

Acoustic phonetic examination of speech from subjects with late primary palate repair

R. Perry^a, L. Ham^a, Á. Martínez^a, L. Bermúdez M.D.^B, A. Lizarraga M.D.^B

^a *Laboratorio de Lingüística, Universidad Nacional de Colombia, Bogotá, Colombia*

^b *Operation Smile Int., Equipo de Investigación Médica*

Summary

Velopharyngeal insufficiency, an important feature in cleft lip and palate (CLP), bears a heavy weight on the (un)intelligibility of speech signals. Standard instruments for phoneticians include facilities for acoustic analyses and others. This paper presents preliminary, non-quantitative, results from the examination of pre- and post- surgery speech samples by CLP patients. The study focused on one sentence type and, in particular, on the production of voiceless stops and the vowel [a]. Our observations show that acoustic cues provide useful information about aspects of the tokens whose qualities improve, whether this be evident or not from an impressionist standpoint. Suggestions are forwarded about the possibility of using these and other phonetic tools to enhance diagnostics and treatment of CLP.

Introduction

As regards the patient with CLP, velopharyngeal insufficiency is the main axis of a series of constraints that hinder the production of normal speech signals. One of its consequences is the persistent disturbance of what should be a set of normal aerodynamic maneuvers within the vocal tract. There are many more consequences, some of which seem relatively unexplored. Seldom are we aware of the fact that CLP subjects experience the phenomena related to bone conduction of their own speech rather differently as compared to the experience of someone without this sort of anomaly.

In this paper, we claim that the treatment team for CLP can beneficially be extended to include phoneticians, at least as concerns current practice within the Spanish-speaking Americas. We will try to illustrate our claim with discussion of some examples of acoustic phonetic analyses of audio samples by CLP patients (see Fig. 1, below).

Those of us who work on phonetics and linguistics were recently invited by our Operation Smile Team to discussions concerning properties of speech samples from

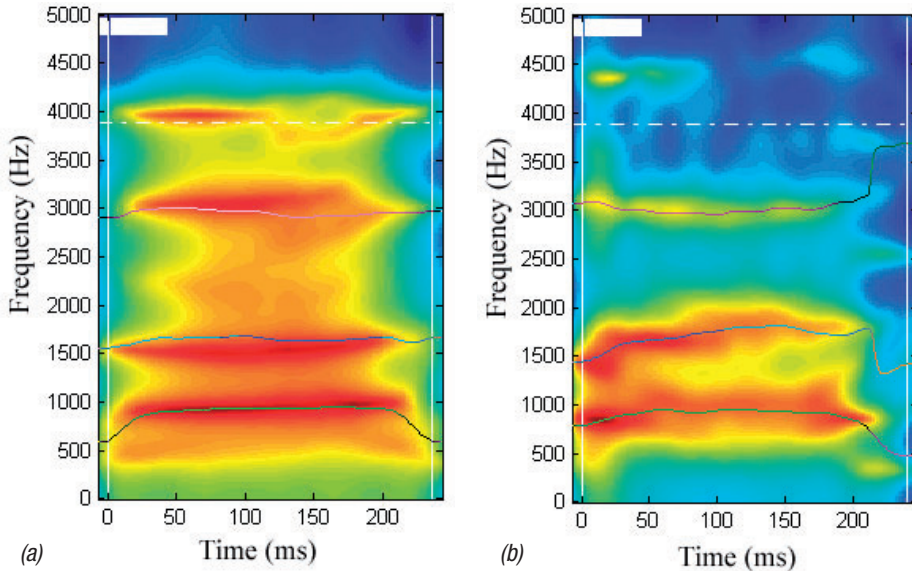


Figure 1. Example of two relatively successful central-frequency-of-formant tracings (slim lines) for two tokens ((a)pre-; (b)post- surgery) of the vowel [a] by one of the subjects. Notice (i) much more accurate tracking in (b); (ii) relative absence of nasal murmur in (b); (iii) profusion of nasal resonances in (a).

the database regularly collected as ground for the CLP patients' history, diagnostics and treatment¹.

Materials and Methods

Pre- and post- (12 mo.) surgery speech tokens (for the type *el pato toca el pito*—‘the duck plays the whistle’²) by a group of 9 CLP patients (from Colombia, Paraguay and Venezuela, age $7 \leq - \geq 13$ ³) were piecemeal inspected and selected for the cleanest audio signals from a pool of tokens for 5 different Spanish sentence types. These had been recorded by speech and hearing pathologists with a Sony Pro Hi-MD MZ-M200 recorder and its accessory microphone, the Sony ECM-DS70P (44.1 kHz sampling rate, 16-bit quantization rate), in fieldwork/office environments.

Stereo channels were mixed into mono tracks, downsampled to 16 kHz (anti-aliasing routines being observed) and processed for analyses^{[1] [2] [3]}.









































































Choice of the tokens was meant to enable analysis of pulmonic voiceless plosive consonants, [p], [t] and [k], in a controlled intervocalic context, and of the vowel [a]. F₀ values were manually measured for each studied sample for the vowel [a] of the words *pato* and *Tato*, respectively, in order to calculate appropriate parameter

1 Hereby, the Laboratorio de Lingüística, Universidad Nacional de Colombia, would like to acknowledge its appreciation of the team's openness and readiness to venture into collaborative work with it.

2 For subjects 4 and 5, the sample elicited pre-surgery was *el pato toca el pito* and the post-surgery one was *Tato toca el pito*—‘Tato plays the whistle’.

3 Informed consent conditions were met.

Table 1.  Spectral information for  Spectral information against  Spectral information neutral

	1	2	3	4	5	6	7	8	9
Loss of nasal formants									
Amplitude decrease of nasal murmur									
Attainment of burst in release (voiceless stops)									
Enhancement of the aerodynamics									
Enhancement of the saliency of oral formants									
Global enhancement of spectral profile									
Drop in acoustic coupling									
Reduced formant bandwidth									

settings for the production of sound spectrograms and power spectra (narrowband FFT, LPC envelope overlaid).

Results

Initially, the protocol for recording audio signals was not meant for acoustical phonetic analyses. Hence, our results do not afford statistics, and are not to be taken as absolute, in any sense.

As for the study of vowels with potential hypernasality, harmonic spacing in the frequency domain for women's and children's voices severely affects precise determination of central frequencies of oral formants or their tracking by (semi)automatic means. Even worse, these tasks and the task of telling oral formants from nasal formants are made much harder as the degree of nasality increases, since increasing excitation of the nasal cavities brings about new formants (poles) and anti-formants (zeros) in the global sound signal, and these elements tend to variously displace the expected resonances for particular oral (non-nasalized) vowels along the frequency scale^{[4] - [9]}. Both the scenario just sketched and other considerations as regards hypernasality effects in the production of stop consonants motivated us to adopt the strategy of observing the signals in 3 different analysis environments^{[1] [2] [3]}, each one assigned to one of 3 different researchers. We subsequently conferred to compare results. Consensus was, then, reached around a series of general, acoustically based, indexes for the determination of improvements concerning hypernasality in the post-surgery audio signals as compared to the pre-surgery ones.

Table 1 shows the collective assignment of subject-by-subject values for 3 different intervals adopted for assessing what, from the standpoint of the production of the vowel [a] and of voiceless stops, can be inferred in relation to the transit from the pre-surgery condition to the post-surgery one.

Study of the quality of the voiceless plosive consonants led to exploratory classification of the subjects in two major groups: those who seem to have already (before surgery) centered their articulatory habits on **place of articulation** (PA), and those who have centered their articulatory habits on **manner/timing of articulation** (MA). This is also reflected in the third row of table 1. Group PA subjects seem to more

rapidly improve, as regards the quality and intelligibility of their speech signals, upon the repair procedure. Group MA subjects tend to keep their manner of articulation habits more resiliently (via predominance of glottal stops); their main immediate improvement seems to emerge as an ability to not keep on so intensely stressing their larynges as before surgery. Except for subject 9, those that show an improvement in these respects are all denoted by green squares.

Conclusions

Collaboration of the CLP treatment team with phoneticians is a promissory undertaking. Phonetics can enhance the team's speech signal recording and analysis techniques ^[10], provide useful guidance as regards sociophonetic issues, new instruments for the examination of the aerodynamics of speech ^[9], and leadership in exploiting current knowledge concerning proprioceptive aspects involved in speech production ^[11].

Acknowledgments

We would like to express our appreciation of Myriam Jiménez and Mónica Ramírez for technical assistance.

References

1. BOERSMA P, WEENINK D, Praat: doing phonetics by computer. Version 5.1, Amsterdam, Netherlands, University of Amsterdam, 2009.
2. SJÖLANDER K, BESKOW J, WaveSurfer, Version 1.8.5. Stockholm, Sweden, Kungliga Tekniska Högskolan, 2005.
3. MORRISON GS, NEAREY TM, FormantMeasurer: Software for efficient human-supervised measurement of formant trajectories, Canberra, Australia, 2008.
4. HATTORI S, YAMAMOTO K, FUJIMURA O, Nasalization of vowels in relation to nasals, *J. Acoust. Soc. Am.*, 30, 267–274, 1958.
5. DANG J, HONDA K, SUZUKI H, Morphological and acoustical analysis of the nasal and the paranasal cavities, *J. Acoust. Soc. Am.*, 96, 2088–2100, 1994.
6. CHEN M, Acoustic correlates of English and French nasalized vowels, *J. Acoust. Soc. Am.*, 102 (4), 2360-2370, 1997.
7. STEVENS KN, *Acoustic Phonetics*, MIT Press, Massachusetts, 2000.
8. JOHNSON K, *Acoustic & Auditory Phonetics*, 2nd Ed., Oxford, Blackwell Publishing Ltd., 2003.
9. NIU X, KAIN A, VAN SANTEN JPH, Estimation of the Acoustic Properties of the Nasal Tract during the Production of Nasalized Vowels, *Proc. EUROSPEECH*, September 2005.
10. GRANQVIST S, SVEC JG, Microphones and room acoustics and their influence on voice signals, *PEVoC 6*, London, 2005.
11. ITO T, TIEDE M, OSTRY DJ, Somatosensory function in speech perception, *Proc Natl Acad Sci U S A*, 106(4), 1245–1248, 2009.