LINGUAL DIFFERENCES IN BRAZILIAN PORTUGUESE ORAL AND NASAL VOWELS: AN MRI STUDY

Marissa Barlaz, Maojing Fu, Julianna Dubin, Zhi-Pei Liang, Ryan Shosted, Brad Sutton

University of Illinois at Urbana-Champaign
{goldrch2, mfu2, jdubin3, z-liang, rshosted, bsutton}@illinois.edu

ABSTRACT

It has been shown that oral and nasal vowel pairs can differ substantially in their oro-pharyngeal configuration (besides the position of the velum), an effect that may lead to the enhancement of certain acoustic features associated with nasalization. We use real-time MRI to compare the lingual configuration of oral and nasal vowels in Brazilian Portuguese at various timepoints during the production of these vowels. Results of area function comparisons and smoothing spline ANOVAs indicate that lingual articulation of nasal vowels increasingly moves towards the center of the oral cavity throughout the duration of the vowel, whereas the lingual articulation of their oral counterparts moves towards the outer extrema. We conclude that the lingual configurations in these targets are in line with the effects of vowel nasalization on formant frequencies. This further reaffirms the possibility that oral articulation in nasal vowels may be employed to enhance the acoustic effects of nasalization.

Keywords: rt-MRI, smoothing spline ANOVA, Brazilian Portuguese, nasal vowels.

1. INTRODUCTION

Articulatory studies increasingly suggest a connection between oro-pharyngeal articulation and the production of phonological nasality. The canonical view of nasal vowels involves lowering of the velum, which, generally speaking, lowers F1 for low vowels and raises F1 for high vowels. Recent work has shown that the oral configuration of nasal vowels may enhance the acoustics of nasalization or at least magnify differences between oral and nasal vowel congeners [21, 3, 25].

A study capturing simultaneous electromagnetic articulography (EMA), aerodynamic and acoustic data on Hindi oral and nasal vowels reveals lingual adjustments to tongue position for nasal vowels relative to their oral counterparts, some of which affect F1 in ways reminiscent of nasality [25]. EMA and MRI studies show that velic lowering coupled with tongue retraction and labialization are consistent characteristics of French nasal vowels. [2, 4, 3].

Brazilian Portuguese (BP) has a vowel inventory containing five oral-nasal pairs:/i-ĩ/, /e-ẽ/, /a-ɐ̃/, /o-õ/, and /u-ũ/ [1, 15]. Acoustic accounts show centralization of nasal vowels in comparison with orals (particularly the raising of the low vowel). The articulatory bases of the resulting observations have not been fully explored. [24] found evidence of tongue-body raising during the nasal vowel /ɐ̃/ and /ũ/ with respect to their oral congeners. [5] found evidence of reduced oral cavity volume in nasal vowels.

Current methodologies allow for non-invasive examination of articulators previously unavailable for viewing during production of natural speech. Real-time (rt-) MRI is presented in this study as a way to explore the lingual features of oral and nasal vowels in BP. rt-MRI has achieved near the limit of temporal and spatial resolution allowed through conventional approaches. Further advances require the integration of advanced sampling schemes and image reconstruction algorithms to enable visualization of soft tissue movement in near-real time.

Additionally, many rt-MRI studies focus on measurements such as average pixel intensity, as a way to identify the presence or absence of tissue within a given region. More advanced statistical methods can be applied to the shapes of articulators, allowing for more detailed understanding of articulation. Smoothing Spline ANOVA (SSANOVA) [13] has been used to compare lingual shapes in EMA and ultrasound data [6, 16, 19], as well as formant tracks [9], though it has been underutilized in MRI studies. The use of SSANOVA with rt-MRI imaging allows for comparison of the entire tongue contour in better spatial resolution than in many previous studies.

2. METHODS

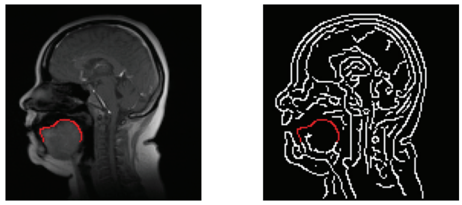
The subject is a female, native BP speaker in her mid-20s, from Fortaleza, Brazil. The data used here constitute part of a larger project on vocal tract imaging. A word list consisting of BP lexical items was used, with the target vowels preceded by obstruents. Test items include: *bafum* [b**afũ**] 'bad smell', *tupã* [t**u**pẽ] 'Tupi god', and *refém* [x**e**fẽ]'hostage'. The items were chosen to minimize the effects of coarticulation as best as possible, and to compare oral and nasal vowel counterparts. While the settings for these vowels do not perfectly control for coarticulation, the use of data from a single speaker controls for the possibility of inter-speaker variation. Target words were placed in the carrier phrase *diga X agora* 'say X now'. Phrases were presented in a randomized order in the 3T Siemens Trio MRI scanner at the Beckman Institute for Advanced Science and Technology, at the University of Illinois. The speaker was instructed to repeat the phrase at a normal rate until the noise of the scanner ceased (about 2 minutes). Due to different speaking rates between trials, an unequal number of tokens was collected for each lexical item: bafum (48), tupã (48), and refém (49).

In this work, we use a powerful approach, the partial separability (PS) model [17], to enable sub-Nyquist sampling by exploiting the inherent spatiotemporal correlations of the object being imaged [10, 11, 12]. This allows for good quality reconstruction at a high nominal frame rate - specifically, 100 frames per second for single slice imaging. A single mid-sagittal slice was used for this study. The image resolution of each slice is 128x128 voxels, and the resolution of each voxel is 2.2 mm x 2.2 mm x 6.5 mm (through-plane depth).

Reconstructed images were converted to black and white edge images in Matlab 2012a [18] using the Canny method of edge detection with the *edge* function at a 0.05 threshold. A region of interest (ROI) was selected around the oral cavity based on inspection of several images. The spectrogram was derived from synchronized, noise-cancelled audio recorded with a MR-compatible headset with an attached optical microphone worn by the subject (Dual Channel-FOMRI, Optoacoustics, Or Yehuda, Israel). The start and end point of each vowel were manually segmented in Praat. Five images $I_{1\text{-}5}$ were chosen at equidistant intervals throughout the vowel with their corresponding contours $C_{1\text{-}5}$. The tongue contours $C_{1\text{-}5}$ in the ROI were extracted using the *bwboundaries* function on the black and white images, as seen in Figure 1. A total of 1450 contours were analyzed.

For each image analyzed, we calculated area under the curve (AUC) as a measure of the amount of tissue within a region [26]. First, the ROI used to select tongue contours was split into two sections, corresponding to anterior and posterior regions of the oral cavity. Using static reference regions permits comparison of fronting and backing of the tongue through evaluation of tissue volume in each region.

Figure 1: Example of the tongue contour extracted for /u/ in the word *tupã*).



If the AUC is greater in the anterior half of the ROI, it indicates of more tissue in that region and therefore tongue fronting. Second, we split the horizontal range of the tongue into three segments, corresponding to tongue tip (TT), tongue middle (TM), and tongue dorsum (TD) regions. Comparing AUC in these three regions allows for comparison of height of these segments of the tongue.

Differences in AUC were compared using multivariate analysis of variance (MANOVA) from the *car* package [7] in R 3.0.2 [20]. All AUC measures were multiplied by 4.84, the squared in-plane voxel resolution. In each analysis, the dependent variables were AUC for each lingual region, and the independent variables were vowel nasality (oral vs. nasal), curve number (1–5), and their interaction.

In order to compare tongue contours, we used SSANOVA with the R package *gss* [14] to compare the splines that best fit the oral and nasal vowel data. Unlike conventional ANOVA, SSANOVA does not generate a test statistic. Rather, Bayesian confidence intervals are built around the splines fit to the data. If the confidence intervals do not overlap, the differences between the curves are considered significant [8].

3. RESULTS

3.1. Area Under the Curve

The results of a one-way MANOVA indicate that AUC differs significantly for the front and back regions of the ROI, based on nasality ($F(2,1439) = 585.88$, $p < 0.0001$, curve number ($F(8,2880) = 361.23$, $p < 0.0001$) and their interaction ($F(8,2880) = 12.83$, $p < 0.0001$). In order to determine the significance of individual factors, multiple one-way ANOVAs were conducted, and a Bonferroni correction was applied. The results hold that the area differences are significantly different ($p < 0.025$), in both the front region and back region of the ROI, as seen in Table 1.

Further individual analysis on $C_{1\text{-}5}$ shows significant difference between the oral and nasal curves in all of the curves within front portion of the oral cavity, but not the back portion.

Table 1: AUC results for one-way ANOVAs on ROI front and back, showing the significance of each independent variable—nasality, curve number ($C_{1\text{-}5}$) and their interactions —in the comparison of area for oral and nasal vowels.

Region	IV	DF	*F*	*p*
Front	Nasality	(1,1439)	1157.084	<0.0001
	CurveNo	(4,1436)	122186.209	<0.0001
	Nasality:CurveNo	(4,1436)	25.714	<0.0001
Back	Nasality	(1,1439)	4.2767	0.0389
	CurveNo	(4,1436)	5309.5875	<0.0001
	Nasality:CurveNo	(4,1436)	0.4602	0.766

The results of a one-way MANOVA indicate that AUC differs significantly for the TT, TM and TD regions of the ROI based on nasality ($F(3, 1438) = 130.654$, $p < 0.0001$, curve number ($F(4,1436) = 180.990$, $p < 0.0001$) and their interaction ($F(4,1436) = 5.346$, $p < 0.0001$). Further one-way ANOVAs were carried out to determine the significant effects on each region individually, as seen in Table 2. These differences remain significant ($p < 0.017$) through all three vowel pairs. The differences become insignificant in the TT and TM regions for all three vowel pairs by C_4, though not in the TD region.

Table 2: AUC results for one-way ANOVAs performed on tongue tip, middle, and dorsum regions, showing the significance of each independent variable—nasality, curve number ($C_{1\text{-}5}$) and their interactions—in the comparison of area for oral and nasal vowels.

Region	IV	DF	*F*	*p*
TT	OralNasal	(1,1439)	284.486	<0.0001
	CurveNo	(4,1436)	28218.234	<0.0001
	OralNasal:CurveNo	(4,1436)	15.457	<0.0001
TM	OralNasal	(1,1439)	127.4043	<0.0001
	CurveNo	(4,1436)	20777.1088	<0.0001
	OralNasal:CurveNo	(4,1436)	3.0618	0.0159
TD	OralNasal	(1,1439)	182.2159	<0.0001
	CurveNo	(4,1436)	48368.7189	<0.0001
	OralNasal:CurveNo	(4,1436)	1.3761	0.2399

3.2. SSANOVA

SSANOVA was performed for each oral/nasal pair, for curves $C_{1\text{-}5}$. Results show differences in tongue shape, especially in lingual height, for all three vowels throughout the first half of the vowel. Nasal vowels exhibit much higher tongue bodies than oral vowels, as seen in Table 3. Tongue shape is more contracted and narrow for nasal vowels than for oral vowels. For the back vowel pairs /a-ɐ̃/ and /u-ũ/, there is a significant difference in the tongue tip shape across all curves within the vowel, with the oral vowel exhibiting a more pronounced point at the tongue tip. While this may be an effect of coarticulation (the preceding segment in the oral item is [t]), the tongue tip points persist throughout the duration of the vowel. A more controlled set of materials is needed to investigate this. The body of the tongue is significantly higher for nasal vowels than for oral vowels in all three vowel pairs, an effect less likely due to coarticulation.

At C_3 and onwards, the body of the tongue of the oral vowel stretches outwards (up for /u/, down for /a/), while the tongue body for the nasal vowel compresses toward the center of the oral cavity. By C_4, the difference in tongue height is no longer significant. The tongue dorsum also collapses downwards in /ũ/. These movements can be seen in Figure 2. In the x direction (corresponding to frontness), the lingual shape moves in opposite directions: oral vowel tongue configurations move backwards, whereas nasal vowels move forwards towards the center of the oral cavity. However, the tongue tip moves backward for nasal vowels, suggesting an overall compression of the tongue in the oral cavity.

For the front vowel pair /e-ẽ/, the opposite pattern is seen. The oral and nasal vowels begin at the same height and frontness, yet the nasal vowel gradually retracts and slightly raises, while the oral vowel remains fronted.

Table 3: Distance between nasal and oral vowel tongue heights, in mm. Distance is taken by finding the maximum tongue height of the nasal vowel and subtracting the height of the oral value at the same x value, then multiplying by 2.2, the in-plane voxel height.

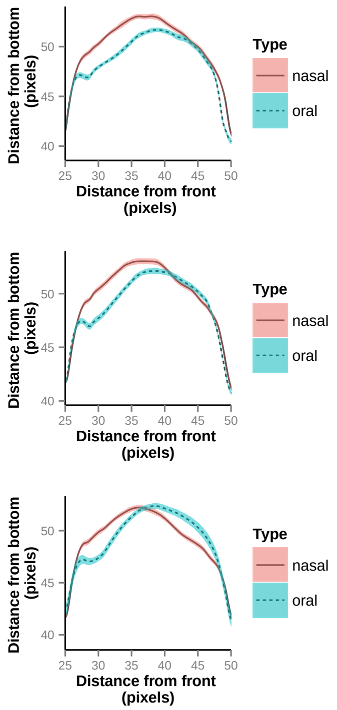
	First	Middle	Last	
a-ɐ̃	4.386	3.806	1.231	
e-ẽ	1.458	3.027	3.141	
u-ũ	3.126	2.241	0.714	

4. DISCUSSION

The results of this study show significant differences in tongue fronting, height, and shape between oral /a, e, u/ and their nasal counterparts /ɐ̃, ẽ, ũ/ in BP. This suggests that oral articulation of vowels differs between nasal and oral vowels, meaning that the configuration of the oral cavity affects the articulation of nasal vowels, as also described in [3].

Analysis of AUC results show there is a difference in height and frontness of the tongue for oral vowels and their nasal counterparts, especially within $C_{1\text{-}3}$, or the first 50% of the vowel. This corroborates the

Figure 2: *Top to bottom*: SSANOVA outputs for the comparison of */u-ũ/* at the beginning, middle and end of the normalized vowel durations. Axes are given in relation to bottom-left corner of the ROI; smaller numbers on the x-axis are closer to the front of the mouth, and smaller numbers on the y-axis are closer to the bottom of the oral cavity.



previous studies that show that oral cavity configuration differs in oral and nasal vowels. The increased height of the tongue in nasal vowels corroborates the findings of [24, 5] which both observe tongue-body raising in BP nasal vowels.

Analysis of tongue shape allows for more detailed understanding of the differences between these vowel pairs. For /a–ɐ̃/, the tongue body raises over time for the nasal, while lowering for the oral. The tongue body also moves front slightly during the nasal vowel, though the tongue tip retracts. The fronting and raising of the tongue imply a lower F1 and a slightly higher F2 for /ɐ̃/ than for /a/. For /u–ũ/, the tongue body raises slightly over time for /u/, while lowering for /ũ/. We speculate that the gesture associated with the nasal vowel suggests that the tongue is contracting to make room for the lowered velum, which may be necessary to avoid epiphenomenal contact between the lowered velum and high tongue dorsum [22, 23]. The tongue also moves forward (by as much as 1 cm) throughout the duration of the nasal vowel. The lowering and fronting of the tongue correlate with a higher F1 for /ũ/ than for /u/, as seen in acoustic studies of BP [1, 15]. For /e-ẽ/, initial tongue positions are very similar. The tongue gradually retracts for /ẽ/, correlating with the lowering of F2. The tongue body of the nasal vowel raises slightly over time, again implying a slightly higher F1, though this phenomenon is not as substantial as in /ɐ̃/ and /ũ/.

The findings show that BP nasal vowels are articulated in ways similar to those of French: the primary oral difference between oral and nasal vowels is contraction of the tongue body towards the center of the oral cavity. As seen previously [2], the overall tongue compression and tongue tip backing in the oral cavity lower F2, a phenomenon observed cross-linguistically in nasal vowels. Oral articulation was also found to modify F1—raising the frequency for high vowels, and lowering it for low vowels. These effects of the oral cavity enhance the acoustic output of the nasalized sound, in line with previous studies of BP vowel acoustics.

Moreover, our findings show that the centralization of nasal vowels builds throughout the production of the vowel, rather than being a static phenomenon. It is increasingly important to study articulatory events over time, rather than at a single instant or by averaging values throughout the duration of a phone, in order to understand the complex articulatory features in the production of a sound.

This study highlights the importance of using high-resolution imaging techniques and advanced statistical modeling to understand complex articulatory configurations within previously hard-to-view regions of the vocal tract. The use of rt-MRI allows for the imaging of the entire vocal tract, including the entire tongue contour, in excellent temporal and spatial resolution. This can give insight into different measures such as tongue height, frontness, and shape. Graphical analysis such as SSANOVA allows for comparison across the entire articulator, as well as further analysis of individual segments.

5. CONCLUSION

We present rt-MRI evidence suggesting that the oral configuration of BP nasal vowels differs from that of their oral counterparts, implying that lingual articulations enhance the effects of velic lowering in nasal vowels. The results show that multiple segments of the tongue differ in shape and position in oral and nasal vowel counterparts, further suggesting that multiple articulators are responsible for the acoustic effects of nasalization on formants. The study also presents the utility of SSANOVA in rt-MRI analysis of difficult problems in the mapping of acoustics and articulation.

6. REFERENCES

[1] Barbosa, P. A., Albano, E. C. 2004. Brazilian Portuguese. *Journal of the International Phonetic Association* 34(2), 227–232.

[2] Carignan, C. 2014. An acoustic and articulatory examination of the “oral” in “nasal”: The oral articulations of French nasal vowels are not arbitrary. *Journal of Phonetics* 46, 23 – 33.

[3] Carignan, C., Shosted, R., Fu, M., Liang, Z.-P., Sutton, B. Accepted. A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *Journal of Phonetics*.

[4] Carignan, C., Shosted, R., Fu, M., Liang, Z.-P., Sutton, B. P. 2013. The role of the pharynx and tongue in enhancement of vowel nasalization: a real-time MRI investigation of French nasal vowels. *INTERSPEECH* 3042–3046.

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

[6] Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America* 120(1).

[7] Fox, J., Bates, D., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Monette, G., Nilsson, H., Ogle, D., others, 2009. car: Companion to Applied Regression, r package version 1.2-16. *http://cran. r-project. org/web/packages/car/index. html (accessed on August 2012)*.

[8] Freeman, V. 2014. Bag, beg, bagel: Prevelar raising and merger in Pacific Northwest English. *University of Washington Working Papers in Linguistics* 32.

[9] Fruehwald, J. 2008. SS ANOVA.

[10] Fu, M., Christodoulou, A. G., Naber, A. T., Kuehn, D. P., Liang, Z.-P., Sutton, B. P. 2012. High-frame-rate multislice speech imaging with sparse samping of (k, t)-space. *Proceedings of the 20th Annual Meeting of International Society of Magnetic Resonance in Medicine* 12.

[11] Fu, M., Zhao, B., Carignan, C., Shosted, R. K., Perry, J. L., Kuehn, D. P., Liang, Z.-P., Sutton, B. P. 2015. High-resolution dynamic speech imaging with joint low-rank and sparsity constraints. *Magnetic Resonance in Medicine* 73, 1820–1832.

[12] Fu, M., Zhao, B., Holtrop, J., Kuehn, D., Liang, Z.-P., Sutton, B. 2013. High-frame-rate full-vocal-tract imaging based on the partial separability model and volumetric navigation. *Proceedings of the 21st Annual Meeting of International Society of Magnetic Resonance in Medicine* 4269.

[13] Gu, C. 2002. *Smoothing Spline ANOVA Models*. Springer Series in Statistics. Springer New York.

[14] Gu, C. 2007. gss: General Smoothing Splines. *http://cran.r-project.org/web/packages/gss/gss.pdf* R package version (2007): 1-0.

[15] Kelm, O. R. 1989. Acoustic characteristics of oral vs. nasalized /a/ in Brazilian Portuguese: Variation in vowel timbre and duration. *Hispania* 72(4), 853–861.

[16] Kochetov, A., Sreedevi, N., Kasim, M., Manjula, R. 2014. Spatial and dynamic aspects of retroflex production: An ultrasound and EMA study of Kannada geminate stops. *Journal of Phonetics* 46, 168–184.

[17] Liang, Z.-P. 2007. Spatiotemporal imaging with partially separable functions. *Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, 2007* 181 – 182.

[18] MATLAB, 2012. *Version 7.14.0.039 (R2012a)*. The MathWorks Inc.

[19] Mielke, J., Olson, K. S., Baker, A., Archangeli, D. 2011. Articulation of the Kagayanen interdental approximant: An ultrasound study. *Journal of Phonetics* 39(3), 403–412.

[20] R Development Core Team, 2007. R: A language and environment for statistical computing. *http://www.R-project.org*.

[21] Rong, P., Kuehn, D. P. 2010. The effect of oral articulation on the acoustic characteristics of nasalized vowels. *Journal of the Acoustical Society of America* 127(4), 2543–2553.

[22] Shosted, R. 2006. Vocalic context as a condition for nasal coda emergence: Aerodynamic evidence. *Journal of the International Phonetic Association* 36, 39–58.

[23] Shosted, R. 2011. Excrescent nasal codas in Brazilian Portuguese: An electropalatographic study. Lee, W.-S., Zee, E., (eds), *Proceedings of the XVIIth International Congress of Phonetic Sciences* Hong Kong. City University of Hong Kong 1834–1837.

[24] Shosted, R. To Appear. Nasal vowels are not [+NASAL] oral vowels. In: Smith, J., Ihsane, T., (eds), *Romance Linguistics 2012: Selected Papers from the 42nd Linguistic Symposium on Romance Languages (LSRL)*. John Benjamins.

[25] Shosted, R., Carignan, C., Rong, P. 2012. Managing the distinctiveness of phonemic nasal vowels: articulatory evidence from Hindi. *Journal of the Acoustical Society of America* 131(1), 455–465.

[26] Takano, S., Honda, K. 2007. An MRI analysis of the extrinsic tongue muscles during vowel production. *Speech Communication* 49(1), 49–58.