Laryngealizations in Cleft and Non-Cleft Speech: Acoustics and Prosodic Considerations

Aveliny Mantovan Lima-Gregio, Plínio Almeida Barbosa

Speech Prosody Studies Group, State University of Campinas, Brazil

avelinylima@gmail.com, pabarbosa.unicampbr@gmail.com

Abstract

This paper describes laryngealizations in vowels in the speech of individuals with (TG) and without cleft palate (CG). The relation of prosodic boundary to frequency and extension of laryngealizations is investigated. Three repetitions of the reading of a text by ten male adults who composed TG and CG are considered for analysis. The results reveal: higher frequency of laryngealization in the region of a prosodic boundary than outside this region for CG; consistent drop of intensity in the laryngealized vowel segment in CG and TG; higher jitter in CG and TG; superior f0 perturbation in TG.

Index Terms: cleft palate, laryngealization, glottalization

1. Introduction

The compensatory articulations realized by the individuals with cleft palate are possibly functional replacements that sound like distortions of the usual speech gestures. In the majority of times the replaced property is the place of articulation (generally altered for more posterior), whereas the manner of articulation is maintained. Such alterations are recruited in response to structures of the vocal tract that are compromised due to cleft palate [12]. The compensatory articulations can coexist with developmental processes and patterns [4], however, both the non-compensatory articulations and the compensatory articulations are not necessary consequences in the speech of the patient with cleft palate [3]. The most common type of compensatory articulation is the glottal stop [ʔ] [11], which is a transient articulation of consonantal nature produced by the sequence of adduction and abduction of the vocal folds with an increased subglottal pressure. The presence of this compensation severely compromises the speech of the individuals with cleft palate, a result, mainly, due to the loss of contrast especially of the place of articulation, considering the inventory of constriction locations that the articulators realize for the consonants in normal conditions.

In a previous pilot study carried out from an analysis of a reading of a Brazilian subject with cleft palate, a high frequency of laryngealization during the vowels was observed. This finding, and the absence, in the literature concerning Brazilian Portuguese (henceforth BP) of an analysis that would explain the reason and the limits of the compensation found, motivated the study presented here.

However, the phenomenon of glottalization has been deeply studied in many languages for non pathological speech [5,6,7,8,9]. In general terms, these studies showed some acoustic changes related to laryngeal modifications: drop of fundamental frequency (f_0), intensity and amplitude, and irregularity of glottal pulses. As regards the link between these modifications and prosody, some studies [5,9] reported higher occurrence of glottalizations: on utterance-final words, on word-initial vowels at the beginning of utterances, with a large influence of full and intermediate intonational phrases boundaries.

In this paper, duration-related phrase stress boundaries are first determined from a semi-automatic segmentation of the read material in VV units (Vowel-to-Vowel units) [1], in order to investigate the rhythmic structure of the utterances. Then, a second step investigates the relation of boundary vicinity to frequency and extension of laryngealizations. Our hypothesis is that the phrase stress boundaries act as attractors to the laryngealizations.

2. Method

The cleft palate subjects of this study are patients of the Laboratory of Experimental Phonetics at the University of São Paulo Hospital for Rehabilitation of Craniofacial Anomalies (USP-HRAC). The cleft data were recorded at that place and the non-cleft data were recorded at the Speech Prosody Studies Group Room. The USP-HRAC's Institutional Review Board approved this study and it has been done according to the principles outlined in the Declaration of Helsinki.

2.1. Subjects

Ten male speakers, aged from 25 to 53 years, and having good reading skills participated in this study. Two groups were formed. The test group (TG) was composed by five men with operated cleft palate, and according to a recent evaluation that is part of a clinical promptuary, they presented glottal stops in substitution of plosives and moderate hypernasality. The control group (CG) was composed by five non-cleft men. Both groups did not present hearing loss, history of recent surgery (last month), cold or hoarseness in the day of the recording.

2.2. Speech Stimuli

The corpus was composed by the reading of the text ("A história do urso preto" – *The story of the black bear* [10]) in BP, with nine sentences containing oral sounds only. This text is used in routine visits of the patients to HRAC-USP, which produces very natural readings.

Text: *A HISTÓRIA DO URSO PRETO*

Esta é a história do urso preto que veio da floresta. Hoje ele vive atrás das grades do zoológico da cidade. Todo domingo ele aparece logo cedo para ver os outros bichos. Gosta de olhar os pássaros e de ouvir o barulho das araras. Outro dia, resolveu sair da jaula. Logo viu que era difícil. Ficou triste, pois descobriu que estava preso. Decidiu que deveria esperar o dia de sua verdadeira liberdade.

In order to introduce a perturbation during the subjects' reading, the word "dia" (*day*) was replaced by the word "domingo" (*Sunday*), containing a nasal and a nasalized vowel. Each subject repeated the reading three times producing 30 readings for analysis.

2.3. Data Recording

The recordings were made in an sound-treated room of the Laboratory of Experimental Phonetics at HRAC-USP (TG) and in the room of the Speech Prosody Studies Group (CG), using the microphone *Shure Beta 58A*, and the audio plate *Sound Blaster Audigy Audio 9000* (*Audigy2*, *Creative*). The recording was made in mono-channel, with 22,050 Hz of sampling rate and amplitude resolution of 16 bits.

2.4. Data Analysis

The software Praat [2] version 5.0.25, was used to obtain the acoustic measures of interest. The data were segmented and organized into four tiers: "VowelOnset" (for VV intervals, that is, intervals between consecutive vowel onsets), "CV" (for segmenting vowels and consonants), "CVglottals" (for annotating the laryngealization phenomena: the laryngealized stretch was marked: 'l' if laryngealization, 'g' if glottal stop), and "Syntat." (for marking the syntactic link between phonological words). The latter tier will not be discussed here.

Acoustic criteria were used to identify laryngealization. Initially, auditory evaluation was made by the first author. Then, both the wave forms and spectrograms were visually analyzed (see fig. 1, 2 and 3) for all excerpts.

The measures of duration of laryngealization, f_0 and jitter were extracted automatically by two scripts, as explained in the next section. For the analyses shown here, two types of boundary region were defined, according to its extension around the VV unit in preboundary position. The narrow boundary region (nBReg) is formed by three VV units, one unit just before the preboundary VV, the preboundary VV itself, and one VV unit to its right. The broad boundary region (bBReg) follows the same procedure, but includes two VV units before the preboundary one, and two after it. Both types are referred here as boundary region (BReg). The positions outside these regions in each case are labeled as outside boundary regions (OBReg). The reason for that is because the automatic algorithm can assign a boundary to a post-stressed VV and the laryngealization can be located in the stressed VV leftwards, phenomenon still related to the boundary. Since words in Portuguese can have two post-stressed syllables, and post-boundary initial syllables are also affected by the boundary to their left, the chosen boundary region can safely capture the phenomena related to boundary neighborhood.

2.4.1. Scripts

Two scripts were used to automatically generate the acoustic measures: "NewGlottalizationAnalysis" to compute the acoustic modifications on the laryngealized segments, and "SGdetector", to detect prosodic boundaries.

The output of the first script is a text file containing for each vowel: I) the duration of vowel in milliseconds (ms); II) the percentage of duration of creaky voice over the total duration of the vowel; III) the percentage of f_0 drop (from modal to laryngealized part); IV) the percentage of difference in jitter (from modal to laryngealized part); V) the drop of intensity (from modal to laryngealized part), in decibels (dB). Besides, the file also indicates the vowel identity, and the type and the position of the laryngealization as regards the vowel, if applicable (on right, left or both). This script also allows corrections in the automatically detected glottal pulses.

The output of the second script is a text file containing for each VV unit: I) the duration of the VV unit (ms); II) a symbol to indicated if the VV unit is in preboundary position (1) or not (0); III) the duration of the stress group, from the extension between two consecutive VV units at prosodic boundary; IV) the number of VV units within the stress group; V) normalized values for VV unit duration (z-scores). It is from the automatic detection of the VV duration z-score maxima that the script assigns a prosodic boundary to a particular VV unit. The procedure is explained in [1]. For this work is enough to say that a particular VV unit is assigned as a boundary, that is the right edge of the stress group ending with this unit, if its normalized duration is a local maximum of the normalized duration consecutive values.

3. Results

A great variety of glottal realizations were observed, for both groups. Figures 1, 2 and 3 show some examples of laryngealizations found in this study.

In Fig.1, laryngealization occurs in both groups (left, CG and right, TG) at the end of the last vowel [ɐ] (pointers) of the word "araras" (*macaws*). It appears with low amplitude and irregular glottal pulses in both cases.

Fig.2 also illustrates the laryngealization in the middle of the word "floresta" (*forest*). Left pointer (CG) indicates the laryngealization during the vowel [ɛ] and right pointer (TG) indicates the laryngealization at the beginning of the vowel [ɛ]. In this case, the phenomenon appears with lower f0 in CG than in TG and with irregular glottal pulses.

In Fig.3, three different laryngealizations happen for a subject of the TG in three different moments between two words: I) at the end of the word-final vowel [e]; II) at the beginning of the word-initial vowel [a]; and III) at the end of the word-initial vowel [a]. These configurations presented low f0 (see how the vertical pulses become more apart in region I), intensity and amplitude, and irregular glottal pulses.

Kruskal-Wallis test comparisons among the CG individuals for the parameters percentage of f0 drop (from Hz), percentage of jitter difference, and drop of intensity (dB) of the laryngealized stretches in BReg and OBReg revealed (at least $p<10^{-3}$) that only the percentage of f0 drop parameter was not significant between the participants. In all subjects no statistical difference between these parameters values when the levels BReg (in boundary region) and OBReg (outside boundary regions) were compared. The same comparisons were made among the TG participants, which showed statistical differences between them for all parameters. However, there were no differences considering boundary regions.

Considering all CG subjects the average values for frequency of laryngealizations in VV units revealed statistical differences between them and also between BReg and OBReg. For TG the statistical analysis showed that only PKB subject was different for frequency of laryngealization and between OBReg vs BReg.

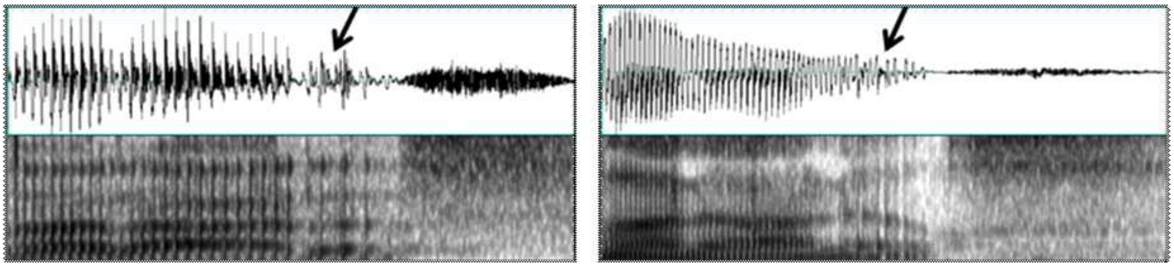

Figure 1: Example of laryngealization during the excerpt “araras”.

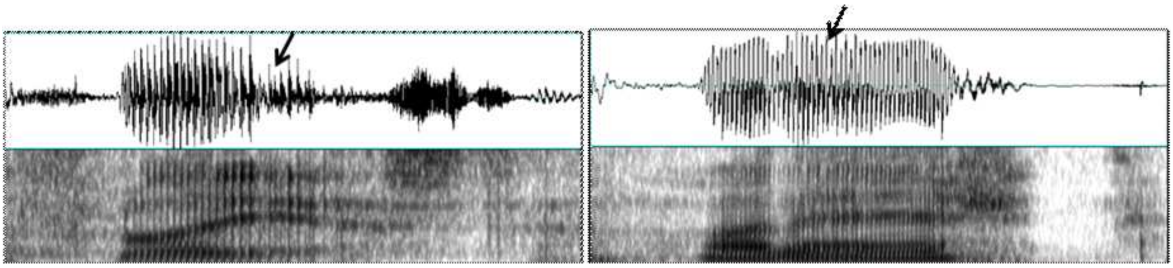

Figure 2: Example of laryngealization during the excerpt “floresta”.

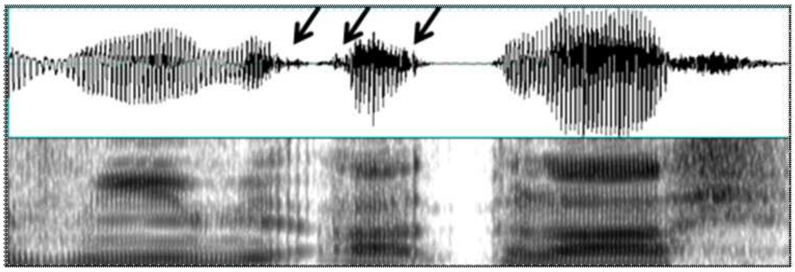

Figure 3: Example of laryngealization during the excerpt “vive atrás”.

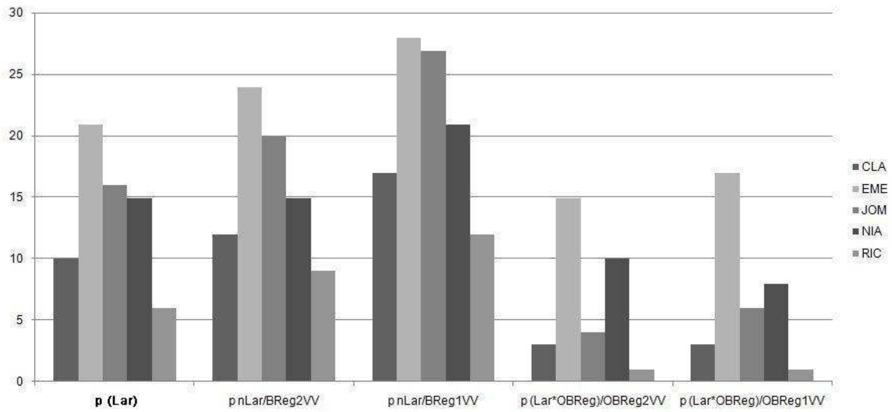

Figure 4: Comparisions among CG individuals for a priori probability and conditional probabilities.

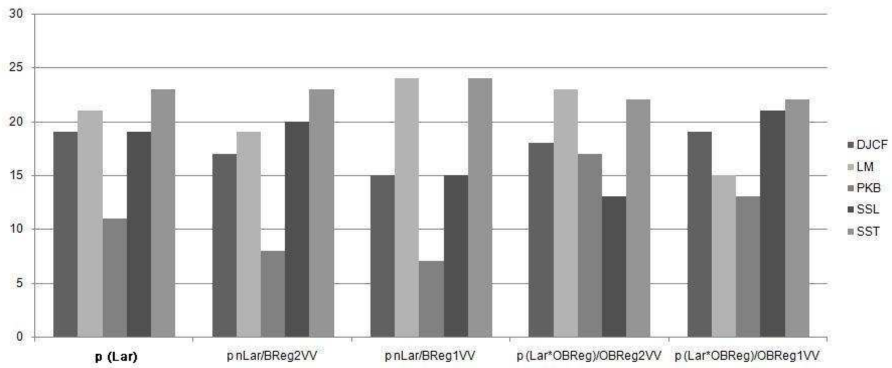

Figure 5: Comparisions among TG individuals for a priori probability and conditional probabilities

Fig.4 illustrates the a priori probability, p(Lar), and conditional probabilities p(Lar/BReg) and p(Lar/OBReg), for CG. It shows that a significantly lower probability of the phenomenon outside the boundary region and that the majority of laryngealization takes place within boundary regions (exception EME with more than 15% outside boundary region).

Fig.5 illustrates the a priori probability, p(Lar), and conditional probabilities p(Lar/BReg) and p(Lar/OBReg), for TG.

The frequency of laryngealization is higher in TG (19%), against (13%) in the CG.

4. Discussion

As regards the acoustic parameters, there is a lot of variability in degree and nature of the modifications. All subjects in both groups have more reduced values of intensity and higher jitter in the laryngealized part of the vowel, though the proportion varies for the two parameters. As regards f_0, it slightly increases in average in the analyzed window (from 1 to 4%) for two out of five subjects (participants DJCF and LM) in TG, whereas drops for the other eight subjects, from 1% to 0.

According to the literature consulted, in non pathological speech, f_0 and amplitude are parameters that drop in laryngealized stretches, whereas in this study, f_0 and glottal pulses observed in boundary regions showed some irregularity intra and inter-subjects. Intensity was the most consistent parameter (always drop) in the analysis intra and inter-subjects, in agreement with literature.

Considering both groups analyzed, the results altogether indicate a relatively similar frequency of the phenomenon of laryngealization, ranging from 2 to about 25% over the VV units, regardless of region type. Nonetheless laryngealization is largely more frequent in the boundary region than outside this region in the CG, indicating that strong prosodic boundaries may favor laryngealization, due to more extreme and more variable tension in the vocal folds around the boundary. This study agrees with non pathological speech literature that reported high occurrence of glottalizations on utterance-final words and word-initial vowels at the beginning of utterances [5,9], and can support our hypothesis that phrase stress boundaries attract laryngealizations.

On the other hand, according to the results, the TG doesn't make difference between BReg and OBReg, which means that cleft speakers present laryngealizations independent on the boundaries. This can be explained by the complexity of glottal system related to the compensatory misarticulations present on cleft palate speech.

Despite this general pattern, there is inter-subject variability: subject EME was an exception in CG, having in average, more than 15% of laryngealization outside the boundary region. In TG, subject PKB has a different behavior, having a high frequency of laryngealizations outside boundary regions. A possible explanation for PKB difference can be the speech therapy for eliminating compensatory articulations that he did ten years ago. Others studies in non pathological speech also suggest an interspeaker variability [7,8,9] for glottal measures, although Hanson and Chuang [8] indicate that male speakers present less variation than female speakers when the same glottal characteristics in both genders are compared.

5. Conclusions

This study suggests that phrase stress boundaries act as attractors to the laryngealizations in non cleft speech. It also suggests that the high inter-subject variability of some of the analyzed acoustic parameters affected by laryngealization in boundary regions and outside this region in cleft speech can be explained by the complex glottal system, which receive tension and overload, related to the articulatory disorders present on cleft palate speech.

6. Acknowledgments

We would like to thank Prof. Maria Inês Pegoraro-Krook and her team at HRAC-USP that kindly ceded the premises of the Laboratory of Experimental Phonetics for recordings and help us to found the patients. This work was supported by State of São Paulo Research Foundation (FAPESP), process number 06/60069-0. The second author acknowledges a grant from National Counsel of Technological and Scientific Development (CNPq), 300371/2008-0.

7. References

[1]Barbosa, P. A., “From syntax to acoustic duration: a dynamical model of speech rhythm production”, Speech Communication, 49:725-742, 2007.

[2]Boersma, P. and Weenink, D. “Praat: doing phonetics by computer”, version 5.0.25, software downloaded: http://www.praat.org, 2008.

[3]Bzoch, K. Communicative Disorders - Related to Cleft Lip and Palate. Austin, 1997, 580p.

[4]Chapman, K. L. “Phonological process in children with cleft palate”. Cleft Palate Craniofacial Journal, 30:64-72, 1993.

[5]Dilley, L.; Shattuck-Hufnagel, S. and Ostendorf, M. “Glottalization of word-initial vowel as a function of prosodic structure”. Journal of Phonetics, 24:423-444, 1996.

[6]Esling, J. H.; Fraser, K. E. and Harris, J. G. “Glottal stop, glottalized ressonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka)”. Journal of Phonetics, 33:383-410, 2005.

[7]Gerfen, C. and Baker, K. “The production and perception of laryngealized vowels in Coatzospan Mixtec”. Journal of Phonetics, 33:311-34, 2005.

[8]Hanson, H. M. and Chuang, E. S. “Glottal characteristics of male speakers: Acoustic correlates and comparison with female data”. J. Acoust. Soc. Am. 106(2):1064-77, 2001.

[9]Redi, L. and Shattuck-Hufnagel, S. “Variation in the realization of glottalization in normal speakers”. Journal of Phonetics, 29:407-29, 2001.

[10]Suguimoto, M. L. C. P. and Pegoraro-Krook, M. I. “Avaliação nasométrica em adultos normais falantes do português brasileiro”. Pró-Fono: Revista de Atualização Científica, 8(1):7-12, 1996.

[11]Trost-Cardamone, J. E. “Diagnosis of specific cleft palate speech error patterns for planning therapy or physical management needs”, in Bzoch, K. R. [Ed], Communicative Disorders - Related to Cleft Lip and Palate, 313-30, Austin, 1997.

[12]Wilbert, K. and Bradley, D. P. “Relationship between perception of error type and spectrographic analysis of cleft palate and developmental speech”. Paper presented at Florida Cleft Palate Association. Tampa, 1986.