number of times the segment occurred in the relevant target sequence category in the nonword stimuli. For example, the consonant /f/ appears in three low-frequency target sequences on the nonword repetition task in the stimuli /auftəgə/, /bufkit/, and /næfkətu/. Therefore, we multiplied the PI accuracy score for /f/ by 3 in calculating a child’s total PI accuracy for consonants in low-frequency sequences. Similarly, the vowel /æ/ occurs in two high-frequency target sequences in the stimuli /mæbəp/ and /bedæg/. Therefore, we multiplied the PI accuracy score for /æ/ by 2 in calculating a child’s total PI accuracy for vowels in high-frequency sequences.

As described above, the PI accuracy score for vowels ranged from 0 to 3 because each vowel was elicited three times in an inter-consonantal position. The total possible PI accuracy score for the vowels in each of the two sequence types therefore is 42, the same as for the total accuracy score for the vowels in the nonword task.

The consonant types that were used in the 8 CC and 14 CV and VC sequence pairs were /p, t, k, b, d, g, m, n, f, v, j, w/. For most of these consonants, the maximum number of correct productions was 9, for the three elicitations of each in word-initial, medial, and final position. The two exceptions were /w/ and /j/, which were elicited only in word-initial position. We multiplied the accuracy scores for these two consonant types by 3 before multiplying each consonant type by the number of times it occurred in the target sequences so that the accuracy scores for /w/ and /j/ would be weighted similarly to the accuracy scores for all of the other consonant types. The highest possible PI accuracy score for consonants was thus 270 — i.e., 9 (for the maximum adjusted PI score) \* 30 (for the 16 consonants in the 8 CC sequences and the 14 consonants in the CV and VC sequences).

## Results

### Nonword accuracy scores for children with PD relative to TD age peers

The first set of analyses compares children with PD to their TD age peers. In particular, we wanted to determine whether sublexical sequence frequency had a greater effect on repetition accuracy for children with PD than for their age peers. Figure 2 shows mean accuracy scores for high- and low-frequency sequences for these two groups of participants. A two-way (frequency by group) repeated measures ANOVA revealed significant effects of frequency ( $F[1,78] = 85.81$ ,  $p < .001$ ,  $\eta^2 = .52$ ) and group ( $F[1,78] = 58.32$ ,  $p < .001$ ,  $\eta^2 = .43$ ). The frequency by group interaction was not significant. That is, children with significantly below age-level articulatory ability produce low-frequency sequences less accurately than they do high-frequency sequences, but they are no different from their typically developing age peers in the size of this effect.

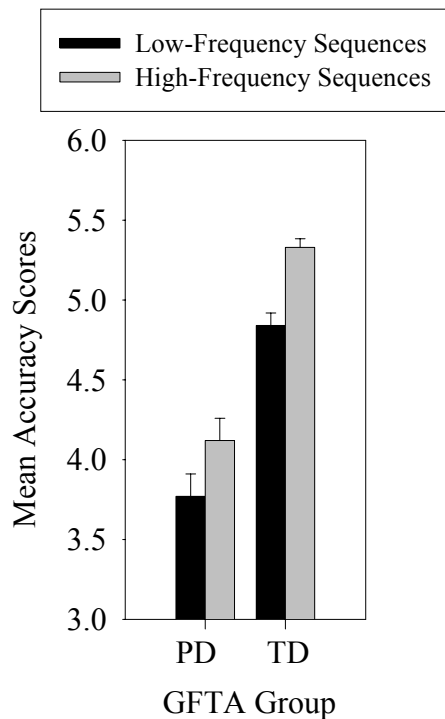


Figure 2. Mean accuracy scores (with error bars) for the low- and high-frequency sequences for children with PD and their TD age controls.

These results suggest that sublexical sequence frequency affects accuracy in the same way for children with PD and their TD age peers. We can contrast this to the results of Edwards et al. (2003), which showed that the effect of sequence frequency is different for children in different age groups. The youngest children, who had the lowest accuracy overall, showed the largest effect of sequence frequency, whereas the oldest children, who had highest overall accuracy, showed only a small difference in between high- and low-frequency sequences. In the current study, the children with PD were less accurate overall, as is to be expected from their GFTA scores, but they were not disproportionately less accurate on the low-frequency target sequences.

On the other hand, it is possible that children with PD still do differ from their TD age peers, by making more prosodic errors on low-frequency sequences relative to their error rate on high-frequency sequences. Therefore, the effect of group and sequence frequency on prosody-changing errors was the focus of the next analysis. The left-hand panel of Figure 3 shows the mean number of errors that were substitution errors for low- versus high-frequency sequences for the two groups of participants. The children with PD made more errors that changed the prosody of the sequence for both the low- and the high-frequency target sequences, whereas the children with TD made proportionally more errors that are mere feature-changing substitutions. A two-way mixed model ANOVA with sequence frequency as the within-subjects factor, and GFTA group as the between-subject factor showed a significant effect of participant-group ( $F[1,75] = 13.5$ ,  $p < .001$ ,  $\eta^2 = .15$ ), but no significant effect of target sequence frequency, and no interaction between sequence frequency and group.

The right-hand panel of Figure 3 shows this same analysis for the three age groups in the larger set of 104 children with TD. The youngest typically developing children, like the children

with PD, made proportionally more errors that change the prosody of the sequence relative to the older children. A two-way mixed model ANOVA yielded similar results. There was a significant effect of age ( $F[3,120] = 5.487, p = .001, \eta^2 = .12$ ), but no significant effect of target sequence frequency, and no interaction. Thus, the youngest children relative to the older children performed similarly to the children with PD relative to their TD age peers. Both the youngest children and the children with PD produced more prosody-changing errors relative to their comparison groups, but the proportion of prosodic errors was not influenced by phonotactic probability in either of these groups or in their comparison groups.

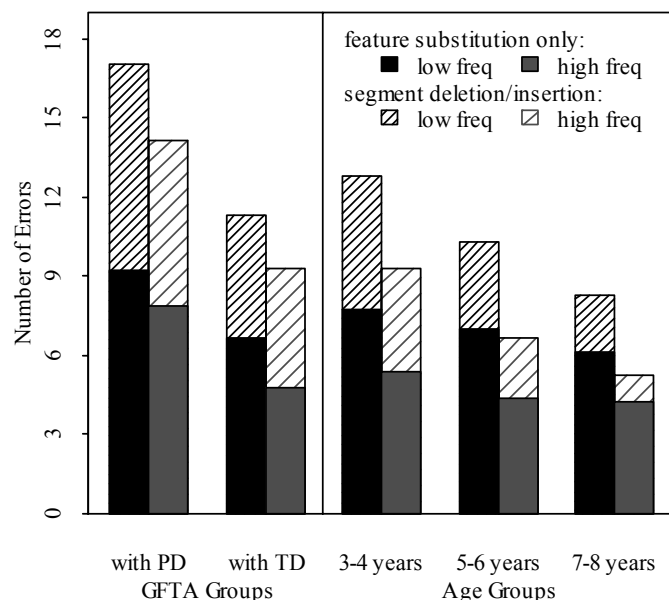


Figure 3. Number of errors that were substitutions versus additions/deletions for the low and high frequency sequences for the children with PD and their TD age controls (left panel) and for the three age groups of children with TD (right panel).

The results in Figures 2 and 3 showed that the effect of sublexical sequence frequency on accuracy was similar for the children with phonological disorders and their typically developing age peers. This suggests that the source of the differences in overall accuracy between these two groups is different from the source of the larger effect of sequence frequency on accuracy for children with smaller vocabularies relative to children with larger vocabularies in the earlier study (Edwards et al., 2003). That is, other attested differences between children with PD and their typically developing age peers, such as the defining difference in GFTA standard score, should be relatively independent of the size of the sequence frequency effect, except insofar as these other differences are themselves correlated with age and vocabulary size. To explore this possibility, we examined the relationship between target sequence frequency and several other factors in predicting the accuracy scores for individual participants. We included all 104 TD children and the 40 children with PD in these analyses. This gave us a wide range of vocabulary sizes and age (the children ranged in age from 3;2 to 8;10) and a wide range of articulatory ability (GFTA scores ranged from below the first percentile to the 99<sup>th</sup> percentile).

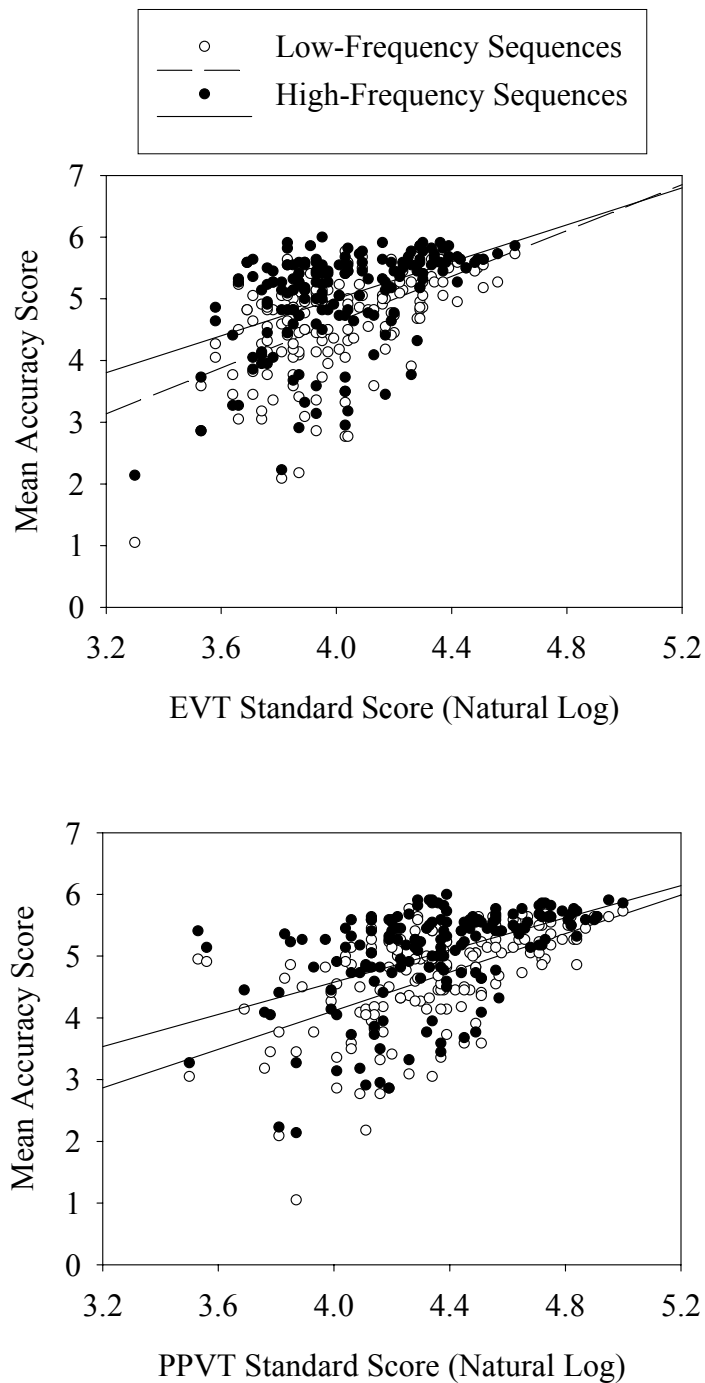


Figure 4. Mean accuracy scores for low- and high-frequency sequences plotted against receptive vocabulary size (PPVT-III, Figure 4a) and expressive vocabulary size (EVT, Figure 4b) for all 104 children with TD and the 40 children with PD.

#### Sequence frequency and vocabulary size

Edwards et al. (2003) observed an interaction between vocabulary size and sublexical sequence frequency. The larger the vocabulary size, the smaller the effect of sublexical sequence

frequency. We wanted to examine this same relationship for the larger group of participants, which included 40 children with PD. Figure 4 plots mean accuracy scores for the high- and low-frequency sequences against the log-transformed raw scores for the PPVT-III and EVT for the whole group of 144 children (104 children with typical phonological development and 40 children with phonological disorders).

As in the earlier study, we examined the relationship between phonotactic probability effects and vocabulary-size effects on accuracy. In this study, we used hierarchical multiple regression (HMR) to examine these effects. We ran two regressions examining the influence of vocabulary size on accuracy. In these regressions, mean accuracy was the dependent measure. Each regression had two independent measures: log-transformed raw score on the EVT or the PPVT-III, and a binary variable dummy-coding sequence frequency. The vocabulary-size measure was always entered into the regression first. In both of the regressions, both of the predictor variables accounted for a significant proportion of variance in the accuracy scores. The log-transformed EVT raw score accounted for 23.6% of the variance in accuracy scores ( $F[1,286] = 88.3, p < .01$  for  $R^2$  change) and the dummy-coded sequence-frequency variable accounted for 6.1% of the variance ( $F[1,285] = 18.6, p < .01$  for  $R^2$  change). The log-transformed PPVT-III raw score accounted for 25.7% of the variance in accuracy scores ( $F[1,286] = 99.1, p < .01$  for  $R^2$  change) and the dummy-coded sequence-frequency variable accounted for 6.3% of the variance ( $F[1,285] = 19.2, p < .01$  for  $R^2$  change). Thus, there is a significant correlation between overall accuracy and vocabulary size as well as a significant main effect of phonotactic probability, just as there was in Edwards et al. (2003).

Moreover, in Figure 4 (as in Figure 3 from Edwards et al. [2003] which plots these same relationships for typically developing children and adults), there is an interaction between vocabulary size and phonotactic probability. That is, the regression lines for the high-frequency and low-frequency sequences are not parallel. Instead, the distance between the two regression lines (which represents the effect of target sequence frequency) decreases as vocabulary size increases. Thus, while the addition of the PD group makes for more scatter, because children with PD have overall lower accuracy but typically do not have smaller vocabularies than their age peers, the pattern of a larger effect of sequence frequency for children with smaller vocabularies is exactly the same as it was in the earlier study that included only the children with TD.

#### Sequence frequency and measures of speech perception and articulation ability

Figure 5 plots repetition accuracy for the low-and high-frequency sequences against two other measures that are like vocabulary size in being correlated with age, but which typically differ between children with PD and children with TD. In Figure 5a, the x-axis variable is the measure of speech perception ability reported in Edwards et al. (2002). In the picture of different types of phonological representation outlined above, this d-prime value is a measure of the robustness of the auditory representations in the lower left corner of Figure 1. In Edwards et al. (2002), this score was significantly lower for children with PD relative to their TD age peers. It was also significantly lower for younger children with TD relative to older children.

Again, we used hierarchical linear regression with a dummy-coded variable representing frequency to examine these relationships. Mean accuracy score was the dependent variable. Whole-word discrimination (entered into the regression first) and the dummy-coded frequency variable were the independent measures. The measure of speech perception ability predicted 22.3% of the variance in the accuracy scores ( $F[1,278] = 81.1, p < .01$  for  $R^2$  change). The

dummy-coded frequency variable accounted for 6.1% of the variance in scores ( $F[1,277] = 18.0$ ,  $p < .01$  for  $R^2$  change).

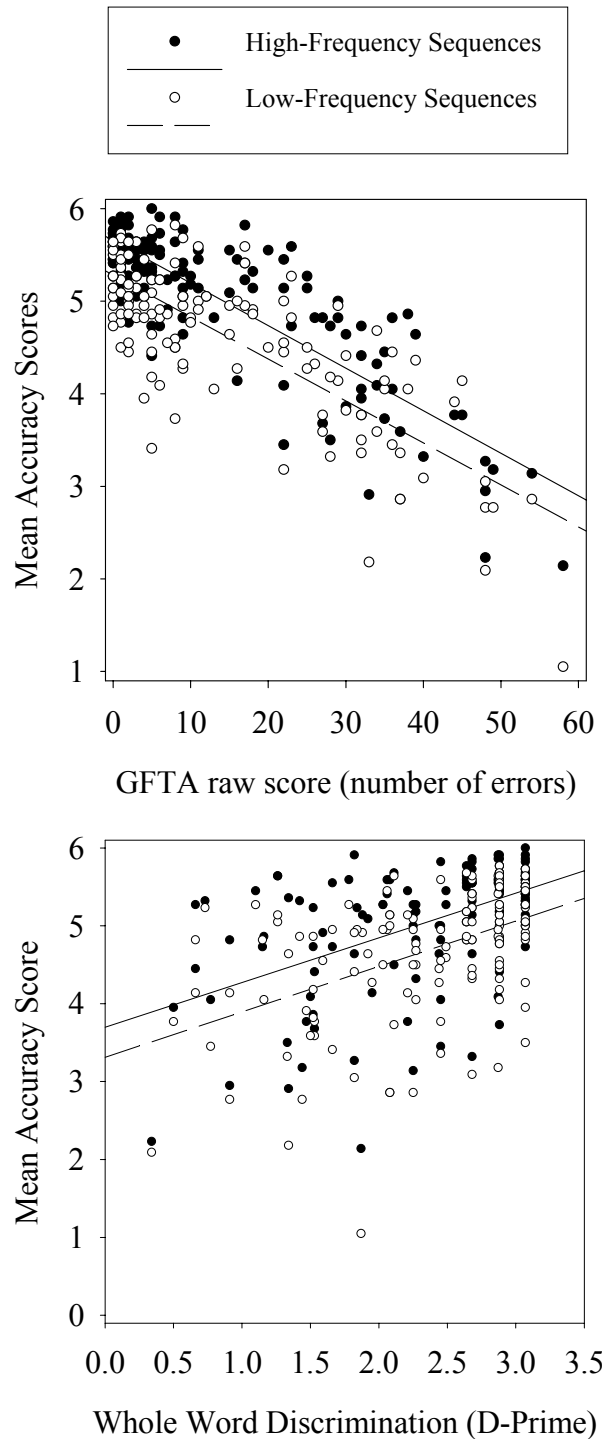


Figure 5. Mean accuracy scores for the low- and high-frequency sequences for all child participants (PD and TD) plotted against d-prime values for whole word (Figure 5a) and against discrimination GFTA raw score (Figure 5b).

Figure 5b plots mean accuracy scores for low- and high-frequency sequences against the number of errors on the GFTA. By using this raw score rather than the standard score, we can interpret performance on the GFTA as a measure of general “articulatory ability.” That is, we can use GFTA raw score as a measure of the robustness of the articulatory representations in the lower right corner of Figure 1. Again, we used hierarchical linear regression with a dummy-coded variable representing frequency to examine these relationships. Mean accuracy score was the dependent variable. GFTA raw score (entered into the regression first) and the dummy-coded frequency variable were the independent measures. The GFTA raw score accounted for 62.6% of the variance in the accuracy scores ( $F[1,286] = 478.3, p < .01$  for  $R^2$  change). The dummy-coded frequency variable accounted for 12.5% of the variance in scores ( $F[1,285] = 40.7, p < .01$  for  $R^2$  change).

Note that in both panels of Figure 5, the regression line for the high-frequency sequences lies above that for the low-frequency sequences. The distance between the lines is the effect of phonotactic probability once the other measure has been factored out. For any given level of speech perception ability as measured by the d-prime score plotted in Figure 5a, each child produces high-frequency sequences with greater accuracy than he or she produces low-frequency sequences. Similarly in Figure 5b, for any given level of articulatory ability, as measured by the GFTA raw score, each child produces high-frequency sequences with greater accuracy than he or she produces low-frequency sequences.

Figure 4 also showed a separation between the regression curves for the low- and high-frequency target sequences, but there is a critical difference between the pattern in Figure 4 and that in Figure 5. While there is a strong correlation of nonword repetition accuracy with both speech perception ability and articulatory accuracy, there is no interaction with target sequence frequency in either case. In both panels in Figure 5, the regression line for the high-probability sequences lies above that for the low-probability sequences, but the two lines do not converge and instead run in parallel across the entire width of each plot. This suggests that, at least by the time that the lexicon has grown to the size that is typical of a three-year-old’s vocabulary, any differences across children in robustness of representation in the two primary phonological domains are independent of whatever it is that accounts for the phonotactic probability effect. Before concluding that the effect must be related to differences in the higher-order representations that link the two domains, however, we must first rule out an alternative explanation for the separation between the regression curves in Figure 5.

As noted earlier, low-frequency target sequences often contain one segment that is itself lower in frequency than the analogous segment in the paired high-frequency sequence. Suppose that the effect of phonotactic probability on repetition accuracy is due simply to the fact that low-frequency phonemes are inherently more difficult to produce. If this is the case, then we should find that performance on a target sound in producing a familiar word predicts performance on the nonword task and that the frequency of the sequence in which the sound is embedded in the nonword stimulus has no additional predictive power. That is, the regression line for the high-frequency sequences should lie on top of the regression line for the low-frequency sequences.

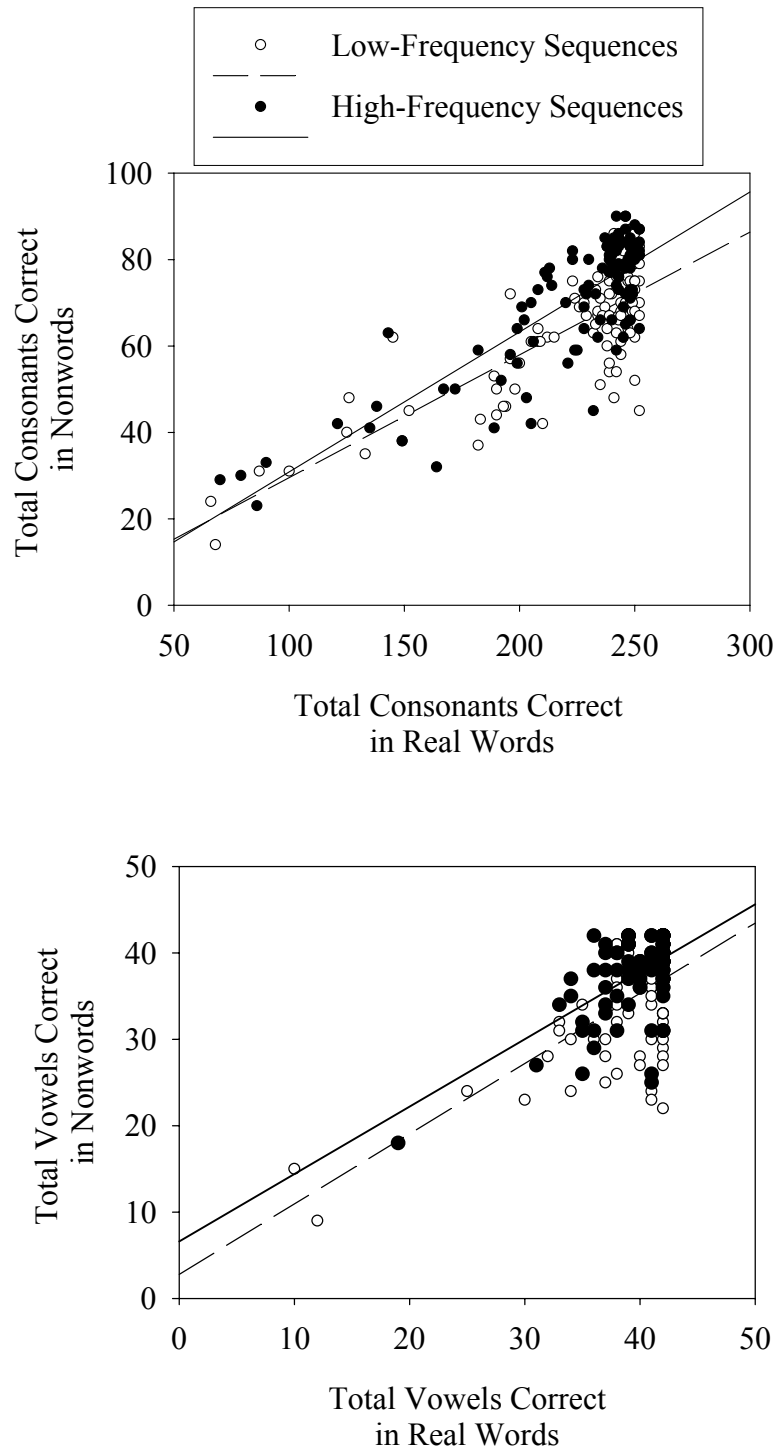


Figure 6. Segmental accuracy scores of segments in low- and high-frequency target sequences in nonwords plotted against accuracy scores for these same segments in familiar words on the PI, for children with PD and children with TD (age six and under). Accuracy scores plotted separately for the consonants in the target sequences (Figure 6a) and the vowels in the target sequences (Figure 6b).

### Sequence frequency and accuracy on individual phonemes

To see whether this was the case, we plotted accuracy for each participant for the consonants and vowels on the nonword repetition task against accuracy for the same sounds in the familiar real words of the PI. Figure 6 shows the two plots. Each y-axis value in the plots is the relevant total segmental accuracy score for a child's productions of the target segments on the nonword repetition task, and each x-axis value is the matching weighted total for the same segments on the PI. Since the PI was not given to children with TD who were seven or older, the relationship between correct production in familiar words and accuracy on the low- and high-frequency target sequences in these plots is based only on the productions by the children with PD and by the children with TD who were six years or younger.

Not surprisingly, there is a high correlation between accuracy on sounds in the nonwords and accuracy on the same sounds in familiar words. Children who produce sounds accurately in familiar words are also likely to produce sounds accurately in nonwords, and vice versa. We evaluated the effect of accuracy in familiar words on accuracy in nonwords using a series of hierarchical multiple regressions. In the regressions examining the influence of phoneme accuracy in real words, the sum of phonemes correct was the dependent measure. The independent measures were a dummy-coded measure representing sequence frequency, and accuracy in familiar words. In the regressions examining the influence of consonant accuracy, 63.2% of the variance was accounted for by accuracy in real words ( $F[1,222] = 381.6, p < .01$  for  $R^2$  change), and the variance accounted for by sequence frequency was 12.2% ( $F[1,221] = 30.6, p < .01$  for  $R^2$  change). In the regression examining the influence of vowel accuracy, 38.4% of the variance was accounted for by accuracy in real words ( $F[1,222] = 138.3, p < .01$  for  $R^2$  change), and the variance accounted for by sequence frequency was 9.4% ( $F[1,221] = 24.1, p < .01$  for  $R^2$  change).

The amount of variance accounted for by the PI accuracy score was smaller for the vowel analysis than for the consonant analysis because there was a much smaller range of accuracy values for the vowels, especially in familiar words. This was not surprising, in that most children, even children with PD, produce most vowels accurately by about age three, the age of the youngest children in this study (Pollock, 1994). Since the PI is a measure of articulatory ability independent of the GFTA score that we used to differentiate the children with PD from the children with TD, this result offers independent support for the relationship between articulatory ability and overall accuracy shown in Figures 2 and 5b.

At the same time, this analysis also supports the interpretation of the difference between Figures 4 and 5. That is, there is a small but consistent effect of sublexical sequence frequency in Figure 6. In both plots, the regression line for the high frequency sequences lies above the regression line for the low-frequency sequences. Moreover, the two lines do not converge for the children with the highest PI scores. Rather, for the vowel productions in Figure 6b, the lines are parallel, and for the consonant productions in Figure 6a, they converge at the other end of the scale, for the four children with the poorest articulatory ability, who have very low accuracy scores for both the PI and the nonword task. Thus, there is an effect of sequence frequency on accuracy, over and above any effect of phoneme frequency. Also, the decrease in the size of the effect with the increase in vocabulary size in older children is independent of any accompanying increase in articulatory ability as measured on the PI.

### Discussion

In this study, which included both children with phonological disorders and typically developing children, there was a complex relationship between vocabulary and nonword repetition accuracy, as in Edwards et al. (2003). In Figure 4, accuracy was correlated with vocabulary size over all of the target sequences, but there was also an interaction between the sequence frequency and vocabulary size — the larger the vocabulary, the smaller the effect of sublexical sequence frequency. Nonword repetition accuracy for both high and low-frequency sequences was also correlated with articulatory ability and speech perception ability, in Figure 5. There was no interaction between the effects of sublexical sequence frequency and articulatory ability or between the effects of sublexical sequence frequency and speech perception ability. This is what we might expect if the emergence of robust articulatory-motor and acoustic-auditory representations depended primarily on maturational factors that are less directly related to vocabulary growth.

Interpreting these two results together, the interaction in Figs. 4a and 4b, suggests a different kind of phonological skill, which we have equated with the emergence of something like phonemes, higher-order phonological categories that allow the child to generalize the mapping between acoustic and articulatory representations from familiar contexts to novel ones. These results support the second of the two explanations proposed in the Introduction for the interaction between phonotactic probability and vocabulary size. That is, we propose that as the child learns more and more words, there emerges an intervening level of symbolic units (phonemes) which mediates between the more primary articulatory and acoustic representation.

This picture of the core of phonology as a set of emergent generalizations about acoustic-articulatory mappings over the words that the child knows predicts the interaction between target sequence frequency and vocabulary size seen in both this study and Edwards et al. (2003). That is, one way to understand the phonological construct of “phoneme” is that it is a prototypical set of correspondences between auditory and articulatory patterns in this core layer of hidden nodes that links together the modality-specific auditory, articulatory, and semantic representations of words. Learning to perceive and say a particular consonant or vowel in many different words establishes it as a strong, stable activation pattern in this core. Novel auditory patterns that are only slightly different from familiar ones, such as the formant transitions and noise spectrum for an /f/ in an /fk/ cluster as compared to an /ft/ cluster, will be assimilated to the more prototypical mapping from that part of the acoustic space to an /ft/ sequence. Thus the /f/ in /fk/ can be parsed in terms of well-practiced articulatory configurations for producing /f/ in attested clusters, enabling the speaker to pronounce the phoneme accurately even in this completely novel segmental context.

The smaller proportion of the prosody-changing errors by the older children with TD relative to the younger children and by the TD age controls relative to the children with PD in Figure 3 is also predicted by this account. If there is a parsing error in the mature system, it is unlikely to be a complete breakdown that changes the prosody by deleting one of the consonants or inserting an epenthetic vowel between the /f/ and /k/. Instead, one or both consonants might assimilate to a similar phoneme, producing a different sequence without changing the prosodic structure of the cluster.

Phonological disorder can also be described within this framework. The greater proportion of prosody-simplifying errors for the children with PD relative to their age controls and the youngest TD children relative to the older children probably reflects a less well-developed articulatory representation of the precise temporal coordination needed to produce a VC rime or

a fluent CC cluster. The fact that this proportion was not affected by target sequence frequency suggests a maturational effect that is independent of the growth of the lexicon.

We also found that the influence of sublexical sequence frequency on production accuracy was similar for both the children with PD and their TD age controls. The fact that accuracy on the same phonemes produced in other segmental contexts in real words could not account for the lower accuracy of phonemes in the low-frequency sequences in the nonwords is also in keeping with this account. That is, children with PD made more errors than their age peers in producing the target phonemes in the real words elicited with the PI, because they have less robust articulatory generalizations, but this is independent of the difficulty that all children have with mapping from acoustics to articulation for the same phonemes in the unfamiliar context in the nonwords.

Together, these results suggest that children with phonological disorders do not have a modality independent deficit in the linguistic capacity for abstracting symbolic categories to map between these two more primary phonological domains. Rather, we propose that phonological disorder in many children is related to subtle perceptual and motor deficits that result in a degraded representation of other talkers' productions and of the child's own productions.

In summary, these results suggest that phonological acquisition involves not only the development of well-practiced articulatory-motor and acoustic-auditory representations, but also the emergence of a symbolic representation that links both of these primary phonological representations in the lexical core. Thus, phonological development sets the stage for the even higher-order generalizations about morpheme structure and phonotactics that characterize the mature phonological system. These generalizations in the adult differ from those in the child's phonology in that they do not usually result in production errors unless the system is stressed. However, for both the adult and the young child, phonological generalization cannot be divorced from the words that the speaker recognizes and produces.

#### Acknowledgements

This work was supported by NIDCD grant 02932 to Jan Edwards and by NIH training grant T32 DC0051 to Robert A. Fox. We thank the children who participated in the study, the parents who gave their consent, and the schools at which the data were collected. For assistance in stimuli preparation, data collection, and analysis, we thank Erin Casey, Lynn Carahaly, Lisa Draper, Melissa Epstein, Heidi Hochstetler, Maryann Holtschulte, Bridgett Isermann, Satako Katagiri, Laurie Vasicek, Amy Vitale, Pauline Welby, and David White.

#### References

- Bailly, G., Laboissiere, R., & Schwartz, J. L. (1991). Formant trajectories as audible gestures. *Journal of Phonetics*, *19*, 9-23.
- Barlow, J. A. & Dinnsen, D. A. (1998). Asymmetrical cluster development in a disordered system. *Language Acquisition: A Journal of Developmental Linguistics*, *7*, 1-49.
- Barlow, J. A. & Gierut, J. A. (1999). Optimality theory in phonological acquisition. *Journal of Speech Language & Hearing Research*, *42*, 1482-1498.
- Beckman, M. E., & Edwards, J. (2000). Lexical frequency effects on young children's imitative productions. In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 207-217). Cambridge, UK: Cambridge University Press.

Bernhardt, B. & Stoel-Gammon, C. (1996). Underspecification and markedness in normal and disordered phonological development. *Children's language*, 9, 33-54.

Bernhardt, B. (1994). The prosodic tier and phonological disorders. *First and second language phonology*. 149-172.

Burgemeister, B. B., Blum, L. H., & Lorge, I. (1972). *Columbia Mental Maturity Scale*. New York: Harcourt Brace Jovanovich, Inc.

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.

Dunn, L. & Dunn, L. (1997). *Peabody Picture Vocabulary Test – III*. Circle Pines, MN: American Guidance Services.

Edwards, J., Beckman, M. E., and Munson, B. (2003). *The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition*. Manuscript submitted for publication.

Edwards, J., Fourakis, M., Beckman, M., and Fox, R. (1999). Characterizing knowledge deficits in phonological disorders. *Journal of Speech, Language, and Hearing Research*, 42, 169-186.

Edwards, J., Fox, R. A., & Rogers, C. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and phonetic inventory size. *Journal of Speech, Language, and Hearing Research*, 45, 231-242.

Edwards, J., Gibbon, F., & Fourakis, M. (1997). On discrete changes in the acquisition of the alveolar/velar stop consonant contrast. *Language and Speech*, 40, 203-210.

Forrest, K., Chin, S. B., Pisoni, D. B., & Barlow, N. N. (1995). *Talker normalization in normally articulating and phonologically delayed children: Methodological considerations* (Research on Spoken Language Processing, No. 19, pp. 229-251). Bloomington: Indiana University Press, Department of Psychology.

Gibbon, F. (1990). Lingual activity in two speech-disordered children's attempts to produce velar and alveolar stop consonants: Evidence from electropalatographic (EPG) data. *British Journal of Disorders of Communication*, 25, 329-340.

Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research*, 42, 382-397.

Goldman, R. & Fristoe, M. (1986). *The Goldman Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Services.

Hoffman, P. R., Daniloff, R. G., Bengoa, D., & Schuckers, G. H. (1985). Misarticulating and normally articulating children's identification and discrimination of synthetic [r] and [w]. *Journal of Speech and Hearing Disorders*, 50, 46-53.

Ingram, D. (1976). *Phonological disability in children*. NY: Elsevier.

Ingram, D. (1988). The acquisition of word-initial [v]. *Language and Speech*, 31, 77-85.

Isermann, B. C. (2001). *Variability and consistency of articulation in children with phonological disorders*. Unpublished master's thesis. Department of Speech and Hearing Science, Ohio State University.

Jordan, M.I. (1990). Motor learning and the degrees of freedom problem. In M. Jeannerod (Ed.) *Attention and Performance XIII* (pp. 796-836). Hillsdale, NJ: Erlbaum.

Kaufman, N. (1995). *Kaufman Speech Praxis Test for Children*. Detroit, MI: Wayne State University Press.

- Martin, C. W. (1989). A comparative study of two deep tests of articulation. *Journal of Communication Disorders*, 22, 383-395.
- Ohde, R. H., & Sharf, D. J. (1988). Perceptual categorization and consistency of correct as cited. synthesized /r-w/ continua by adults, normal children, and /r/-misarticulating children. *Journal of Speech and Hearing Research*, 31, 556-568.
- Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: a distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381-415). Mahwah, NJ: Erlbaum.
- Pollock, K. (1994). Assessment and remediation of vowel misarticulations. *Clinics in Communication Disorders*, 4, 23-37.
- Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with functional articulation disorder. *Journal of Speech and Hearing Research*, 54, 193-208.
- Shriberg, L. D. & Kwiatkowski, J. (1994).. Developmental phonological disorders: I. A clinical profile. *Journal of Speech and Hearing Research*, 37, 1100-1126.
- Smit, A.B., Hand, L., Freiling, J.J., Bernthal, J.E., & Bird, A. (1990). The Iowa articulation norms and its Nebraska replication. *Journal of Speech and Hearing Disorders*, 55, 779-798.
- Stemberger, J. P., & Bernhardt, B. H. (1999). The emergence of faithfulness. In B. MacWhinney (Ed.), *The emergence of language* (pp. 417-446). Mahwah, NJ: Lawrence Erlbaum.
- White, S. D. (2001). *Covert Contrast, Merger, and Substitution in Children's Productions of /k/ and /t/*. M.A. thesis, Ohio State University.
- Williams, K. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Services.
- Winitz, H. (1969). *Articulatory Acquisition and Behavior*. New York: Appleton-Century-Crofts.