Towards the use of ultrasonography to study aging effects in vowel production

F. Barros, S. Silva, L. Albuquerque, A. R. Valente, A. Teixeira, P. Martins, C. Oliveira

Universidade de Aveiro, Portugal

1. Introduction

The aging process is expected to play an important role in speech production. Although the acoustic properties of aging speech have been extensively studied, and the main goal of a line of research we have been pursuing for several years, its underlying articulatory details have not been well understood. Our previous acoustic results were not conclusive, there are vowels that presented a different pattern of formant frequencies variation with age and gender. We observed that vowel formants tend to decrease mainly in females and to centralize in older males with aging [1,2]. So, age related changes in vowels production have been measured mainly indirectly using formants frequencies, and F1 and F2 variation might be related to specific articulatory adjustments of the older speakers during speech.

Due to the limited knowledge of the articulatory basis of previously acoustic findings concerning speech aging, ultrasounds (US) tongue imaging synchronized with audio can be used to investigate the physiological differences between elderly and young adult speech. One of the major challenges of the US image, refers to the processing, visualization and analysis of the data collected. To overtake this, automatic tools needs to be developed to detect tongue measures (e.g. tongue height and tongue advancement). Then, the main objective of the ongoing study is to investigate a different automatic method to detect tongue measures to study the age effects on the European Portuguese (EP) vowels production.

2. Research Question

Established the relevance of performing production studies of aging effects using data acquisition methods such as US, before a large scale analysis of a reasonable number of subjects of different ages, two very important questions must be answered: what parameters to use instead of the commonly used formants in acoustic studies and how to obtain them in an automatic and reproducible manner to allow analysis of a large set of subjects?

3. Method

Corpus acquisition and data processing: composed by 9 EP oral vowels [i], [e], [ɛ], [a], [o], [ɔ], [u], [ɐ] and [ɨ] in isolated context and in a disyllabic sequence C**V**.CV (in stop and fricative consonant context). The stimuli were embedded in a carrier sentence. Synchronous acquisition of ultrasound images and speech sounds using Articulate Assist took place in a quiet room, using an endocavitary probe (65EC10EA) with 90º field of view positioned under the participants' chin using a stabilization helmet. Ultrasound was collected using a Mindray DP6900 at a frame rate of 60Hz. Audio was collected with a Philips SBC ME400 microphone connected to an external sound system (UA-25 EX USB). Corpus acquisition began with the production of the sequence /tatatata/ to assess sound and image synchronization and with swallowing saliva for hard palate delineation. Each participant repeated each sentence 3 times. The recorded data was first automatically segmented at phoneme level using WebMAUS.

<u>Image processing</u>: Based on the acoustic midpoint of the vowels, the corresponding images were selected and processed using an unsupervised method to extract points-of-interest in the tongue (e.g., tongue height). The method uses a radial sweep approach and, for each radial (steps of 5º of angular distance [3]), collects all the pixel intensities. The highest intensity point is extracted for each radial sweep (Fig. 1), its coordinates automatically extracted, and only the highest *y* coordinate was considered to represent the highest point of the tongue body. The *x*-coordinate reflects the front-back position of the tongue in the *y* coordinate.

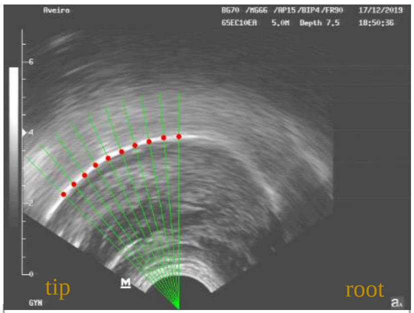

Fig.1 Illustrates an example of 10 radial sweeps and the extracted points on the tongue for vowel [a].

4. Fist results

These results reflect the production of the central EP vowels [a], [ɐ] and [ɨ] (9 repetitions of each isolated vowel per speaker). Central vowels were selected based on age-related changes in the acoustical vowel space, mainly in the reduction of the F1 space that seems to be a result of tongue elevation alterations. Fig. 2, shows, as a representative example, the data for a young female speaker, and reveals that vowels [ɐ] and [a] present similar tongue height, but different anterior-posterior tongue position. Moreover, vowels [ɐ] present a large dispersion in *x* coordinate, maybe due to the angular distance used. As expected, [ɨ] present a higher tongue elevation and higher tongue advancement.

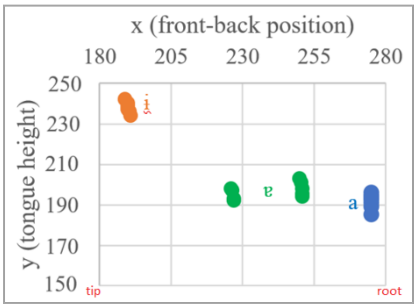

Fig.2 The graph dispersion of the *x y* coordinates of the highest point of the tongue in central EP vowels.

5. Conclusion

The present data are a starting point for a larger ongoing project concerning the analysis of the relationship between tongue measures collected by US and formant frequencies across the lifespan. To the best our knowledge data related to articulatory changes with aging are residual in the literature. So, using the proposed method, that seems to be able to separate the vowels under study with the parameters selected, we intend to contribute to increase the feasibility of the articulatory vowel study in lifespan.

Acknowledgments

This research was financially supported by the projects VoxSenes POCI-01-0145-FEDER-03082 (COMPETE2020 under POCI and FEDER, and by national funds (OE), through FCT/MCTES) and SOCA – Smart Open Campus CENTRO-01-0145-FEDER-000010 (Portugal 2020 under POCI and FEDER). And by the grant SFRH/BD/115381/2016, and by IEETA (UID/CEC/00127/2019).

References

[1] L. Albuquerque, C. Oliveira, A. Teixeira, P. Sa-Couto, and D. Figueiredo, “Age-related changes in European Portuguese vowel acoustics,” in *INTERSPEECH*, 2019, pp. 3965–3969.

[2] L. Albuquerque, C. Oliveira, A. Teixeira, P. Sa-Couto, J. Freitas, and M. S. Dias, “Impact of age in the production of European Portuguese vowels,” in *INTERSPEECH*, 2014, pp. 940–944.

[3] L. Ménard, C. Toupin, S. R. Baum, S. Drouin, J. Aubin, and M. Tiede, “Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults,” J. Acoust. Soc. Am., vol. 134, no. 4, pp. 2975–2987, 2013.