PHONOLOGICAL AND PHONETIC ANALYSIS OF CLEFT PALATE SPEECH

Bernhardt, B. & Doan, A., University of British Columbia, Canada
Stoel-Gammon, C., University of Washington, USA

ABSTRACT

Speakers with repaired cleft palates may continue to have atypical speech development. Speech pathologists are often asked to comment on the need for, and effectiveness of, followup medical or prosthetic interventions. Few clinicians have access to technology for such determinations. We present articulatory-phonetic and nonlinear phonological analytic procedures which can assist in these determinations, using data from a child on a speech bulb reduction program.

INTRODUCTION

Although surgical repairs have become very effective for cleft palates in recent years, over 30% of persons with repaired cleft palate persist in having atypical speech development. Developmental substitutions (e.g. use of glides for liquids), compensatory substitutions (e.g. palatals, glottals, pharyngeals, or nasal snorts), and/or imbalances between oral and nasal resonance may continue to reduce speech intelligibility. A number of perpetuating factors may result in this persistent speech disorder. Surgeries may have been only moderately successful in achieving palate closure and lengthening. Chronic otitis media may have affected a child's perceptual bases for speech production. General developmental delay, ill health, or social interaction difficulties may have had a general negative impact on communication development. Even if none of these factors exist, habituation to the original characteristics of the oral mechanism (both structurally and functionally) can result in adherence to the original (deviant) phonological and phonetic system.

As part of the cleft palate team, speech-language pathologists, are not only responsible for direct speech intervention, but also are asked to make judgments as to the need for further surgical or prosthetic intervention, or the effectiveness of such procedures. Many useful technologies are currently available to assist the clinician in making such judgments. The velar port functioning can be assessed with tools such as multiview videofluoroscopy, nasoendoscopy, or nasometry. The speech signal can be evaluated through acoustic analysis. Electropalatography can inform about the placement of the tongue with respect to the palate. In selected hospital centers, such technologies have become useful adjuncts to perceptual judgments and analyses. However, many clinicians not working in (well-funded) hospital settings do not have such technologies available. They need to rely on phonetic transcription, general intelligibility assessment, and phonetic and phonological analyses to determine needs and effectiveness. As Howard (1993) [1] shows, detailed phonetic transcription and phonological analysis can lead to the judgment that a person with a cleft palate may have a well-developed phonological system with particular deficits in the articulatory realization of the phonological contrasts.

Inherent in this dichotomous interpretation of speech production is the assumption that intent and phonological representation are separate from the actual phonetic implementation of a word. In terms of phonology, we cannot confirm intent, nor determine the nature of underlying representation. All we can do is make inferences about a person's phonological system based on observation of *contrasts* and/or systematic sound *patterns* or *changes* in *phonetic* output. The more coherent and psychologically real our phonological theory, the more reliable and valid the inferences we make. In this paper we utilize aspects of current nonlinear (multilinear) phonological analysis. This theoretical framework allows detailed determination of patterns at various levels of representation: from the word, foot and syllable levels, to subsegmental featural levels (see below for elaboration). We also assume that phonological constraints operate on output, and that phonological constraints are often phonetically grounded (or motivated), following Archangeli & Pulleyblank (1994) [2]. Thus, the interrelationships between phonetics and phonology are assumed to be very close.

Phonological interpretation depends crucially on reliable and detailed phonetic transcription. The end-product of speech production processing is the articulatory-phonetic output. The more narrowly a listener is able to transcribe that output reliably, the more valid the data. Transcription skill aside, phonetic data tends to be generally "noisy," with even random variability with respect to exact location and timing of discrete phonetic units. A speaker with a repaired cleft palate may have a tendency to variability, as she or he accommodates to both developmental and induced mechanism changes, in the context of possible fluctuating hearing acuity or other perpetuating/predisposing factors. There may or may not be an attempt to reduce variability through more or less rigid adherence to the patterns of the early-established phonological and phonetic systems. Quantification of degree and type of variability is thus a relevant aspect of analysis.

The final determination of relative intactness of phonological and phonetic systems requires a balance between articulatory-phonetic and phonological analyses. This can be approached by attributing sufficient emphasis to phonetic detail and variability, while abstracting away from detail to infer general patterns. In this paper we will outline aspects of nonlinear phonological frameworks that lead to a detailed analysis of phonological systems, and utilize phonetic data in a variety of speech sample conditions to observe variability and consistency. The child whose analysis we use as an example had a speech bulb, the effectiveness of which was being evaluated one month after she started to wear it. We will utilize the nonlinear and phonetic analyses to outline similarities and differences with the speech bulb in and out, and between single words and connected speech.

SUBJECT

Tia (a pseudonym) was 5;11 at the time of this evaluation. She was born with a bilateral cleft lip and palate. Lip repair was done at 4 months of age, and initial palatal repair at 22 months of age. At the time of the palatoplasty, she also had a bilateral myringotomy to reduce otitis media. At 4;5 she had an orticochea pharyngoplasty to assist closure of the velopharyngeal port. (Orticochea pharyngoplasty is a type of pharyngoplasty in which a sphincter is created using the lateral and posterior pharyngeal walls.)

She and her family have received speech and language counselling or intervention services since she was an infant. Until about 3;6, she had recurrent otitis media with a fluctuating mild conductive hearing loss, a moderate delay in language comprehension, and a severe delay in language production. At the time of assessment, hearing status and language skills were within normal limits.

However, speech was and continues to be moderately unintelligible.

Her speech is characterized by hypernasality, nasal air emission and turbulence, and a moderate degree of use of compensatory and developmental phonological/phonetic patterns. Most frequent compensatory substitutions are glottals and pharyngeals, with some use of palatals for lingual consonants, nasals for oral stops, and nasal snorts. The glottal and pharyngeal substitutions are sometimes doubly articulated with oral place consonants. The most prevalent developmental pattern is the use of glides for liquids. Other developmental patterns include some word-initial and word-final omissions (for both singletons and clusters) assimilations, and occasional use of stops for fricatives. Generally, she matches the adult target more often word initially than in other word positions.

A nasoendoscopy after the orticochea pharyngoplasty revealed little movement of the palate or lateral pharyngeal walls. (Nasometry revealed an oral-nasal resonance imbalance, although we will not focus on the resonance issues in this paper.) Because of the velopharyngeal incompetence, and limited success of speech therapy in reducing the deviant speech patterns, the cleft palate team decided to implement a speech bulb reduction program.

PROCEDURES

Speech bulb program

A speech bulb is a dental appliance with a velar extension which closes off the velar port maximally (although as we see for Tia, some nasal emission still occurred, meaning complete occlusion was not achieved). Speech therapy continues crucially after the child receives the bulb in order to eliminate compensatory patterns while the mechanism is being normalized through use of the appliance. At appropriate intervals, the speech bulb is reduced minimally in size, with the hope that the child will initiate use of the lateral pharyngeal walls to achieve closure, in order to continue to have a normalized mechanism. Before such reductions are made, however, notable improvement in speech production needs to occur.

The data

Tia had worn her speech bulb for about a month when the videotaped data for this analysis were collected (a typical period for a post-bulb evaluation). In her case, three sets of data were available for the evaluation analysis: a small set of data with the bulb out, and a longer data set with the bulb in, both in single words and connected speech.Sufficient pre-bulb data and speech bulb out data were not available from the hospital to make a complete pre-post comparison. However, sufficient data are available to comment on the variable effects of the bulb in single words and connected speech. This paper is not an evaluation of the speech bulb program for this child, but an example of analytic procedures. We would of course recommend the same type of pre-analysis for clinical evaluation of the speech bulb for a given client.

The general analysis framework

Following phonetic transcription of the three data sets (speech bulb out, and speech bulb in, single words versus connected speech), data was analyzed in four ways. Syllable and word structure characteristics were defined, and three types of feature description were made, two of which related to nonlinear feature geometry, and one to actual phonetic output. Proportional matches with the adult target were calculated for each of the nonlinear codings, and cumulatively across all three feature conditions for phonetic coding, the rationale being that phonetic realization is the end-product of phonological processing, and therefore subsumes all previous nonmatches. We elaborate on the nonlinear descriptions below.

Nonlinear analysis

Nonlinear phonological theory focuses on the hierarchical nature of relationships among phonological elements. Two major levels of phonological organization are associated in principled ways: the prosodic level and the segmental level. The prosodic level includes all structure above the level of the segment and ultimately ties in with stress and intonation patterns. The lowest level of prosodic structure is subsyllabic structure, i.e., the onset or rhyme ([b] onset versus [æt] rhyme of <u>bat</u>). Progressively larger units of structure include the syllable, the foot (incorporating strong and weak syllables), the word, and ultimately, the phrase. The segmental level is described in most accounts as an hierarchy of features, each of which has some autonomy within the constraints of the hierarchical relationships. At the highest level of the hierarchy, the Root Node is the sum total of all of the features and this node connects to the prosodic tiers "above." The features immediately dominated by the Root Node are the manner and sound class features. In this account, following Bernhardt & Stoel-Gammon (1994) [3], we use:

a) [+consonantal] to refer to true consonants
b) [+sonorant] to refer to vowels and glides
c) [+continuant] to refer distinctively to fricatives (including /h/)
d) [+consonantal]-[+sonorant] to refer to liquids
e) [-continuant] to refer to stops, and
f) [+nasal] to refer distinctively to nasals.

These categories were sufficient for the speaker in question.

Also dominated by the Root Node are the Laryngeal Nodes and Place Nodes, which dominate their own set of features. Laryngeal Node features used in this study are:

a) [+voice] and [-voice]
b) [+spread glottis] for /h/
c) [+constricted glottis] for [ʔ]

Place Node features used in this paper required some elaboration beyond what is typically needed for English, due to the compensatory substitutions involving place:

a) Labial, referring to all labial articulations, including /p/, /b/, /m/, /w/, /f/, /v/, and /r/ (the latter in combination with Coronal)
b) Coronal, referring to all consonants using the tip and blade of the tongue, and including /t/, /d/, /s/, /z/, /ʃ/, /ʒ/, /tʃ/, /dʒ/, /θ/, /ð/, plus /j/ and other palatals in combination with Dorsal
c) Coronal [-anterior], referring only to blade articulations: /ʃ/, /ʒ/, /tʃ/, /dʒ/
d) Coronal [+distributed], referring to interdentals /θ/ and /ð/
e) Dorsal, referring to articulations with the body of the tongue and pharynx. This includes both the typical dorsals of English (/k/, /g/, and /ŋ/) which are [+high] and uvulars and pharyngeals which are [+low], two of her compensatory substitutions.
f) Place combinations: Labial-Coronal, for /r/, Dorsal-Coronal for palatals (also involved in compensatory substitutions), and [+constricted glottis]-Other Place for doubly articulated consonants with glottal and other places, Dorsal-[+low]-Other Place for consonants with both pharyngeal or uvular and other places of articulation.

Another subtheory of nonlinear phonological theory important to the feature designation used here is the theory of underspecification. According to this minimalist theory, only the unpredictable features of a segment are present in underlying representation, other predictable features assumed to be encoded during processing. In phonetic output, then, more features are assumed to be present than underlyingly, but even in output, a minimalist description of relevant distinctive features is assumed. Thus,when we describe the phonetic output for labial stops, we will only refer to the minimum number of features which distinguish them unless we have reason to include others: [-continuant], [voice] features, and [Labial]. Where categories overlap, because of phonetic implementation or phonological aberrations, additional features will be added. Thus, if the child is having trouble differentiating nasals and stops, the feature [-nasal] may become a relevant comment for that child on labial stop production, even though it would not normally be needed in phonological and phonetic description.

As noted above, levels of representation have some autonomy in terms of phonological constraints and processes, but also interact with one another in principled ways. Features are combined into segments, and segments are combinations of features, and can be viewed componentially. To exemplify:

a) The feature [Labial] may be established in a phonological system, but not in combination with all manner or voice features. Thus, /m/ and /w/ may be possible segments, but /f/ and /b/ may not be possible segments, surfacing as /m/ or voiceless [m]. Thus, the feature [Labial] is intact, but it can only cooccur with [+nasal], and not with [+continuant] or [-nasal].

b) Alternately, the feature [Labial] may not occur at all: /m/, /b/, etc., all surfacing with some other place of articulation.

c) The feature [+continuant] may be present in the system and realized in combination with [+spread glottis] for /h/, but may not cooccur with oral Place features, which all surface as [h]. Thus, the feature [+continuant] is present in output, but not in combination with other features.

d) Alternately, no fricatives, either oral or glottal may appear, the feature [+continuant] thus not occurring at all.

e) There may be voicing constraints. For example, a child may produce /k/ for both /k/ and /g/, but produce /t/ and /d/. When this is the case, the features [Dorsal] and [-continuant] are present, but the feature [+voice] does not cooccur in output with the [Dorsal] and [-continuant] features, only in combination with [Coronal] and [-continuant].

f) Alternately, no [+voice] obstruents may occur, in which case the feature [+voice] is not yet established in the system.

These constraints on feature presence or cooccurrence may apply across word positions, or only to one word position. In any event, we are able to describe the output data phonologically in such ways, deriving a perspective on major and minor problem areas for a child: whether in terms of feature establishment per se, or in terms of feature cooccurrence.

Final phonetic product

In the analysis we present, phonological features are examined in terms of their presence and their combinatorial power in the various speech conditions. However, the end-product of phonological processing is the articulatory-phonetic realization. Thus, both the phonological mismatch types are subsumed in this last category, and further details of phonetic deviance are included. Note that we are not in any of these cases making inferences about phonological representation per se. We assume that phonological constraints operate on output, and thus can be observable in the phonetic output. In the final production, the feature [Labial] may cooccur with the appropriate features, but have some component of difference which infers a phonetic realization problem. In other words, phonological contrast is evident, and features combined, but phonetic realization results in some overlap or imprecision. Note that in the features we described above, details of implementation are not generally included. For example, [Labial] does not distinguish between bilabials and labiodentals. In English, labial fricatives are labiodental, and hence that detail of phonetic implementation is not considered a phonologically relevant distinction, even though it is phonetically important. The [+anterior] coronals in English are alveolar in phonetic place of articulation, but in other languages [+anterior] coronals are dental, another example of a relevant phonetic and irrelevant phonological distinction. These differences are important in the examination of deviant speech, where mechanism abnormalities can result in an "irrelevant phonetic distinction" becoming both "phonologically" and "phonetically relevant." Examples of such end-product differences include:

a) Nasal emission superimposed on oral articulation, indicating phonological intent to produce an oral consonant, with insufficient velar port closure

b) Weak stop productions that have a fricative-like aspect, but that are still distinguishable from true fricatives in the child's output

b) Place of articulation that is within the region of the intended articulation (and phonologically "accurate") but is phonetically different. Examples of such deviations may include:

i) Labiodentals produced with the lower teeth and upper lip

ii) All labials produced with the teeth and the lips, but with consistent manner differences to distinguish them from each other and from the /f/ and /v/

iii) Coronals produced in the palatal region, provided they are consistently distinguished from dorsal target consonants

iv) Dorsals produced consistently as palatals or uvulars in a way that consistently distinguishes them from coronals

c) Voicing that distinguishes between voiced and unvoiced segments, but not with the expected adult target values of the language. For English, this may imply a failure to aspirate word-initial stops.

In summary, then, the data were examined in the three conditions: speech bulb out, and speech bulb in, single words and connected speech. The phonetic realizations were coded in terms of:

a) feature presence (using the features described above, and including the concept of underspecification)

b) feature cooccurrence, and

c) final phonetic output.

Tia's speech production: Syllable and word structure

As we commented earlier, syllable and word structure was reasonably well-developed. Thus, phonological constraints for word production were minimal, except with respect to consonant clusters, which were still developing. Across the conditions, she was able to produce up to 3-syllable words consistently. Word shape matches were better in the connected speech condition than the single-word condition. This will be seen to be a trend opposite to that for feature production, and may reflect the particular words sampled rather than a better ability to produce clusters in running speech. In any event, phonological output at the prosodic structure level was adequate with the speech bulb in and out.

Tia's speech production: Features

There is not room in this paper for a complete examination of the features in terms of occurrence, cooccurrence and final phonetic realization. Hence, we will summarize the relevant procedures and results here, and present a table which exemplifies the analysis.

Feature occurrence

The major feature occurrences of concern were [continuant] (for stops and fricatives), [+consonantal]-[+sonorant] (for liquids), Coronal, and Dorsal. Tongue placement and oral pressure are implicated for all those but the liquids (which could have been developmental). Hence the speech bulb program should have had an impact on feature production. This was in fact the case for the Root (manner) features involving obstruents. In the 'bulb in-single words' condition, [-continuant] was 96% accurate in terms of occurrence, and [+continuant] was 85% accurate. Performance was somewhat worse in the connected speech condition (93% and 68% respectively), but this was an improvement over the 45% accuracy for [-continuant] in the bulb-out condition. Coronal and Dorsal were approximately the same across conditions for feature occurrence, which is important in the evaluation of the speech bulb. Compensatory articulations often affect place of articulation, and if they did not change noticeably in the month, she was not yet ready for bulb reduction, and it can be assumed that phonological habituation was strong for place. The artificial occlusion of the velar port, however, did increase her ability to produce high-pressure obstruents, indicating that phonologically, those features were well-established, and that mechanism was affecting phonetic realization.

Feature cooccurrence

Additional concerns can become apparent when examining feature cooccurrence. Either new features can show up as problematic, or features already of concern in terms of occurrence can decrease in accuracy. Cooccurrence accuracy implies that all Root, Laryngeal, and Place features of a segment are accurate, unless separate tabulation is made for Root-Laryngeal, Root-Place, and Laryngeal-Place combinations (which we have not done here in terms of brevity). Feature cooccurrence here is thus a measure of *segmental* phonological accuracy.

Decreases in accuracy were noted for [-continuant] and [Dorsal] across the 'bulb out' and 'bulb in-connected' speech conditions. For example, in the 'bulb out' condition, accuracy for [-continuant] decreased from 45% to 33%, and in the connected speech condition, from 93% to 77%. However, in the 'bulb in-single words' condition, accuracy was maintained for these features at a fairly high level. For example, [-continuant] was 96% and 95% across the two conditions. The differences between the single words and connected speech conditions for these features does indicate some continuing concern for speech bulb reduction however, particularly since the dorsal stop substitutions involved compensatory palatal and uvular articulations, and other manner errors involved nasal substitutions for stops.

No changes were noted for Coronal or liquids.

New features involved in the feature cooccurrence analysis were [Labial] and [+voice]. Hence, although the [Labial] and [voice] features were reasonably well-established in terms of occurrence, cooccurrence restrictions affected them also across conditions, although least in the 'bulb in-single words' condition.

Phonetic end-product

The final analysis was of the phonetic variants. As the sum total of *all* deviations, decreases in performance can be expected, but may not necessarily occur. Further decreases in accuracy were noted for:

a) [-continuant] in all conditions
b) [+continuant] in the 'bulb in' conditions
c) Labial and Coronal in 'bulb out' and Coronal in 'bulb-in' conditions
d) [-voice] in the 'bulb out' and connected speech conditions.

These tendencies further emphasize the phonetic difficulty with high-pressure obstruents and nasal emission, suggesting that occlusion was not sufficient with the particular bulb being used. Again, the place features continued to show deviations of nasalization and precision of placement, showing phonetic difficulty and compensation not yet corrected by the appliance and the therapy program. See Table 1.

CONCLUSION

An example of an analysis methodology differentiating types and degrees of phonological and phonetic features was presented for a child with a repaired cleft palate who was on a speech bulb reduction program. By dividing phonological features into occurrence and cooccurrence categories, and also separating out a phonetic category, relative robustness of features was identified. Coronal and Dorsal place were problematic across typologies and bulb-in/out conditions. Labial and [voice] features became implicated in the cooccurrence and phonetic conditions, showing they were less robust than phonological occurrence indicated. Furthermore, obstruent manner features responded positively to the speech bulb, although less so in connected speech. Overall the speech bulb program was having some influence on obstruent manner production (because of occlusion of the nasal cavity), but not sufficient influence in connected speech. Place and voice remained problematic, and hence the compensatory substitutions were not yet sufficiently diminished for reduction to be done. The methodology is thus seen to have potential for use in clinical situations, particularly when technological assessment is not available.

Table 1. Place Node Features: Bulb-in, connected speech

	Present?	Cooccurring?	Phonetic accuracy
Labial	95%	87%	84%
Coronal	79%	71%	63%
Dorsal	98%	77%	74%

REFERENCES

[1] Howard, S. 1993. Articulatory constraints on a phonological system. *Clinical Linguistic and Phonetics*, 7, 299-317.
[2] Archangeli, D. & Pulleyblank, D. 1994. *Grounded Phonology*. Cambridge, MA: MIT Press.
[3] Bernhardt, B. & Stoel-Gammon, C. 1994. Nonlinear phonology. *JSHR*, *37*, 123-143.