HHS Public Access

Author manuscript

Clin Linguist Phon. Author manuscript; available in PMC 2020 August 04.

Published in final edited form as:

Clin Linguist Phon. 2019 ; 33(10-11): 885–898. doi:10.1080/02699206.2019.1584247.

Quantifying Phonological Knowledge in Children with Phonological Disorder

Philip N. Combiths,

San Diego State University / University of California, San Diego, CA, USA

Jessica A. Barlow,

San Diego State University, CA, USA

Emilie Sanchez

San Diego Unified School District, CA, USA

Abstract

Generative phonologists use contrastive minimal pairs to determine functional phonological units in a language. This technique has been extended for clinical purposes to derive phonemic inventories for children with phonological disorder, providing a qualitative analysis of a given child's phonological system that is useful for assessment, treatment, and progress monitoring. In this study, we examine the single-word productions of 275 children with phonological disorder from the Learnability Project (Gierut, 2015) to confirm the relationship between phonemic inventory---a measure of phonological knowledge---and consonant accuracy---a quantitative, relational measure that directly compares a child's phonological productions to the target (i.e. adult-like) form. Further, we identify potential percent accuracy cutoff scores that reliably classify sounds as in or out of a child's phonemic inventory in speech-sound probes of varying length. Our findings indicate that the phonemic function of up to 90% of English consonants can be identified for preschool-age children with phonological disorder when a sufficiently large and thorough speech sample is used.

Keywords

phonology; phonological disorder; assessment; phonemic inventory; accuracy

Introduction

Phonological disorder (PD) is one of the most prevalent communication disorders in young children, with prevalence estimates in the rangeof 7−−11% for children at five years of age (Law, Boyle, Harris, Harkness, & Nye, 2000). This developmental impairment of unknown aetiology occurs independently of another primary motivating condition and functionally impairs the development, manipulation, and production of the phonological units of

Contact details: Philip Combiths, School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, CA 92182. USA, pcombiths@sdsu.edu.

Declaration of interest

A portion of this research was funded by NIH NIDCD F31 DC017697–01. The authors have no other declarations of interest to report.

Author Manuscript Author Manuscript Author Manuscript

language. Consequently, this form of developmental communication disorder prevents acquisition of phonological skills in a timely manner, which can impact communication and literacy skills and later academic, socio-emotional, and occupational outcomes (Beitchman, Wilson, Brownlie, Walters, Inglis, et al., 1996; Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Felsenfeld, Broen, & McGue, 1992, 1994; Lewis et al., 2016; Peterson, Pennington, Shriberg, & Boada, 2009). Given the prevalence and impact of this impairment, researchers and clinicians alike are continually exploring techniques for assessment and progress monitoring to better capture the phonological abilities of children with PD. We begin our discussion of this topic by describing and comparing two such measures below.

Phonemic inventory analysis

Speech-language pathologists (SLPs) and researchers utilize a variety of measures to assess productive knowledge of speech sounds in children with PD. One such measure used to qualitatively describe a child's functional phonological knowledge is the phonemic inventory. This unique, linguistically motivated measure ostensibly captures a child's functional use of contrastive speech sounds (i.e. phonemes). The methodology for deriving a phonemic inventory originated in a generative linguistics framework and has been used to provide phonological descriptions of fully formed adult languages (e.g. Voegelin, 1957). In order to characterize those phonemes that are used contrastively to distinguish words among speakers of a given adult language, phonologists require semantically distinct word pairs with minimal phonological contrast (i.e. minimal pairs) to demonstrate a speech sound's phonemic function (e.g. Chomsky & Halle, 1968). For instance, the words 'car' /kɑɹ/ and 'tar' /tɑɹ/ form a minimal pair in English because they differ by only a single segment, demonstrating that the contrast between /k/ and /t/ is sufficient to distinguish words. Thus, /k/ and /t/ are contrastive phonemes in English.

In discussion of phonemes and phonemic inventories within a generative phonology framework, it is important to emphasize that a phoneme is an abstraction that describes a categorical representation of a functional, contrastive unit within a word. These abstract units can have a number of phonetic expressions, all recognized as variants of the same categorical phoneme by speakers of the same language. For instance, the categorical phoneme /p/, in English, can be produced as [p=] (e.g. [sp=un] 'spoon'), [p^{h}] (e.g., [p^{h}ɑt] 'pot'), or [p˺] (e.g. [stɑp˺] 'stop')---any of these productions would be understood as productions of the phoneme /p/ by a native speaker of English. Furthermore, although the presence of minimal pairs is the primary evidence for establishing phonemic status, it is only one part (albeit an important one) of a larger process involving a comprehensive description of the phonetic environments in which a given sound occurs and its phonetic similarity (or dissimilarity) to other sounds of that language.

The phonemic inventory and its corresponding methodology have subsequently been extended for clinical purposes to derive the phonemic inventories of individual children with PD, although the criterion for phonemic status has been simplified such that the primary requirement is the presence of two minimal pairs (i.e. four words in total) to establish a speech segment as a phoneme in a given child's phonemic inventory (Barlow & Gierut, 2002; Dinnsen, 1984; Gierut, Simmerman, & Neumann, 1994). Because children in the process of language development demonstrate unique and dynamic phonological systems (Fry, 1967; Jakobson, 1968), their phonemic inventories are likewise varied and subject to change over time. Importantly, the phonemic inventory is also considered an independent measure because it neutrally describes the functional phonemes a child uses, without reference to their *accuracy* or *correctness*.

An independently constructed phonemic inventory therefore provides a snapshot of the child's phonological knowledge; however, a child's inventory can also be compared to the inventory of the target adult language to provide additional information. This comparison generates a relational measure of those phonemes that are *missing* from the individual's phonemic inventory. Thus, an inventory of the segments that are 'in' a child's phonemic inventory is an independent measure, and an inventory of the phonemes that are 'missing' from a child's phonemic inventory is a relational (i.e. comparative) measure because labelling phonemes as 'missing' requires a comparison to the adult target inventory (Dinnsen, 1984). A description of the phonemes that are either 'in' or 'missing' from the inventory provides simultaneously a neutral snapshot of a given child's phonological system and a comparative indication of the weaknesses or gaps in their phonological knowledge. This information is used to determine the presence or severity of PD (e.g. Gierut et al., 1994), monitor change over time (e.g. Gierut, 1992), and to guide the selection of appropriate speech-sound targets and goals for intervention (e.g. Barlow & Gierut, 2002; Morrisette, Farris, & Gierut, 2006).

Despite its unique informativeness, phonemic inventory analysis is not commonly employed by practicing SLPs (McLeod & Baker, 2014). Perhaps this is due to the abstract nature of the knowledge it captures or the opacity of the underlying generative assumptions from which this measure is derived, but there are clear logistical barriers as well. Ferguson and Farwell (1975) and Gierut et al. (1994) describe these obstacles to phonemic inventory analysis in children, including the variability of their word productions and the difficulty of obtaining sufficient words to serve as minimal pairs. Certainly, the descriptive process required to identify minimal pairs and generate a child's phonemic inventory requires collection of a thorough speech sample strategically designed to capture contrastive minimal pairs, which may be time-prohibitive for many practicing clinicians.

Production accuracy

Whereas phonemic inventories provide insight into a child's functional phonological knowledge, other frequently employed speech sound measures eschew underlying knowledge and instead capture production accuracy. A consonant accuracy measure compares each consonant segment produced by the child to its corresponding target (i.e. adult-like) form. This pairwise comparison requires no assumption of underlying phonological function, and it generates a percent accuracy score that is relational, quantitative, and immediately interpretable. The most commonly used segmental accuracy measure is Percentage of Consonants Correct-Revised (PCC-R; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997), which collapses across all consonants to provide a single accuracy percentage from a given sample. However, an SLP may also choose to examine consonant accuracy for each consonant separately to provide more nuanced accuracy information. Furthermore, with the advent of computer-assisted analysis, including freely available software (e.g. Phon; Rose & Hedlund, 2017), SLPs can calculate consonant accuracy measures consistently and relatively quickly (Byun & Rose, 2016).

Despite their differences, phonological knowledge and accurate production are presumed to be related to one another and are often discussed jointly (e.g. Gierut, Elbert, & Dinnsen, 1987). However, it is notable that measures of one can contrast with the other. One such instance is discussed in Dinnsen and Barlow (1998) and Dinnsen, Green, Gierut, and Morrisette (2011). In their example, a child uses the consonant /θ/ phonemically to contrast words, yet this same child never uses /θ/ accurately due to a chain-shift substitution pattern, such that [θ] is produced exclusively as a substitution for /s/ (i.e. dentalisation), and every instance of target /θ/ is produced as [f] (i.e. labialisation). These patterns result in productions, such as [fʌm] for ‘thumb’ and [θʌm] for ‘some’, which serve as a minimal pair for both /f/ and /θ/. Consequently, /θ/ would be considered phonemic and thus ‘in’ the child’s phonemic inventory, despite the child’s 0% accuracy for production of /θ/. Given the potential for divergence, there is motivation to better evaluate the relationship between phonemic inventory (a measure of phonological knowledge) and consonant accuracy (a measure of adult-like production).

Current study

To better understand these different phonological assessment measures, the purpose of the investigation described here is twofold. Our first goal is to identify the relationship between phonemic inventory---a qualitative, linguistically motivated measure of phonological knowledge---with consonant accuracy---a quantitative, relational measure that directly compares a child’s phonological productions to the target form. By identifying a relationship between these two measures, we improve our understanding of how production accuracy reflects phonological knowledge in children with PD. Our second goal is to determine if the relationship between the measures would permit identification of a percent accuracy cutoff score (or cutoff range) that reliably classifies sounds as ‘in’ or ‘out’ of a child’s phonemic inventory in speech-sound probes of varying length. A percent accuracy cutoff suggestive of the phonemic function of a given consonant could provide useful information about phonological knowledge without the time-consuming process of identifying minimal pair contrasts.

Method

Participants

Data for this study were drawn from 275 children between 3 and 8.5 years old (mean age = 4;4), whose single-word productions were transcribed as part of the Learnability Project (Gierut, 2015b). Participants in the Learnability Project were monolingual, English-speaking children residing in the Midwestern United States who presented with functional PD, determined by performance > 1 SD below the mean on the first or second edition of the Goldman-Fristoe Test of Articulation (GFTA/GFTA-2; Goldman & Fristoe, 1986, 2000) and a reduced phonemic inventory, missing at least 6 target English consonants. Furthermore, all participants had normal hearing, no documented history of motor or otherwise organic disorders, no indication of cognitive delay, and normal oral-motor function. Participating children received experimental speech intervention; however, data in this study come only from the children's pre-treatment samples. Additional demographic information for participants in the Learnability Project can be found in Gierut (2015b)[1].

Data transformation

Data were phonetic transcriptions of children's productions of words in the Phonological Knowledge Probe (PKP; Gierut, 1986), a single-word probe eliciting 293 words, and the GFTA or GFTA-2, eliciting 44 and 53 words, respectively. Original archival transcriptions included the orthography of the target word and transcription of the child's production in IPA notation. Reliability for 10% of archival consonant transcriptions was reported at 93% (Gierut, 2015a).

To facilitate analyses using Phon (v2.2; Rose & Hedlund, 2017), data were translated from their archival Excel format to Phon-readable Unicode text using a Python script. Non-standard notation conventions were translated to standard IPA notations compatible with Phon. For instance, the US English rhotic consonant, transcribed as [r] in the archival data, was translated to the standard IPA notation [ɹ]. Some diacritic symbols in the original data, such as [$^{\wedge}$β], were not available as characters in Phon. In these instances, the symbol was changed to a similar diacritic (e.g. [^{b}]) and appended with [^{s}] (e.g. [deɪ$^{\wedge}$β] became [deɪ^{bs}]) to document the change during translation. These diacritic differences between the original archival transcription and the translated format did not impact our analyses, as these changes were implemented consistently across the data. Furthermore, diacritic symbols were ignored during consonant accuracy calculations, as described below.

Orthographic and IPA transcriptions translated directly from archival data were sufficient for extracting each child's phonemic inventory, as this measure did not require comparison to the target production of each word. However, calculation of consonant accuracy required transcription of the target, adult-like form of each word. Given the scope of the data to be analysed (approximately 93,000 words or 243,000 consonants), we generated a single, representative set of target transcriptions for all sampled word productions to permit relational analyses. A two-step process generated these target transcriptions for comparison to the children's productions. First, broad target transcriptions were generated for each word in the PKP, GFTA, and GFTA-2 from the English IPA dictionary in Phon. Second, two research assistants (undergraduate and graduate students of speech-language pathology or linguistics) reviewed archival transcriptions extracted from 200 participants and compared these to the dictionary-generated targets to arrive at consensus for a single target transcription for each word deemed to best capture the dialect spoken by these children and the transcription conventions of the archival data. These transcriptions were also reviewed by the first author. Once confirmed, these target transcriptions were aligned to each child's

[1]Archival data were retrieved from the Gierut / Learnability Project collection of the IUScholarWorks repository at https://scholarworks.iu.edu/dspace/handle/2022/20061 The archival data were original to the Learnability Project and supported by grants from the National Institutes of Health to Indiana University (DC00433, RR7031K, DC00076, DC001694; PI: Gierut). The views expressed herein do not represent those of the National Institutes of Health, Indiana University, or the Learnability Project. The author(s) assume(s) sole responsibility for any errors, modifications, misapplications, or misinterpretations that may have been introduced in extraction or use of the archival data.

transcribed productions using the English syllabification and alignment algorithms in Phon and then compared to generate the relational consonant accuracy measure used in this study.

To validate the generated target forms, alignment, and our automated percent consonant accuracy measure, percent consonant accuracy for word productions in the PKP was calculated manually by research assistants for 20% of participants. Procedures for manual calculation followed those outlined for calculation of PCC-R. On average, manually calculated accuracy deviated 3.9% (SD = 3.7%) from automatically generated values. Correlation between manually derived accuracy and automated accuracy measures was 0.96.

Variables

Two primary measures were derived for each of 23 American English consonants (excluding /ʒ/ due to limited sampling of this consonant), for each child. The first measure was a binary, categorical designation of phoneme status. For a given child, if two contrastive minimal pairs were identified for a given consonant, that consonant was deemed phonemic and coded as ‘in’ the phonemic inventory. For many children, one or more non-ambient sounds (e.g. /wʴ/ or /ʔ/) were also used contrastively; however, only the phonemic status of ambient phonemes was recorded because the corresponding accuracy measure is only derivable for ambient consonants. When two minimal pairs were not identified, that consonant was coded as ‘out’ of the child’s phonemic inventory. Minimal pairs were identified, and phoneme status was confirmed using the AutoPATT plugin for Phon (Combiths, Amberg, & Barlow, 2016). Data used to calculate this measure were participants’ word productions from the 293-item PKP (rather than the shorter GFTA or GFTA-2) to obtain sufficient opportunities for two contrastive minimal pairs for each of 23 English phonemes.

The second measure was a quantitative measure of consonant production accuracy. For a given child, accuracy was calculated for their production of each English consonant in Phon by comparing each target consonant with its corresponding segment in the child’s production. This comparison was automated with a consonant accuracy query in Phon. In order to provide an accuracy calculation that is easily replicable and robust to varied ages and severities of impairment, we followed the same procedures used for the global measure of PCC-R (Shriberg et al., 1997). Unlike the original Percentage of Consonants Correct measure (Shriberg & Kwiatkowski, 1982), PCC-R ignores distortions in its calculation. Although distortion patterns are diagnostically informative, their absence from these calculations make the measure simpler, less prone to error, and appropriate for a more diverse population of children. Furthermore, PCC-R is a well-attested and reliable consonant accuracy measure (see Shriberg et al., 1997). Following PCC-R procedures, to be coded as correct, the child’s production was required only to match the base target phone. For instance, production of [s] for target /s/ was considered correct; however, phonemic substitution, such as [f] for target /θ/, or omission of a target consonant were considered incorrect. By these criteria, each child was designated a percentage accuracy for each of the 23 English consonants. Because consonant accuracy may be more robust to varying sample length than phonemic inventory (which requires multiple minimal pair opportunities for each phoneme), the accuracy measure was calculated separately for productions in the PKP and the GFTA/GFTA-2. This permitted comparison between probe types.

Additional variables used in the analyses were participant age, sample type (PKP, GFTA/GFTA-2), and normative age of acquisition (early, middle, late) for each English consonant, as categorized in (Shriberg, 1993).

Analyses

Logistic regression determined the ability of percent consonant accuracy to predict the phonemic inventory measure, including the mitigating impacts of sample length, child age, and consonant age of acquisition. Receiver-operating characteristic (ROC) curve analysis determined cutoff accuracy values with optimal sensitivity and specificity according to elicited sample length, child age, and age of acquisition. Regression models, ROC curves, and optimal cutoff value estimation were conducted in R (R Core Team, 2013) using pROC, OptimalCutpoints, and visreg packages.

Results

Descriptive statistics

Mean percent accuracy for all PKP consonants ‘out’ of the children’s phonemic inventories was 12.0% SD 23.3%. Mean percent accuracy for all PKP consonants ‘in’ the children’s phonemic inventories was 74.0% SD 26.4%. Mean accuracy for ‘out’ GFTA/GFTA-2 consonants was 12.1% SD 25.0%. Mean accuracy for ‘in’ GFTA/GFTA-2 consonants was 70.9% SD 31.1%. Thus, phonemic consonants were produced with greater accuracy than non-phonemic consonants, and mean accuracies for phonemic and non-phonemic consonants were similar across probes. Means and standard deviations for PKP consonants according to phonemic inventory classification, normative age of acquisition, and participant age are displayed in table 1.

Logistic regression

Phoneme status of 23 American English consonants, excluding /ʒ/, was predicted by consonant accuracy ($p < 0.01$) and classification as an early-, middle-, or late-acquired consonant ($p < 0.01$). As expected, consonants with higher accuracy and those that are earlier-acquired are more likely to be used as phonemes by the child. The main effect of child age on phoneme status was not significant ($p = 0.33$). Significant Consonant Accuracy × Child Age ($p < 0.01$) and Consonant Accuracy × Age of Acquisition ($p < 0.01$) interactions also emerged, such that consonant accuracy was most predictive of phoneme status in younger children and for middle- and late-acquired consonants. These interactions are displayed in figure 1.

Receiver operating characteristic curve analysis

The ability of a percent consonant accuracy cutoff to classify the phoneme status of 23 English consonants was quantified with ROC curve analysis, using several paradigms to determine the optimum cutoff value. For the larger 293-item PKP, a consonant accuracy of 20.4% was the most efficient cutoff, correctly classifying the phoneme status of 90.0% of

English consonants for all 275 children (sensitivity = 94.9%; specificity = 83.1%). Other potential percent accuracy cutoff values, derived from various methods for determination of optimal classification, including maximum efficiency (i.e. most accurate classification; Galen, 1986; Greiner, 1996), Youden's Index (Greiner, Pfeiffer, & Smith, 2000; Youden, 1950), and closest to ROC plot point 0,1 (Metz, 1978; Vermont et al., 1991), are displayed in table 2.

Furthermore, sample length impacted potential cutoff score classification accuracy. Classification accuracy, sensitivity, and specificity of several potential consonant accuracy cutoff values are displayed for data from the 293-item PKP and the 44−−53-item GFTA/GFTA-2 in table 3. When data are drawn from a larger (ostensibly more thorough) sample, optimal consonant accuracy cutoff values are lower, and sensitivity, specificity, and classification accuracy are higher than when data are drawn from a smaller sample. Accordingly, the area under the receiver operating characteristic curve (AUC/AUROC) was higher for the PKP data (AUROC = 0.94, 95% CI = [0.934, 0.947]) than for the GFTA/GFTA-2 data (AUROC = 0.90, 95% CI = 0.895, 0.911). Note that an AUROC closer to 1 has better overall classification ability. ROC curves for phoneme classification based on consonant accuracy data from the PKP and GFTA/GFTA-2 are displayed in figure 2. These findings are discussed in the next section.

Discussion

In this study, a qualitative measure of phonological knowledge and function (i.e. phonemic inventory) was compared to a quantitative accuracy measure (i.e. percent consonant accuracy) in young, monolingual English-speaking children with PD. A strong relationship emerged between a given consonant's percent accuracy and its contrastive, phonemic use. Furthermore, ROC curve analyses indicated that a relatively low consonant accuracy cutoff (approximately 20−−30%) can correctly classify up to 90% of English consonants as either ‘in’ or ‘out’ of a given child's phonemic inventory using data from the 293-item PKP. In other words, when a child produces a given English consonant with a percent accuracy above the cutoff of 20–30%, it is more likely that this child already uses the consonant phonemically to contrast words. Our analyses also found that sensitivity was higher than specificity for this cutoff range, indicating that a percent accuracy cutoff is better at correctly including phonemic consonants in the inventory than correctly excluding non-phonemic consonants. Finally, the ability of percent consonant accuracy to predict a child's phonemic use of a given consonant in these data was also mitigated by several factors, including speech sample length, child age, and normative age of acquisition of the consonant in question.

A percent consonant accuracy cutoff was poorer at classifying a child's phonemic inventory when consonant accuracy was derived from the GFTA/GFTA-2 (i.e. a short sample of 44−−53 words). The highest classification accuracy of 90% was only achievable with productions from the PKP, a larger speech sample of 293 words. In addition to sampling size, differences in the predictive power of consonant accuracy across the two probes may also have been related to qualitative differences between them. The PKP was designed to capture phonological knowledge in a research context and, consequently, was constructed with many production opportunities for each consonant (at least 5 in each word position) and to allow opportunities to demonstrate contrastive minimal pairs (Gierut, 2015c). Conversely, the GFTA/GFTA-2 is a standardized testing instrument designed for rapid administration. In typical usage, an examiner derives a single score, collapsed across all consonants, for comparison to a normative database. Although the GFTA/GFTA-2 is markedly shorter than the PKP, it is also likely that differences in the depth and breadth of these samples contributed to differences in their ability to provide accuracy calculations sufficient to predict a consonant's phonemic status.

Finally, percent consonant accuracy was also better able to predict the phonemic status of consonants produced by younger, preschool-age children than those produced by older, school-age children, and phonemic classification based on consonant accuracy was more accurate for normatively late-acquired sounds (/ʃ, s, θ, ð, ɹ, z, l/) than for normatively middle-acquired (/t, ŋ, k, ɡ, f, v, tʃ, dʒ/) and especially for early-acquired sounds (/m, b, j, n, w, d, p, h/; Shriberg, 1993). Older children are more advanced developmentally and, thus, less likely to demonstrate variable accuracy rates and phoneme usage. Similarly, earlier-acquired consonants are more likely to be produced accurately and used phonemically. These ceiling effects could be applicable to the broader population of children with PD, but they are also likely confounded by distributional limitations of the study sample, and this will be discussed more below.

Implications for assessment

The relationship between consonant accuracy and phonemic inventory identified in this study has potential implications for phonological assessment. By describing and quantifying the relationship between a qualitative, descriptive measure of phonological knowledge and a quantitative, relational measure of accurate production, we confirm the informativeness of both measures and highlight similarities between them. Although seemingly a simple comparison between two measures of speech-sound usage, phonemic inventory and percent consonant accuracy represent divergent conceptualizations of phonological skill. A phonemic inventory is intended to describe a child's contrastive speech-sound units (i.e. phonemes), and, as such, several generativist linguistic assumptions are required for meaningful interpretation of this measure. The term 'phonemic inventory' was first derived from linguistic descriptions of fully formed adult languages spoken by an entire community of speakers. The extension of this measure to the analysis of child phonology borrows from the descriptivist tradition, requiring instances of minimal pairs to confirm the contrastive role of a given speech-sound phoneme. On the other hand, consonant accuracy does not require any assumption of a consonant's underlying phonemic function. Rather, it relies on a direct comparison of each consonant in the child's word productions to its corresponding target consonant in the adult-like form of the intended word. The higher a child's percent accuracy for a given consonant, the more closely the child's production of that consonant coincides with target, adult productions.

Despite the inherent differences between these two measures, the findings of this study confirm that there is considerable overlap in the useful information they capture. Further still, the quantitative, relational information provided by percent consonant accuracy may provide a relatively accurate estimate of a preschool-age child's functional phonemic usage of later-developing consonants, but only when this accuracy measure is derived from a sufficiently thorough sample. Consequently, the relationship between measures of qualitative phonological knowledge and quantitative production accuracy should be considered in assessment and subsequent treatment goal selection and progress monitoring---especially given the increasing availability of (often quantitative) computer-assisted measures that have the potential to shift the assessment landscape.

Limitations and future directions

Limitations in the archival data used in this study likely contributed to the poorer predictive power of consonant accuracy for earlier acquired sounds and older children. The participants in this study, as expected, demonstrated greater mastery of early-acquired consonants, as indicated by less variable, generally higher accuracy rates and more frequent inclusion in their phonemic inventories. This ceiling effect likely impacted the ability to predict phonemic inventory inclusion from consonant accuracy for these consonants. It is possible that phonemic use of early-acquired sounds could be predictable from consonant accuracy given data with more variable early-acquired consonant accuracy rates and phonemic use, such as with younger children or those with more severe impairment. The poorer predictive power of consonant accuracy to categorize phonemic function for school-age children is also likely impacted by the participants' age distribution in these data. The majority of 275 participants were preschool-aged, with only 57 school-aged children in the sample. Consequently, poorer predictions for school-aged children may simply reflect the limited sampling of school-aged children in these data.

Future work examining the relationship between measures of phonemic function and quantitative accuracy could address these sampling limitations through prospective data collection involving younger children or those with more severely impacted phonological systems as well as a greater number of school-aged children. Although the role of child age and normative acquisition trajectories require further investigation, the relationship between phonemic inventory inclusion and consonant accuracy identified in the current data remain most robust for late-acquired sounds in preschool-aged children with PD.

Finally, the lower sensitivity of percent accuracy in determining phonemic status suggests that, when classification error does occur, it is more likely to result in over-identification of phonemic consonants. Future work should identify and compare the clinical impact of over- and under-estimation of phonological knowledge to determine which of these error types is most important to minimize. This type of work could guide modification of optimal percent accuracy cutoff values or the development of other criteria to improve the clinical utility of estimates of phonological knowledge. As our understanding of the relationship between phonemic inventory and consonant accuracy measures improves, clinicians may eventually be able to infer information about functional phonological knowledge from a quantitative accuracy measure, which could streamline assessment, treatment target selection, and progress monitoring for children with PD.

References

Barlow JA, & Gierut JA (2002). Minimal pair approaches to phonological remediation. Seminars in Speech and Language, 23, 57–68. [PubMed: 11938491]

Beitchman JH, Wilson B, Brownlie EB, Walters H, Inglis A, & Lancee W (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. Journal of the American Academy of Child and Adolescent Psychiatry, 35(6), 815–825. [PubMed: 8682763]

Beitchman JH, Wilson B, Brownlie EB, Walters H, & Lancee W (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. Journal of the American Academy of Child and Adolescent Psychiatry, 35(6), 804–814. [PubMed: 8682762]

Byun TMA, & Rose Y (2016). Analyzing clinical phonological data using Phon. Seminars in Speech and Language, 37(2), 85–105. [PubMed: 27111269]

Chomsky N, & Halle M (1968). The Sound Pattern of English. New York: Harper & Row.

Combiths PN, Amberg R, & Barlow JA (2016). Phonological analysis software in assessment and treatment of speech-sound disorders. Paper presented at the Convention of the American Speech-Language-Hearing Association, Philadelphia, PA.

Dinnsen DA (1984). Methods and empirical issues in analyzing functional misarticulation In Elbert M, Dinnsen DA, & Weismer G (Eds.), Phonological theory and the misarticulating child (ASHA Monographs No. 22) (pp. 5–17). Rockville, MD: ASHA.

Dinnsen DA, & Barlow JA (1998). On the characterization of a chain shift in normal and delayed phonological acquisition. Journal of Child Language, 25, 61–94. [PubMed: 9604569]

Dinnsen DA, Green CR, Gierut JA, & Morrisette ML (2011). On the anatomy of a chain shift. Journal of Linguistics, 47(02), 275–299. [PubMed: 22389522]

Felsenfeld S, Broen PA, & McGue M (1992). A 28-year follow-up of adults with a history of moderate phonological disorder: Linguistic and personality results. Journal of Speech and Hearing Research, 35, 1114–1125. [PubMed: 1280310]

Felsenfeld S, Broen PA, & McGue M (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: Educational and occupational results. Journal of Speech and Hearing Research, 37, 1341–1353. [PubMed: 7877292]

Ferguson CA, & Farwell CB (1975). Words and sounds in early language acquisition: English initial consonants in the first fifty words. Language, 51, 419–439.

Fry DB (1967). The phonemic system in children's speech. British Journal of Disorders of Communication, 3(1), 13–19.

Galen RS (1986). Use of predictive value theory in clinical immunology In Manual of Clinical Laboratory Immunology (3rd edition ed., pp. 966–970). Washington, DC: American Society for Microbiology.

Gierut JA (1986). On the relationship between phonological knowledge and generalization learning in misarticulating children [Doctoral Dissertation, 1985]. Bloomington, IN: Indiana University Linguistics Club.

Gierut JA (1992). The conditions and course of clinically induced phonological change. Journal of Speech and Hearing Research, 35, 1049–1063. [PubMed: 1447917]

Gierut JA (2015a). Experimental designs and protocols In Learnability Project Working Paper. Bloomington, IN: Indiana University.

Gierut JA (2015b). Participant eligibility and demographics In Learnability Project Working Paper. Bloomington, IN: Indiana University.

Gierut JA (2015c). Phonological protocols In Learnability Project Working Paper. Bloomington, IN: Indiana University.

Gierut JA, Elbert M, & Dinnsen DA (1987). A functional analysis of phonological knowledge and generalization learning in misarticulating children. Journal of Speech and Hearing Research, 30, 462–479. [PubMed: 3695441]

Gierut JA, Simmerman CL, & Neumann HJ (1994). Phonemic structures of delayed phonological systems. Journal of Child Language, 21, 291–316. [PubMed: 7929683]

Goldman R, & Fristoe M (1986). Goldman-Fristoe Test of Articulation. Circles Pines, MN: American Guidance Service.

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

Greiner M (1996). Two-graph receiver operating characteristic (TG-ROC): update version supports optimisation of cut-off values that minimise overall misclassification costs. Journal of Immunological Methods, 191, 93–94. [PubMed: 8642206]

Greiner M, Pfeiffer D, & Smith RD (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Preventive veterinary medicine, 45(1–2), 23–41. [PubMed: 10802332]

Jakobson R (1968). Child Language, Aphasia and Phonological Universals. The Hague, Netherlands: Mouton.

Law J, Boyle J, Harris F, Harkness A, & Nye C (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. International Journal of Language & Communication Disorder, 35(2), 165–188.

Lewis BA, Patton E, Freebairn L, Tag J, Iyengar SK, Stein CM, & Taylor HG (2016). Psychosocial co-morbidities in adolescents and adults with histories of communication disorders. Journal of Communication Disorders, 61, 60–70. [PubMed: 27032038]

McLeod S, & Baker E (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. Clinical Linguistics & Phonetics, 28(7–8), 508–531. [PubMed: 25000375]

Metz CE (1978). Basic principles of ROC analysis. Paper presented at the Seminars in Nuclear Medicine.

Morrisette ML, Farris AW, & Gierut JA (2006). Applications of learnability theory to clinical phonology. Advances in Speech-Language Pathology, 8, 207–219.

Peterson RL, Pennington BF, Shriberg LD, & Boada R (2009). What influences literacy outcome in children with speech sound disorder? Journal of Speech, Language, and Hearing Research, 52(5), 1175–1188.

R Core Team. (2013). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing Retrieved from http://www.R-project.org/

Rose Y, & Hedlund G (2017). Phon (Version 2.2) [Computer software]. Retrieved from https://www.phon.ca/phontrac

Shriberg L (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. Journal of Speech and Hearing Research, 36, 105–140. [PubMed: 8450654]

Shriberg L, Austin D, Lewis BA, 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.

Vermont J, Bosson J, Francois P, Robert C, Rueff A, & Demongeot J (1991). Strategies for graphical threshold determination. Computer methods and Programs in Biomedicine, 35(2), 141–150. [PubMed: 1914452]

Voegelin CF (1957). Six statements for a phonemic inventory. Journal of American Linguistics, 23(2), 78–84.

Youden WJ (1950). Index for rating diagnostic tests. Cancer, 3(1), 32–35. [PubMed: 15405679]

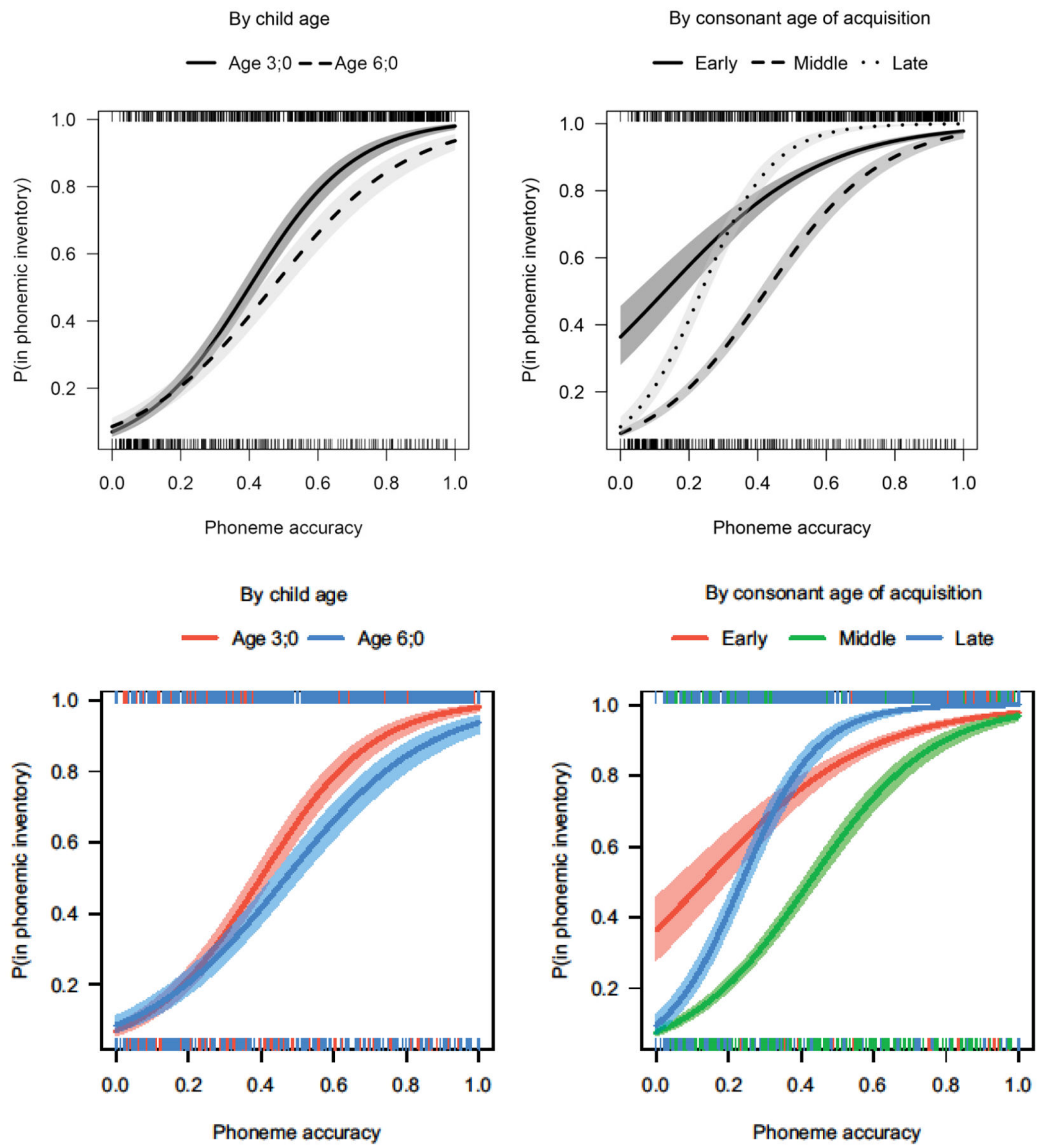

Figure 1.

Phonemic inventory classification.

Note. Regression trends are graphed on a logarithmic scale, displaying the probability of phonemic inventory inclusion according to phoneme accuracy. To illustrate the Consonant Accuracy × Child Age interaction, data are shown for children at 3 and 6 years of age (reflective of preschool and school-age groups, respectively). Steeper curves indicate better predictive power. Ticks along the upper and lower plot borders indicate actual data points.

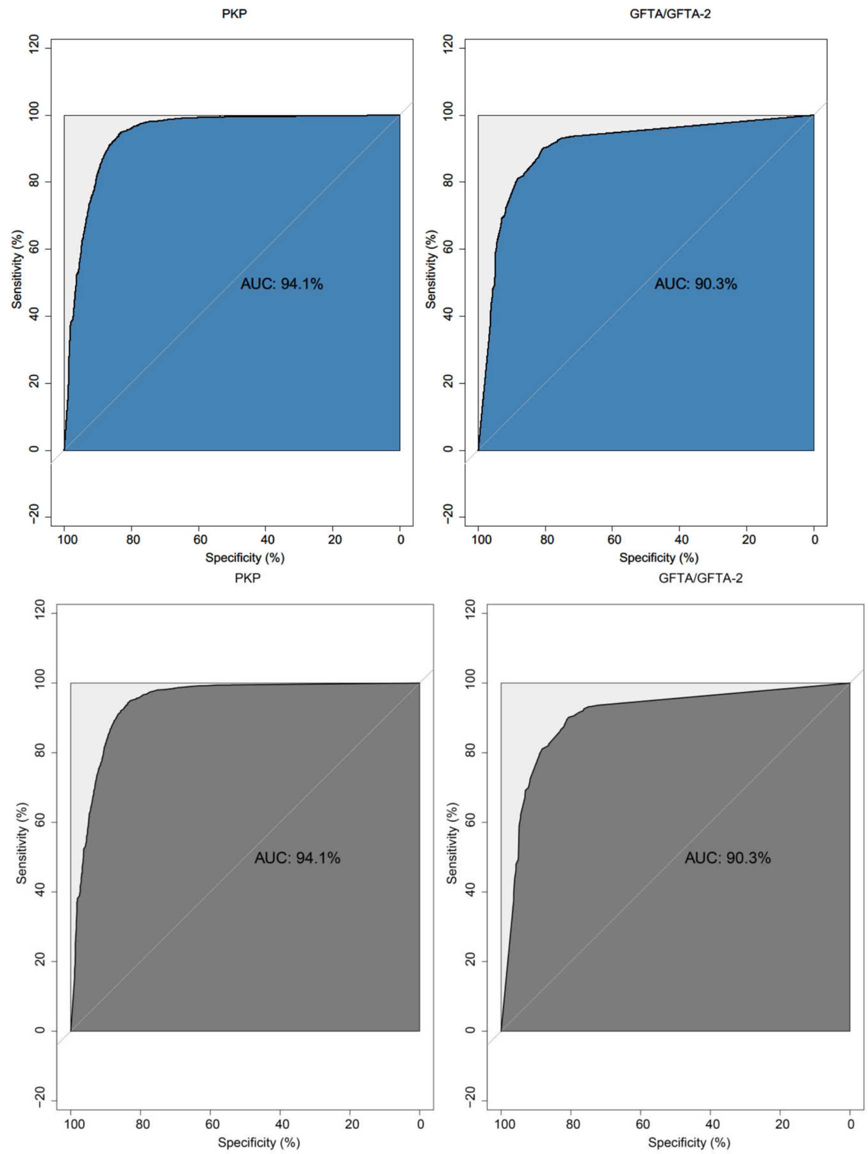

Figure 2.

ROC curve for phoneme classification via consonant accuracy by sample type.

Table 1.

Mean consonant accuracy by phonemic inventory, age of acquisition, and participant age.

	Age of acquisition	
	'Out' consonants	***'In' phonemes***
Early	52.8% (35.6%)	83.9% (19.4%)
Middle	8.4% (17.6%)	68.0% (26.3%)
Late	9.2% (19.3%)	55.6% (30.1%)
	Participant age	
	'Out' consonants	***'In' phonemes***
<5 Years	11.0% (21.9%)	72.8% (26.7%)
≥5 Years	16.6% (28.3%)	78.3% (24.7%)

Note. Consonant age of acquisition classification based on Shriberg (1993). Standard deviations in parentheses.

Table 2.

Potential consonant accuracy cutoff values predictive of phoneme status

	Maximum Efficiency	Youden's Index	Closest to ROC (0,1)
Cutoff	20.4%	21.1%	30.2%
Class.	90.1%	90.0%	89.1%
Sens.	94.9%	83.6%	90.9%
Spec.	83.1%	88.9%	86.6%

Note. Cutoff = optimum percent accuracy cutoff value. Class. = classification accuracy. Sens. = sensitivity. Spec. = specificity.

Table 3.

Phoneme classification metrics for potential accuracy cutoff values by sample type.

	20% Cutoff			**30% Cutoff**			**40% Cutoff**			**50% Cutoff**		
	Sens.	***Spec.***	***Class.***	***Sens.***	***Spec.***	***Class.***	***Sens.***	***Spec.***	***Class.***	***Sens.***	***Spec.***	***Class.***
PKP	0.95	0.83	0.90	0.91	0.87	0.89	0.85	0.89	0.87	0.79	0.91	0.84
GFTA	0.90	0.80	0.86	0.86	0.83	0.85	0.81	0.88	0.84	0.73	0.92	0.80

Note. Sens. = sensitivity. Spec. = specificity. Class. = classification accuracy.