DISTINCT DEVELOPMENTAL PROFILES IN TYPICAL SPEECH ACQUISITION

Jennell C. Vick¹, Thomas F. Campbell², Lawrence D. Shriberg³, Jordan R. Green⁴, Hervé Abdi⁵, Heather Leavy Rusiewicz², Lakshmi Venkatesh¹, and Christopher A. Moore¹

¹ Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington, USA
² Center for Research on Acquired and Neurologic Communication Disorders in Children, Children’s Hospital of Pittsburgh; and The University of Pittsburgh, Pittsburgh, Pennsylvania, USA
³ Waisman Center, University of Wisconsin - Madison, Madison, Wisconsin, USA
⁴ Department of Special Education and Communication Disorders, University of Nebraska - Lincoln, Lincoln, Nebraska, USA
⁵ School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, Texas, USA

Running Head: DEVELOPMENTAL PROFILES IN TYPICAL SPEECH

Corresponding Author: Jennell C. Vick
Department of Psychological Sciences
Case Western Reserve University
Cleveland Hearing & Speech Center
11635 Euclid Avenue - Room 333
Cleveland, Ohio 44106-7154
216-368-3352
216-368-6078 (fax)
jennell@case.edu

Copyright © 2012 by the American Physiological Society.
ABSTRACT

Three- to five-year-old children produce speech that is characterized by a high level of variability within and across individuals. This variability, which is manifest in speech movements, acoustics, and overt behaviors, can be input to subgroup discovery methods to identify cohesive subgroups of speakers, or to reveal distinct developmental pathways or profiles. This investigation characterized three distinct groups of typically developing children and provided normative benchmarks for speech development. These speech development profiles, identified among 63 typically developing preschool-aged speakers (ages 36 - 59 months), were derived from the children’s performance on multiple measures. These profiles were obtained by submitting to a \(k\)-means cluster analysis 72 measures that composed three levels of speech analysis: behavioral (e.g., task accuracy, percentage of consonants correct), acoustic (e.g., syllable duration, syllable stress), and kinematic (e.g., variability of movements of the upper lip, lower lip and jaw). Two of the discovered group profiles were distinguished by measures of variability, but not by phonemic accuracy; the third group of children was characterized by their relatively low phonemic accuracy, but not by an increase in measures of variability. Analyses revealed that, of the original 72 measures, eight key measures were sufficient to best distinguish the three profile groups.

KEYWORDS: Child Speech Production, Speech Motor Control, Subgroup Discovery, Articulatory Kinematics
Typical development of speech motor control from three to five years is both promoted and constrained by a wide range of developmental factors, including linguistic, phonologic, social, phonetic, phonemic, phonotactic (Gildersleeve-Neumann et al. 2008), physiologic, cognitive, and communicative influences. These challenges in speech development are compounded by the rapid and non-uniform growth of the vocal tract and of craniofacial musculoskeletal structures (Fitch and Giedd 1999; Vorperian and Kent 2007; Vorperian et al. 2005). Although this complex developmental process exhibits a broad range of commonalities across children and across languages (e.g., early development of bilabials), the principle of motor equivalence and the essential variability of speech production would seem to make it unlikely that there is a singular motor solution across typically developing children to these influences and constraints. Nevertheless, it does seem likely that some solutions would provide advantages with respect to enhanced motor control or speech output, such that the distribution of speech production variability may be non-uniform and non-monotonic.

Despite the well-described development of speech through highly predictable milestones, typical mature speech production is characterized by ubiquitous variability (MacNeilage 1970) across observational domains within and across speakers. This striking paradox in motor development (i.e., a highly generative and variable motor behavior that develops universally with predictable characteristics and milestones) presents an opportunity to evaluate the interaction of widely divergent developmental influences in achieving a conceptually singular behavior: spoken communication. It may be that speech motor development is supported by the child’s
access to motor stereotypies (Thelen et al. 1987) that underlie oral motor and respiratory behaviors, which s/he is able to access, fractionate, and remodel to achieve stabilized speech production (Green et al. 2002; Grillner 1980; Moore and Ruark 1996). Anatomic and physiologic constraints and growth patterns may similarly drive patterns of early speech development. A well-ordered path to speech production might be expected from these regularities in the physiologic characteristics of developing speech. Alternatively, children, like adults, may take advantage of the motor systems’ capacity to achieve a singular target (i.e., a speech sound) through a continuously variable range of movement solutions (Hebb 1949).

The high variability inherent to speech development can be used to reveal potential developmental trajectories across levels of observation (e.g., acoustic, kinematic, behavioral) of oral motor behaviors in children. For example, there are instances within and across children for which variability can be expected to be relatively low (e.g., early-acquired speech sounds and words). In contrast, increased variability, within or across children, may reflect the instabilities entailed by coordinative development, craniofacial growth, or the demands of increasing lexical and phonetic complexity. Although each child’s developmental context is unique, such that the developmental profile s/he exhibits will reflect high idiosyncratic variation, the ontogenetic and phylogenetic influences to develop speech may be so great that children are driven along a common pathway that reflects the trade-offs of motor capacities and behavioral (e.g., communicative, lexical) objectives. Identifying characteristic moments of changing coordinative stability that are common among subsets of children may reveal these developmental pathways to mature speech motor control, and further characterize children’s growing foundation for speech production. It would also enhance our understanding of the constraints and goals in speech development.
The near-universality of the development of speech motor control is evident in children’s early productions of fluent, relatively intelligible speech, despite the absence of a number of later-appearing speech sounds. Acquisition of speech sounds follows a fairly predictable order, though the timing of acquisition varies (Shriberg and Kwiatkowski 1994). Moreover, when children produce words containing sounds that are absent from their repertoires, they generate errors that are predictable and consistent with their developmental stage, including common substitutions, omissions, and distortions of yet-to-be acquired speech sounds. These anticipated patterns of disruption reveal vulnerabilities in this developmental pathway and suggest the operation of potent motor stabilities (e.g., Thelen, 1991), even in the presence of pervasive individual variability.

A range of physiologic measures, including articulator displacement or muscle activation patterns, supports a developmental framework specific to speech. These measures have already provided insight into the emergence and development of speech production, revealing, for example, that early speech and vocal behaviors are distinct from other, earlier appearing, nonspeech oral motor behaviors (e.g., electromyographic patterns differentiate vocal and non-vocal behaviors, even in infants; (Moore and Ruark 1996). Investigations of oral motor development have further revealed differentiation of speech and nonspeech behaviors over the course of several months in infants and young children (Green et al. 2000; Steeve and Moore 2009). The developmental trajectory of lip and jaw coordination during babbling and speech is nonlinear, and exhibits differences in coordinative control across articulators through the first six years of development (Green et al. 2000; Green et al. 2002; Steeve and Moore 2009). These group descriptions do not, however, address the children’s individual paths to mature speech motor control; these previous observations of relatively small groups of children precluded the
discovery of distinct developmental paths or stages. The present investigation, using a much larger sample of children and measures, was designed to evaluate the range of individual pathways to speech motor development, and to reveal distinct subgroups among these pathways. 

Refinement of Variable Speech Motor Control

Children exhibit articulatory movements that are smaller, slower, and more variable than adults (Goffman and Smith 1999; Maner et al. 2000; Smith and Goffman 1998; Smith et al. 1995; Smith and Zelaznik 2004; Smith and Gartenberg 1984; Walsh and Smith 2002). This variability is consistently found for point measures of temporal (e.g., onset of lower lip depression) and spatial (e.g., maximum lower lip depression) measures of displacement (Goffman and Smith 1999; Maner et al. 2000; Smith et al. 1995; Smith and Gartenberg 1984), and for whole-trajectory measures of stability (e.g., spatiotemporal index), which incorporate the entire movement gesture for a word or phrase (Goffman and Smith 1999; Goffman et al. 2008; Maner et al. 2000; Smith and Goffman 1998; Smith and Zelaznik 2004; Walsh and Smith 2002). This increased movement variability in children compared to adults is exacerbated by increased utterance complexity, reflecting the effects of cognitive processing demands on motor control of speech movement (Maner et al. 2000; Smith and Zelaznik 2004; Walsh et al. 2006).

It is unclear how variability and instability in the coordinative infrastructure for speech are manifest in the child’s speech output. It may be, for example, that the developing coordinative framework is most apparent only at the limits of the child’s speech capabilities. The predictable segmental errors in children with speech sound disorders may reflect the rate-limiting effects of motor development (Campbell et al. 2003; Davis 2005), and more generally, motor development may be deterministic in the rate and sequence of speech development. These influences have not been studied thoroughly however, as most studies of articulatory kinematics
in children include only perceptually correct productions of target words or phrases. This omission leaves the full range of effects (e.g., cognitive, social, linguistic) unaddressed, since physiologic constraints must be evident in erred as well as in correct productions. An important exception is the finding by Goffman, Gerken and Lucchesi (2007) of few significant correlations between kinematic and segmental measures in typically developing children or in those with specific language impairment. This finding, that articulatory variability has no direct relationship to segmental errors, supports a model in which the developing speech communication system is decoupled from and thus not entirely determined by the physiologic constraints of motor development. The present investigation was designed to further explore the relationships across levels of observation (i.e., behavioral, kinematic, and acoustic) in three- to five-year-old children. 

**Experimental Objectives**

The present investigation was designed to take advantage of the perspective on speech motor development afforded by children progressing through a very rapid and well-defined period of communicative growth. These observations were designed to isolate some of the behavioral elements of mastery for speech production (e.g., intelligibility; phonetic inventory) in conjunction with other key elements of coordinative development (e.g., acoustic or kinematic features). This investigation was also intended to reveal commonalities across children, such that those who exhibit shared relationships across variables could be identified as following a common developmental path, and that alternative developmental paths could also be distinguished. Accordingly, this experiment quantified and modeled a sample of physiologic and behavioral characteristics of speech development in young children to identify different solutions to the development of speech motor control, and to identify those measures that best differentiate among these subgroups of typical speech development. The long-term aim of this effort is to
elucidate the most common means by which typically developing children manage the inherent variability afforded by their developing speech motor and cognitive systems so that typical variation can be distinguished from the approaches to speech motor control used by children with speech sound disorders.

The search for developmental pathways shared by multiple children (i.e., comprising a subgroup) was investigated using a multivariate approach designed to reveal distinct groupings, stages, or continua in the measures obtained. The emergence of discreet statistical groups through the convergence of a broad array of empirical measures (i.e., through cluster analysis) required a relatively large number of children \( (n = 63) \) and a comparably large number of measures \( (n = 72) \) distributed across levels of observation (i.e., behavioral, acoustic, kinematic).

Following cluster analysis, discriminant analysis was used to identify the smallest effective set of specific key variables that could uniquely discriminate and characterize the emergent subgroups of typical children. Sampling across a range of ages (three to five years) during which speech production is being refined allowed for a range of maturational profiles to emerge.

The preliminary descriptive statistics obtained also provided a comprehensive quantitative characterization of the physiologic development of the labial, mandibular, and phonatory systems during speech and nonspeech behaviors in children with typical speech development aged 36-60 months. This data set comprises a normative reference set for typical speech development during early childhood. As a practical consideration, incorporating a broad range of measures is critical to the statistical description of speech production in children of this age, but the eventual utility of this approach requires the reduction of this large preliminary array of measurements to those that contribute most to the variance in performance across- and within-subjects.
METHODS

Participants

Eighty-three participants between 36- and 57-months of age with presumed typical speech development were initially enrolled in this study. Twenty participants were excluded: nine (10.7%) were found to have speech delay (i.e., according to the Speech Disorders Classification System); four (4.8%) dropped out of the study voluntarily; and the data from seven participants (8.4%) were not acquired due to technical problems that precluded further analyses (e.g., inadequate movement tracking). All protocols and analyses were completed for the remaining 63 participants. All children were monolingual speakers of English from the greater metropolitan area of Pittsburgh, Pennsylvania. Inclusion criteria also included parental report of typical speech and language development with no history of treatment by a speech-language pathologist or audiologist. Exclusionary criteria included histories of developmental, neurologic, or significant medical disorders.

>Insert Table 1 about here<

The 63 participants were characterized by the categorical demographic variables listed in Table 1. Continuous demographic variables (age, maternal age and education) are listed in indices 53-55 in Table 3. Participating children were between the ages of 36 and 57 months with an average age of 46.1 months. Fifty-seven percent of the participants were female (i.e., 36 of 63), 43% were male (i.e., 27 of 63), and 8% of the sample (5 of 63) reported a positive family history (at least one first-degree relative affected) of speech or language disorders. The age and years of education for each participant’s mother (mean age: 33.6 years; range 21-45 years; mean years of education: 16.4; range 10-22 years. See Indices 54 and 55 in Table 3) were also
included in the statistical model following earlier predictors of disordered speech development by Campbell and colleagues (2003).

Inclusion Criteria

Participants exhibited speech development within normal limits, as determined from a 15-minute conversational sample with a trained assistant. Narrow transcription of these speech samples were analyzed using the Programs to Examine Phonetic and Phonological Evaluation Records (PEPPER: Shriberg et al. 2001), which generated an error profile for each sample. These error profiles were compared to age-matched children in a lifespan reference database to confirm each child’s inclusion with respect to typical speech development.

Each participant also exhibited: receptive language skills within normal range (scaled score of 7 or above) as screened by the Linguistic Concepts subtest of the Clinical Evaluation of Language Fundamentals – Preschool (CELF-P; Wiig et al. 1992); oral structures within normal limits as evaluated by the Oral/Speech Motor Control Protocol (Robbins and Klee 1987); and normal hearing sensitivity, screened at 25 dB HL at 1, 2, and 4KHz.

Experimental Tasks

The remaining experimental tasks were elicited during a second experimental session, which usually occurred within two weeks of the first test session and included the targeted array of speech and nonspeech tasks. These tasks spanned a range of speech production complexity so that each child was expected to complete at least some of the tasks before reaching a level that exceeded his/her developmental capacities. Each type of task is represented in the data stream depicted in Figure 1.

Speech Sample (PEPPER)
Each child engaged in a play session to elicit a 15-minute conversational speech sample which was audio recorded for later transcription and analysis.

**Lexical Stress Task**

Target productions included five repetitions each of three trochaic (i.e., initial syllabic stress; e.g., BAba) and three iambic (i.e., second syllable stress; e.g., baBA) bisyllables. The consonant varied among /b/, /p/, and /m/ (i.e., BAba, baBA, PApa, paPA, MAma, maMA), but the vowel was always /a/. Five imitations of each target were produced by the child in response to a recorded adult female model, yielding up to 30 imitations of contrastive lexical stress productions (i.e., five repetitions each of six targets) for each child. Production of distinct lexical stress contrasts have been shown previously to reveal developmental differences among children in this age range (Shriberg et al. 2003). Occasionally, more or fewer than five repetitions of a target word were elicited from a child. In either case, all viable repetitions were included in the analysis, which yielded 1840 productions across all speakers (i.e., 937 trochees; 903 iambs; see Table 2).

>Insert Table 2 about here<

**Nonword Repetition Task**

Five repetitions of each of four nonword stimuli were elicited imitatively using a recorded female model. The four nonword stimuli were a subset of the nonwords included in the Syllable Repetition Task (SRT; Shriberg et al. 2009), were two- and three-syllables in length, and had primary stress on the first syllable (“bama,” “bada,” “bamana,” and “manaba”). These stimuli varied in difficulty and accordingly yielded an exaggerated range of variability; the systems and processes critical to repetition of nonwords include auditory perception, representation, formation, memory storage and retrieval, and motor planning (Shriberg et al. 2006). The range
of difficulty presented by the increasing syllabic complexity of these stimuli, like the Lexical Stress Task, provided a means to titrate the production abilities of children in this age range. Occasionally, more or fewer than five repetitions of each target nonword were elicited from a participant; in either case all repetitions from all participants were included, yielding a total of 1,362 productions. Details are shown in Table 2.

Nonverbal Tasks

The nonverbal tasks included two behaviors, voluntary vertical jaw oscillations and chewing. The vertical jaw oscillations were produced imitatively in response to a live model (provided by the experimenter) elevating and depressing the mandible cyclically (i.e., opening and closing the mouth) through five cycles whenever possible. This task was repeated five times. The second task, chewing, was sampled over two trials, during each of which the child chewed a single Goldfish™ cracker. Measures of performance on nonverbal tasks were included to distinguish speech-specific movement characteristics from those of more general orofacial behaviors.

Data acquisition

Children were seated in a Rifton positioning chair fitted with a table during the experimental session. They were instructed to sit upright and to keep their hands on the table, holding a soft toy to avoid hand movements that might interfere with the acquisition of speech movement data.

Audio Data

Audio recordings of the session were obtained using a lapel-style wireless microphone (Shure model UI-UA) affixed to the child’s forehead with surgical tape. This placement provided a fixed microphone-to-mouth distance. In a few instances, when the child would not
tolerate this placement, the microphone was taped to the headrest of the chair. The signal from
the microphone was amplified using a Mackie 12-channel mixer (model 1202-VLZ Pro). The
amplified signal was recorded with a video recorder (Panasonic, AG-1980), and then antialiased
and digitized with the video signal at a sampling rate of 44.1 KHz.

**Video (Articulatory) Data**

Vertical movement records of the upper lip, lower lip, and jaw were extracted from front
view video recordings of the child’s face. An infrared camera and light source (Burle, TC351A)
were used to illuminate small reflective markers attached in the midline of the child’s upper lip,
lower lip, and jaw (above the mental symphysis). Additional markers were placed on the tip of
the nose and the forehead to provide landmarks for correction of head movement by motion
tracking software. After each videotape was reviewed and logged, targeted task events were
digitized for subsequent parsing and analysis. The sampling rate for the kinematic data was 60
Hz. These movement records were low-pass filtered (\(f_{lp} = 15\) Hz) forward and reverse with a
digital, zero-phase shift, third-order Butterworth filter. In addition, the best straight-line linear
trend was removed from each displacement record to correct for very low frequency artifact
(e.g., baseline drift).

Two independent computer-based movement tracking systems were used to extract the
position of each marker in the frontal plane (i.e., vertical and lateral positions) in Cartesian
coordinates from the digitized video recordings. The first movement tracking system was
version 6.05 of Motus (Peak Performance). The second system, which offered significantly
improved processing speed, was DS-MTT version 2 by Henesis®, a custom MATLAB® routine
created for movement tracking. Concordance across the two systems was greater than 90% on
an overlapped sample of 15% of the total data set.
Data Processing and Standardizing

Acoustic Parsing

Each acceptable token was parsed using the audio signal from the video camera as a reference. The experimenter listened to the segment and roughly parsed the onset and offset of the production. The algorithm subsequently added 50 milliseconds to the beginning and end of the parsed signal to ensure inclusion of complete acoustic information for later perceptual judgments.

Analyses specific to the vocalic portion of each production required extraction of these segments from the word produced. The vowels of individual syllables were closely parsed and analyzed using custom algorithms in MATLAB®. The vowel onset was defined as the first positive-going zero crossing in the waveform when the waveform became periodic (i.e., vocalic). The vowel offset was defined as the final negative-going zero crossing in the periodic signal associated with the vowel. Audio playback also assisted the experimenter when completing fine acoustic parsing.

Jaw, Upper Lip, and Lower Lip Kinematic Parsing

Kinematic signals were parsed with reference to the vertical displacement of the jaw marker. The first order derivative of vertical jaw position (i.e., velocity) was superimposed on the position trace to parse the vertical jaw displacement onset and offset boundaries (i.e., using zero crossings, as described in (Green et al. 2000). The last zero crossing appearing in the velocity trace before the jaw depression for the vowel operationally defined the onset of the opening movement; the first velocity zero crossing during jaw depression for the final syllable marked the offset of final jaw elevation (See Figure 2). These onset and offset points were used to parse the displacement trajectories from the upper and lower lip.
Because the movement of the jaw contributes substantially to the displacement of the lower lip, the raw record of mandibular position included the movements of both the lower lip and the jaw. Accordingly, the jaw displacement signal was subtracted, sample-by-sample, from the lower lip displacement signal. The resulting trajectory represented the net lower lip movement (Green et al. 2000).

Nonverbal Task Parsing

Analysis of the nonverbal tasks (i.e., chewing and voluntary jaw oscillation) was limited to jaw marker movement (i.e., no other markers were analyzed). Samples that included fewer than three cycles or exhibited movement artifact (e.g., obscured markers) were excluded from the analyses. First and last chewing cycles were removed from each chewing trial, as well. Because the analyses for these tasks evaluated cycle-to-cycle variability, each cycle was demarcated. Jaw elevation-depression-elevation (open-close) cycles were parsed algorithmically, demarcating each cycle by its peak elevation (identified by the associated zero velocity point). Because of the irregular displacement signal associated with molar contact during chewing, numerous velocity zero crossings occurred during some instances of jaw elevation. In these cases, the algorithm specified the latest occurring zero-crossing as the cycle onset/offset boundary.

Conversational Speech Sample Transcription and Analysis

The examiners followed a standard SDCS protocol to obtain the speech samples; all data reduction was accomplished using the Phonology Project Laboratory Manual (unpublished). Narrow phonetic transcription of the continuous speech samples was completed by two experienced transcribers using consensus transcription procedures using a set of diacritical notations that classified subphonemic differences in articulatory place, manner, voicing, and
duration (Shriberg et al. 1984). Calculation of point-to-point interjudge (86.7%) and intrajudge
(91.8%) agreement for consonants was based on observations of 10 children who were part of a
larger sample that included the children from the present study as well as children with speech
delay of unknown origin. These procedures were described in a prior report (Shriberg et al.
2010).

Perceptual Analyses

Perceptual analyses were completed using the parsed audio signals, with separate
perceptual judgment tasks for the lexical stress and the nonword repetition tasks. In the lexical
stress perceptual task, audio files for each child’s productions of each of the bisyllables were
presented in randomized order including blocks of ten children (i.e., about 300 audio files in each
test block). Two listeners assessed whether the item was produced with the intended phonemic
target and identified which syllable was stressed (first, second, or both, when even stress was
perceived). All 1,840 contrastive stress productions were judged by two listeners. Joint
probability concordance between the two listeners was 89.1%. Phonemic accuracy was similar
between trochees (91.5%; 855/937; see Table 2) and iambs (88.7%; 801/903); however, imitative
stress was markedly better for trochees (95.7%; 897/937) than for iambs (88.6%; 800/903).

The phonemic accuracy of the nonword repetition task productions was judged blindly
(i.e., both to production target and participant) using an open-set, broad phonetic transcription
task. The presented items were randomized by participant and production type. Three listeners
judged each production. Transcription was finalized using a “best two of three” criterion (i.e.,
agreement by at least two of the three judges, which was reached on 95.2% (1297/1362) of the
productions. For the remaining 65 productions, the unedited video file was presented to enhance
transcription task. Only when all syllables in a single token were produced accurately was the
token scored as correct. By this criterion, 73% of the nonwords (998 of 1362) were judged to be
phonemically correct. Analysis by syllable length revealed decreasing accuracy from the first-
attempted, two-syllable productions to the final, three-syllable productions in the experimental
paradigm, with a drop in accuracy from the two-syllable (88.1%; 616 of 699 attempted) to the
three-syllable (57.8%; 383 of 683) productions (bada: 90.4%; 309 of 342; bama: 85.7%; 306 of
357; bamana: 63.1%; 210 of 333; manaba: 52.4% 173 of 330). Details are included in Table 2.

>Table 3 about here<

Construction and derivation of the modeled data set

Successful cluster analysis depends on adequate variances of the elements of a statistical
model such that individual cases (participants) can converge into distinct and reproducible sub-
populations. The present set comprised 69 continuous and three categorical variables spanning
three orthogonal domains of observation: behavioral, acoustic, and kinematic. Multiple
measures within each domain ensured that changes in performance within and across children
would be observed (e.g., ceiling and floor effects were avoided by including observations across
articulators of simpler and more difficult tasks, which appear earlier and later in motor
development). In addition, while numerous measures have been investigated, it was not known
which measures would be most sensitive to distinctions among emergent groups of preschool
children with typical speech acquisition. A final benefit of this large set of measures is its future
use as a benchmark data set for subsequent developmental studies of typical and disordered
speech. Mean and range values for each measure are reported in Table 3.

Behavioral Measures

Proportion Attempted (Index 1 in Table 3). Although most participants attempted all of the
verbal tasks, some children declined to imitate some of the presented tokens. Because the
resultant empty data cells would result in these children self-selecting out of the analyses
(perhaps for reasons unrelated to speech development) and contributing inappropriately and
immeasurably to cluster identification, the overall proportion of tasks attempted by each
participant was entered into the statistical model as an experimental measure. The overall mean
for this measure, 0.94 (ranging from 0.16 to 1.00; see Table 3) reveals that most children
completed all of the tasks presented.

Proportion of productions with all phonemics correct (Index 2 in Table 3). Perceptual
analyses yielded the overall proportion of verbal productions transcribed that matched the
intended target for each participant. Data from all attempted verbal productions (i.e., trochees,
iambs, and nonwords) were collapsed into a single measure from each participant.

Proportion of each lexical stress type with Correct Stress (Indices 3 and 4 in Table 3). The
proportions of attempted trochees and iambs produced with correct stress were included
separately for each stress type and each child as input measures to the statistical model.

Proportion Used (Index 5 in Table 3). This value reports the proportion of all tokens produced
by a child that were produced with both phonemic and lexical stress accuracy and had useable
kinematic trajectories (i.e., free of movement artifact or markers obscured by hand motion).
Children sometimes were minimally constrained and sometimes produced spurious movements
out of frustration with a particular task, which reduced the number of usable tokens. Tokens
produced with accurate phonemics and stress were included in acoustic analyses, but were
omitted from the kinematic analyses when necessary. Overall, 70% of the attempted verbal
productions by the participants in the study were suitable for all levels of measurement.
The Programs to Examine Phonetic and Phonological Evaluation Records (PEPPER: Shriberg et al. 2001) were used to generate error profiles for each child from a 15-minute conversational speech sample. The PEPPER software environment was used a) to classify participants’ inclusion status as described previously and b) to produce percentage scores on all of the competence variables (see below). It yielded a number of speech performance metrics that were included as input to the classification model. Each measure is intended to provide a metric of a unique speech domain within speech production.

**PCC:** Percentage of Consonants Correct; measure of phonetic accuracy, designed to quantify common and uncommon speech sound distortions, as systematized in the Appendix to Shriberg et al (1997).

**PCCR:** Percentage of Consonants Correct, Revised; this index, which does not count distortions (allophonic variation) as errors, provides a measure of phonemic accuracy. This measure provides greater distinction between children with normal speech acquisition and those with speech delay in lifespan reference data. It reflects the acquisition of the phonemes in the child’s ambient community (Shriberg et al. 1997).

**PVC:** Percentage of Vowels Correct.

**PVCR:** Percentage of Vowels Correct, Revised; parallel to PCCR, does not count distortions as errors.

**PPC:** Percentage of Phonemes Correct.

**PPCR:** Percentage of Phonemes Correct, Revised; does not count distortions as errors.

**II:** Intelligibility Index; proportion of unintelligible words to total words produced, expressed as a percentage.

**AWU:** Average Words per Utterance; an estimate of the child’s expressive language
function and verbal productivity, which varies greatly within typically speaking children in any given continuous speech sample. AWU is very highly correlated ($r > .90$) with mean length of utterance in preschool-aged children (Shriberg and Kwiatkowski 1994). This output measure reflects general lexical productivity, not grammar.

---

**Acoustic Measures**

**Mean Acoustic Area Ratio – Trochees and Iambs (2 measures; Indices 6 & 7 on Table 3).**

For the lexical stress task the primary acoustic analysis used was the Acoustic Area Ratio (adapted from Xie et al. 2004), which expresses the relative energy of the first syllable relative that of the second syllable. For each syllable, the mean amplitude was multiplied by the vowel duration to provide a measure of acoustic area (i.e., under the rectified and smoothed acoustic signal). Because the audio signals were not calibrated, amplitude was measured using the root mean square (RMS) of the signal in 20ms windows. The mean amplitude for each participant and task was expressed as an average of these values. The Acoustic Area Ratio is the quotient of acoustic area of the first syllable and the second syllable; values greater than 1.0 indicate that primary stress was on the first syllable, those less than 1.0 indicate that primary stress was on the second syllable. The measure is similar to a metric used in the identification of subgroups of children with suspected apraxia of speech, the Lexical Stress Ratio (Shriberg et al. 2003).

Overall, as expected and reported in Table 3, mean acoustic area ratio was greater for trochees (1.89) than for iambs (0.62). Individual data varied with ranges of 0.6 – 4.72 for trochees and 0.22 – 1.31 for iambs.
Coefficient of Variation of Acoustic Area Ratio – Trochees and Iambs (2 measures; Indices 8 & 9 on Table 3). Variability in lexical stress production was evaluated using the coefficient of variation. This measure, the quotient of the standard deviation and the mean, was used to provide a normalized comparison of variability across widely divergent means and units of measure, and was comparable for trochees (43.73) and iambs (40.81).

Word Duration – two- and three-syllable nonwords (2 measures; Indices 10 & 11 on Table 3). Word duration was measured for the two- and three-syllable nonwords from the burst of the /b/ or the initial glottal pulse of the /m/ to the final glottal pulse of the final /a/; glottal offset was operationally defined as the final negative-going zero crossing in the periodic signal associated with the vowel. Durations were collapsed within word length (i.e., across two-syllable nonwords: “bama” and “bada”, and across three-syllable nonwords: “bamana” and “manaba”). Duration was substantially greater across children for three-syllable productions (1.45 seconds) than for the two-syllable productions (0.77 seconds).

Coefficient of Variation of Word Duration – two- and three-syllable nonwords (Indices 12 & 13 on Table 3). These measures reflected the normalized variability of word duration for each participant. As seen in Table 2, the mean values for the group were comparable for the two- (11.78) and three-syllable (10.32) nonwords.

Kinematic Measures

Mean Maximum Displacement – upper lip, lower lip, jaw both for verbal (Indices 14 – 16) and nonverbal (Indices 56 & 57) tasks. The maximum vertical displacement of the first syllable of all productions was measured for the upper lip, lower lip, and jaw. For nonverbal productions, the maximum displacement of the first cycle was measured for the jaw marker only.
This measure was extracted automatically using custom scripts in MATLAB® that identified the local minimum (jaw and lower lip) or maximum (upper lip) adjacent to the first vowel onset. Mean maximum displacement of the upper and lower lips produced by the group on verbal tasks were similar (0.22cm and 0.24cm, respectively), whereas the mean measure for the jaw was much larger (0.72cm). Mean values for jaw maximum displacement were slightly larger for nonverbal tasks, including voluntary jaw oscillations (1.08cm) and chewing (1.00cm).

Coefficient of Variation for Maximum Displacement – upper lip, lower lip, jaw both for verbal (Indices 17-19) and nonverbal tasks (Indices 57 & 58). The normalized variability for first-syllable (verbal tasks) and first cycle (nonverbal tasks) maximum displacement was measured using the coefficient of variation. For the group, variability was similar for the upper and lower lips (39.3 and 38.0, respectively) and the jaw (31.2). Mean values for the jaw were similar for the nonverbal tasks, including voluntary jaw oscillation (38.0) and chewing (31.0).

Spatiotemporal Index – Trochees (Indices 20, 23, 26), iambic (Indices 21, 24, 27), 3-syllable nonwords (Indices 22, 25, 28), and nonverbal tasks (Indices 60, 61); Upper lip, lower lip and jaw (9 measures). Measures of movement stability were derived from the whole-word movement records for trochaic, iambic, and three-syllable nonword productions for each single articulator. The spatiotemporal index (STI), developed by Smith and colleagues (Smith et al., 1995), uses time- and amplitude-normalization of movement records for repeated whole words to reflect movement stability. Amplitude normalization of individual trajectories was accomplished by dividing each displacement record by its standard deviation. Linear temporal normalization involved interpolation of each amplitude normalized trajectory to 1,000 points using a cubic spline fitting algorithm in MATLAB®. The standard deviation was calculated across repetitions.
DEVELOPMENTAL PROFILES IN TYPICAL SPEECH

For each set of 20 consecutive points (i.e., 2% of the normalized sample per calculation); the sum of these 50 standard deviations yielded the STI.

For both the lexical stress task and the three-syllable nonword repetition task, STI was calculated for the upper lip, lower lip, and jaw markers' vertical displacements during each target production; only correct repetitions were included in these analyses. Overall, STI values for the upper lip and lower lip were higher than for the jaw; and values for the three-syllable nonwords were higher than for the two-syllable trochee and iamb productions (See Table 3).

A variant of the STI was used for the nonverbal tasks: the cyclic spatiotemporal index (cSTI), which quantifies the stability of repeated individual movement cycles (van Lieshout and Moussa 2000). Calculation of the cSTI used the parsed open-close cycles from the jaw vertical displacement records which were time- and amplitude-normalized and aligned by start and end times. The sum of the 50 standard deviations calculated at every 20th sample constitutes the cSTI. This measure of variability was higher for chewing (26.52) than for voluntary jaw oscillations (15.19).

**STI for Lip Aperture and Jaw/Lower Lip Aperture – Trochees, iambs, and 3-syllable nonwords (Indices 29-34; 3 measures).** Two additional measures of combined articulatory movement stability were calculated for whole words. Analysis of the net articulatory effect arising from the movements of the upper and lower lips reflects the underlying articulatory goal and its associated functional synergies (Smith & Zelaznik, 2004). The STIs of lip aperture trajectories (Kleinow and Smith 2006) were derived from the difference in trajectories for the upper and lower lip for each correct production. A similarly derived STI was calculated for the normalized difference trajectories between the lower lip and the jaw displacements. Lip aperture
STI values shown in Table 3 (Indices 29 – 31) were generally lower (reflecting lower variability) than lower lip / jaw STI values (Indices 32 – 34).

**Convergence Index – Trochees, iambs, and 3-syllable nonwords; Upper lip, lower lip and jaw (Indices 35 – 43 on Table 3; 9 measures).** Following Goffman and colleagues (2007) a Convergence Index (CI) was calculated. This measure was derived from all attempted productions, including most notably those that contained phonemic and lexical stress errors, as well as those produced accurately. This index provided a measure of kinematic stability, even for children who produced few, if any, repetitions with phonemic or lexical stress accuracy.

**CI / STI Ratio - Trochees, iambs, and 3-syllable nonwords; Upper lip, lower lip and jaw (Indices 44 – 52 on Table 3; 9 measures).** The ratio of the CI and STI was derived as a reflection of the relative stability of movement patterns across correct and incorrect productions. The CI/STI ratio has a value of 1.0 when only correct productions were obtained; higher ratios occurred when the variability was higher for the incorrect productions than the correct productions. This measure, for example, will reveal speakers whose variability increased dramatically for incorrect phonemic or lexical stress productions. For the group, the mean values for these ratios were all close to one, although the ranges were large in some cases (e.g., 0.46 – 3.00 for UL on 3-syllable productions – Index 46). Ratios less than one reveal that some repeated, incorrect productions were produced with less variability than correct productions for individual speakers.

**Consideration of Age and Sex as Covariates**

Because of well-described developmental and sex-related differences among many of these measures, it was essential to evaluate explicitly these potential effects. The linear relationship between age and each of the other 69 continuous measures was evaluated by
calculating the Pearson correlation for each measure with age. Although all of the significant
correlations were low using Cohen’s criterion (i.e., \( r < .5 \)), the highest linear relationship with
age was noted in variables derived from the PEPPER conversational speech sample analysis.
Specifically, the Percentage of Consonants Correct (PCC), the Percentage of Consonants
Correct, Revised (PCCR), the Percentage of Phonemes Correct (PPC), and the Percentage of
Phonemes Correct, Revised (PPCR) each had correlations greater than 0.4 \((p \leq 0.001)\). This
result was highly anticipated, of course, as these age-normed measures were explicitly developed
to be especially sensitive to speech development.

Sex differences were tested using a two-way between-groups analysis of variance; these
potential effects were readily dismissed as deterministic for each of the continuous variables.
There were no significant differences between males and females on any of the measures \((\alpha =
.05)\).

**Data Reduction and Analysis**

An obvious challenge associated with modeling a data set of 72 variables is narrowing
consideration to those variables that are orthogonal, most representative of group behaviors, and
sufficiently variable to inform interpretation of the model. Figure 2 provides a schematic of
these data reduction steps and subsequent analyses.

**Multiple Factor Analysis**

Multiple Factor Analysis (MFA; Abdi and Valentin 2007; Escofier and Pagès 1990;
Pagès and Husson 2005) is a variant of Principal Component Analysis (PCA); it is used to
integrate multiple data sets obtained from a single set of observations. MFA was used in the
present approach to reduce the dimensionality of the data preparatory to cluster analysis.
Conceptually, MFA organizes variables into tables of similar variables, with each table matrix
normalized prior to calculation of a global PCA. MFA is especially appropriate for applications like the present one in which the analysis of many different groups of variables are used to describe the same set of observations (Ding and He 2004; Gan et al. 2007).\(^1\)

Prior to MFA, each variable was first centered so that its mean was zero, then normalized so that the sum of its squared elements was equal to one. The MFA analysis was performed with XLSTAT (v. 2008; Addinsoft). The inclusion of factors in subsequent analyses was based on the cumulative proportion of the variance associated with each factor (Abdi and Williams 2010; Fielding 2007; Jolliffe et al. 1986). Subsequent analyses were based on the first 24 of 62 resulting factors, which accounted for 90.99% of the variance in the original data set. The remaining factors individually account for very small amounts of the total variance (i.e., 9% divided among the 38 remaining factors) and, therefore, could be dropped from further consideration.

**Cluster Analyses**

The factors derived from the MFA were subjected to cluster analyses for post hoc identification of subgroups, using XLSTAT (v. 2008; Addinsoft) and two independent clustering algorithms. The first algorithm, hierarchical agglomerative clustering, was used to generate hypotheses about the number of clusters within the data; this step enhanced the subsequent cluster analysis approach, \(k\)-means, which requires the user to specify the number of clusters \textit{a priori}. Using Euclidean distance to measure dissimilarity and Ward’s method for agglomeration, the participant factor scores for each of the first 24 factors were submitted to a hierarchical agglomerative cluster analysis of all 63 children. The results of this clustering algorithm suggested that a six cluster solution was optimal, although two of these clusters had only single
cases. When these single cases were merged with adjacent clusters, the total number of groups was reduced to three, suggesting that a three-cluster solution was optimal.

Confirmatory clustering for this data set used $k$-means clustering, which is an iterative partitioning method of clustering that minimizes the within-cluster variation for a specified number ($k$) of clusters. The result is the creation of a family of clusters in which each entity belongs to just a single cluster (Fielding 2007; Mirkin 2005). This common clustering algorithm (Gan et al. 2007) has been used effectively to classify children with specific language impairment (Conti-Ramsden et al. 1997) and children with speech delay of unknown origin (Venkatesh 2007).

To refine the clusters discovered using agglomerative hierarchical clustering, $k$-means clustering was run with two, three, four, five, six and seven cluster solutions, using the groups found by the agglomerative clustering as initialization seeds. To validate and confirm the optimal cluster solution, external validity indices were computed using the Cluster Validity Analysis Platform (CVAP, v3.4; Wang 2007), a custom graphical-user interface in MATLAB®. Internal validity of a cluster solution uses only features inherent to the dataset; indices typically measure the cohesion of clusters and the separateness of the clusters in the solution. Two indices were applied to the present solution, the Silhouette Index and the Calinski-Harabasz Index. The Silhouette Index evaluated the overall quality of the cluster solution by measuring each cluster’s tightness and separation. The optimality of three-cluster solution was further supported by its highest average silhouette value. The Calinski-Harabasz Index was used to confirm these results using a pseudo-$F$ statistic, evaluating the cluster solution by evaluating the cohesion and separation of clusters (Calinski and Harabasz 1974). Again, the three-cluster solution was determined to be optimal.
Mean values for each of the 69 continuous raw measures for each of the three clusters are reported in Table 3 in the columns labeled Clst 1, Clst 2, and Clst 3.

ANOVA

To identify significant differences in the continuous raw variables among clusters, a one-way, unbalanced ANOVA (XLSTAT; v. 2008, Addinsoft) was used to evaluate each of the variables independently as a regressor on the three clusters, applying a Bonferroni correction for pairwise comparisons. Pairwise comparisons are reported in the last three columns of Table 3.

Descriptive Discriminant Analysis

A forward stepwise descriptive discriminant analysis was performed to determine what minimal combination of the original raw measures would best differentiate the three cluster groups. The non-difference of the three clusters’ covariance values was confirmed with a Box Test ($\chi^2 = 89.60; df = 182; p = .99$), which indicated that the data were in agreement with the assumptions of a discriminant model. The probabilities specified in the discriminant analysis were proportional to the cluster group sizes, as there was no a priori expectation of different probabilities for membership in a given cluster. The result was two discriminant functions (formed from linear combinations of eight of the original 72 raw variables) that were found to maximize the differences between the three clusters of participants. Because stepwise methods can be associated with inflated chance associations and significance rates, jackknife resampling was used to cross-validate the discriminant function to decrease the probability of overestimating the inter-cluster distinctiveness (Tan et al. 2006). The hit rate on the cross-validated sample with resulting discriminant functions was 95.2%.

>Figure 3 about here<
Figure 3 depicts the three clusters of individuals in the two-dimensional statistical space created by the two discriminant functions. The clusters’ centroids appear as black stars and the cases are surrounded by 95% confidence ellipses. The ellipses indicate the certainty of the positions of the clusters in the statistical space, with a probability of 1 - \( \alpha \) (95%) and assuming a bivariate normal distribution. The minimal overlap of these ellipses indicates that these groups are distinct.

**RESULTS**

Inspection of the raw measures that contributed most substantially to the two dimensions (factors) of the discriminant space elucidated the most salient and distinct characteristics of the individuals comprising each of the three clusters. Pearson correlation statistics for the two factors and the eight most significant measures appear in Table 4. For each measure, the correlation with Factor 1 and Factor 2 are included (i.e., \( \rho_{F1} \), \( \rho_{F2} \)) with the predominant correlations in boldface. A high correlation value on Table 4 suggests that the measure accounted for a large proportion of the variance for that factor. In other words, Factor 1 and Factor 2 can be described best using the measures with which they had the highest correlation values. In addition, these correlations can be used to describe the individuals who score at the extreme ends of Factor 1 and Factor 2 in Figure 3. For instance, individuals in Cluster 2 scored on the high end of Factor 1, suggesting that they also score high on those measures that were positively correlated with Factor 1 and score low on those measures that were negatively correlated with Factor 1.

>Table 4 about here<
Factor 1 (F1; x-axis in Figure 3) accounted for 64.6% of the variance in the data and is positively correlated with:

- Acoustic area ratio of the first syllable to the second during trochaic productions (Index 6 on Table 3)
- Variability of acoustic marking of stress in iambic productions (Index 19)
- Jaw STI during iambic productions (Index 27)
- Variability of jaw maximum displacement in the first syllable of all productions (Index 9)
- Cycle-to-cycle variability of jaw movements during chewing (Index 61)
- Variability of lower lip movements during all productions that were meant to be trochees, regardless of accuracy (convergence index; Index 38)

Participants with high scores for Factor 1 (i.e., Cluster 2; see Figure 3) exhibited greater variability across behaviors (verbal and nonverbal) and across domains (kinematic and acoustic). This factor was labeled in Figure 3 as “Variability” to highlight the association of higher scores on Factor 1 with greater variability across a range of measures.

Factor 2 (F2; y-axis) accounted for 35.3% of the variance in the data and was positively correlated with:

- Percentage of Consonants Correct, Revised (Index 63)
- Maximum displacement of the jaw during chewing (Index 57)

Children with high scores on Factor 2 (i.e., Cluster 1 and Cluster 2) had high scores on the PCCR and lowered their jaws more during chewing. This factor was labeled in Figure 3 as “Phonemic Accuracy” to highlight the association of higher scores on Factor 2 with higher scores on measures of phonemic accuracy.
Subgroup Descriptions

The members of Cluster 1 \((n = 21;\ 10\ \text{males}\ \text{and}\ 11\ \text{females})\) generated measurements that were significantly different from those from members of either Cluster 2 or Cluster 3 on 40 different measures (see pairwise comparisons in the last three columns of Table 3). Specifically relative to members of Cluster 2, members of Cluster 1 produced all of the tasks, both verbal and nonverbal, with less variability (i.e., lower coefficients of variation for acoustic and kinematic metrics and lower spatiotemporal and convergence indices: Indices 8, 9, 17-20, 23-36, 38-41, 43 and 61 on Table 3). These results were confirmed by the gathering of members of this group at the low end of Factor 1 in the discriminant analysis, which was associated with low measures of variability. In addition, compared to members of Cluster 2, a greater proportion of the speech tasks produced by members of Cluster 1 were usable for kinematic analyses (i.e., more productions with accurate phonemics, lexical stress, and viable articulator tracking; Index 5 on Table 3). Compared to members of Cluster 3, members of Cluster 1 were significantly older (Cluster 1 mean = 49.3 months, sd = 5.6; Cluster 3 mean = 42.9 months, sd = 5.0; Index 53 on Table 3); there was no significant difference in age between Clusters 1 and 2. Members of Cluster 1 produced distinguishing speech and nonverbal tasks with less kinematic variability than members of the other clusters (e.g., Indices 18, 29, 41 on Table 3). Accordingly, the Cluster 1 profile group was designated as \textit{High Stability}.

Members of Cluster 2 \((n = 25;\ 6\ \text{males}\ \text{and}\ 19\ \text{females})\) included a greater proportion of females (.76) than the members of the Cluster 1 (.52) or Cluster 3 (.35). Cluster 2 members were more variable in their acoustic realizations of iambic stress tokens compared to members of the other clusters (Index 9 on Table 3). Both point and dynamic measures of articulatory variability were significantly greater for members of this group compared to the others (Indices 18, 26, 29,
Accordingly, the Cluster 2 profile group was designated as *High Variability*. Members of this cluster also distinguished themselves from members of the High Stability profile group on the two measures of acoustic area ratio (Indices 6 and 7 on Table 3). For both trochaic and iambic productions, values for the acoustic area ratio were significantly higher, suggesting that first syllables were produced with greater emphasis (i.e., louder and longer) by the High Variability group compared to those produced by members of the High Stability group.

Members of Cluster 3 ($n = 17$; 11 males and 6 females) were significantly younger (42.9 months) than members of the High Stability profile group (49.3 months), but no significant age difference was found when they were compared to members of the High Variability group (45.6 months; Index 53 on Table 3). Their low scores on the Phonemic Accuracy dimension of the discriminant analysis (see Figure 4) were consistent with the pairwise comparisons of the individual raw measures. Specifically, members of Cluster 3 scored lower on many of the PEPPER output variables (i.e., PCC, PCCR, PPC, PPCR; Indices 62 - 67) compared to members of the other two profile groups. For this reason, this profile group was designated as *Low Phonemics*. In addition, members of the Low Phonemics profile group had jaw and lower lip maximum displacements that were significantly lower than those produced by members of the other two groups when chewing (Index 57) and lower than the High Stability profile group on verbal tasks (Indices 15 & 16).

Figure 4 presents standardized performance of the three profile groups relative to the whole sample mean on each of the eight measures that were found to best discriminate the groups. Significant pairwise comparisons from the ANOVA are indicated with braces and asterisks. The three profile groups were discriminated across levels of observation including
behavioral (i.e., PCCR), phonatory (i.e., trochaic and iambic acoustic measures), and articulatory (i.e., verbal and nonverbal measures of displacement and variability).

In summary, characteristics of members of each of the three groups identified by cluster analyses were identified using pairwise post hoc ANOVA comparisons and placement in the statistical space created by discriminant analysis. The High Stability profile group included 21 members whose performance was characterized by low levels of acoustic and kinematic variability across both verbal and nonverbal tasks and high phonemic accuracy. In contrast, the 25 members of the High Variability profile group produced comparatively greater variability across acoustic and kinematic measures, but had similar levels of phonemic accuracy to the High Stability profile group. The 17 members of the Low Phonemics group scored comparatively lower on measures of phonemic accuracy, such as the Percentage of Consonants Correct Revised and other output measures from the PEPPER analysis. While these scores were significantly lower than the scores of the other two profile groups, they were still within the range of children with normal speech acquisition. Overall, significant pairwise comparisons between profile groups existed on 41 of the 72 raw measures. Results from a forward-stepping discriminant analysis suggested eight of the measures could be used to best discriminate the three profile groups. The eight measures included PCCR, five measures of acoustic and kinematic variability, an acoustic measure of lexical stress production, and a measure of jaw maximum displacement during the chewing task.

**DISCUSSION**

Analyses of this large sample of 3-to-5-year-olds’ speech revealed three distinct developmental clusters. This finding alone, that the processes of speech development are non-uniformly distributed within and across children, weakens support for a model of speech
production characterized by gradual, monotonic refinement and stabilization of early-appearing movement patterns and behaviors. Rather, the present findings supported the alternate hypothesis that the many potent influences driving speech development, within and across children, elicit a range of behavioral characteristics, and that the speech production solutions available to young children are manifest in distinct kinematic/acoustic/behavioral profiles.

Beyond the identification and description of emergent clusters among these typically developing children, the present approach identified those measures that most specifically and accurately characterized each group. The initial modeling results revealed that the most distinctive descriptors included a combination of measures related to phonemic accuracy and articulatory/phonatory stability. Forward-stepping descriptive discriminant analysis was used to identify the eight most discriminating measures among the original 72. These eight distinguishing variables are shown in Figure 4. Of these eight, four were related to articulator movement variability: (1) the coefficient of variation of the maximum displacement for the jaw, (2) the spatiotemporal index of the jaw movements for iambic productions, (3) the convergence index of the lower lip movements for trochaic productions (like the STI, but including productions in error), and (4) the cyclic variability of the movement of the jaw during chewing. The fifth discriminant measure was (5) phonemic accuracy (Percentage of Consonants Correct, Revised; PCCR), which enhances the ecological validity of these measures as sensitive to differences in speech output. The sixth and seventh measures were derived from speech acoustics: (6) the acoustic area ratio for production of trochees, and (7) the coefficient of variation of the acoustic area ratio for iambic productions. Finally (8) the maximum displacement of the jaw during chewing distinguished the Low Phonemics group from the other two groups, which was consistent with several other measures of maximum displacement for the
Low Phonemics group (i.e., during word production, jaw and lower lip maximum displacements were significantly smaller in this group than in the High Stability group; upper lip maximum displacements were significantly smaller than those of the High Variability group). It may be that the smaller jaw maximum displacements observed in this cluster may simply reflect morphological differences associated with this group being the youngest overall of the three identified groups. Maximal mandibular displacement has been found to be correlated with age and stature in children (Landtwing 1978), and with facial morphology in adults (Fukui et al. 2002), although the size or orofacial structures has been found to have little relationship with the scale of speech movements (Riely and Smith 2003). It is likely that these kinematic differences are a distinguishing characteristic of individuals in this group.

The three emergent speech profiles may have resulted from each group of children’s differential weighting of widely divergent, yet powerful, influences that are in effect during speech development (e.g., communicative intent, craniofacial morphology, level of language or motor development, phonetic repertoire). These three identified clusters were broadly characterized by their (1) relatively high production stability, (2) relatively high production variability, or (3) relatively low phonemic accuracy. The statistical support and ecological validity of these profiles were strengthened by a modeling approach (Multiple Factor Analysis) and cluster analyses designed to maintain the interpretability of the findings throughout the process. The group characteristics relied directly on the surface measures (e.g., acoustic area ratios) so that each profile could be distinguished by established speech descriptors. The High Stability group was characterized by low values for variability measures across both acoustic and kinematic domains, and by high performance measures of phonemic accuracy (e.g., Percentage of Consonants Correct); the High Variability group was characterized by significantly higher
variability on both acoustic and kinematic variability measures, and with performance measures
of phonemic accuracy that were similarly high to the High Stability group; and the Low
Phonemics group was characterized by low performance on measures of phonemic accuracy.
Each of these clusters represented a distinct behavioral response or profile for the speech
production tasks by each child.

During early speech motor development, children expand and refine their phonetic,
phonemic, and lexical repertoires while experiencing rapid neurological, behavioral, and
morphological growth. Prior studies have indicated, not unexpectedly, that these tumultuous
conditions yield a high degree of heterogeneity in observed speech motor control and resulting
phonetic behaviors (Goffman 1999; Holm et al. 2007; Smith and Goffman 1998; Smith and
Zelaznik 2004; Smith and Gartenberg 1984). The present results revealed that, within these
broad ranges of variation, it is possible to identify distinct subgroups within children, which
suggested the possibility of alternative pathways to mature speech production. These distinct
solutions to the control problems presented by speech development may reflect discrete
individual responses to necessary trade-offs in early speech production, including such factors as
articulatory accuracy versus phonotactic complexity, or word familiarity/mastery/frequency
versus lexical expansion and word novelty. Children might be expected to react to these
competing influences idiosyncratically, although the range of potential solutions appears to be
limited to those exhibited by the three identified clusters. Adoption of specific speech
production styles in response to specific goals (e.g., opting for high movement stability over
phonotactic complexity) may lead to parallel developmental pathways, manifest in the present
experiment as distinct clusters, such that the clusters identified comprise distinct groups of
children who have taken separate pathways to the goal of mature speech production.
Another possible interpretation of these findings of distinct developmental clusters is that each child passes sequentially through each or some of the identified groups, perhaps even repeatedly cycling through these clusters, during speech development. The present observations are consistent with the suggestion that stages of relative stability in developing speech motor control are interleaved with transitional periods of relative instability. It seems likely that speech development entails periods of relatively high or low movement stability, and of relatively high or low phonemic accuracy. The present sample, which revealed children in each of these states, comprised a snapshot of each child’s development, and leaves open the alternative hypothesis that observing their productions at another time or under different conditions would reveal his or her membership in one of the other two groups. It seems unlikely that such shifts could happen very quickly however, for two compelling reasons: (1) task effects were not observed in the present data set, and (2) it seems especially unlikely that a child’s PCCR scores would exhibit the levels of volatility needed for a child to move abruptly from membership in one cluster to another. Thus these clusters appear to have emerged from developmental differences, rather than from task-responsive effects. Frequent longitudinal sampling would be necessary to test this suggestion empirically.

A sequential pathway involving these three profiles might be envisioned by proposing that each cluster represents a distinct period in the developing coordinative infrastructure of the speech production systems (e.g., increasing accuracy in predictive mapping of vocal tract shapes and movements to speech acoustics). Performance by members in the Low Phonemics profile group, for example, may reflect a state in which the child generates relatively stable speech production at the expense of a limited phonetic repertoire. Developmental influences promoting an expansion of this phonetic repertoire would promote lexical expansion and enhanced
intelligibility, but would likely come at the cost of transitioning through a state of relative instability, like that measured in the High Variability profile group. Mastery of this expanded speech inventory would be expected to give rise to a High Stability profile, until developmental influences once again promote transitioning through another period of instability to a higher-yet stage of speech development. Even though age effects did not determine cluster membership, this progression is supported by one age-related post hoc result: children in the Low Phonemics group (mean age: 42.9 months) were significantly younger than children in the High Stability group (mean age: 49.3 months), although the ages of children in the High Variability group (mean age: 45.6 months) were not statistically different from those in either of the other groups.

It is critical to incorporate the likely effects of oral motor development into this scenario, however; differences in motor variability among the three profile groups were present across both verbal and nonverbal tasks, which suggested that these varying motor coordinative stabilities were not specific to the speech production system.

In any case, it cannot be determined from the present results whether membership in a cluster occurs in parallel across children (i.e., each child expresses the speech characteristics of a single cluster throughout development), sequentially within each child (i.e., that each child will manifest the characteristics of each cluster in order through different periods of development), or whether developmental periods elicit the characteristics of a specific cluster.

It is also important to note that, although measures of phonemic accuracy were correlated positively with age, specific tests of this relationship revealed no group effect for age and any of the remaining measures. Similarly, there was no significant relationship observed between sex and any of the reported measures. These findings make it very unlikely that the observed clusters were driven by chronological age or sex effects. By contrast, Smith and Zelaznik (2004)
found significant differences in phrase-level kinematic variability between four- and five-year-olds, as well as between males and females (males were more variable). A potential reason for this difference in findings was the result of the level of analysis: single nonwords were analyzed in the present study compared to phrases in the study by Smith and Zelaznik.

The present design also provided a description of the relationships of observations across speech production domains, generating a broad characterization of the physiologic, behavioral, and acoustic dimensions of speech motor development of young children, and provided a test set for post hoc questions of speech development. For example, these data facilitated an evaluation of the hypothesis that children who exhibited the highest variability in production measures would produce the greatest number of phonemic errors. The CI/STI ratios (Indices 44 – 52 on Table 3; the convergence index (CI) and spatiotemporal index (STI) for each articulator and target production) compare variability of incorrect and correct productions with those of correct productions only, and were used to evaluate this question. No statistically significant finding was obtained for this test, though it exemplifies the wide range of questions that remain to be addressed in this data set.

Broader Implications

The present results have raised several immediate questions, have important implications for the design of future studies of speech production of children at these ages, and also inform our interpretation of past findings. Research on speech development has most often relied on small sample sizes using measures in a single domain (e.g., speech kinematics) derived from a single subsystem (e.g., mandible). These prior investigations have framed our understanding of typical development, disordered speech production, and clinical practice. For example, it is widely understood that young children produce speech that is more variable than that of older
children and adults (cf.: Chermack and Schneiderman 1986; Eguchi and Hirsh 1969; Kent 1976; 
Kent and Forner 1980; Smith 1978; Tingley and Allen 1975). The present results narrowed 
these findings to specific types of variability for these target utterances.

A study of short phrase production by eight four- and eight seven-year-old children by 
Smith and Goffman (1998), for example, reported mean lower lip plus jaw spatiotemporal index 
(STI) values of 24.1 for the four-year-olds, which was significantly lower than the 18.5 STI 
obtained for the seven-year-olds. The present results included comparable measures for the STI 
of the jaw, but during trochaic bisyllables: children in the High Variability group had mean STI 
values of 24.7, children in the High Stability group had mean STI values of 18.6, and those in the 
Low Phonemics group had mean STI values of 20.7. The results of the present High Variability 
group replicate those reported by Smith and Goffman for four-year-olds, though the High 
Stability group exhibited STI values that were significantly lower than the High Variability 
group. Despite the difference in context (i.e., nonword bisyllables in the present study vs. 
phrases in the Smith and Goffman study) the present results may provide a refined context for 
interpretation of these findings, demonstrating that the ubiquitous variability of speech 
production can be understood in a more detailed way, with some children adopting a speech 
production style that reduces this variability.

The present investigation has also demonstrated the feasibility and necessity of 
experimental designs that use samples that are sufficiently large to differentiate real, but subtle, 
differences in speech production. Moreover, multivariate analysis and clustering techniques 
were shown to provide interpretable results that will provide a more detailed understanding of 
speech production, developmental pathways, and disorders. Nevertheless, the current dataset, 
though large by current research practices in developmental speech physiology, is still too small
to provide a high resolution description of the various developmental trajectories that result in mature speech production. For example, the full distribution and shape of the variations in developmental profiles could only be understood with a very large, developmental sample that includes frequent longitudinal sampling.

Limitations

In addition to the limits posed by the cross-sectional design of the present investigation, these findings were also affected by empty cells in the data set. Lost data sometimes resulted from the failure of the system to track the optical targets consistently; automated identification of the flat reflectors was easily disrupted by extraneous movements. It is also worth noting that movement of the mandible was transduced and projected only in the coronal plane, which neglected other degrees of freedom, particularly rotation around the horizontal axis, and anterior-posterior translation (Westbury 1988). Multi-camera movement tracking systems are generally sufficient to overcome these limitations. An additional source of error is that associated with the use of a fleshpoint (e.g., the midline chin placement of the jaw marker) as an indicator of skeletal movement (Green et al. 2007); elimination of this error is challenging in young children, of course.

Conclusions

The primary aim of this study was to evaluate the speech production profiles of young children using a large battery of measures from physiologic, acoustic, and behavioral domains. Of the 72 measures obtained, eight were found to be sufficient to distinguish three clusters of children. These findings suggested that the high variability that is typical of speech production is not uniformly observed across children; different groups of children exhibit significantly different degrees of kinematic, acoustic, and behavioral variability. This finding supported the
suggestion that developing speakers employ distinct developmental profiles in producing speech. The present data set was not designed to determine whether these differences were exclusive across children (i.e., with different groups of children developing along parallel paths to achieve mature speech production) or within children (i.e., with each child progressing along a pathway with separable, distinct periods). It remains for future investigations to resolve this question by sampling these behaviors longitudinally with relatively frequent observations.

Prior studies have established that variability characterizes measures of articulatory kinematics and acoustics at this period of development. The multivariate statistical techniques employed in this study enhanced the resolution of our understanding of speech motor development at these ages. High variability was observed in one of the identified groups of children, although the phonemic accuracy of the children in this group was not significantly different from that observed in the group of children with the lowest level of articulatory and acoustic variability. A third group was identified whose articulatory and acoustic variability was no different than the other two groups, but whose phonemic accuracy was significantly lower.

These results have also further elaborated the interactions of speech production systems across levels of analysis, providing a better understanding of how observations at one level (i.e., labiomandibular movement) do not reliably predict observations at another level of analysis (i.e., phonemic productions). In the present sample, children in the High Variability group produced speech sounds that were recognized as acceptable tokens of the target sounds, yet were quantitatively different from those produced by children in the High Stability group.
Because of the large number of potentially interrelated measures in the present investigation, it was necessary to first transform the entire data set to reduce its dimensionality while maintaining as much of the variation as possible (Ding & He, 2004; Gan, Ma, & Wu, 2007). Typically, individual variables (vectors) are normalized prior to analysis, using vector normalization, such as Z-score. The problem with such an approach in a very large dataset with measures of diverse scope and variability is that blocks of measures with the greatest variance will dominate the outcome of the analysis. MFA mitigates this problem by normalizing all elements within the measure, thereby balancing the between-measure variance contributed to the analysis. Each group of similar variables, represented as a matrix, is normalized by dividing all elements by the first singular value of the matrix (i.e., the matrix equivalent of a standard deviation). This normalizing value for the measure matrix is calculated using a PCA: the square root of the first eigenvalue is the first singular value for the matrix. Once each table of measures is normalized, the matrices are then concatenated and the resulting matrix is submitted to a grand PCA. The resulting principal components, or factors, linearly integrate the variables that contribute similarly to the variance of the data set (See Abdi & Valentin, 2007) for a detailed presentation on this technique).
Acknowledgments

Funding for this project was provided by the National Institute on Deafness and Other Communication Disorders (NIDCD R01 DC00822 to Dr. Moore) and the American Speech Language Hearing Foundation (New Century Scholars Doctoral Scholarship to Dr. Vick). We are especially grateful to all of the participants and their families who graciously volunteered their time for this project. In addition, we would like to acknowledge the work of the individuals whose work was critical for participant recruitment, data acquisition, data extraction, and computer programming: Tammy Nash, Jill Brady, Dayna Pitcairn, Denise Balason, Stacey Pavelko, Mitzi Kweder, Sharon Gretz, Kevin Reilly, Roger Steeve, Kathryn Connaghan, Yumi Sumida, Alyssa Mosely, Rossella Belli, Ettore Cavallaro, Jeanette Wu, Jenny Morus, Kelsey Moore, Mary Reeves, Nicholas Moon, Dennis Tang, Adam Politis, Andrea Kettler, Laura Worthen, Dara Cohen, and Greg Lee. We are also grateful to the associate editor, Dr. Catherine Carr, and three anonymous reviewers whose comments have greatly improved the focus of this paper.

Dr. Vick is currently affiliated with Case Western Reserve University Department of Psychological Sciences. Dr. Campbell is currently affiliated with the Department of Behavior and Brain Sciences at The University of Texas at Dallas in Richardson, Texas. Dr. Moore is currently affiliated with the Department of Veterans Affairs in Washington, DC. Dr. Venkatesh is currently affiliated with Sweekaar Rehabilitation Institute for Handicapped in Secunderabad, Andhra Pradesh, India. Dr. Rusiewicz is currently affiliated with the Department of Speech-Language Pathology at Duquesne University in Pittsburgh, Pennsylvania.
REFERENCE LIST


Figure Legends

Figure 1: Schematic of data reduction stream.

Figure 2: Example of kinematic parsing using velocity zero crossings. The top graph plots the acoustic waveform for a child’s production of “Baba.” The bottom graph plots the vertical displacement of the mandibular marker. Vertical lines represent the velocity zero crossings. Onset and offset decisions are marked.

Figure 3: Clusters in two-dimensional discriminant function space with each clusters’ centroid as stars. Cases for each cluster are encircled by 95% confidence interval ellipses. Axes are bordered by explanations for each of the two factors.

Figure 4: Standard scores of each group on each of the eight variables found to maximally distinguish the groups. Significant pairwise comparisons are indicated by brackets (*: \( p < 0.017 \); **: \( p < 0.001 \)).
Table 1. *Categorical demographic variables*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Name</th>
<th>Category</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male</td>
<td>1</td>
<td>27</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>2</td>
<td>36</td>
<td>57.1</td>
</tr>
<tr>
<td>Family</td>
<td>Negative</td>
<td>0</td>
<td>58</td>
<td>92.1</td>
</tr>
<tr>
<td>History</td>
<td>Positive</td>
<td>1</td>
<td>5</td>
<td>7.9</td>
</tr>
<tr>
<td>Race</td>
<td>Caucasian</td>
<td>1</td>
<td>51</td>
<td>81.0</td>
</tr>
<tr>
<td></td>
<td>African-American</td>
<td>2</td>
<td>9</td>
<td>14.2</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>3</td>
<td>3</td>
<td>4.8</td>
</tr>
</tbody>
</table>
Table 2. Perceptual analysis of lexical stress and nonword repetition tasks. Includes number attempted and judged accuracy of phonemics and stress. “Overall” refers to the productions that were able to be used for all measures in the statistical model, which required correct phonemics and stress, as well as viable kinematics (i.e., without movement artifact).

<table>
<thead>
<tr>
<th>Type</th>
<th>Word</th>
<th>Attempted N</th>
<th>Correct N</th>
<th>%</th>
<th>Stress Correct N</th>
<th>%</th>
<th>Overall Correct N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trochee</td>
<td>baba</td>
<td>325</td>
<td>284</td>
<td>87.4</td>
<td>321</td>
<td>98.9</td>
<td>233</td>
<td>71.7</td>
</tr>
<tr>
<td></td>
<td>mama</td>
<td>308</td>
<td>285</td>
<td>92.5</td>
<td>292</td>
<td>94.7</td>
<td>237</td>
<td>76.8</td>
</tr>
<tr>
<td></td>
<td>papa</td>
<td>304</td>
<td>286</td>
<td>94.1</td>
<td>284</td>
<td>93.4</td>
<td>237</td>
<td>78.1</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>937</td>
<td>855</td>
<td>91.2</td>
<td>897</td>
<td>95.7</td>
<td>707</td>
<td>75.4</td>
</tr>
<tr>
<td>Iamb</td>
<td>baba</td>
<td>307</td>
<td>259</td>
<td>84.4</td>
<td>303</td>
<td>98.8</td>
<td>209</td>
<td>68.1</td>
</tr>
<tr>
<td></td>
<td>mama</td>
<td>295</td>
<td>274</td>
<td>92.9</td>
<td>232</td>
<td>78.5</td>
<td>183</td>
<td>62.0</td>
</tr>
<tr>
<td></td>
<td>papa</td>
<td>301</td>
<td>267</td>
<td>88.7</td>
<td>266</td>
<td>88.4</td>
<td>201</td>
<td>66.9</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>903</td>
<td>801</td>
<td>88.7</td>
<td>801</td>
<td>88.7</td>
<td>593</td>
<td>65.7</td>
</tr>
<tr>
<td>Nonword</td>
<td>bada</td>
<td>342</td>
<td>309</td>
<td>90.4</td>
<td>n/a</td>
<td>n/a</td>
<td>291</td>
<td>85.0</td>
</tr>
<tr>
<td></td>
<td>bama</td>
<td>357</td>
<td>306</td>
<td>85.7</td>
<td>n/a</td>
<td>n/a</td>
<td>304</td>
<td>85.1</td>
</tr>
<tr>
<td></td>
<td>bamana</td>
<td>333</td>
<td>210</td>
<td>63.1</td>
<td>n/a</td>
<td>n/a</td>
<td>188</td>
<td>57.0</td>
</tr>
<tr>
<td></td>
<td>manaba</td>
<td>330</td>
<td>173</td>
<td>52.4</td>
<td>n/a</td>
<td>n/a</td>
<td>160</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>1362</td>
<td>998</td>
<td>73.3</td>
<td>n/a</td>
<td>n/a</td>
<td>943</td>
<td>69.2</td>
</tr>
<tr>
<td>GRAND</td>
<td></td>
<td>3202</td>
<td>2654</td>
<td>82.9</td>
<td>n/a</td>
<td>n/a</td>
<td>2243</td>
<td>70.0</td>
</tr>
</tbody>
</table>
Table 3. 69 continuous variables basic statistics for the overall sample (Grand Mean), as well as for the three clusters. F statistics and p-values are reported for significant pairwise comparisons. Pairs with significant differences are reported.

<table>
<thead>
<tr>
<th>Index</th>
<th>Measure</th>
<th>Min</th>
<th>Max</th>
<th>CV</th>
<th>Grand Mean</th>
<th>Clst 1 Mean</th>
<th>Clst 2 Mean</th>
<th>Clst 3 Mean</th>
<th>F</th>
<th>P</th>
<th>Sig. pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Proportion attempted</td>
<td>0.16</td>
<td>1.00</td>
<td>0.17</td>
<td>0.94</td>
<td>0.96</td>
<td>0.88</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Proportion with phonetics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct</td>
<td>0.10</td>
<td>1.00</td>
<td>0.23</td>
<td>0.80</td>
<td>0.87</td>
<td>0.74</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Proportion with stress correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>0.00</td>
<td>1.00</td>
<td>0.28</td>
<td>0.82</td>
<td>0.84</td>
<td>0.77</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Proportion with stress correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iambs</td>
<td>0.00</td>
<td>1.00</td>
<td>0.35</td>
<td>0.72</td>
<td>0.87</td>
<td>0.60</td>
<td>0.73</td>
<td>8.03</td>
<td>0.001</td>
<td>1:2</td>
</tr>
<tr>
<td>5</td>
<td>Proportion with no phonetic errors and good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.10</td>
<td>0.98</td>
<td>0.25</td>
<td>0.70</td>
<td>0.80</td>
<td>0.63</td>
<td>0.68</td>
<td>6.33</td>
<td>0.003</td>
<td>1:2</td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>P</td>
<td>pairs</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>kinematics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Mean acoustic area ratio–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>0.60</td>
<td>4.72</td>
<td>0.46</td>
<td>1.89</td>
<td>1.48</td>
<td>2.22</td>
<td>1.79</td>
<td>5.12</td>
<td>0.009</td>
<td>1:2</td>
</tr>
<tr>
<td>7</td>
<td>Mean acoustic area ratio–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iambs</td>
<td>0.22</td>
<td>1.31</td>
<td>0.38</td>
<td>0.62</td>
<td>0.51</td>
<td>0.71</td>
<td>0.63</td>
<td>4.08</td>
<td>0.022</td>
<td>1:2</td>
</tr>
<tr>
<td>8</td>
<td>CV of acoustic area ratio–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>17.26</td>
<td>112.51</td>
<td>0.43</td>
<td>43.73</td>
<td>34.89</td>
<td>47.41</td>
<td>45.20</td>
<td>3.61</td>
<td>0.033</td>
<td>1:2</td>
</tr>
<tr>
<td>9</td>
<td>CV of acoustic area ratio–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iambs</td>
<td>11.58</td>
<td>94.56</td>
<td>0.35</td>
<td>40.81</td>
<td>33.43</td>
<td>48.92</td>
<td>37.70</td>
<td>8.63</td>
<td>&lt;0.0001</td>
<td>2:3</td>
</tr>
<tr>
<td>10</td>
<td>2-syll. word duration</td>
<td>0.58</td>
<td>1.08</td>
<td>0.13</td>
<td>0.77</td>
<td>0.79</td>
<td>0.75</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3-syll. word duration</td>
<td>0.94</td>
<td>1.80</td>
<td>0.12</td>
<td>1.45</td>
<td>1.47</td>
<td>1.47</td>
<td>1.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Grand</td>
<td>Clist 1</td>
<td>Clist 2</td>
<td>Clist 3</td>
<td>F</td>
<td>P</td>
<td>pairs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>CV of 2-syll. word duration</td>
<td>6.02</td>
<td>21.91</td>
<td>0.30</td>
<td>11.78</td>
<td>11.52</td>
<td>12.32</td>
<td>11.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>CV of 3-syll word duration</td>
<td>2.48</td>
<td>27.73</td>
<td>0.52</td>
<td>10.32</td>
<td>10.23</td>
<td>10.80</td>
<td>9.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Mean maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– UL</td>
<td>0.09</td>
<td>0.51</td>
<td>0.41</td>
<td>0.22</td>
<td>0.22</td>
<td>0.25</td>
<td>0.18</td>
<td>3.49</td>
<td>0.037</td>
<td>2:3</td>
</tr>
<tr>
<td>15</td>
<td>Mean maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– LL</td>
<td>0.09</td>
<td>0.45</td>
<td>0.38</td>
<td>0.24</td>
<td>0.27</td>
<td>0.25</td>
<td>0.19</td>
<td>4.56</td>
<td>0.014</td>
<td>1:3</td>
</tr>
<tr>
<td>16</td>
<td>Mean maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Jaw</td>
<td>0.28</td>
<td>1.22</td>
<td>0.29</td>
<td>0.72</td>
<td>0.80</td>
<td>0.73</td>
<td>0.59</td>
<td>5.89</td>
<td>0.005</td>
<td>1:3</td>
</tr>
<tr>
<td>17</td>
<td>CV of maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– UL</td>
<td>24.81</td>
<td>90.44</td>
<td>0.26</td>
<td>39.26</td>
<td>34.40</td>
<td>43.07</td>
<td>39.46</td>
<td>4.65</td>
<td>0.013</td>
<td>1:2</td>
</tr>
<tr>
<td>18</td>
<td>CV of maximum displacement</td>
<td>18.26</td>
<td>90.53</td>
<td>0.35</td>
<td>38.00</td>
<td>30.28</td>
<td>46.11</td>
<td>35.61</td>
<td>10.62</td>
<td>&lt;0.0001</td>
<td>1:2</td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>P</td>
<td>Sig.</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>displacement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– LL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>CV of maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Jaw</td>
<td>15.54</td>
<td>55.02</td>
<td>0.24</td>
<td>31.21</td>
<td>26.52</td>
<td>34.53</td>
<td>31.62</td>
<td>8.17</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>UL STI –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>17.39</td>
<td>41.35</td>
<td>0.21</td>
<td>27.67</td>
<td>25.26</td>
<td>30.44</td>
<td>26.47</td>
<td>5.34</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>UL STI - lambs</td>
<td>11.89</td>
<td>40.79</td>
<td>0.22</td>
<td>26.69</td>
<td>24.82</td>
<td>28.31</td>
<td>26.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>UL STI – 3-syll.</td>
<td>11.33</td>
<td>40.79</td>
<td>0.31</td>
<td>24.33</td>
<td>23.19</td>
<td>26.68</td>
<td>21.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>LL STI –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>11.21</td>
<td>45.44</td>
<td>0.22</td>
<td>30.88</td>
<td>26.34</td>
<td>34.30</td>
<td>31.01</td>
<td>10.25</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>LL STI - lambs</td>
<td>15.34</td>
<td>47.08</td>
<td>0.22</td>
<td>29.81</td>
<td>25.96</td>
<td>32.16</td>
<td>30.68</td>
<td>6.42</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>LL STI – 3-syll.</td>
<td>11.50</td>
<td>48.09</td>
<td>0.26</td>
<td>30.22</td>
<td>24.58</td>
<td>34.47</td>
<td>30.26</td>
<td>12.74</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Jaw STI –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>11.25</td>
<td>33.22</td>
<td>0.21</td>
<td>21.56</td>
<td>18.59</td>
<td>24.71</td>
<td>20.75</td>
<td>14.74</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Jaw STI - lambs</td>
<td>9.80</td>
<td>33.48</td>
<td>0.26</td>
<td>18.33</td>
<td>15.93</td>
<td>20.22</td>
<td>18.63</td>
<td>5.38</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Grand Mean</td>
<td>Clst 1 Mean</td>
<td>Clst 2 Mean</td>
<td>Clst 3 Mean</td>
<td>F</td>
<td>P</td>
<td>Sig. pairs</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------</td>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td>28</td>
<td>Jaw STI – 3-syll.</td>
<td>10.63</td>
<td>44.60</td>
<td>0.24</td>
<td>27.74</td>
<td>24.71</td>
<td>25.40</td>
<td>25.40</td>
<td>8.62</td>
<td>0.001</td>
<td>1:2</td>
</tr>
<tr>
<td>29</td>
<td>Lip aperture STI – Trochees</td>
<td>11.41</td>
<td>32.59</td>
<td>0.21</td>
<td>19.98</td>
<td>17.59</td>
<td>18.92</td>
<td>11.47</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>2:3</td>
</tr>
<tr>
<td>30</td>
<td>Lip aperture STI – lambs</td>
<td>8.99</td>
<td>32.25</td>
<td>0.27</td>
<td>17.52</td>
<td>15.23</td>
<td>19.21</td>
<td>17.71</td>
<td>4.44</td>
<td>0.016</td>
<td>1:2</td>
</tr>
<tr>
<td>31</td>
<td>Lip aperture STI – 3-syll.</td>
<td>8.14</td>
<td>38.12</td>
<td>0.25</td>
<td>23.35</td>
<td>21.08</td>
<td>21.30</td>
<td>6.85</td>
<td>6.85</td>
<td>0.002</td>
<td>2:3</td>
</tr>
<tr>
<td>32</td>
<td>Jaw/LL aperture STI –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trochees</td>
<td>16.65</td>
<td>45.44</td>
<td>0.20</td>
<td>31.41</td>
<td>26.34</td>
<td>31.69</td>
<td>17.83</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>1:3</td>
</tr>
<tr>
<td>33</td>
<td>Jaw/LL aperture STI –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lambs</td>
<td>15.34</td>
<td>47.08</td>
<td>0.22</td>
<td>29.81</td>
<td>25.96</td>
<td>30.67</td>
<td>6.41</td>
<td>6.41</td>
<td>0.003</td>
<td>1:2</td>
</tr>
<tr>
<td>34</td>
<td>Jaw/LL aperture STI –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-syll.</td>
<td>11.50</td>
<td>48.09</td>
<td>0.26</td>
<td>30.15</td>
<td>24.57</td>
<td>34.33</td>
<td>30.21</td>
<td>12.39</td>
<td>&lt;0.0001</td>
<td>1:3</td>
</tr>
<tr>
<td>35</td>
<td>UL CI – Trochees</td>
<td>15.14</td>
<td>41.62</td>
<td>0.21</td>
<td>28.20</td>
<td>25.32</td>
<td>30.75</td>
<td>27.80</td>
<td>5.25</td>
<td>0.008</td>
<td>1:2</td>
</tr>
<tr>
<td>36</td>
<td>UL CI – lambs</td>
<td>11.89</td>
<td>40.79</td>
<td>0.21</td>
<td>27.26</td>
<td>24.61</td>
<td>29.08</td>
<td>27.46</td>
<td>3.95</td>
<td>0.024</td>
<td>1:2</td>
</tr>
<tr>
<td>37</td>
<td>UL CI – 3-syll.</td>
<td>11.44</td>
<td>48.88</td>
<td>0.27</td>
<td>30.75</td>
<td>27.65</td>
<td>33.08</td>
<td>30.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Clst 1 Mean</td>
<td>Clst 2 Mean</td>
<td>Clst 3 Mean</td>
<td>F</td>
<td>P</td>
<td>Sig. pairs</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>38</td>
<td>LL CI – Trochees</td>
<td>16.65</td>
<td>46.16</td>
<td>0.19</td>
<td>31.51</td>
<td>26.39</td>
<td>35.16</td>
<td>32.19</td>
<td>19.47</td>
<td>&lt;0.0001</td>
<td>1:2</td>
</tr>
<tr>
<td>39</td>
<td>LL CI – Iambs</td>
<td>15.34</td>
<td>47.08</td>
<td>0.20</td>
<td>29.99</td>
<td>25.91</td>
<td>32.53</td>
<td>31.08</td>
<td>8.51</td>
<td>0.001</td>
<td>1:2</td>
</tr>
<tr>
<td>40</td>
<td>LL CI – 3-syll.</td>
<td>11.91</td>
<td>48.09</td>
<td>0.23</td>
<td>32.33</td>
<td>27.66</td>
<td>34.61</td>
<td>34.24</td>
<td>7.17</td>
<td>0.002</td>
<td>1:2</td>
</tr>
<tr>
<td>41</td>
<td>Jaw CI – Trochees</td>
<td>11.25</td>
<td>33.19</td>
<td>0.20</td>
<td>22.03</td>
<td>18.72</td>
<td>24.90</td>
<td>21.95</td>
<td>15.94</td>
<td>&lt;0.0001</td>
<td>2:3</td>
</tr>
<tr>
<td>42</td>
<td>Jaw CI – Iambs</td>
<td>9.80</td>
<td>33.48</td>
<td>0.25</td>
<td>18.70</td>
<td>15.83</td>
<td>20.01</td>
<td>19.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Jaw CI – 3-syll.</td>
<td>12.04</td>
<td>44.60</td>
<td>0.22</td>
<td>29.29</td>
<td>26.00</td>
<td>32.33</td>
<td>28.59</td>
<td>6.22</td>
<td>0.004</td>
<td>1:2</td>
</tr>
<tr>
<td>44</td>
<td>UL CI / STI ratio – Trochees</td>
<td>0.87</td>
<td>1.50</td>
<td>0.08</td>
<td>1.02</td>
<td>1.00</td>
<td>1.01</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>UL CI / STI ratio – Iambs</td>
<td>0.93</td>
<td>1.48</td>
<td>0.09</td>
<td>1.03</td>
<td>0.99</td>
<td>1.04</td>
<td>1.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>UL CI / STI ratio – 3-syll.</td>
<td>0.46</td>
<td>3.00</td>
<td>0.34</td>
<td>1.07</td>
<td>1.14</td>
<td>0.98</td>
<td>1.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>P</td>
<td>pairs</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>----</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>47</td>
<td>LL CI / STI ratio – Trochees</td>
<td>0.86</td>
<td>3.00</td>
<td>0.25</td>
<td>1.05</td>
<td>1.00</td>
<td>1.08</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>LL CI / STI ratio – Iambs</td>
<td>0.86</td>
<td>1.55</td>
<td>0.08</td>
<td>1.01</td>
<td>1.00</td>
<td>1.02</td>
<td>1.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>LL CI Index / STI ratio – 3-syll.</td>
<td>0.51</td>
<td>2.18</td>
<td>0.49</td>
<td>1.11</td>
<td>1.11</td>
<td>1.15</td>
<td>1.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Jaw CI / STI ratio – Trochees</td>
<td>0.93</td>
<td>1.56</td>
<td>0.08</td>
<td>1.03</td>
<td>1.01</td>
<td>1.01</td>
<td>1.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Jaw CI / STI ratio – Iambs</td>
<td>0.80</td>
<td>1.64</td>
<td>0.13</td>
<td>1.03</td>
<td>1.00</td>
<td>0.99</td>
<td>1.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Jaw CI / STI ratio – 3-syll.</td>
<td>0.69</td>
<td>1.60</td>
<td>0.17</td>
<td>1.08</td>
<td>1.07</td>
<td>1.04</td>
<td>1.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>Age in months</td>
<td>36.00</td>
<td>57.00</td>
<td>0.13</td>
<td>46.17</td>
<td>49.29</td>
<td>45.64</td>
<td>42.94</td>
<td>5.82</td>
<td>0.005</td>
<td>1:3</td>
</tr>
<tr>
<td>54</td>
<td>Mother’s age in years</td>
<td>21.00</td>
<td>45.00</td>
<td>0.17</td>
<td>33.63</td>
<td>32.65</td>
<td>32.29</td>
<td>36.61</td>
<td>3.65</td>
<td>0.032</td>
<td>1:2</td>
</tr>
<tr>
<td>55</td>
<td>Mother’s years of education</td>
<td>10.00</td>
<td>22.00</td>
<td>0.15</td>
<td>16.45</td>
<td>16.16</td>
<td>16.08</td>
<td>17.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>P</td>
<td>Sig.</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>56</td>
<td>Jaw maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td>0.31</td>
<td>1.08</td>
<td>1.25</td>
<td>1.03</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– vertical jaw oscillations</td>
<td></td>
<td></td>
<td></td>
<td>2.45</td>
<td>0.41</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>Jaw maximum displacement</td>
<td>0.21</td>
<td>1.41</td>
<td>0.25</td>
<td>1.00</td>
<td>1.07</td>
<td>1.06</td>
<td>0.82</td>
<td>6.92</td>
<td>0.002</td>
<td>1:2</td>
</tr>
<tr>
<td></td>
<td>– chewing</td>
<td></td>
<td></td>
<td></td>
<td>1.25</td>
<td>0.55</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>CV of jaw maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td>0.07</td>
<td>1.16</td>
<td>0.64</td>
<td>0.38</td>
<td>0.38</td>
<td>0.31</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>– oscillations</td>
<td></td>
<td></td>
<td></td>
<td>1.16</td>
<td>0.64</td>
<td>0.38</td>
<td>0.38</td>
<td>0.31</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>CV of jaw maximum displacement</td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
<td>0.63</td>
<td>0.32</td>
<td>0.31</td>
<td>0.31</td>
<td>0.29</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>– chewing</td>
<td></td>
<td></td>
<td></td>
<td>0.63</td>
<td>0.32</td>
<td>0.31</td>
<td>0.31</td>
<td>0.29</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>cSTI – oscillation</td>
<td>5.83</td>
<td>24.43</td>
<td>0.26</td>
<td>15.19</td>
<td>13.88</td>
<td>15.15</td>
<td>16.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>cSTI – chewing</td>
<td>13.89</td>
<td>37.66</td>
<td>0.22</td>
<td>26.52</td>
<td>22.74</td>
<td>28.96</td>
<td>28.34</td>
<td>9.74</td>
<td>0.001</td>
<td>1:2</td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>P</td>
<td>Sig.</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>62</td>
<td>Percentage of consonants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct</td>
<td>71.40</td>
<td>98.25</td>
<td>0.07</td>
<td>88.78</td>
<td>91.69</td>
<td>90.35</td>
<td>83.09</td>
<td>13.36</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>Percentage of consonants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct, revised</td>
<td>80.53</td>
<td>99.07</td>
<td>0.05</td>
<td>92.37</td>
<td>94.98</td>
<td>93.23</td>
<td>88.02</td>
<td>19.09</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>Percentage of vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct</td>
<td>88.54</td>
<td>100.00</td>
<td>0.03</td>
<td>96.75</td>
<td>97.26</td>
<td>96.96</td>
<td>95.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>Percentage of vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct, revised</td>
<td>91.72</td>
<td>100.00</td>
<td>0.02</td>
<td>98.08</td>
<td>98.33</td>
<td>98.29</td>
<td>97.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>Percentage of phonemes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>80.76</td>
<td>98.73</td>
<td>0.05</td>
<td>91.92</td>
<td>93.90</td>
<td>92.96</td>
<td>88.11</td>
<td>12.22</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>Measure</td>
<td>Min</td>
<td>Max</td>
<td>CV</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>P</td>
<td>pairs</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>67</td>
<td>Percentage of phonemes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct, revised</td>
<td>86.78</td>
<td>99.24</td>
<td>0.03</td>
<td>94.63</td>
<td>96.30</td>
<td>95.23</td>
<td>91.76</td>
<td>17.58</td>
<td>&lt;0.0001</td>
<td>1:3</td>
</tr>
<tr>
<td>68</td>
<td>Intelligibility index</td>
<td>87.65</td>
<td>100.00</td>
<td>0.03</td>
<td>97.54</td>
<td>98.09</td>
<td>97.71</td>
<td>96.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>Average words per utterance</td>
<td>1.48</td>
<td>6.42</td>
<td>0.31</td>
<td>3.43</td>
<td>3.51</td>
<td>3.79</td>
<td>2.86</td>
<td>4.29</td>
<td>0.018</td>
<td>2:3</td>
</tr>
</tbody>
</table>
Table 4. Correlations between the eight variables identified by discriminant analysis and Factor 1 (F1) and Factor 2 (F2) from the discriminant analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>F1 $\rho$</th>
<th>F2 $\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trochees - Acoustic Area Ratio</td>
<td>0.404</td>
<td>0.120</td>
</tr>
<tr>
<td>Chewing - Jaw Max. Displacement</td>
<td>-0.087</td>
<td>0.500</td>
</tr>
<tr>
<td>Iambs - Acoustic Area Ratio CV</td>
<td>0.475</td>
<td>0.226</td>
</tr>
<tr>
<td>PCCR</td>
<td>-0.279</td>
<td>0.669</td>
</tr>
<tr>
<td>Iambs - Jaw STI</td>
<td>0.428</td>
<td>0.026</td>
</tr>
<tr>
<td>Jaw Max. Displacement CV</td>
<td>0.508</td>
<td>0.026</td>
</tr>
<tr>
<td>Chewing - cSTI</td>
<td>0.532</td>
<td>-0.121</td>
</tr>
<tr>
<td>Trochees - Lower Lip CI</td>
<td>0.689</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Strong correlations are boldfaced.
Jaw vertical displacement with velocity zero crossings

Time (seconds)

Voltage (volts cm)

Jaw vertical displacement with velocity zero crossings

Time (seconds)

Onset

Offset
Phonemic Accuracy

Cluster 1
Cluster 2
Cluster 3
Centroids

Variability

Factor 1 (64.8%)
Factor 2 (44.3%)

Phonemic Accuracy