Interactions of voice quality settings

Zuleica Camargo[1], Albert Rilliard[2]

[1]Integrated Acoustic Analysis and Cognition Laboratory (LIAAC), Pontifical Catholic University of São Paulo, Brazil

[2] Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), University Paris-Saclay, France

Abstract

This paper aims at investigating the possible interactions of the voice quality settings (VQS) related to vocal tract (supralaryngeal) and phonatory (laryngeal) activities in the Vocal Profile Analysis Scheme (VPAS), based on Laver's model. The corpus was composed of semi-spontaneous speech samples and repetitions of key-sentences (designed for voice quality evaluation) produced by 71 speakers. The database was perceptually annotated by means of the VPAS, adapted to Brazilian Portuguese (BP), indicating the vocal tract and phonatory VQS. Hierarchical clustering analysis was run. Results reinforced the interactions between the vocal tract and phonatory VQS, having the general muscular tension (vocal tract and laryngeal hyper and hypofunction) adjustments explained the main grouping of those VQS.

Key Words: phonetics, voice quality, auditory perception

Introduction

Since voice quality has linguistic, paralinguistic and extra-linguistic functions, many approaches are available to the evaluation in terms of its perceptual, acoustic and physiological characteristics.

Classifying voice qualities from the perceptual perspective is still a difficult task and relying on a phonetic description of voice quality as provided by Laver´s model (1980) and on the perceptually based Vocal Profile Analysis Scheme, the VPAS (Laver, MackenzieBeck, 2007), have proven useful for many fields, covering crosslinguistic and sociolinguistic variants, clinical demands, forensic and speech technologies applications, and expressive effects in speaking and singing voices (French et al, 2015; Mackenzie-Beck, Schaeffler, 2015; Rilliard, d'Alessandro, Evrard, 2018).

The adaptation of the system was carried out in Brazilian Portuguese context, referred to as the BP-VPAS (Camargo, Madureira, 2008). The basic analytical unit is the voice quality setting (VQS), a long-term muscular tendency in the vocal apparatus (vocal tract and phonatory levels). This system overcomes a traditional limitation of voice quality evaluations: the impressionistic descriptions.

Some of our previous investigations also challenged the demands in processing statistical analysis for the parameters generated by the VPAS

annotations (Camargo et al, 2015). For the moment, we are concerned with the interactions between the elements in VPAS system, trying to find instances of correlation among the VQS. This paper addresses the possible interactions of the VQS related to vocal tract (supralaryngeal) and phonatory (laryngeal) levels.

Methods

The corpus was composed of semi-spontaneous speech samples and repetitions of 10 key-sentences (designed for voice quality evaluation) read by 71 speakers (52 female and 19 male, ranging from 17 to 58 years old). The database was perceptually annotated by means of the BP-VPAS by two expert judges, generating the vocal profiles (the combination of VQS) for 1,608 utterances.

The VPAS annotations of the collected utterances may be considered as a large contingency table, counting the occurrences of each feature. This shows it is a complex set of data, especially because many dimensions of VPAS are, by construction, uncorrelated. However, some relations exist between profiles – especially because the speakers tend to recruit several VQS together, generating compound VQS (Laver, 1980, Laver, Mackenzie-Beck, 2007). This is captured by the analysis presented hereafter.

A correspondence analysis was run on the (raw) VPAS annotations for all speakers and all speaking styles (semi-spontaneous and read). The 20 first dimensions of the analysis explained more than 80% of the variance. They were used as an input for a hierarchical clustering procedure. The number of clusters was determined so as to minimize a criterion of inertia gain that leads to 11 categories built on the patterns of VPAS features, just exposed on results.

Results and discussion

Results reinforced the importance of interactions between vocal tract and phonatory VQS, specially by the indication of the 11 clusters generated: Cluster #1 (n=15): Raised Tongue Body and Harsh Voice VQS; Cluster #2 (n=22): Advanced Tongue Body, Raised Larynx and Falsetto VQS; Cluster #3 (n=665): Retracted Tongue Body, Retracted Tongue Tip, Laryngeal Hyperfunction and Harsh Voice VQS; Cluster #4 (n=79): Labiodentalization, Minimized Range of Lips and Tongue Body, Pharyngeal Expansion, Vocal Tract Hypofunction and Creaky VQS; Cluster #5 (n=15): Audible Nasal Air Escape and Nasal VQS; Cluster #6 (n=409): Advanced Tongue Tip, Minimized Range of Tongue Body and Jaw, Nasal, Denasal, Laryngeal Hypofunction, Breathy Voice and Modal VQS; Cluster #7 (n=35): Spread Lips, Extended Range of Tongue Body, Minimized Range of Jaw and Creaky Voice VQS; Cluster #8 (n=123): Spread Lips, Closed Jaw, Lowered Tongue, Pharyngeal Constriction, Vocal Tract Hyperfunction VQS; Cluster #9 (n=144): Advanced and Lowered Tongue Body, Open Jaw, Vocal Tract Hypofunction, Lowered Larynx, Creaky and Breathy Voices VQS; Cluster #10 (n=63): Pharyngeal Expansion, Lowered Larynx and Modal VQS; Cluster #11 (n=36): Lowered

Tongue Body, Extended Range of Jaw and Lips, Vocal Tract Hyperfunction and Breathy Voice VQS.

The Clusters #3, #6, #9 and #8 were the largest ones and reinforced the links between vocal tract and phonatory mechanisms, by respective muscular tension settings, in which: laryngeal hyperfunction was combined with vocal tract hyperfunction VQS (pharyngeal constriction, retracted tongue body, raised larynx) and, instead, laryngeal hypofunction was combined with vocal tract hypofunction VQS (advanced tongue tip, open jaw, pharyngeal expansion, lowered tongue body and larynx).

Laver's general muscular tension settings (hyper-hypofunction scale) seemed to explain the main grouping of vocal tract and phonatory VQS. Tense VQS were also related to non-modal (harsh voice) and lax VQS were closely related to modal or creaky voices.

These findings tended to associate an essential element linking vocal tract and phonatory voice events, i.e., the vertical position of the larynx (Fink, Basek, Epanchin, 1956). In hyperfunction muscular VQS, raised larynx was found and on hypofunction VQS, lowered larynx was registered. These indications were drafted in some of our previous explorations of VQS.

Additional remarks were related to tense falsetto linked to raised larynx and advanced tongue body (#2). The findings were related to a recurrent female setting characterized by anteriorization supralaryngeal articulations and breathy voice (#6); sounding "smaller", cf. cliché of "young girl" (Léon 1993). Some sources of variation of VQS are related to the organic structure of the vocal apparatus (the anatomy) and the phonetic usages (Laver, MackenzieBeck, 2007).

Conclusion

Results reinforced the interactions between vocal tract and phonatory VQS, based on muscular tension adjustments related to hyper and hypofuncional states.

Acknowledgements

References

Camargo, Z., Madureira, S. 2008. Voice quality analysis from a phonetic perspective: Vocal Profile Analysis Scheme Profile for Brazilian Portuguese (BP-VPAS). In: Barbosa, P.A., Madureira, S.; Reis, C. (Ed.), Proc. 4th Conference on Speech Prosody, 57-60, Campinas, Brazil.

Camargo, Z, Gomes, P.C., Madureira, S., Rusilo, L.C. 2015. Voice quality from a phonetic perspective: supralaryngeal and muscular tension settings. In: The Scottish Consortium for ICPhS 2015 (Ed.). Proc.18th International Congress of Phonetic Sciences, paper number 0198.1-5, Glasgow, Scotland.

Fink, B.R., Basek, M., Epanchin, V. 1956. The mechanism of opening of the human larynx. Laryngoscope 66(4), 410-425.

French, P., Foulkes, P., Harrison, P., Hughes, V., San Segundo, E., Stevens, L.C. 2015. The vocal tract as a biometric: output measures, interrelationships, and efficacy. In: The Scottish Consortium for ICPhS 2015 (Ed.). Proc.18th International Congress of Phonetic Sciences, paper number 0817.1-5, Glasgow, Scotland.

Laver, J.1980. The phonetic description of voice quality. Cambridge University Press.

Laver, J.; Mackenzie-Beck J. 2007. Vocal Profile Analysis Scheme-VPAS. Speech Science Research Centre, Queen Margaret University College, 2007.

Léon P. 1993. Précis de phonostylistique-parole et expressivité. Paris: Nathan.

Mackenzie-Beck, J., Schaeffler, F. 2015. Voice quality variation in Scottish adolescents: gender versus geography. In: The Scottish Consortium for ICPhS 2015 (Ed.). Proc.18th International Congress of Phonetic Sciences, paper number 0737.1-5, Glasgow, Scotland.

Rilliard, A., d'Alessandro, C., Evrard, M. 2018. Paradigmatic variation of vowels in expressive speech: acoustic description and dimensional analysis. The Journal of the Acoustical Society of America 143(1), 109-122.