Relationships Among Consistency/Variability and Other Phonological Measures Over Time

Ann A. Tyler, PhD; Kerry E. Lewis, PhD

This article explores selected phonological measures, their relationships to one another, and how groups differentiated by such measures change over time during intervention. Relationships among global quantitative measures of severity (percent consonants correct), measures of variability/consistency, and measures of whole-word complexity and syllable shape from 40 children with speech sound disorders were examined. All relationships were strong prior to intervention and remained stable during the course of intervention. Groups of 10 were differentiated with the measure of variability so that a variable group had many different error substitutions and the consistent group had few different error substitutions across the system. For these 2 groups, comparison of change at 3 points during the course of a 24-week intervention showed trends that were markedly similar in their linearity. There was steady improvement in percent consonants correct scores over time and a graduated decrease in the variability of errors on target sounds. The lack of a discernable difference between the consistent and variable groups in their response to the same intervention is seen as evidence to suggest that such groups may not need different types of intervention. **Key words:** consistency, ECI, intervention, phonology, variability

ALTHOUGH there are numerous qualitative indices for characterizing aspects of phonological development and impairment, there are fewer quantitative measures. Researchers have often developed measures to suit their theoretical orientation, quantifying different aspects of a child’s phonological system, such as underlying representations, phonological processes, or severity of impairment. As an example, percent correct underlying representations (Dinnsen & Chin, 1995) is calculated by quantifying underlying representations, phonological rules, positional constraints, and inventory constraints in a child’s system. Adhering to phonological process theory and analysis, the Process Density Index is used to measure the average number of phonological processes used per word in a sample (Edwards, 1992; Wolk, Edwards, & Conture, 1993). Some of these measures encompass multiple characteristics about the child’s phonological system, or cross an entire word, whereas others reflect measurement of individual phonemes, just as many standardized tests of articulation measure the accuracy of individual phonemes. This article will explore selected quantitative measures, their relationships to one another, and how groups differentiated by such measures change over time during intervention.

QUANTITATIVE MEASURES

The recently developed phonological mean length of utterance (PMLU; Ingram & Ingram, 2001) is intended as a measure of whole-word complexity and was designed to capture developmental growth in a child’s phonological
system in the same way that mean length of utterance reflects grammatical development with age. PMLU is calculated by assigning one point for each segment (consonant or vowel) of each word in a sample, additional points for each correct consonant, and determining an average per word for the entire sample. In addition, PMLU is the pivotal measure for comparing the complexity of a child’s productions with that of the target words attempted to determine whole-word proximity (Ingram & Ingram, 2001, p. 8). PMLU has not been validated as a measure that systematically changes with phonological growth, although Ingram and Ingram proposed different stages corresponding to ranges in PMLU scores and it is included in computerized analyses such as PROPH (Long, Fey, & Channel, 2002).

Another measure used to characterize a child’s phonological system is percent word match (PWM). It reflects the adherence to syllable shapes for target words in a sample (Bernhardt & Stemherger, 2000). A match is recorded when the child produces a syllable shape that matches the target syllable shape, regardless of whether the individual phonemes are produced correctly. For example, if the child produces a CVC for a target word that has a CVC shape, a match is recorded. However, if the child produces a CV for a CVC shape, there is no match. The number of matches is divided by the total number of target words and multiplied by 100 to obtain a percentage score. Both PMLU and PWM measures account for the correspondence between the number of segments in the child’s production and the target, thus reflecting additions and omissions regardless of accuracy.

Severity is another commonly used metric to characterize the extent of speech sound disorder. More often than not, severity is qualitative in nature. Hodson’s Assessment of Phonological Processes—Revised is a test that produces a severity rating based on a score (the Phonological Deviancy Score), reflecting common error patterns (Hodson, 1986). Perhaps the most familiar quantitative phonological measure is percent consonants correct (PCC; Shriberg & Kwiatkowski, 1982), which was designed to quantify severity of involvement through a measure reflecting the constructs of intelligibility, disability, and handicap. PCC is the percentage of target consonants that are correctly articulated by a child in a conversational speech sample and thus measures the accuracy of individual phonemes. Shriberg and Kwiatkowski (1982, p. 265) developed the following severity classifications: mild = 85%-100%; mild-moderate = 65%-85%; moderate-severe = 50%-65%; and severe = less than 50%. Although PCC was originally designed to be calculated from continuous speech samples, it is routinely calculated by clinicians and researchers from samples taken from articulation tests or other nonconversational, single-word samples. PCC and PMLU are similar in their accounting for correct consonant productions and are thus affected by incorrect productions or omissions.

Measures have also been developed to quantify consistency/variability in developing and disordered phonological systems. Variability has been of interest because it may be related to severity, as well as hold implications for change. Variability may be manifested by the occurrence of some correct productions for an error phoneme, or it may be exhibited by the occurrence of multiple different substitutes within and across word positions for an error phoneme. Measures that quantify variability/consistency have been developed to examine a child’s production of the same word multiple times, as well as productions of the same target phonemes in different contexts. Dodd’s 25-Word Test for Inconsistency involves three repetitions of multisyllabic words, with the criterion of 40% inconsistent incorrect productions designated as indicative of inconsistent phonological disorder (Dodd, 1995). A similar measure proposed to capture variability is the Proportion of Whole-Word Variation (Ingram, 2002). For this measure, multiple spontaneous productions of the same word are examined; the number of different versions is divided by the number of attempts, for each word produced multiple times in a sample, and an average
value for the sample is calculated. Values close
value to 1.00 would be indicative of increased vari-
bility, whereas those close to 0 would indi-
cate consistency in repeated productions of
individual words. The error consistency index
(ECI) is another metric designed to measure the
overall consistency of error substitutions
within a child's phonological system (Tyler,
2002; Tyler, Lewis, & Welch, 2003). The ECI is
a raw number that is calculated by summing
the total number of different substitutions that
a child makes, in each word position, for each
of the 23 consonant phonemes. A low ECI indi-
cates fewer different substitutions across a
smaller number of consonants, and a high ECI
indicates a greater number of different substi-
tutions across a larger number of consonants.
This measure was found to range from 12 to
70 in our sample of 40 preschool-age children
with speech-language disorders.

The quantitative measures discussed thus
far, PMLU, PWM, PCC, and variability mea-
sures such as ECI, may have the potential to
serve as global indices of change over time as
the result of intervention. In clinical settings,
change is often documented through the at-
tainment of long-term and short-term goals for
accurate sound productions in specified con-
texts. Change may also be monitored through
the comparison of standardized test scores,
although these are not ideally suited to mea-
suring intervention effectiveness (Paul, 2001).
In intervention efficacy studies, change is
often monitored through the use of single-
word probes designed specifically to sam-
ple target and generalization sounds in mul-
tiple contexts. Although these probes allow
for documentation of change on certain in-
tervention targets, they do not account for
global change in the system. Intelligibility rat-
ings may account for global change, although
these are problematic because they are influ-
enced by numerous variables and have been
shown to be quite unreliable even among ex-
perienced practitioners (Gordon-Brannan &
Hodson, 2000). Quantitative measures that re-
fect whole-word characteristics and/or ac-
count for the accuracy of all phonemes are
advantageous because they more closely rep-
resent all elements of the system as needed
in conversational speech. If quantitative mea-
sures are to be useful for documenting
change, they should first be investigated to es-
tablish their validity as well as their relation-
ship to one another.

CONSISTENCY/VARIABILITY AND
SUBGROUPS

The consistency/viability of error substi-
tutions has also recently been of interest be-
cause it may provide characterization of a
unique subgroup of children with phonologi-
cal disorder. Subgroups have been identified
on the basis of etiology (Shriberg, 2004) as
well as symptomatology. For example, Dodd
(1995; Dodd & Bradford, 2000) subgrouped
children with speech disorders according to
surface error patterns and hypothesized the
following groups: (1) delayed development,
(2) consistent nondevelopmental error pat-
terns, and (3) inconsistent errors. Although
such subgroups have received little system-
atic investigation, Dodd viewed variable substi-
tutions are viewed negatively and thought
by some to be reflective of an underlying
deficit in assembly, storage, or retrieval of the
phonological output plan (Dodd, 1995; Dodd
& Bradford, 2000). If unique subgroups exist,
these may hold important implications for as-
essment and intervention.

What is unclear is whether or not the group
with highly variable error substitutions sim-
ply corresponds to the severe end of the con-
tinuum in a population where phonological
knowledge varies, likely as a result of a vari-
ety of underlying deficits in the speech pro-
cessing system. Tyler, Williams, and Lewis (in
press) examined 20 children in each of two
groups representing the extreme ends of the
ECI score distribution. The two groups were
compared on a variety of quantitative mea-
sures to determine if their phonological sys-
tems differed in other primary ways. The re-
sults revealed significant group differences on
PCC, inventory size, and the number of er-
or substitutes on target sounds selected for
treatment. In all cases, the variable group
performed lower than did the consistent group. These differences suggest that children who have multiple substitutes both within and across word positions for a number of consonants (i.e., a high ECI) likely are at the more severe end of the continuum of children with speech sound disorders, perhaps with multiple underlying deficits. Similarly, in a group of 46 children with speech sound disorder, Isermann (2001) found no distinct subgroup of children with variable substitutions, using a measure of the percent of variable substitutions for phonemes in error across all sounds. However, when groups of 15 with the highest and lowest percent inconsistent substitutions were compared on standardized test scores, they differed significantly on the Goldman Fristoe Test of Articulation (GFTA; Goldman & Fristoe, 1986), with the more variable group scoring lower (Isermann, 2001).

**CONSISTENCY/VARIABILITY AND CHANGE**

From a developmental perspective, variability in behavior may hold important clues to patterns of change. For example, in the latter half of the prelinguistic period, increased phonetic variability is reflective of a transitional state presaging the transition from pre-word to early word behavior (Robb & Macrae, 2004; Thelen & Smith, 1993). Variability in different error substitutes may similarly be interpreted as instability in the representation of a phoneme and behavioral indication of reorganization. Results of several studies suggest that children with variable error substitution patterns respond differently to treatment than do those with consistent error substitutions, necessitating different intervention approaches for the variable group (Dodd & Bradford, 2000; Forrest, Dinnsen, & Elbert, 1997; Forrest, Elbert, & Dinnsen, 2000). On the basis of the findings that children with variable substitutions had greater difficulty learning their target sound and generalizing it to untreated word positions than did children with consistent substitutions, Forrest and Elbert (2001) suggested that treatment should be modified because children must learn that the same sound occurs in distinct environments. In some studies, however, pretreatment variability and learning outcomes were examined only for a specific set of phonemes, those that were missing from the children's inventories. However, variability for phonemes not present at all may be quite different from the variability of incorrect and correct productions for phonemes that do occur in the inventory. In fact, Isermann (2001) found that, across all phonemes, the greatest variability in error substitutes was for sounds that were in the middle range for PCC, suggesting they were sometimes correct, but subject to multiple substitutes. There was less variability for sounds that were never correct, and thus not in the inventory. When phonemes are present but are inconsistently correct and substituted, variability may be interpreted as a precursor to behavioral reorganization. This would suggest that the system is in an unstable state and might foreshadow a positive response to treatment. Assumptions that treatment must be somehow different for a presumed subgroup of children with variable error substitutes may be premature.

Recent findings regarding consistency/variability of error substitutions in children with speech sound disorder suggest a number of avenues for further exploration. First, there is a need to determine if measures of consistency/variability, such as the ECI, are related to other quantitative phonological measures, in particularly PCC, as a metric of severity. Although a variety of measures have been developed, many of them have not been sufficiently investigated to establish their validity or their relationship to one another. It would also be important to examine relationships between measures over the course of intervention to determine if they remain stable.

The purpose of this article is to first examine the relationships among several commonly used quantitative measures, including our measure of consistency/variability, the ECI. There is also a need to further explore the performance of groups identified for their extremely consistent or variable error
substitutions during the course of intervention to determine if there are similarities or differences over time in patterns of change for quantitative measures such as the PCC. It would also be important to examine change in the consistency/variability for targets of intervention. Second, we wished to examine change in the consistent and variable groups in our sample of 40 children by using global quantitative measures, such as PCC, and a measure of variability for only the intervention target sounds.

PARTICIPANTS AND PROCEDURES

For examination of relationships among measures, as well as examination of change over time for groups differentiated by their consistency/variability, we used a sample of 40 children who participated in a larger study conducted at the University of Nevada, Reno. These 40 preschoolers with speech and language impairments ranged in age from 3 years to 5 years 11 months (mean age of 4 years 2 months). Children selected for the study met specific selection criteria and were enrolled in the Washoe County School District early childhood programs. Participants were required to score at least one standard deviation below the mean on the Bankson-Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990); the mean standard score for Word Inventory was 69, for Consonant Inventory was 68, and for Phonological Process Inventory was 70. Participants were also required to score at least one standard deviation below the mean in expressive language on the Preschool Language Scale-3 (PLS-3; Zimmerman, Steiner, & Pond, 1992) or on the Clinical Evaluation of Language Fundamentals—Preschool (CELF-P; Wiig, Secord, & Semel, 1992), or to have a mean length of utterance in morphemes greater than one and a half standard deviations below the mean based on Leadholm and Miller's normative data (Leadholm and Miller, 1993). The expressive language mean standard score was 77, and the receptive language mean standard score was 86; the mean length of utterance was 2.70 in morphemes. Criterion for nonverbal cognitive functioning required that the participants be within one and one half standard deviations of the mean on the Columbia Mental Maturity Scale (CMMS; Burgemeister, Blum, & Lorge, 1972); the mean score was 103. All participants were reported to have neurological, behavioral, hearing, and motor skills within normal limits and demonstrated normal functioning on oral motor assessment (Robbins & Klee, 1987).

As part of the larger study in which these children were participants, speech and language assessments were performed three times; prior to treatment, midway through the treatment period, after 12 weeks, and at the conclusion of the treatment period, after 24 weeks. Speech samples were collected from administration of the BBTOP and a list of 15 supplemental words. With the addition of the 15 supplemental words, the sample provided at least three opportunities for production of each of the 24 consonants in both the word-initial and word-final positions. Administration of the BBTOP was performed by graduate assistants or certified speech-language pathologists in small, quiet rooms at the child's school or at the University of Nevada Speech and Hearing Clinic. The samples were audiotaped with two external lapel microphones using either a Marantz PMD 230 or 430 audio recorder.

The children's productions were transcribed online using broad phonetic transcription symbols from the International Phonetic Alphabet and then were checked by a senior researcher from audio replay of the session. Quantitative measures were calculated from transcriptions, which were also analyzed using the Interactive System for Phonological Analysis (ISPA; Masterson & Pagan, 1993) computer program to calculate PCC. ISPA also determines the percentage of occurrence of phonological processes, the frequency of occurrence of phones in the phonetic inventory, and consonant substitution analysis. PMLU, PWM, and ECI were calculated manually; secondary judges calculated these
measures on 20% of the samples to measure interjudge agreement. Derived interjudge agreement for the three measures was 99%. PWM was calculated by comparing each production from the BBTOP and 15 supplemental word samples to the target words. Using this sample, PMLU was calculated following Ingram and Ingram's procedure (Ingram & Ingram, 2001). ECI was calculated by examining ISPA printouts to identify sound substitution errors for each phoneme across all word positions and totaling the number of different substitutions made for all 23 phonemes (Tyler, Lewis & Welch, 2003).

**Intervention program**

All 40 children described previously participated in a 24-week intervention period, after the initial preintervention assessment. Goal selection procedures and intervention strategies are described in detail elsewhere (Tyler, Lewis, Haskill, & Tolbert, 2003); thus only the major features will be summarized here. The children attended therapy twice a week in their local early childhood programs for a 30-min individual session and a 45-min group session. Intervention was provided by graduate student interns supervised by early childhood program speech-language practitioners.

Each child had four phonology goals and four morphosyntax goals, which were cycled in different schedules across the 24 weeks of intervention. Goals were selected individually for each child on the basis of initial assessment results, and each goal was addressed for a total of 3 weeks (six sessions). Language goals addressed finite grammatical morphemes (e.g., past tense—ed, third person singular regular, copula and auxiliary be) because these have been shown to be particularly vulnerable in children with language impairments (Bedore & Leonard, 1998). For phonology targets, each child usually had three different initial and/or final sound targets and several cluster targets chosen by considering his/her phonetic inventory, phonological processes used (>35%–40% of the time), sound classes affected, word/syllable structure, and positional constraints. The target sounds were from among the obstruents /s, z, š, ±, št, k, g/ and liquids /l, r/. It is important to note that the different intervention sequences in the larger study did not lead to significant differences in phonological change, $F(3, 35) = 0.12, p = .95$.

Phonological intervention procedures primarily involved production practice activities, after the session's initial 5-10 min of auditory awareness activities. These involved listening to books containing numerous models of the target sound to heighten awareness. Drill-play activities were designed to establish production of sounds in new contexts. Naturalistic activities were used to facilitate practice in spontaneous communicative situations. Incorrect productions were immediately recast by the clinician in naturalistic activities, whereas in drill-play activities, clinicians responded to incorrect responses by providing a model and requesting an imitation. All clinicians were provided with the same books, stimulus pictures, and naturalistic scripts, accompanied by written directions for their use, to ensure reliable implementation across settings.

**RESULTS AND DISCUSSION**

**Relationships among quantitative measures over time**

We first wished to examine the relationships over time among several commonly used quantitative measures, including our measure of consistency/variability, the ECI. Measures of PCC, PWM, PMLU, and ECI were obtained for the 40 children on three occasions: preintervention (Time 1), after 12 weeks of intervention (Time 2), and after 24 weeks of intervention (Time 3). Pearson product–moment correlation coefficients illustrating the relationships among the four phonological measures at each time point are shown in Table 1.

At Time 1, a pattern of strong associations was found among the variables (Table 1). The strength of the associations among the variables was maintained in Time 2 data and
Table 1. Bivariate correlations (r) among variables PCC, PWM, PMLU, and ECI at preintervention and after 12 and 25 weeks of intervention (N = 40)

<table>
<thead>
<tr>
<th></th>
<th>PCC</th>
<th>PWM</th>
<th>PMLU</th>
<th>ECI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preintervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCC</td>
<td>.87</td>
<td>.97</td>
<td>-.88</td>
<td></td>
</tr>
<tr>
<td>PWM</td>
<td>.89</td>
<td>.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMLU</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 12 weeks</td>
<td>.93</td>
<td>.99</td>
<td>-.87</td>
<td></td>
</tr>
<tr>
<td>PCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PWM</td>
<td>.95</td>
<td>.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMLU</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 24 weeks</td>
<td>.84</td>
<td>.97</td>
<td>-.89</td>
<td></td>
</tr>
<tr>
<td>PCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PWM</td>
<td>.90</td>
<td>.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMLU</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PCC = percent correct consonant; PWM = percent word match; PMLU = phonological mean length of utterance; ECI = error consistency index.

Time 3 data. There were marked positive relationships among PCC, PWM, and PMLU and a marked inverse relationship between ECI and the other three measures. The very high correlation between PCC and PMLU reflects their similarity in requiring a tally of correct consonant productions. Calculation of PWM is similar, in part, to calculation of PCC and PMLU in that syllable shape matches are affected by additions and omissions. The equally strong negative correlations between ECI and each of the other measures are a reflection of the fact that ECI scores increase as substitution or omission errors increase, whereas PCC, PWM, and PMLU scores decrease as substitution or omissions errors increase. Thus, PCC, PWM, and PMLU increase as speech improves, whereas ECI decreases as speech improves.

The measures examined here were originally intended to quantify different aspects of a child’s phonological system. PCC was designed as a measure of severity of involvement, PWM as a measure of adherence to syllable shapes, ECI as a measure of overall error variability, and PMLU as a measure of whole-word complexity and an overall indicator of a child’s level of phonological development. Even with these seemingly different objectives, the calculations for all measures are based on similar facets of a child’s productions. As part of their calculations, the measures include, either intentionally or inadvertently, the element of phoneme accuracy and, particularly, consonant accuracy. While each measure may purport to capture a unique aspect of the phonological system, none has been sufficiently studied to establish construct validity. As one aspect of construct validity, discriminant validity is affirmed to the extent that different measures of different, but related, constructs are not so strongly associated that they appear to be measures of the same construct. The present findings of markedly strong relationships among the variables at three different points in time (Table 1) strongly suggest that the commonality among the four measures is of much greater magnitude than their uniqueness. In all probability, these measures should be viewed as four different measures of the same construct until further evidence challenges that conclusion.

Measures for consistent and variable groups over time

We also sought to examine the pattern of change over time for two groups of children, those with highly consistent or variable substitution errors. Children whose preintervention ECI scores fell at the extreme consistent and variable ends of the distribution of ECI scores from all 40 participants made up the two subgroups. Using the upper and lower quartiles of the ECI distribution, this resulted in two groups of 10 children. The upper quartile represented the variable group, in which ECI scores were greater than 44.75. The lower quartile represented the consistent group, in which ECI scores were less than 22.25.

For these two groups, we examined PCC and a quantitative measure of variability for each child’s three intervention targets from the sounds /s, z, f, j, tʃ, k, g, l, r/. This was a score similar to the ECI, but it represented
the total number of different sound substitutions/omissions across word positions for each child’s targets only. For example, if a child’s target sounds were /f, z, tʃ/ and he/she substituted [b, o, k for /z/, [p] for /f/, and [j] for /tʃ/, the target consistency score would be 5. Tyler, Williams, and Lewis (in press) found that, similar to ECI, this target consistency index (TCI) differed significantly at pretreatment for the consistent and variable groups.

Figure 1 shows group mean PCC scores obtained at Time 1, Time 2, and Time 3 for the two groups of children. Very similar trends can be seen for the two groups of children. Both the consistent and variable groups showed steady improvement in PCC across the three measurement points. For the consistent group, the mean PCC score was 73.9 (SD = 11.13) at Time 1, 76.9 (SD = 14.11) at Time 2, and 82.6 (SD = 10.6) at Time 3. For the variable group, the mean PCC score at Time 1, Time 2, and Time 3 was 38.6 (SD = 6.8), 50.0 (SD = 10.6), and 59.4 (SD = 12.4), respectively.

Figure 2 shows mean TCI scores at each of the three measurement points for the two groups of children. Both consistent and variable groups showed a smoothly graduated decrease (making fewer substitution errors) in TCI across the measurement points. For the consistent group, the mean TCI score was 5.0 (SD = 4.2) at Time 1, 4.2 (SD = 2.1) at Time 2, and 2.5 (SD = 1.9) at Time 3. For the variable group, the mean TCI score for the three measurement points was 9.4 (SD = 1.6) at Time 1, 7.4 (SD = 1.6) at Time 2, and 4.9 (SD = 2.7) at Time 3.

Visual inspection of Figures 1 and 2 suggests that the means for both the consistent group and the variable group were most accurately described by linear trends, trends with mildly positive slope for the PCC data and trends with moderately negative slope for the TCI data. As an additional check on the pattern of scores over time in the two groups, polynomial fitting was applied to the PCC data to make sure the trends for both groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time. Results indicate that the positive slope trends for both the groups were best described as linear as opposed to quadratic. If, for example, one group had a delayed response to intervention, a quadratic solution would result with the best fit line being relatively flat across the first two measurement points and then rising between points 2 and 3. Neither group displayed a quadratic trend in PCC over time.

As was the case with the trends found for PCC, neither of the two groups’ TCI data resulted in a quadratic solution. The variable group showed a significantly linear trend in TCI scores (F = 21.07, p = .001). While the linear solution for the consistent group did not reach significance (F = 3.4, p = .10), linear trend was a considerably more apt description of the data than was quadratic trend.
Relationships Among Consistency/Variability and Other Phonological Measures

\[ F = 0.857, p = .38 \]. This lack of statistical significance for the linear solution may have been due to the relatively large scatter (SD = 4.2) found for the consistent group at Time 1. However, from inspection of the means, it is clear that, like the variable group, the trend exhibited by the consistent group’s TCI scores was linear and not quadratic.

Based on results from both descriptive statistics and polynomial fitting, it seems clear that consistent and variable groups’ response trends were markedly similar in their linearity, producing a graduated increase in PCC scores over time and a graduated decrease in TCI scores over time. With respect to the data of the present investigation, there was no discernable difference between the consistent-error subjects and the variable-error subjects in their response to intervention over time.

Our findings showed that the two groups differentiated for their extreme consistency or variability in error substitutions changed similarly over the course of a 24-week intervention period on both global and more discrete quantitative phoneme measures. Both groups showed a steady improvement in PCC, a global measure of consonant accuracy from pre-intervention, to the 12-week post-intervention and 24-week post-intervention samples. On the discrete measure of the number of different error substitutes for just the children’s intervention targets (TCI), both groups also showed similar linear patterns of change as these scores decreased across the three measurement points. The linear trends evident for both groups indicate that neither had a delayed response or lack of response to intervention.

These results reflect an initial attempt to quantify change in groups differentiated by their error consistency/variability across all phonemes with not only a broad measure of consonant accuracy but also one for consistency of intervention targets. As such, the findings do not confirm those obtained in studies of individual children who have consistent versus variable substitutions for sounds missing from their inventories. Using a single-subject design, Forrest et al. (2000) found that four children with variable substitutions required more treatment sessions than did four with consistent substitutions and still did not generalize the target sound to untreated word positions. For example, children with variable substitutions had not reached a 50% generalization criterion in 20 sessions, whereas children with consistent substitution patterns terminated treatment after 9–16 sessions because that criterion had been met. On the basis of these previous findings, we might have expected our variable group to display a different response, perhaps delayed, to intervention as compared to the consistent group. The present finding of markedly similar trends for change in both PCC and variability for target sounds (TCI) for both groups suggests that they were similarly responsive to the same intervention conditions.

These present group data call into question the assumption that children with variable substitutions need a different type of intervention (Dodd & Bradford, 2000; Forrest & Elbert, 2001). Children in the variable group were just as responsive as children in the consistent group, after 12 weeks of intervention, which primarily involved structured, drill-play production activities for three target sounds presented in a cyclical format. These three sounds were not always absent from the children’s inventories, although they were from among the group of sounds most frequently cited as difficult for children with speech disorders (Rvachew, Nowak, & Cloutier, 2004). A similar number of children in each group (4) had targets absent from their inventories, although the children in the variable group had a larger percentage of targets absent from their inventories (37%) than children in the consistent group had (13%). Because measures were obtained for multiple targets in the present study, this may mask potential differences in learning for just one target sound. On the other hand, the quantitative measures reported here reflect change across the entire system. There were no differences between the consistent and variable groups when change on their targets and generalization of those targets to untrained word positions were compared from the beginning to end of intervention (Tyler et al., in press).
The larger number of targets absent from the inventories of children in the variable group is simply a reflection of their increased severity level. Similarly, increased variability appears to be a feature of increased severity. What was striking about the variable group was not its initial severe status and compromised systems, but its responsiveness to intervention, exhibited by a marked increase in PCC accompanied by a decrease in variability. This would suggest that variability quantified as multiple different substitutes across all phonemes in the system may be hypothesized to reflect instability in the representation of numerous phonemes. This unstable state may transcend individual phonemes and, in turn, render the entire system more amenable to reorganization, particularly when several phonemes are targeted in a cyclical fashion. Contrary to the belief that variability is a liability and requires a unique intervention, it may be a characteristic making children most suitable for a cyclical approach that targets multiple phonemes and error patterns.

Clinical implications

The two sets of findings reported here hold important implications for assessment and intervention. First, the strong relationships among quantitative measures of severity (PCC), variability/consistency (ECI), and whole-word complexity (PMLU) suggest that they may be measuring the same construct. Thus, children with severely compromised phonological systems are likely to be more variable in their error substitutions and produce less complex syllable shapes. Even over time during intervention, however, these relationships hold. The stability of the relationships among quantitative measures over time suggests that any of them may be useful as global indicators of change during the course of intervention. Because PCC is the most familiar, it shares the commonality of measuring consonant accuracy with the other measures, and it has received more validating investigation, it may be the measure of choice. If, however, clinicians are interested in documenting changes in variability of error substitutions, quantifying these with the target consistency measure used here appears to provide a useful way to monitor change on a discrete group of sounds. We provide a note of caution that our sample was controlled so that there were similar opportunities for phonemes in various positions and such a sampling technique would be necessary. Although a citation sample such as this is clinically preferable, there is a need to validate the measures we obtained with those from conversational samples as well.

Second, if groups are selected for their extreme behavior on a global phonological measure, such as variability/consistency, this does not necessarily imply that they need a different type of intervention. We may simply be sampling two ends of a continuum of disorder that respond equally well to behavioral interventions targeting their prominent deficits when change is indexed with quantitative measures at a global level. As evidence, the variable and consistent groups examined here showed similar linear trends in PCC change and increased consistency for their target sounds during the course of a 24-week intervention period. Given the significant differences in the preintervention capabilities of the two groups, our findings lend support to the possibility that additional factors having to do with motivation or learning style must be considered in explaining children’s responsiveness to intervention (Kwiatkowski & Shriberg, 1998).

REFERENCES

Bankson, N. W., & Bernthal, J. E. (1990). Bankson-Bernthal Test of Phonology-3BTOP. Austin, TX: Pro-Ed.


nonlinear phonology for clinical application. Austin, TX: Pro-Ed.


