**Variability in the dynamic of nasal vowels in European Portuguese**

Conceição Cunha[1], Samuel Silva[2], Arun Joseph[3] António Teixeira[2], Catarina Oliveira[2], Paula Martins[2],Jens Frahm[3])

[1]*IPS, LMU Munich, Germany* [2] *Aveiro University, Portugal* [3]*Max Planck Institute for biophysical Chemistry, Germany*

The main aim of this paper is to describe the dynamic of nasal vowels in two segmental contexts and two prosodic conditions. Further issues are to relate different temporal dynamics to the size of the phonological inventories (see Oliveira, 2009, Martins, 2014) and to better understand the relationship between acoustic and articulatory data. The main hypothesis to be investigated is whether the coordination of oral and nasal gestures is delayed in Portuguese (Oliveira, 2009, Meireles et al, 2015).

The analysed corpus consists of minimal pairs containing all stressed oral [i, e, ɛ, a, ɔ, o, u] and nasal vowels [ɐ̃, ẽ, ĩ, õ, ũ] in one and two syllable words. All words were randomized and repeated in two prosodic conditions embedded in one of three carrier sentences alternating the verb and adverb as follows (*Diga*| *ouvi*| *leio* 'Say| I heard| I read' and *baixinho*|*depois 'gentle, after'*) as in '*Diga ponte, diga ponte baixinho*' ('Say bridge, Say bridge gently'). So far, this corpus has been recorded from fifteen native speakers of European Portuguese (8m, 7f) of EP and the work presented here considers the articulatory data already processed for two speakers.

Real-time magnetic resonance imaging (RT-MRI) recordings were conducted using a 3 Tesla Siemens Prisma Fit MRI System equipped with a 64-channel head coil (Niebergall et al. 2013, Frahm et al. 2014). MRI acquisitions involved a low-flip angle gradient-echo sequence with radial encodings and a high degree of data undersampling. Speech sound was synchronously recorded by means of an optical microphone (Dual Channel-FOMRI, Optoacoustics). Additional data (20 speakers) was recorded without scan noise. The RT-MRI data was processed using the semi-automatic methods proposed by Silva and Teixeira (2015) resulting in sequences of vocal tract contours (see Fig 1.). The analysis framework (Silva and Teixeira, 2016) proposes the comparison of different vocal tract configurations by dividing them in multiple regions with articulatory relevance (e.g., velum, tongue tip, tongue back, tongue dorsum) and applying measures that provide a normalized quantitative assessment of regional differences. Static analysis, as in Fig 2, is introduced for a clarification of the framework capabilities. Here, we can see differences on the vocal tract (VT) between the production of [a] and [u]. The diagram answer the question: "What happens to the vocal tract when I move from an [a] to an [u]?". The major differences between the VT configurations of both vowels were quantified and are visually presented, in the right panel, and include the movement of the lips, with a reduction of the lip aperture (LA) for the rounding of the [u]; the tongue dorsum (TD) moves up; the tongue tip (TT) position changes slightly back and down; and the tongue back (TB), as a result of a change in tongue height, moves further from the pharynx wall. As expected, no relevant differences in the velar region are observed.

For a first overall dynamic analysis of nasal vowels, to get a general insight on the major aspects of their production, Fig. 3 shows, as an example, the evolution of the vocal tract configuration along the production of [ẽ], i.e., what changes happen in the VT for the time interval annotated as an [ẽ]. The production of [ẽ] is essentially characterized by velum movement, with a notable change around 33% of the vowel duration, and some variation for the tongue dorsum and tongue tip, observed for the end of the vowel, and a possible consequence of coarticulation effects. Acoustically, the entire first two formants (F1 and F2) were compared for each nasal vowels and oral counterpart in order to identify major modifications in formants trajectories. Although the analysis is still on going, preliminary results confirm that the velum closure starts during the first half of the acoustic vowel and it is strongly influenced by vowel quality and prosodic context.

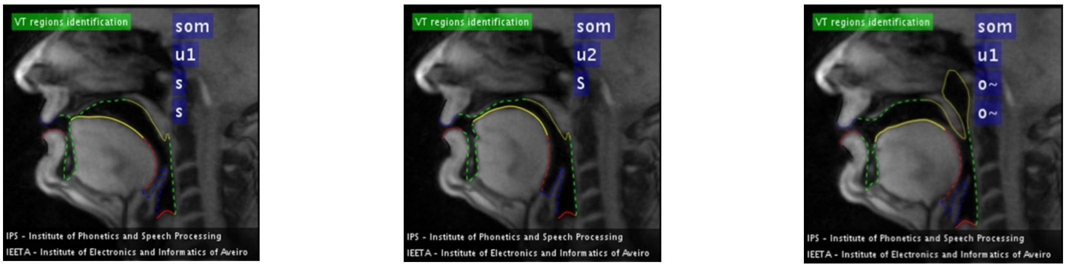


Figure 1. *Vocal tract contours of the alveolar [s] and postalveolar [ʃ] sibilants and the nasal vowel [õ] in the word som 'sound' for comparison reasons.*

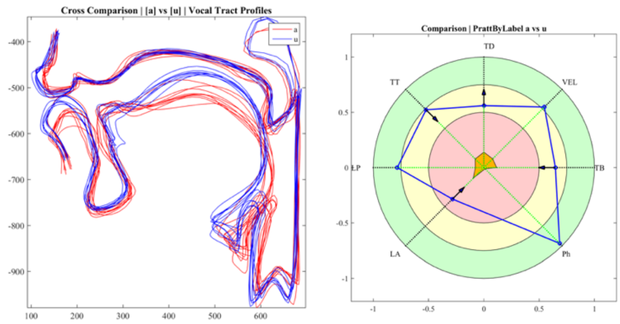


Figure 2. *Crosscomparison of the vowels [a] an [u] : Vocal tract contours in the left panel and quantified normalized regional differences between the two vowels, in the right panel. On the right, a dot in the yellow or red circular coronas signals notable differences for the correspondig region, with a small arrow hinting the direction of the displacement for that articulator.*

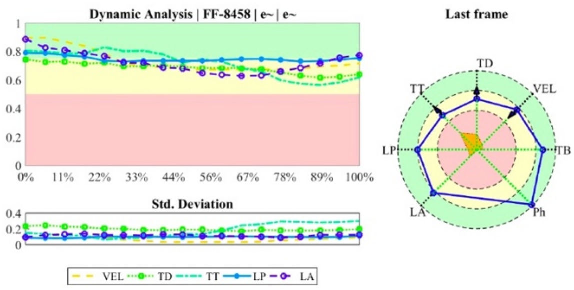


Figure 3. *Movement to the different articulators along the production of [ẽ] gathering data (i.e., occurrences) for multiple contexts. The variability observed, at the end of the vowel, is probably caused by coarticulatory effects*

**Selected references**

Frahm, J., Schätz, M. Untenberger, M., Zhang, S., Voit, D., Merboldt, K. D. Sohns, J. M., Lotz, J. and Uecker, M. 2014. On the temporal fidelity of nonlinear inverse reconstructions for real-time MRI–the motion challenge. *The Open Medical Imaging Journal*, 8, 1–7.

Niebergall, A., Zhang, S., Kunay, E., Keydana, G., Job, M., Uecker, M. and Frahm, J. (2013). Real-Time MRI of speaking at a resolution of 33 ms: Undersampled radial FLASH with nonlinear inverse reconstruction. *Magnetic Resonance in Medicine*, 69, 477-485

Silva, S. and Teixeira, A. 2015. Unsupervised segmentation of the vocal tract from real-time MRI sequences, *Computer Speech and Language*, 33/ 1, 25–46.

Silva, S. and Teixeira, A., 2016, Quantitative Systematic Analysis of Vocal Tract Data, *Computer Speech and Language*, 36/C, 307-329