INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

On the role of oral configurations in European Portuguese nasal vowels

Conceição Cunha[1], *Samuel Silva*[2] *António Teixeira*[2], *Catarina Oliveira*[3], *Paula Martins*[3], *Arun Joseph*[4], *Jens Frahm*[4]

[1]IPS, LMU Munich, Germany
[2]DETI/IEETA, Universidade de Aveiro, Portugal
[3]Health School/IEETA, Universidade de Aveiro, Portugal
[4]Max-Planck-Institut für Biophysikalische Chemie, Germany

cunha@phonetik.uni-muenchen.de, sss@ua.pt

Abstract

The characterisation of nasal vowels is not only a question of studying velar aperture. Recent work shows that oropharyngeal articulatory adjustments enhance the acoustics of nasal coupling or, at least, magnify differences between oral/nasal vowel congeners. Despite preliminary studies on the oral configurations of nasal vowels, for European Portuguese, a quantitative analysis is missing, particularly one to be applied systematically to a desirably large number of speakers. The main objective of this study is to adapt and extend previous methodological advances for the analysis of MRI data to further investigate: how velar changes affect oral configurations; the changes to the articulators and constrictions when compared with oral counterparts; and the closest oral counterpart. High framerate RT-MRI images (50fps) are automatically processed to extract the vocal tract contours and the position/configuration for the different articulators. These data are processed by evolving a quantitative articulatory analysis framework, previously proposed by the authors, extended to include information regarding constrictions (degree and place) and nasal port. For this study, while the analysis of data for more speakers is ongoing, we considered a set of two EP native speakers and addressed the study of oral and nasal vowels mainly in the context of stop consonants.

Index Terms: Nasal vowels, real-time MRI, image processing, quantitative methods

1. Introduction

The indirect estimation of the vocal tract configuration from the acoustic characteristics is difficult to infer for nasal vowels due the acoustic effects of the velopharyngeal coupling. Given such complexity, articulatory properties have been studied over the last decades using more desirable methods of assessing the shape of the entire vocal tract during the production of nasal vowels. These instrumental techniques, such as MRI, allow direct comparison of vowel tract's shape between nasals and their oral congeners. A growing number of articulatory studies on languages with phonemic nasal vowels argued for a set of articulatory adjustments on lingual and lips configuration. In French, pairs of nasal/oral vowels showed differences in tongue height, labial aperture and pharyngeal constriction in addition to velopharyngeal coupling [1, 2, 3, 4, 5]. Data collected by Shosted et al. [6] reveal lingual adjustments in the production of Hindi nasal vowels comparing to their oral congeners. Using cineradiographic data for Brazilian Portuguese (BP), Matta Machado [7] found evidence of tongue body raising during the nasal vowel [6~] and [u~] with respect to their oral congeners [1]. A recent MRI study [8] showed significant differences in tongue fronting, height, and shape between oral /a, e, u/ and their nasal counterparts /a~/, /e~/, and /u~/ in this variety.

European Portuguese has a vowel inventory containing five nasal vowels ([6~], [e~], [i~], [u~], [o~). The articulation of EP nasal vowels has been studied in some detail focusing mainly on the dynamic pattern of the velum as provided by electromagnetic articulography (EMA) [9, 10, 11]. More recently, velar movement [12] and the lingual configuration of EP oral and nasal vowels was investigated using both static and real-time magnetic resonance imaging [13, 14]. EP nasal vowels and their oral counterparts have been shown to manifest subtle oropharyngeal distinctions. Nasal vowels [6~] and [o~] exhibited more articulatory adjustments with respect to oral congeners than the high vowels [i~] and [u~].

The line of research motivating this study aims to propose and perform a comprehensive and quantitative analysis of the oropharyngeal configuration of EP vowels in order to clarify whether oral articulatory adjustments support or not the lowering of the velum to convey the [nasal] contrast in EP. Following this purpose, this study extends previous systematic methods applied to RT-MRI sequences of the vocal tract to compute changes on the oral configurations of nasal vowels after velopharyngeal coupling. This will improve our knowledge regarding: (1) the existence and role of oropharyngeal adjustments; (2) the extent of how velar changes affect oral configurations of nasal vowels; (3) the position of the articulators for nasal vowels when compared with their oral counterparts; and (4) the closest oral counterpart for each nasal vowel in terms of oral configuration.

2. Method

High sample rate RT MRI images (50fps) [15] are subject of automatic processing in order to extract vocal tract contours, position and configuration for the different articulators.

2.1. MRI acquisition

RT-MRI recordings were conducted at the Max Planck Institute for Biophysical Chemistry, Göttingen, Germany, using a 3 T Siemens Magnetom Prisma Fit MRI System equipped with high performance gradients (Max ampl=80 mT/m; slew rate = 200 T/m/s). A 64-channel head coil was used. The speaker was lying down and was instructed to read the required sentences. RT-MRI measurements were based on a recently de-

[1]The Speech Assessment Methods Phonetic Alphabet (SAMPA) is adopted throughout this article.

http://dx.doi.org/10.21437/Interspeech.2019-2232

veloped method, where highly under-sampled radial FLASH acquisitions are combined with nonlinear inverse reconstruction (NLINV) providing images at high spatial and temporal resolutions [16]. Acquisitions were made at 50 fps. Speech was synchronously recorded using an optical microphone (Dual Channel-FOMRI, Optoacoustics, Or Yehuda, Israel), fixed on the head coil, placed directly against the speaker's mouth. All volunteers provided informed written consent and filled an MRI screening form. None of the participants had any known language, speech or hearing problems. They were compensated for their participation.

2.2. Corpus

The analysed corpus consists of minimal pairs containing all oral [i, e, E, a, O, o, u, 6] and nasal vowels [6~, e~, i~, o~, u~] in the following words 'pote' 'ponte' 'pôde' 'pato' 'panto' 'patinha 'pito' 'pinto' 'peta' 'pêta' 'pente' 'puto' 'punto'. All words were randomized and repeated in two prosodic conditions embedded in one of three carrier sentences alternating the verb as follows: (diga ('Say'); ouvi ('I heard'); leio ('I read')) as in 'Diga **pote**, diga **pote** baixinho' ('Say **pot**, Say **pot** gently'). The data considered for the analysis presented in this article (two males, 31 and 50 years old, central region of Portugal) is part of a larger experiment. So far, this corpus has been recorded from 16 native speakers (8 male and 8 female) of EP.

2.3. Vocal Tract Data Processing and Analysis

The vocal tract outlines were extracted adopting the method proposed by Silva et al. [17], resulting in contours identifying the different regions of interest, as depicted in Figure 1.

The comparison among vocal tract configurations was performed adopting and extending a previously proposed framework (for the sake of brevity, additional details can be found in [18, 17, 19]) enabling normalized quantification of differences, for different articulators/regions of the vocal tract and their visualization (in a diagram as illustrated in Fig. 1: velum (VEL), tongue dorsum (TD), tongue back (TB), tongue tip (TT), lip protrusion (LP), lip aperture (LA) and pharynx (Ph). For each, the comparison yields a score, from 1 (no difference) to 0 (strong difference), which is represented over the unitary circle (Fig. 1). The green circular corona highlights the values above 0.75 deemed as "no major difference" [18].

Additionally, the work presented here also considered constrictions, characterized by location (CL) and sagittal distance (CD) at three regions: tongue tip (TTCD and TTCL), tongue dorsum (TDCD and TDCL), velum (VELCD, VELCL), and tongue back (TBCD and TDCL). Illustrative examples of outputs are presented in Fig. 2.

3. Results

Due to space limitations we will focus on the CD and CL, starting by a general analysis of the relevant differences in articulators and constrictions.

3.1. Articulation and constrictions

The first question addressed is the comparison of articulators and constrictions of nasal vowels with the oral congeners. Representative plots are presented in Fig. 2.

Table 1 shows the average differences in articulators and constrictions (location and distance) for nasals and their oral counterparts, for both speakers. Relevant differences are highlighted in gray. Major differences when comparing oral and nasal vowels were observed for TT and LA. Tongue backs differ mainly for the [6~]/[6] pair and no relevant differences were reported for LP and Ph.

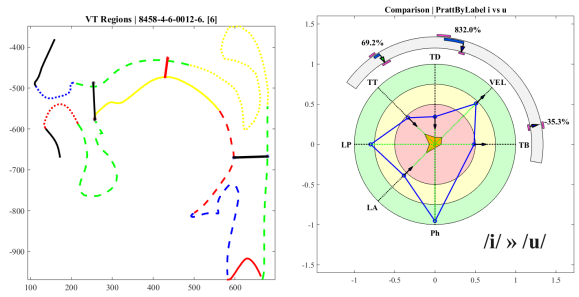

Figure 1: *On the left, vocal tract contour illustrating the different regions identified along with the considered constriction distances; on the right, a representation of the difference found between /i/ and /u/ for the different tract regions (as proposed in [18]).*

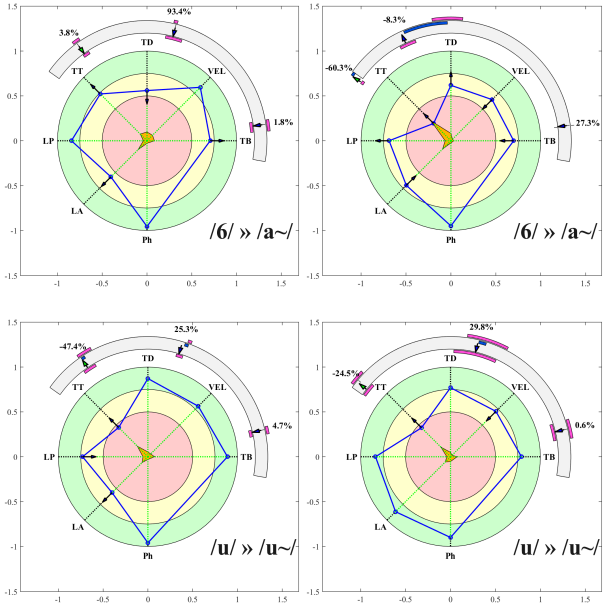

Figure 2: *Some representative results for [6] vs [6~] and [u] vs [u~]. Each column presents results for one of the speakers.*

Not considering TT, mostly affected by coarticulation effects, region and velum, the differences observed between nasal/oral vowels are as follows:

- [6~] differs both from [6] and [a]. It has a less constricted passage at tongue dorsum than [6] and more constricted than [a]; is less constricted than both [6] and [a] in pharyngeal region, causing an increase in TBCD;
- [e~] is very similar to [e] and differs from [E] in TDCD. The nasal vowel presents a TDCD higher than [E] which means a lowered tongue dorsum;
- [i~] also only differs from [i] in TDCD, having a more raised tongue dorsum;
- [o~] differs from [o] by a higher TDCD and from [O] by

Table 1: *Average differences between nasal vowels and their oral counterparts, using data from the 2 speakers. Relevant differences are marked with gray background.*

a) Relative changes in articulators (differences considered relevant when smaller than 0.75 (see [18]):

	Nasal V	Oral V	ΔTB	ΔVEL	ΔTD	ΔTT	ΔLP	ΔLA	ΔPh
1	6˜	6	0.74	0.80	0.68	0.57	0.82	0.70	0.95
2	6˜	a	0.75	0.80	0.68	0.50	0.79	0.59	0.97
3	e˜	e	0.81	0.80	0.94	0.77	0.82	0.86	0.97
4	e˜	E	0.82	0.83	0.75	0.65	0.86	0.77	0.98
5	i˜	i	0.90	0.76	0.93	0.76	0.84	0.79	0.97
6	o˜	o	0.90	0.73	0.74	0.43	0.83	0.64	0.96
7	o˜	O	0.90	0.75	0.77	0.50	0.82	0.68	0.96
8	u˜	u	0.84	0.76	0.82	0.46	0.78	0.72	0.93

b) Constrictions results: distances (CD) and locations (CL):
(relevant differences considered when $p \leq 0.001$ in one-sample Student's statistical test with null hypothesis $H_0 : \mu = 0$)

	Nasal V	Oral V	ΔTTCD	ΔTTCL	ΔTDCD	ΔTDCL	ΔTBCD	ΔTBCL	ΔVELCD	ΔVELCL
1	6~	6	-19.55	0.61	20.76	4.34	10.12	0	2.00	2.11
2	6~	a	-30.62	1.29	-14.21	18.07	22.90	0	-2.73	16.36
3	e~	e	-1.61	1.03	1.23	1.67	-4.75	0	-0.37	-34.27
4	e~	E	-1.57	0.52	-18.59	3.17	5.18	0	-1.18	-27.15
5	i~	i	-8.22	0.06	-7.20	-0.97	2.17	0	18.94	-51.43
6	o~	o	-45.84	1.84	21.56	-0.74	1.17	0	13.95	-0.48
7	o~	O	-36.94	0.50	-5.98	7.08	9.39	0	-2.66	12.15
8	u~	u	-34.34	1.38	14.88	3.62	1.63	0	2.87	33.10

higher TBCD, not being clear which one is the closest oral vowel;

- [u~] differs from [u] by having a higher TDCD, corresponding to a lowered tongue dorsum.

3.2. Closest oral counterpart

In order to investigate which oral vowel corresponds to the nasal counterpart, we compared [6~], [o~] and [e~] with their two candidate oral counterparts [6,a], [o, O] and [e, E] using all the parameters obtained from the pipeline related to constrictions (Table 1.b). Several statistics were calculated (sum of the absolute value of the differences; maximum value of the differences; and mean of the absolute values). The obtained values were summarised in Table 2. Values corresponding to the closest oral based on the metric are highlighted with gray background.

While [6] clearly presents the closest values to [6~], with the exception of the maximum for speaker 2. For [e~], the results are speaker dependent. For [o~] the closest is [o] as the mean is smaller, for both speakers.

3.3. Pharyngeal adjustments

From Table 1, it is clear that Ph does not seem to play a role. This can be confirmed in Fig. 3, presenting the observed differences for TBCL and TBCD. Constriction location is not affected and distance differences appear mainly for the pair [6~]/[6]. Even though [6] is the closest to [6~], the CD has a higher value, for the nasal vowel, meaning a less retracted tongue. Additionally, there is a significant difference of TBCD between [O] and [o~], showing [o~] yields an higher constriction distance and, hence, a less retracted TB.

3.4. Effects for the velar region

Fig. 4 shows the distribution of values for VELCL and VELCD, as function of nasal vowel and speaker (top) and as function of

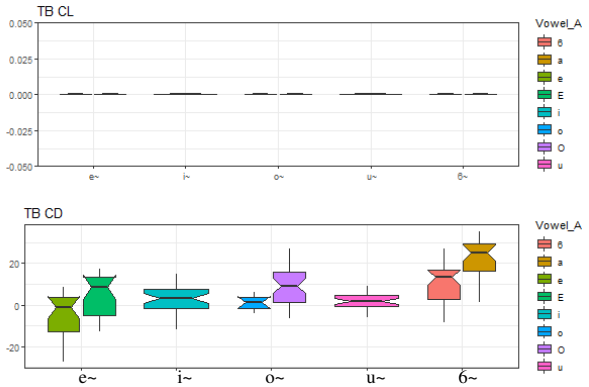

Figure 3: *Differences between nasal vowels and oral counterparts in TB constriction. The CL and CD are presented as function of the oral vowel.*

the oral vowel serving as basis of comparison (bottom plots). In general, the results show the same trend for both speakers, with statistical differences only for CD in [o~] and CL in [o~] and [i~]. Overall, differences in CD are close to zero and, therefore, are not relevant. The only exception is [i~] with significant differences for both speakers (second plot from the top, in Figure 4). Regarding CL, there are several significant difference from 0, with decreases for [e~] and [i~] and increases for [6~] and [u~].

4. Discussion

The results from this study reveal significant adjustments mainly in tongue height, between oral vowels and their nasal counterparts, in EP. These findings could confirm previous research using both static and dynamic MRI data [14, 13], showing that nasal vowels [6~], [o~] exhibit more articulatory ad-

Table 2: *Values for several metrics based in the differences in constriction location and distance: sum of the absolute value of the differences in the four location and constriction pairs; maximum value of the differences; and mean of the absolute values. Values are presented separately, for each speaker, and as the combined results, for both speakers.*

	Speaker 1			Speaker 2			Both Speakers		
	Sum(abs)	Max	Mean	Sum(abs)	Max	Mean	Sum(abs)	Max	Mean
6	64.2	31.0	8.0	60.7	30.0	7.6	124.9	31.0	7.8
a	215.6	55.1	27.0	68.7	20.8	8.6	284.3	55.1	17.8
e	96.1	46.3	12.0	42.6	28.2	5.3	138.7	46.3	8.7
E	54.9	19.8	6.9	71.6	34.4	9.0	126.5	34.4	7.9
i	100.3	62.9	12.5	84.1	43.8	10.5	184.4	62.9	11.5
O	195.2	72.5	24.4	85.6	45.8	10.7	280.8	72.5	17.6
o	138.7	74.0	17.3	58.0	28.1	7.2	196.7	74.0	12.3
u	85.3	34.3	10.7	98.4	45.7	12.3	183.7	45.7	11.5

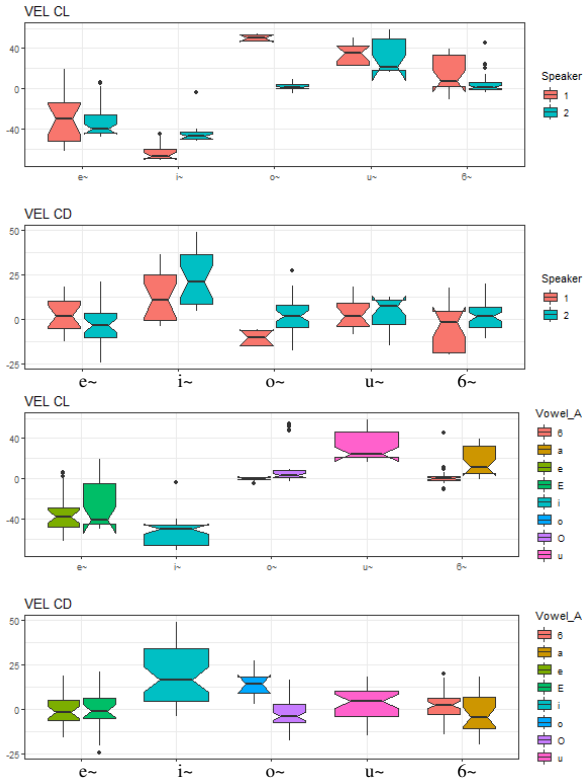

Figure 4: *CD and CL for the velar region: effect of vowel, for the 2 speakers*

justments with respect to their oral congeners than [e~], [i~], and [u~]. Comparisons between /6~/ and /6/–/a/ showed more substantial oropharyngeal adjustments than in other oral/nasal vowels pairs. The nasal vowel displayed an intermediate tongue blade position between the oral vowels /a/ and /6/ and a wider pharyngeal region compared to both oral vowels. Fronting and raising of the tongue for [a~] with respect to [a] have been also pointed out for BP [8] and seems to suggest that the distinction between these vowels implies oropharyngeal articulation after velar lowering [8]. Nasal vowel /u~/ showed a lower tongue dorsum position compared to /u/, as with the pair /o/–/o~/. In the latter, the nasal vowel tends to show a tongue dorsum position similar to [O]. This result for back nasal vowels confirms previous descriptions for BP [8], and is likely due to spatial limitation in the oral cavity, i.e. tongue dorsum moves forward to avoid epiphenomenal contact between lowered velum and high tongue dorsum. For /i/–/i~/, a slightly higher tongue dorsum position for the nasal vowel was observed compared to /i/, while for /e/–/e~/, the tongue configurations were very similar.

In contrast with previous analysis for languages with phonemic nasal vowels (e.g., French) [2, 1, 4], the results of the present study do not provide evidence that nasalisation in EP is associated with pharyngeal variations in addition to velar lowering (except for the pair /6/–/6~/), suggesting that TD may be responsible for maintaining or enhancing the phonemic oral vowel distinction.

5. Conclusion

In this paper, we propose a new quantitative method to systematically evaluate articulatory information from high sample rate RT-MRI. Specifically, the method was applied to the study of oropharyngeal configurations of EP nasal vowels and allowed a direct articulatory comparison between nasal vowels and their oral congeners. Overall, the proposed method improved the description of nasal vowels, and allowed new insights into oral adjustments in nasal vowels, relevant in the field of phonetics and articulatory phonology. Furthermore, it is able to deal with the increasing amount of data provided by recent MRI techniques and it allows the analysis of data obtained for a high number of speakers.

Future work will include the processing and analyses of the full database to address this and other relevant research questions, namely the dynamic pattern of nasal vowel production and gestural timing. Improvements in the visualisation and introduction of other metrics are also of great interest to develop the proposed framework.

6. Acknowledgements

This work is partially funded by the German Federal Ministry of Education and Research (BMBF, Project 'Lautwandel', 01UL1712X), by IEETA Research Unit funding (UID/CEC/00127/2013), by Portugal 2020 under the Competitiveness and Internationalization Operational Program, and the European Regional Development Fund through project SOCA – Smart Open Campus (CENTRO-01-0145-FEDER-000010) and project MEMNON (POCI-01-0145-FEDER-028976). We thank Philip Hoole for the scripts for noise supression and all the participants of the experiment for their time and voice.

7. References

[1] V. Delvaux, T. Metens, and A. Soquet, "French nasal vowels: acoustic and articulatory properties," in *ICSLP2002 - INTERSPEECH 2002*, 2002.

[2] O. Engwall, V. Delvaux, and T. Metens, "Interspeaker variation in the articulation of french nasal vowels," in *Proceedings of the Seventh International Seminar on Speech Production*, 2006.

[3] C. Carignan, R. K. Shosted, C. Shih, and P. Rong, "Articulatory compensation for nasality: An ema study of lingual position during nasalized vowels," *J. Phonetics*, vol. 39, pp. 668–682, 2011.

[4] C. Carignan, "An acoustic and articulatory examination of the 'oral' in 'nasal': The oral articulations of french nasal vowels are not arbitrary," *J. Phonetics*, vol. 46, pp. 23–33, 2014.

[5] C. Carignan, R. K.Shosted, M. Fu, Z.-P. Liang, and B. P.Sutton, "A real-time mri investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of french," *J. Phonetics*, vol. 50, pp. 34–51, 2015.

[6] R. Shosted, "Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from hindi," *J Acoust Soc Am*, vol. 131, no. 1, pp. 455–465, 2012.

[7] M. Da Matta Machado, "Fenômenos de nasalização vocálica em português: estudo cineradiográfico," *Cadernos de estudos linguisticos*, vol. 25, pp. 113–127, 1993.

[8] M. Barlaz, R. K. Shosted, M. Fu, and B. Sutton, "Oropharyngeal articulation of phonemic and phonetic nasalization in brazilian portuguese," *J Phon*, vol. 71, pp. 81–97, 2018.

[9] C. Oliveira and A. Teixeira, "On gestures timing in european portuguese nasals," in *ICPhS*, 2007, pp. p. 405 – 408.

[10] C. Oliveira, "Do grafema ao gesto. contributos linguísticos para um sistema de síntese de base articulatória," Ph.D. dissertation, Universidade de Aveiro, 2009.

[11] C. Oliveira, A. Teixeira, and P. Martins, "Speech rate effects on european portuguese nasal vowels," in *INTERSPEECH 2009*, 2009.

[12] P. Martins, C. Oliveira, S. Silva, and A. Teixeira, "Velar movement in european portuguese nasal vowels," in *Proceedings of IberSPEECH 2012 - VII Jornadas en Tecnología del Habla and III Iberian SLTech Workshop, Madrid, Spain*, 2012.

[13] S. Silva, A. Teixeira, C. Oliveira, and P. Martins, "Segmentation and analysis of vocal tract from midsagittal real-time mri," in *International Conference Image Analysis and Recognition*. Springer, 2013, pp. 459–466.

[14] C. Oliveira, P. Martins, S. Silva, and A. Teixeira, "An MRI study of the oral articulation of european portuguese nasal vowels," in *INTERSPEECH 2012*, 2012.

[15] C. Cunha, S. Silva, A. Teixeira, C. Oliveira, P. Martins, A. Joseph, and J. Frahm, "Exploring Advances in Real-time MRI for Speech Production Studies of European Portuguese," in *iberSpeech*, Barcelona, Spain, November 2018.

[16] M. Uecker, S. Zhang, D. Voit, A. Karaus, K.-D. Merboldt, and J. Frahm, "Real-time mri at a resolution of 20 ms," *NMR in Biomedicine*, vol. 23, no. 8, pp. 986–994, 2010.

[17] S. Silva and A. Teixeira, "Quantitative systematic analysis of vocal tract data," *Computer Speech & Language*, vol. 36, pp. 307 – 329, 2016.

[18] ——, "Critical articulators identification from RT-MRI of the vocal tract," in *Proc. Interspeech*, Stockholm, Sweden, 2017, pp. 626–630.

[19] ——, "Unsupervised segmentation of the vocal tract from real-time MRI sequences," *Computer Speech and Language*, 2015, doi:10.1016/j.csl.2014.12.003.