Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Laryngealization, Gender and Speakers' Distinctiveness in Brazilian Portuguese

Julio Cesar Cavalcanti[1], *Luciana Lucente*[2], *Plinio Barbosa*[3]

[1]Campinas State University, Brazil
[2]Federal University of Alagoas, Brazil
[3]Campinas State University, Brazil
juliocesarcavalcanti.o@gmail.com, lucente.luciana@gmail.com

Abstract

This work aims to analyze how the occurrence of laryngealization in Brazilian Portuguese can contribute to a speaker characterization, through the analysis of laryngealization rates and its occurrence in vowel and consonant segments, in order to verify which measures would be more representative of a personal speech style. This work also aims at analyzing the influence of gender on the laryngealization rates. The corpus consists of semi-spontaneous speech records of 10 speakers, five men and five women, who speak the same dialect, ages ranging from 20 to 26 years old, all of them with a high school degree. These recordings are composed by the retelling of a story titled "Pear Film", a 6-minutes short film. Speech data were segmented and analyzed using the software Praat. Laryngealization was identified by hearing in vocalic and consonant segments by the first author, a speech therapist, and confirmed by waveform and spectrogram inspection. Results show a significant distinction in the occurrence of laryngealization between speakers, which may suggest that laryngealization rates could be relevant for speaker comparison. The results related to gender have revealed higher laryngealization rates for females.

Index Terms: laryngealization, brazilian portuguese, speaker comparison.

1. Introduction

Forensic phonetics is the application of general phonetic knowledge to legal problems, for example, in the identification or characterization of a speaker. Moreover, the tasks that a forensic phonetician can perform include portraying a set of voices, elaborating the vocal profile of a speaker, comparing one or several speech samples of an unknown speaker with one or several samples of a known speaker [1], among others.

According to [2], the main goal of the task of speaker comparison is to answer the question: How similar are the properties observed in the known and questioned speech samples under the hypothesis that the questioned speech sample has the same origin as the known sample in contrast to the hypothesis that they have a different origin?

It is well known that the phonatory singularity includes several aspects of variation, such as different ways of pronouncing vocalic and consonant segments, individual rhythm characteristics, relative durations, intonational patterns, range and tonal variation, as well as variation in voice quality. While contrasts between speech sounds are limited by the scope of variations in the phonological system, voice or voice quality may exhibit higher degrees of variation [3].

In this regard, laryngealization is a type of non-modal phonation characterized by low intensity, low fundamental frequency and irregular glottal pulses [4]. It has been pointed out as a potential clue for speaker characterization, because forensic phonetic research suggests a consistency as to the particularities of its occurrence among different speakers [5].

The literature points out that the phonological environments where this phenomenon is recurrent are prosodic boundaries of medial and final intonational phrases, followed or not by a silent pause [4], [6].

According to [6], the presence of a preceding silent pause or a phrasal boundary directly influences laryngealization. Moreover, the occurrence of laryngealization in a particular place is particularly interesting in the absence of a prosodic boundary in its vicinity, due to its low rate of occurrence.

In American English, [4] found a large variation in laryngealization rate among individuals without articulatory impairment during a reading task with statistically significant differences in most cases.

Furthermore, in an experiment designed to verify what type of linguistic information is preserved when a speaker tries to imitate the speech of another person, [7] observed that the speakers are rarely unable to imitate a pitch accent and/or an intonational phrase boundary. Nevertheless, the non imitation of phonetic cues is frequent, as well as the non reproduction of the duration of pauses and of irregular glottal periods associated with prosodic edges and pitch accents in American English. The results may suggest that, even in an imitation task, speakers are unlikely to perform laryngealization under the same rate.

Many studies have analyzed the influence of gender on the occurrence of laryngeal phenomenon in American and British English, revealing that there is no consensus as whether men or women produce higher laryngealization rates [4]. As for Brazilian Portuguese (BP), there is no evidence of the influence of this variable on laryngealization rate yet.

10.21437/SpeechProsody.2018-141

Therefore, this paper aims to analyze whether the occurrence of laryngealization can contribute to characterize a speaker. This is done by analyzing laryngealization rates in vocalic and consonant segments in order to verify whether these measures can characterize an individual through consistent variation of laryngealization rates across different speakers.

The analysis was performed considering male and female groups for the purpose of observing how gender may be related to the occurrence of the phenomenon.

2. Method

2.1. Corpus

The corpus consisted of a retelling recording after subjects watched a 6-minute movie entitled "Pear Film", produced at the University of California at Berkeley (1975). The movie consists of a sequence of unreported actions without dialogues. The movie script is self-explanatory through the logical succession of the scenes that compose it.

Only the first author was responsible for instructing and monitoring the data acquisition procedures, eliminating the effect caused by more than one instructor.

2.2. Speakers

Ten speakers, five men, M1, M2, M3, M4, M5, and five women, F1, F2, F3, F4, F5 were analyzed. The subjects were speakers of the same dialect (Maceió-AL, Northeastern Brazil), all graduates or undergraduates, ages ranging from 20 to 26 years.

2.3. Analysis

The identification of laryngealization events was based on the auditory analysis of all vowels and consonants produced in the speech and annotated in a tier by using the Praat software [8]. This first identification was confirmed by the waveform and the spectrogram, according to common practice [4], [6], [9]. All segments were labeled as modal (M) or laryngealized (L), which allowed computing laryngealization rates as the proportion of laryngealized segments in vowels and in consonants, separately and as a general measure irrespective of occurring in vowels or consonants for each subject (see examples of laryngealization events in Figures 1 and 2).

Inter-subjects rates were compared in order to verify which measure (vocalic, consonant or general laryngealization) reveals a larger variation between different speakers.

A proportion test was performed to compare the rates between subjects and between gender groups. Bonferroni corrections were used to adjust the level of significance due to multiple comparisons.

2.4. Apparatus

The audio material was acquired in a quiet environment using the Opus 55 Mk II condenser microphone (beyerdynamic) with an external audio interface. The recordings presented an average time of 1 min and 27 s, with a standard deviation of 0.05 s.

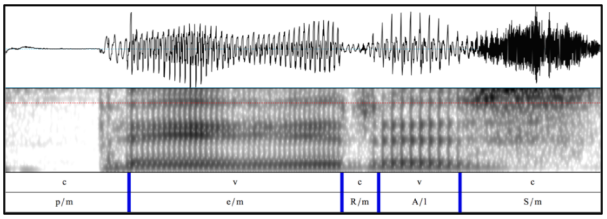

Figure 1: *Example of vocalic laryngealization during the last vocalic segment of the word "peras"(pears). In the first tier, the labels classify the segments in vowel or consonant; in the second one the segment was identified as modal (/m) or laryngealized (/l). The vowel 'a' (A/l) shows irregularly spaced glottal pulses in the spectrogram.*

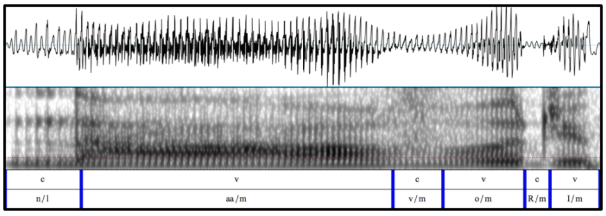

Figure 2: *Example of consonant laryngealization during the first segment of the phrase "na árvore" (on the three). In the first tier, the labels classify the segments as vowel or consonant; in the second one the segment was identified as modal (/m) or laryngealized (/l). The consonant 'n' (n/l) shows an irregularity in the acoustic signal, reinforced by the occurrence of irregularly spaced glottal pulses in the spectrogram.*

3. Results

In this work, we have analyzed 3,133 consonant and vowel segments, and the occurrence of 477 laryngealization events. It can be seen by comparing rates in Figure 3 with those in Figure 4 that a higher incidence of the phenomenon is found in vowel segments (a range from 5 to 24 %) in comparison with those in consonant segments (from 0.3 to 4 %). This difference between the proportion of laryngealization in vocalic and consonant segments was statistically significant for $\alpha = 5\%$, with $p = 0.006$.

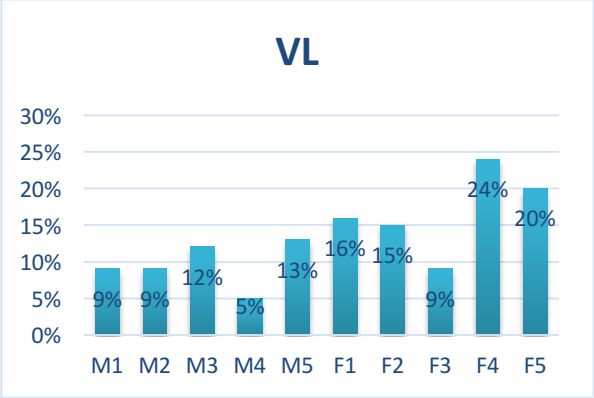

Figure 3: *Vowel laryngealization rates across subjects.*

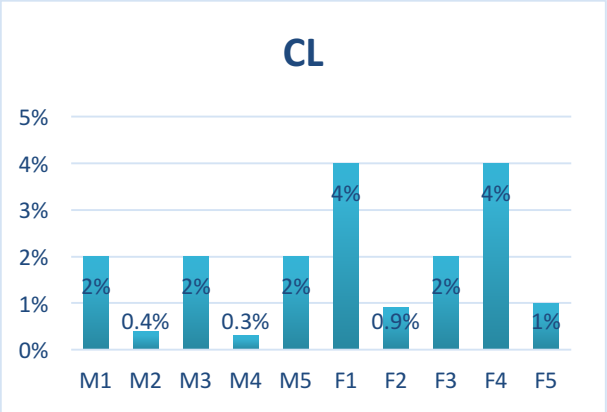

Figure 4: *Consonant laryngealization rates across subjects.*

Regarding the degree of variation in the rates across subjects, a greater variation in the production of laryngealization was found for vowel segments, as shown in these figures: a range from 5 to 24 % for vowels ($p < 0.001$) and from 0.3 to 4 % for consonants ($p < 0.001$).

Considering the general frequency of laryngealization (TOTAL-L) across subjects, a range from 3% to 14% was observed ($p < 0.001$). See Figure 5.

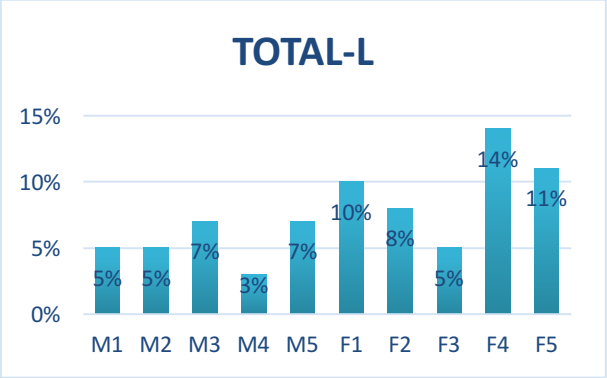

Figure 5: *General laryngealization rates across subjects.*

Table 1 shows the intra-gender pairs that were significant or marginally significant for vowel laryngealization rate. In this condition, 10 pairs of subjects were distinct, three intra-gender and seven across gender (M3/F4; F4/M5; F4/M1; F4/M4; F4/M2; F2/M4; M4/F1).

The incidence of the phenomenon in consonant segments was considerably less effective for differentiating the speakers. Only two inter-gender pairs were significant or marginally significant (F1/M4; F4/M4).

In contrast, when general laryngealization rates are used instead 16 pairs of subjects were statistically significant or marginally significant. Five intra-gender pairs, as shown in Table 2 for TOTAL-L, and eleven across-gender pairs (F4/M3; F4/M1; F4/M5; F4/M4; F4/M2; F2/M4; F1/M4; F1/M2; F5/M1; F5/M4; F5/M2) showed statistical significance.

Regarding vowel laryngealization rates and total laryngealization rates, these results show that laryngealization has a higher intra-gender distinction in the female group. In this group: two significantly different pairs were found for the vocalic laryngealization rates (VL) and three significantly or marginally significant pairs for the total laryngealization rates (TOTAL-L), whereas in the male group only one significant or marginally significant pair was found for the first measurement and two for the second.

Table 1: *Significant and marginally* significant pairs for vowel laryngealization rate.*

Combination	X^2	p-value
F3 / F4	20.721	< 0.001
F3 / F5	15.074	< 0.001
M5 / M4	10.628	0.001*

Table 2*: Significant and marginally* significant combinations for total laryngealization rate.*

Combination	X^2	p-value
F4 / F2	10.425	0.001*
F4 / F3	20.891	< 0.001
F5 / F3	10.638	< 0.001
M3 / M4	10.223	0.001*
M5 / M4	13.087	< 0.001

The results also pointed to differences of laryngealization rates across gender: women exhibit higher laryngealization rates than men for vowel laryngealization ($p < 0.001$) and consonant laryngealization ($p = 0.01$).

4. Discussion

Vowels exhibit the highest rates of laryngealization. Not only was vocalic laryngealization more frequent but it was also the measure that showed the greatest variation across subjects, which could be useful for forensic analysis.

Based on the assumption that laryngealization is a product of voice production in the larynx, it is expected that its effects would be more evident in speech segments involving vocal fold vibration, the so-called "sonorant" elements. On the other hand, in the speech segments where vocal fold vibration does not occur, it is expected that these "unvoiced" elements remain relatively unchanged [3].

In this present study, the statistical analysis suggests that vowel laryngealization rate and total laryngealization rate may have implications for the forensic speaker comparison task in BP, especially those related to intra-gender comparison, the most common ones in forensic analysis. Unless one is dealing with a case of vocal disguise, the ability of a measure to distinguish between same-sex speakers seems imperative.

According to [3], in general, an effective index for the speaker comparison depends on: (a) low intra-speaker parameter variation, and (b) high inter-speaker parameter variation, which provides a greater consistency in the comparison task.

In order to verify the degree of intra-speaker variation for laryngealization rates, a new methodological design should be

developed in the future with recordings of the same speakers in two different moments, with the purpose of evaluating laryngealization rates as a consistent index for the forensic speaker comparison.

The analysis of a single characteristic or aspect of speech has a reduced value if analyzed in an isolated way. That is why the analysis of a set of parameters is recommended for speaker comparison [10]. It is in this sense that the analysis of laryngeal events could contribute to speaker comparison procedures.

In this study a difference in the occurrence of the laryngeal phenomenon across gender was also found. According to [4], the factors that contribute to the differences across gender concerning the rate of laryngealization may be anatomical, sociolinguistic or perhaps a combination of these variables.

[4], [6] and [11] have also showed the tendency of female speakers to produce higher rates of laryngealization in American English when compared to male speakers.

According to [12] sociocultural motivations for the production of laryngealization by women may vary depending on the types of values that each society aspires or as a result of intrinsic cultural aspects.

On the other hand, [13] have stated that the difference observed between male and female groups may lie in the physiology of vocal production itself. The authors speculate that female speakers may not need to contract the vocal fold structure with as much intensity to reach the state of grouping or constriction as men do, because their shorter vocal fold length naturally contributes to deformation and thickness required to produce the laryngeal type of phonation.

5. Conclusion

This work has shown a significant variation in the occurrence of laryngealization among some speakers, and suggests that vowel laryngealization and total laryngealization rates can potentially be used as a parameter in forensic applications. Future investigations should evaluate the degree of variation of laryngealization rates intra-subject and between subjects, in order to verify its effectiveness for the forensic speaker comparison task. The results also pointed to differences of laryngealization rates across gender: women exhibit higher laryngealization rates than men, for all types of laryngealization rates analyzed in Brazilian Portuguese.

6. Acknowledgements

The authors would like to thank the National Council for Scientific and Technological Development (CNPq) for providing the financial resource to carry out this research.

7. References

[1] E. San Segundo, H. Alves, H, And M, F, Trinidad, "CIVIL Corpus: Voice quality for speaker forensic comparison," Procedia - Social and Behavioral Sciences. doi:10.1016/j.sbspro, 2013.

[2] G, Morrison, *Vowel inherent spectral change in forensic voice comparison*. In G.Morrison & Assmann, P. (Eds.) Vowel inherent spectral change, pp. 263-283. Heidelberg, Germany: Springer-Verlag, 2013.

[3] A, Hirson and M, Duckworth, "Glottal fry and voice disguise: A case study in forensic phonetics," Journal of Biomedical Engineering, 15, pp. 193-208, 1993.

[4] L. Redi And S. Shattuck-Hufnagel, "Variation in the realization of glottalization in normal speakers," Journal of Phonetics, vol. 29, pp. 407-29, 2001.

[5] P. French, *An overview of forensic phonetics with particular reference to speaker identification*. Forensic Linguistics 1: pp. 169-181, 1994.

[6] L. Dilley, S. Shattuck-Hufnagel, M. Ostendorf, "Glottalization of word-initial vowel as a function of prosodic structure," Journal of Phonetics, vol. 24, pp. 423-444, 1996.

[7] J. Cole, And S, Shattuck-Hufnagel, "The phonology and phonetics of perceived prosody: What do listeners imitate?," in *Proceedings of the Interspeech 2011*. Florence, Italy, 21-31 August, pp. 969-972, 2011.

[8] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 5.4, retrieved 4 October 2014 from http://www.praat.org/

[9] A. M. Lima-Gregio, "Oclusiva glotal e laringalização em sujeitos com fissura palatina: um estudo segundo abordagem dinamicista" (Doctoral Thesis). Campinas (SP): Universidade Estadual de Campinas; 2011.

[10] H, F, Hollien, *The acoustics of crime: The new science of forensic phonetics.* New York: Plenum, 1990.

[11] J. Ingle, "Pacific Northwest Vowels: A Seattle Neighborhood Dialecto Study." *Paper presented at the149th Meeting of the Acoustical Society of America*, Vancouver, B.C., May 16–20, 2005.

[12] H. Tajfel, "Social identity and intergroup Behaviour", Social Science Information 13.2: pp. 65–93, 1974.

[13] C. Henton and B. Anthony, "Creak as a Sociophonetic Marker," in Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin, ed. Larry M. Hyman and Charles N. Li, pp. 3–29. London: routledge, 1988.