An MRI study of the oral articulation of European Portuguese nasal vowels

Catarina Oliveira[1], *Paula Martins*[1], *Samuel Silva*[2], *António Teixeira*[2]

[1]ESSUA/ IEETA, University of Aveiro (Aveiro, Portugal)
[2]DETI/ IEETA, University of Aveiro (Aveiro, Portugal)

coliveira@ua.pt, pmartins@ua.pt, sss@ua.pt, ajst@ua.pt

Abstract

There is increasing evidence that, in addition to velopharyngeal coupling, lingual position may also change during production of phonemic nasal vowels. In order to investigate differences in oral articulation between European Portuguese (EP) nasal vowels and oral counterparts, imaging data (both static and real-time MRI) of several EP speakers (male and female) are used. Superimposition of outlines of the vocal tract profiles, semi-automatically extracted from MRI images, were used to compare the position of tongue and lips during nasal and oral vowel production. The results suggest that lingual and labial differences between nasal vowels and their oral counterparts are quite subtle in EP. Nasal vowels [ẽ], [õ] exhibited more articulatory adjustments with respect to oral congeners than [ĩ] and [ũ].

Index Terms: nasal vowels, Magnetic Resonance Imaging (MRI), European Portuguese

1. Introduction

Acoustic, articulatory and aerodynamic properties of nasal vowels have been studied in great depth over the last decades using a variety of data and techniques. Nevertheless, the vast majority of the literature on vowel nasality has been mostly concerned with the activity of the velum and the effects of coupling between naso-pharyngeal and oral tracts. It is often assumed that the contrast between the nasal vowels and the oral congeners lies essentially in the lowering of the velum, with no additional articulatory modifications. However, a growing body of articulatory studies on languages with phonemic nasal vowels [1, 2, 3, 4, 5, 6, 7, 8] indicates a set of articulatory adjustments on lingual and lips configuration.

Oral articulation of French nasal vowels is one of the best documented in the literature [1, 2, 3, 4]. In addition to the lowering of the velum, Zerling [3] observed consistent tongue retraction and labialization in the production of the French nasal vowels [ɑ̃] and [ɔ̃]. Recent magnetic resonance imaging (MRI) data [1, 2] show that, although the four Belgian French speakers have used different articulatory strategies to contrast the oral and nasal vowels due to differences in vocal tract anatomy, the tongue was more retracted and raised towards the velum during nasal vowels. The results of Carignan [5] also corroborate the previous lingual and labial findings for French. Articulatory, acoustic, and nasal airflow data collected by Shosted et al. [8] reveal oral articulatory adjustments in the production of Hindi nasal vowels relative to their oral congeners.

The articulation of European Portuguese (EP) nasal vowels has been studied in some detail by our team [9, 10] with emphasis on the dynamic pattern of the velum. However, the oral configuration of EP nasal vowels has not been systematically analyzed until now. The only studies available concern the Brazilian Portuguese (BP) variety [7, 11]. Master et al. [11] compared changes in vocal tract during BP oral and phonemic nasal vowels using tracings from xeroradiography of one speaker. Articulatory differences for /i/ versus /ĩ/ were minimal, but for /a/ versus /ã/ several oral adjustments were found. In a more recent MRI study of one BP speaker, Gregio [7] also reports differences in lingual position between nasal back vowels and oral equivalents according to the following progression [ã]>[õ]>[ũ].

The purpose of this study is to investigate the oral articulation of EP nasal vowels. Imaging data of several speakers are used in order to observe lingual and labial differences between nasal vowels and their oral congeners. As stated above, in many languages where nasalization is distinctive, such as French, nasal vowels are somewhat different from their phonological oral counterparts. In this work, we test the hypothesis that the same tendency might also be observed for EP. In other words, we will try to determine whether oral articulatory adjustments support or not the lowering of the velum to convey the [nasal] contrast in EP. Since Portuguese nasality is typically incremental over the vowel [12, 13], the articulatory analysis is based not only on static MRI data of three speakers, but also on real-time (RT) MRI productions of three other speakers, both obtained in separate moments. In general, MRI gives adequate information on the position of the articulators, but some authors [14] pointed out that sustained productions are hyperarticulated compared to real-time productions. Thus, static MRI have to be complemented with other methods, such as real-time MRI, to correctly mimic articulatory movements.

This paper is organized as follows: section 2 details methods of MR image acquisition and describes the tools used for data analysis; section 3 provides some results; finally in section 4 we briefly discuss our results and summarize directions for future work.

2. Methods

2.1. Speakers and corpus

Three native speakers of European Portuguese (two men - AND, JHM - and a woman - RAQ) aged between 22 and 36, participated in the static MRI study. For the RT-MRI experiment, three EP female speakers (SV, CM, CO), aged between 21 and 33, were recorded. None of the speakers reported having hearing or speech disorders. All had some knowledge in the area of Linguistics and Phonetics, and the speaker AND also had some singing training. All speakers come from the centre of Portugal. An MRI screening form and informed consent was obtained before the study to comply with security and ethics rules.

The static MRI corpus considered in this study represents a subset of a large database collected for the study of the EP sounds (refer to [9] for more details). In this paper only EP nasal and oral vowels were considered. The target sounds were sustained during the period of image acquisition, no sound recordings were performed. Each item was produced only one time, a target word was prompted orally to instruct the speaker (e.g say [ɐ̃] as in [kɐ̃pu]).

The speakers engaged in the RT-MRI experiment produced nonsense words containing EP nasal ([ɐ̃], [ẽ], [ĩ], [õ], [ũ]) and oral ([ɛ], [e], [i], [a], [ɐ], [ɔ], [o], [u]) vowels. Nasal vowels were uttered in three prosodic conditions: word-initial, word-internal and word-final (e.g. [ɐ̃pɐ], [pɐ̃pɐ], [pɐ̃]). The oral vowels appeared in CV_1CV_2 sequences (e.g. papa [papɐ], pupa [pupɐ]), where C is a voiceless bilabial plosive and V1 is the target vowel. A second stimulus set included the five nasal monophthongs and the seven oral vowels to be pronounced in isolation by the speakers. Speakers produced 3-4 repetitions of each vowel, depending on the pre-defined duration of the scan.

2.2. MRI Data collection

2D Static MRI – MRI images were collected using a 1.5 Tesla (Magneton Simphony, Maestro Class, Siemens, Erlanger, Germany) scanner equipped with Quantum gradients. Neck and brain phased array coils were used. The 2D corpus was acquired in the midsagittal plane using a TSE T1 Weighted sequence (Slice thickness = 5 mm, FOV = 200 mm, ETL = 15). The acquisition time was 5,6 seconds. The MRI protocol used to acquire the images has already been described in detail in Martins et al. [9].

RT-MRI – The RT- MRI experiment was conducted at IBILI/Coimbra. The images were acquired on an unmodified 3.0 T MR scanner (Magneton Tim Trio, Siemens, Erlanger, Germany) equipped with high performance gradients (Gmax = *45mT/m*, rise time = 0.2*s*, slew Rate = 200 *T/m/s*, and FOV = 50 *cm*). A 12-channel head and 4-channel neck phased-array coils were used for data acquisition. Parallel imaging (GRAPPA 2) and magnetic field gradients operating at FAST mode were used to speed up the acquisition. After localization images, a T1 W 2D-midsagittal MRI slice of the vocal tract was obtained, using an Ultra-Fast RF-spoiled Gradient Echo (GE) pulse sequence (Single-Shot TurboFLASH), with a slice thickness of 8 *mm* and the following parameters: TR/TE/FA = 72*ms*/1.02*ms*/5°, Bandwidth = 1395 *Hz/pixel*, FOV(mm^2)= 210 x 210, reconstruction matrix of (128 x 128) elements with 50 % phase resolution, in-plane resolution (mm^2) = 3,3 x 1,6, yielding a frame rate of 14 images/second. The acquisition of each sequence took about 5 seconds, resulting in 75 midsagittal images. Each stimulus set was prompted orally by one of the experimenters over the intercom system (e.g., "Say [ɐ̃pɐ, pɐ̃pɐ, pɐ̃]"). Audio was recorded simultaneously with the RT-images inside the MR scanner, at a sampling rate of 16000 Hz, using a fiberoptic microphone. A TTL pulse generated from the MRI scanner allowed the synchronization between MRI images and audio recordings. For both acquisitions, the subjects lay supine in the MR scanner, while producing the stimuli and wore headphones to protect the ears from the noise (see [15]).

3. Image Segmentation and Analysis

All the 2D static images were segmented using a semi-automatic technique (live-wire), based on the deformable models framework , allowing the extraction of the vocal tract outlines [16].

A framework for segmentation of the different RT-MRI image series was developed using MeVisLab (`http://www.mevislab.de`). Segmentation starts with the definition of a region of interest (ROI) roughly encompassing the vocal tract (refer to figure 1 for an outline of the processing pipeline). This ROI just needs to be defined over one of the image frames (and is then replicated over the entire image series). Its main purpose is to avoid the segmentation to go beyond the lips and the nasal cavity and above the hard palate. A region growing algorithm is then applied using a point inside the vocal tract as seed. Due to spatial coherence between the different images, a single seed allows segmentation of the vocal tract for all image frames. The result of the region growing can be adjusted by changing the used intensity interval (or adding more seeds).

Lip segmentation often requires different (narrower) intensity intervals. For improved lip segmentation a separate ROI is defined for the lips and region growing is applied to it using different parameters.

Since the speakers did not move during the image acquisition session, the regions of interest for vocal tract and lips are defined only once for each speaker and then re-used throughout. Furthermore, seed position and intensity interval for the region growing are roughly the same for all image series. Segmentation editing is also possible, e.g., to add regions to the vocal tract by defining a contour around it using live wire.

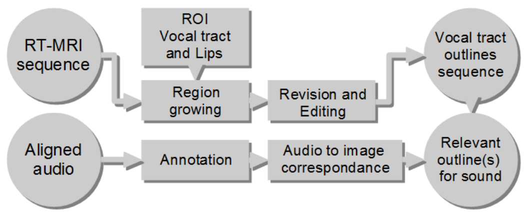


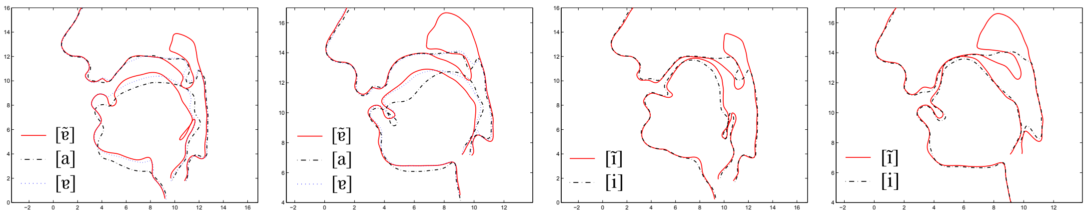
Figure 1: Image processing and analysis pipeline.

The segmentation of a time series (75 image frames) took one to five minutes depending if corrections were needed.

The audio corresponding to each RT-MRI sequence was manually annotated identifying the position of different sounds. Based on this annotation the image frames associated with each vowel were identified.

For comparison of oral and nasal vowels the relevant outline inside the corresponding image frame interval needed to be chosen. As stated by different authors, e.g., [12, 13], the course of nasality in Portuguese is typically incremental, i.e. nasal vowels have an initial oral segment with medial-final nasality. Furthermore, a nasal tail may sometimes emerge after the vowel. Those different phases can be clearly distinguished in Figure 2. Thus, for nasal vowels, the frame where both the velum was clearly lowered and the lips opened was chosen, so as to enable oral/nasal comparisons. For the oral vowels the centre image frame was used.

In each case, only one of the repetitions available in the corpus was selected, since the superimposition of the outlines of the vocal tract during the production of the different vowel repetitions by each speaker proved to be very consistent.



(a) AND (b) RAQ (c) AND (d) RAQ

Figure 3: Superimposed outlines of the vocal tract profiles of speakers AND and RAQ during the production of EP oral and nasal vowels [ɐ̃]-[a]-[ɐ] (left) and [ĩ]-[i] (right).

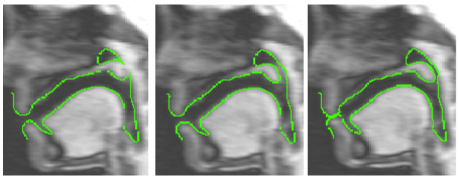


Figure 2: RT-MRI midsagittal images and vocal tract outlines of speaker SV during the production of nasal vowels [ɐ̃] (word [pɐ̃pɐ])

The selected outlines for the nasal and oral vowels were then superimposed for comparison. To study the evolution of the vocal tract along time, for nasal vowels, all corresponding image frames were superimposed.

4. Results

In what follows a summary of the main findings for the analysis of oral/nasal configurations using vocal tract outlines derived from static and RT-MRI data are presented.

4.1. 2D static profiles

In general, superimposed outlines of the vocal tract profiles of three speakers (AND, RAQ, JHM) show small differences in the position of the oral articulators between phonemic nasal vowels and their oral congeners, in addition to the lowering of the velum.

[a]-[ɐ]-[ɐ̃] – As shown in Figure 3, the nasal vowel [ɐ̃] is produced by all three speakers with a significantly more fronted tongue body than oral [ɐ] and [a], with a direct effect in the size of the pharyngeal airway. For two of the speakers (AND, RAQ), the tongue is also raised in [ɐ̃]. Nevertheless, tongue and labial configuration of the nasal vowel is closer to the oral [ɐ] than [a].

[ɔ]-[o]-[õ] – Two of the three speakers (JHM, RAQ) produced a somewhat fronted and lowered [õ] with respect to the oral counterparts, though effects in pharynx were insignificant. Moreover, lingual adjustments were more evident for the male speakers. Results showed also lesser lip protrusion in the realization of [õ].

[ũ]-[u] – As regards this oral-nasal pair, tongue position was quite the same for JHM, though AND and RAQ produced a slightly lowered [ũ] with respect to oral [u].

[ẽ]-[e]-[ɛ] – The articulatory profiles of [ẽ] and [e] were almost indistinct, for all three speakers, except for the lowering of the velum. As expected, [ẽ] and [ɛ] differ essentially in vowel height.

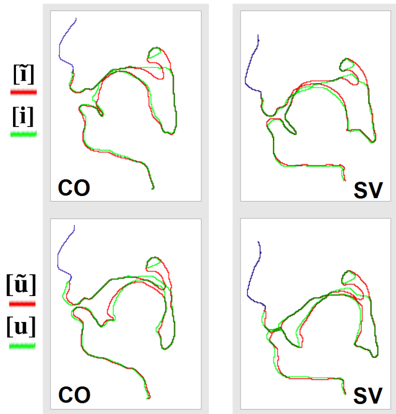


Figure 4: Superimposed outlines of the vocal tract profiles (RT-MRI) of speakers SV and CO during the production of EP oral and nasal vowels [ĩ]-[i] and [ũ]-[u].

[ĩ]-[i] – Lingual configuration is very similar for oral and nasal vowel, as shown in Figure 3, although the tongue was slightly raised for RAQ and somewhat retracted for AND.

4.2. RT-MRI profiles

RT-MRI provides a full midsagittal view of the moving vocal tract, with reasonable temporal resolution, which makes it suitable to examine not only velum lowering over the production of the vowel, but also possible adjustments of the tongue and lips.

Oral-nasal vowel comparison – Comparisons between oral vowels and nasal congeners produced by the three speakers in word-internal environment (e.g. [papɐ] versus [pɐ̃pɐ]) suggest lingual and labial differences even less pronounced than those reported above. All the speakers produced the oral/nasal pairs [ũ]-[u] and [ĩ]-[i] in a very similar way, except for the position of the velum (see Figure 4). In comparison with oral congeners, vowels [ɐ̃], [õ] and [ẽ] denote oral variations slightly more accentuated, as shown in Figure 5.

Nasal vowels over time – Superimposed outlines obtained along the production of the same nasal vowel in [p] environment showed a gradual lowering of the velum and only tiny variations in tongue position. However, for the nasals produced in isolation, differences on the tongue posture along the vowel were somewhat more noticeable, at least for [ɐ̃], [õ] and [ẽ]. As shown in Figure 6 for one of the speakers, the onset of the nasal vowel is oral, as regards both velum position and oral configuration. Gradually, tongue raises and moves forward, while velum lowers (cf. Figure 5).

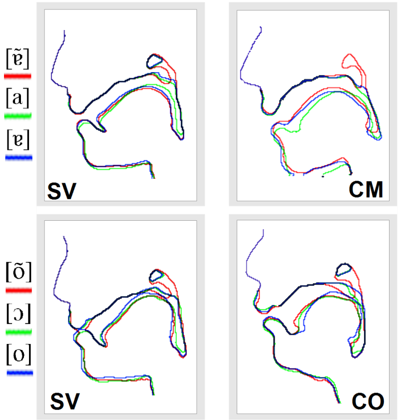


Figure 5: Superimposed outlines of the vocal tract profiles (RT-MRI) of speakers SV and CM (top) and SV and CO (bottom) during the production of vowels [a]-[ɐ]-[ɐ̃] and [ɔ]-[o]-[õ].

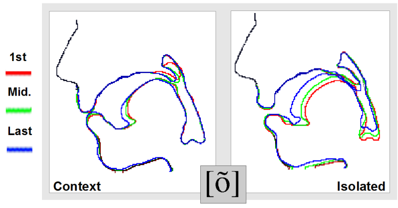


Figure 6: Superimposed outlines of the vocal tract profiles of speaker CM during the production of nasal vowel [õ] in the word [põpɐ] and in isolation.

5. Discussion and Conclusions

The results from this imaging study suggest that lingual and labial differences between nasal vowels and their oral counterparts are quite subtle in EP. Despite some inter-speaker variability in the position of oral articulators, that was observed in both static and dynamic data, nasal vowels [ɐ̃], [õ] (and also [ẽ] in dynamic profiles) exhibited more articulatory adjustments with respect to oral congeners than [ĩ] and [ũ]. The anteriority of [ã] with respect to [a] have been pointed for BP [17, 7]. These results are also consistent with 3D data [9], as the comparison of speaker area functions did not show evidence of substantial changes in the oral articulation between nasal and oral vowel pairs.

Whenever lingual differences between nasal vowels and oral congeners were observed, those were more visible in sustained than in real-time productions. Furthermore, for RT- MRI these lingual changes were more clearly observed in nasal vowels produced in isolation than in word environment, as they are longer in the former context. Therefore, we speculate that RT-MRI temporal resolution might be insufficient to fully capture tongue movements.

Acquisition and analysis of acoustic signals of these particular speakers would be important to fully understand the acoustic and articulatory characteristics of vowel nasality and whether the slight oral adjustments observed imply acoustic changes.

In contrast with previous findings for languages with phonemic nasal vowels (e.g. French) [1, 2, 5, 8], results of the present study do not provide evidence that in EP additional articulatory modifications support the lowering of the velum to convey the nasal contrast. Since oral and nasal vowels are produced with no significant articulatory differences, but the naso-pharyngeal coupling, nasality may be signalled through other traces, possibly the emergence of a final nasal tail.

6. Acknowledgements

This research was partially funded by FEDER through the Program COMPETE and by National Funds (FCT- Foundation for Science and Technology) in the context of HERON II (PTDC/EEA-PLP/098298/2008) and IEETA Research Unit funding FCOMP-01-0124-FEDER-022682 (FCT-PEst-C/EEI/UI0127/2011). The second author acknowledges PhD grant (FCT-SFRH/BD/65183/2009). The authors also thank IBILI Director.

7. References

[1] O. Engwall, V. Delvaux, and T. Metens, “Interspeaker variation in the articulation of French nasal vowels,” in *7th International Seminar on Speech Production*, Ubatuba, Brazil, 2006.

[2] V. Delvaux, T. Metens, and A. Soquet, “Propriétés acoustiques e articulatoires des voyelles nasales du français,” in *Journées d’Étude sur la Parole*, Nancy, 2002, pp. 348–352.

[3] J. P. Zerling, “Phénomenes de nasalité et de nasalization vocaliques: Étude cinéradiographique pour deux locuteurs,” *Travaux de l’Institut de Phonétique de Strasbourg*, vol. 16, pp. 241–266, 1984.

[4] S. Maeda, “Acoustics of vowel nasalization and articulatory shifts in French nasal vowels,” in *Nasals, Nasalization, and the Velum*, M. Huffman and R. Krakow, Eds. Academic Press, 1993, pp. 147–167.

[5] C. Carignan, “Oral articulation of nasal vowels in french,” in *Proceedings of the XVII International Congress of Phonetic Sciences*, Hong Kong, 2011.

[6] M. Matta Machado, “Fenômenos de Nasalização Vocálica em Português. Estudo Cine-radiográfico,” *Caderno de Estudos Linguísticos*, vol. 25, pp. 113–117, 1993.

[7] F. N. Gregio, “Configuração do trato vocal supraglótico na produção das vogais do português brasileiro: dados de imagens de ressonância magnética,” Master thesis, PUC/SP, Brasil, 2006.

[8] R. Shosted, C. Carignan, and P. Rong, “Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi,” *JASA*, vol. 131, no. 1, pp. 455–465, 2012.

[9] P. Martins, I. Carbone, A. Pinto, A. Silva, and A. Teixeira, “European Portuguese MRI based speech production studies,” *Speech Communication*, vol. 50, no. 11-12, pp. 925–952, 2008.

[10] C. Oliveira, P. Martins, and A. Teixeira, “Speech rate effects on European Portuguese nasal vowels,” in *InterSpeech*, Brighton, 2009.

[11] S. Master, P. Pontes, and M. Behlau, “Configuração do trato vocal na emissão das vogais nasais do português brasileiro,” *Acta AWHO*, vol. 10, no. 2, pp. 67–71, 1991.

[12] A. Lacerda and B. Head, “Análise de sons nasais e sons nasalizados do português,” *Revista do Laboratório de Fonética Experimental de Coimbra*, vol. 6, pp. 5–70, 1966.

[13] R. Sampson, *Nasal Vowel Evolution in Romance*. Oxford University Press, 1999.

[14] O. Engwall, “A revisit to the application of MRI to the analysis of speech production - testing our assumptions,” in *International Seminar on Speech Production*, Sydney, 2003.

[15] A. Teixeira, P. Martins, C. Oliveira, C. Ferreira, A. Silva, and R. Shosted, “Real-Time MRI for Portuguese. Database, methods and applications,” in *PROPOR*, Coimbra, 2012.

[16] A. Teixeira, P. Martins, C. Oliveira, and A. Silva, “An MRI study of consonantal coarticulation resistance in Portuguese,” in *International Seminar on Speech Production*, Montreal, 2011.

[17] J. M. Câmara, *Estrutura da Língua Portuguesa*, 4th ed. Petrópolis: Editora Vozes Limitada, 1973.