INTERSPEECH 2009 BRIGHTON

Speech Rate Effects on European Portuguese Nasal Vowels

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

[1]Universidade de Aveiro
[2]DETI/IEETA, Universidade de Aveiro, Portugal
coliveira@ua.pt, pmartins@ua.pt, ajst@ua.pt

Abstract

This paper presents new temporal information regarding the production of European Portuguese (EP) nasal vowels, based on new EMMA data. The influence of speech rate on duration of velum gestures and their coordination with consonantic and glottal gestures were analyzed. As information on relative speed of articulators is scarce, the parameter stiffness for the nasal gestures was also calculated and analyzed. Results show clear effects of speech rate on temporal characteristics of EP nasal vowels. Speech rate reduces the duration of velum gestures, increases the stiffness and inter-gestural overlap.

Index Terms: speech rate, dynamics, nasal vowels, gestures, (European) Portuguese.

10.21437/Interspeech.2009-168

1. Introduction

Portuguese has a rich system of nasal vowels with five monophthongs ([ɐ̃], [ẽ], [ĩ], [ũ], [õ]) and several diphthongs. One of the most important findings as far as Portuguese nasality is concerned refers to the manifestation of nasality in the temporal dimension. The course of nasality in the vowels is typically incremental, with a movement from oral to nasal. Many researchers took also notice of a final consonantal segment after the nasal vowel. The dynamic pattern of Portuguese nasality was first mentioned in [1]. Recent articulatory studies [2, 3] provided additional evidence for the existence of different phases in the production of nasal vowels: 1) oral onset; 2) vocalic nasal portion; and 3) nasal consonantal segment/ nasal tail. Albano [4], within the Articulatory Phonology framework, and based on acoustic data, provides a gestural explanation for the nasal vowels in Brazilian Portuguese: the velum gesture is somewhat delayed with respect to the tongue body gesture, inducing an oral onset, and depending on the degree of overlap between the velum gesture and the following oral gesture (as in "canto"), a transitional/intrusive consonantal segment can appear.

Some supporting evidence for this temporal organization of articulatory gestures in the production of nasal vowels is provided by a previous EMMA study [5], which established that: 1) the velum gesture starts after the release of the previous oral gesture but before the release offset, approximately halfway between the two; and 2) the velum opening and the next oral gesture are articulated sequentially.

As shown by several studies, timing relationships seem to be systematic affected by many factors such as place of articulation, word and syllable position, lexical characteristics of words or speech rate. Supported by kinematic data, Oliveira and Marin [6] found, for Brazilian Portuguese, that in sequences like "partindo" ("leaving") the velum gesture is sequentially coupled with the oral gesture. However, at fast rates, a spontaneous shift from sequential coordination to synchronous coordination might occur.

Accordingly, in the present study we investigate the influence of speech rate on the production of European Portuguese nasal vowels, using new articulatory (EMMA) data and a wider range of production conditions. Specifically, our work focuses on the influence of speech rate not only on the inter-gestural timing, but also on other dynamical aspects such duration of velum gestures and velum stiffness. This latter parameter is a "context-independent measurement of the velocity profile of articulatory movement" [7].

The paper is organized as follows: in section 2, the experimental setup, materials and analysis are presented; the results are given separately for the duration, stiffness and intergestural timing, in sections 3, 4 and 5; and discussion and main conclusions (sections 6 and 7), conclude the paper.

2. Methodology

Data were collected using a two-dimensional electromagnetic transduction device (Carstens AG100). The acoustic signal was recorded and synchronized with the movements signals. Kinematic data were sampled at 500 Hz and acoustic speech signal was sampled at 20 kHz. Standard calibration and post-processing procedures were applied. Sensors were attached to upper and lower lip, jaw, tongue tip, tongue body and velum. Additional sensors served as reference for the signal acquisition system. In this study we focus on data from the velum (V), the upper and lower lip (UL & LL) and the tongue tip (TT).

2.1. Corpus and speakers

Two native speakers of European Portuguese (one male and one female), with normal hearing and no speech deficits, participated in corpus recording. Here we only report the results of the female speaker.

The data are a subset of a more comprehensive corpus, designed to study several aspects of the Portuguese nasality.

The corpus included nonwords with the five nasal consonants ([ɐ̃], [ẽ], [ĩ], [ũ], [õ]) in three prosodic conditions: word-initially, word-medially and word-finally (e.g. [ˈɐ̃.pɐ], [ˈpɐ̃.pɐ], [ˈpɐ̃]). The consonants preceding and following the nasal vowels are labial and alveolar stops ([b], [p], [t], [d]) and the labiodental fricative ([f]). The test items were embedded in the carrier sentence "Diz...três vezes" ("Say...three times"). Two speaking rate conditions were employed: subjects were instructed to read the sentences at a normal, comfortable rate, and to produce each stimulus using a fast rate.

2.2. Data Analysis

The data were analyzed using MATLAB routines to identify important landmarks in the displacement and velocity signals of EMMA recordings. For each gesture the following points were labeled: movement onset (S), achievement of the target (T), target release (R) and release offset (T2) (see Fig. 1). The landmarks were algorithmically identified on the basis of a velocity threshold. For alveolar stops ([t],[d]), the tangential velocity of TT was used; for labial consonants ([p],[b],[f]), vertical velocity of lip aperture was used; for nasal vowels, the velocity signal of the vertical movement of velum was evaluated.

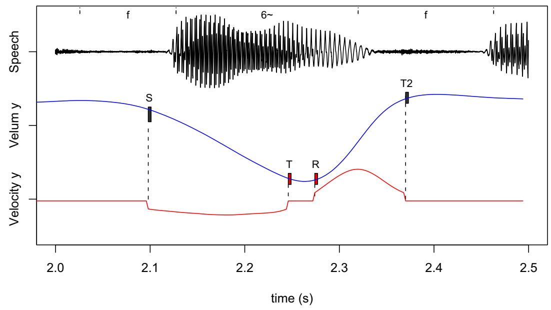

Figure 1: Automatic labeling of the word [fẽf]. Top panel: acoustic signal; middle panel: velum y with events labeled; bottom panel: y-velocity of velum coil.

3. Velum Duration

The values of duration were obtained automatically, on the basis of the landmarks identified and labeled in the velum signal. The results are given in table 1.

Table 1: Mean duration (ms) of velum movements (total cycle, opening phase, plateau, closing phase) by position and rate.

Rate	Position	Total	Opening	Plateau	Closing
Normal	INITIAL	340	133	42	122
	INTERNAL	300	143	24	112
	FINAL		141		
Fast	INITIAL	200	74	15	97
	INTERNAL	220	106	15	79
	FINAL		165		

There was a great decrease of all duration measures with speech rate increase. The only exception is the mean duration of the opening cycle in final position. This difference could be due to problems in the labeling of articulatory movements in the fast speech condition.

Overall, there are clear duration differences between the opening and the closing phases of the velum, in both rates: closing is on average shorter than opening. A notable exception is the initial position in fast rate, where opening is almost 20 ms smaller than the closing phase.

The reported differences were evaluated with a separate three-way analysis of variance (ANOVA) on each dependent variable (total duration, opening duration, plateau duration and closing duration). The independent variables are speech rate, position and nasal vowel.

The total duration of velum gesture shows a significant influence of the rate [F(1,77)=152.88, p<0.001] and nasal vowel [F(4,77)=2.65, p=0.039].

Regarding opening duration, an effect of rate, position and vowel was found.

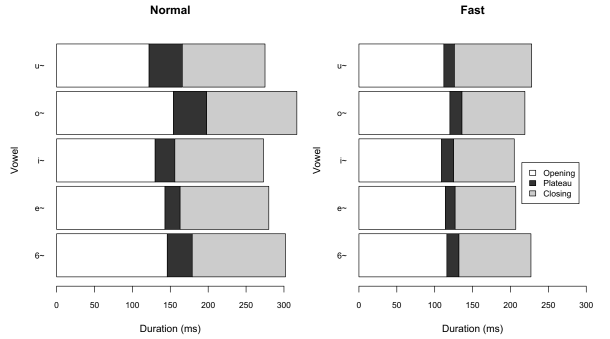

Figure 2: Mean duration of velum gestures by rate and nasal vowel.

The statistical analysis showed that the rate factor had influence on plateau duration [F(1,77)=20.89, p<0.001]. The position was also significant [F(1,77)=4.93, p=0.029].

In a similar way, there was a main effect of rate [F(1,78)=40.23, p<0.001] and position [F(1,78)=9.40, p=0.003] on closing duration.

4. Velum Stiffness

Stiffness was calculated for the opening and the closing phases of velum movement. Following [7], to get the stiffness, the peak velocity (cm/sec) was divided by the maximum displacement (cm). The results, by speech rate and position, are shown in Table 2.

Table 2: Mean values of *stiffness* for opening and closing phases of velum, by speech rate and position.

Speech Rate	Position	Opening	Closing
Normal	INITIAL	10.2	12.5
	INTERNAL	9.6	13.1
	FINAL	9.7	—
Fast	INITIAL	14.6	15.1
	INTERNAL	12.4	16.8
	FINAL	9.4	—

We can observe an increase of stiffness mean values from normal to fast rate, in all the positions, with exception of the final position.

Besides, for both rates, the opening phase had lower stiffness than the closing phase.

We conducted an ANOVA with speech rate, position and nasal vowel as independent variables and stiffness as dependent variable. Due to labeling problems, only data from internal position were included.

There was a main effect of rate [F(1,77)=23.03, p<0.001] and vowel [F(4,77)=2.58, p=0.044] on the opening stiffness.

The stiffness of the closing phase is significantly affected by the rate [F(1,59)=56.51, p<0.001].

The results indicated as significantly smaller the value for the *stiffness* for the opening gesture [t(71)=-10.90, p<0.001]. On average, the difference is between -3.94 and -2.72.

5. Inter-gestural timing

The analysis of inter-gestural timing involved several measurements (see Fig. 3). Gestural coordination was quantified as the time between gestural landmarks, algorithmically derived from local velocity profiles of the relevant receivers in the vocal tract.

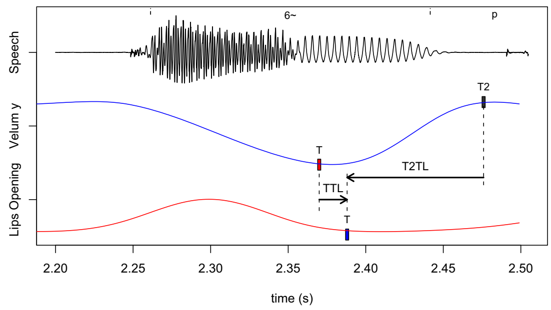

Figure 3: Example of articulatory measurements: TTL and T2TL.

As an index of the temporal overlap between the velum gesture and the following oral consonant (C2), two measures were evaluated: 1) *Target-to-Target Lag* (TTL) - the interval between the achievement of velum opening target and the achievement of oral target (C2) - this parallels [6]; and 2) *Target-to-Target Lag* (T2TL), which was calculated as the time between the achievement of velum closing target and the achievement of target for C2.

As in a previous work [5], the degree of overlap between the previous oral consonant (C1) and the velum gesture was assessed by measuring SRL (*Start-to-Release Lag*), i.e., the interval between onset of velum movement and release for the oral constriction.

Finally, to measure the gestural timing between velum and glottis, the interval between the achievement of velum closing target and opening of the glottis (measured from acoustic signal) was used.

5.1. TTL

The mean values for TTL are presented in Table 3.

Table 3: Mean values (ms) of TTL (95% confidence interval), by speech rate and position.

Speech Rate	Position	Mean	CI 95 %
Normal	INITIAL	-57.1	[-74.6 .. -39.6]
	INTERNAL	-25.8	[-31.9 .. -19.8]
Fast	INITIAL	-38.8	[-42.6 .. -35.0]
	INTERNAL	-21.5	[-27.6 .. -15.5]

The TTL values are all in negative range. Thus, the nasal (i.e. velum) gesture is sequentially coupled with the oral gesture. The velum target precedes the oral target.

At fast rates, relative timing between achievement of targets becomes shorter and the TTL values are more close to zero. The achievement of the oral target is only slightly delayed with respect to the achievement of velum target.

We can observe also a difference in the mean values of the TTL as a function of the lexical position: in initial position, the lag between achievement of velum target and oral constriction is higher than in internal position, in both speech rates.

The results obtained for TTL were further tested with a three-factor ANOVA. The three independent variables are position, rate and oral articulator (tongue tip or lips). The results reveal a significant effect of position [F(1,81)=25.99, p<0.001] and oral articulator [F(1,81)=7.83, p=0.006] on TTL values.

5.2. T2TL

Table 4: Mean values (ms) of T2TL (95% confidence interval), by speech rate and position.

Speech Rate	Position	Mean	CI 95 %
Normal	INITIAL	106.6	[95.5 .. 117.6]
	INTERNAL	110.8	[102.4 .. 119.1]
Fast	INITIAL	61.6	[56.7 .. 66.4]
	INTERNAL	72.8	[66.3 .. 79.2]

As shown in Table 4, there is a positive lag between the velum and the oral gestures, which means that the velum closing gesture follows the oral target in time.

Note the considerable difference in T2TL values between the two speech rates: the interval between the two targets decreases from normal to fast rates.

As illustrated in Fig. 4, the speech rate [F(1,84)=99.58, p<0.001] and oral articulator [F(1,84)=10.42, p=0.002] had significant effect on T2TL. The C2 with tongue tip involvement showed greater lag than the labial consonants.

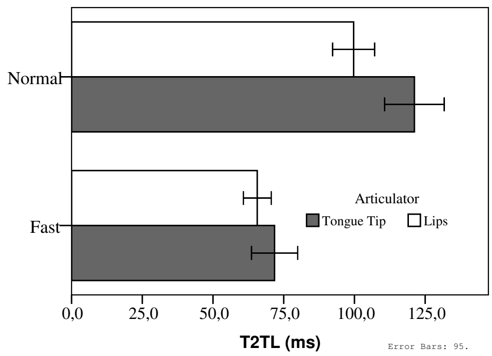

Figure 4: Mean T2TL (in ms), by rate and oral articulator (error bars show the standard deviation).

5.3. Velum-to-Glottis lag

The results for this timing interval are presented in Table 5.

Table 5: Time (95% confidence interval) from achievement of target of closing velum gesture to glottis gesture, by rate and position.

Speech Rate	Position	Mean	CI 95 %
Normal	INITIAL	58.5	[49.7 .. 66.4]
	INTERNAL	53.4	[46.3 .. 60.5]
Fast	INITIAL	46.4	[41.4 .. 51.4]
	INTERNAL	50.2	[45.6 .. 54.7]

The lag between the gestures is positive, i.e., the glottis opens before the end of the velum gesture. We can observe also a rate effect, evidenced in the smaller lag in fast rate compared to normal rate.

We conducted an ANOVA with rate, position and nature of C2 as independent variables. The results confirmed the influence of rate [F(1,55)=6.05, p=0.017] and nature of C2 [F(2,55)=11.37, p<0.001] on articulatory lag. As the results of the post-hoc Tamhane test revealed this was due to differences between the alveolar consonant [t] and the labial consonants [p f].

5.4. SRL

As far as SRL is concerned, the Table 6 shows that, at normal rate, velum onset occurs on average after the release of the previous labial consonant.

Table 6: Mean values (ms) of SRL (95% confidence interval), by speech rate and position (only labial gestures).

Speech Rate	Position	Mean	CI 95 %
Normal	FINAL	21.8	[10.42 .. 33.15]
	INTERNAL	32.9	[22.70 .. 43.01]
Fast	FINAL	-5.2	[-19.35 .. 9.02]
	INTERNAL	-1.6	[-12.57 .. 9.42]

There is an overlap increase from normal to fast rate, in both positions. The C1 release and velum onset gestures tend to be roughly simultaneous.

Effects of rate, position and nature of C1 were quantified by three-way ANOVA. Results confirmed that the rate had influence on SRL [$F(1,48)=27.65$, $p<0.001$].

6. Discussion

Looking first at the duration of the complete cycle of the velum movement (between 200 ms and 300 ms), the results obtained in this study confirm earlier data for Portuguese [5] and French [8, 9]. In spite of that, the duration of the opening and closing gestures is different for the two languages: in French, the two gestures have similar durations, but in Portuguese, the opening movement is longer than the closing gesture, remaining the relation at fast speech rate. Our results showed also a marked rate effect on all duration measures, in line with previous findings, cf. [10].

The duration of articulatory movements is closely related with the stiffness of those movements: gestures with higher stiffness result in shorter duration movements. Our data revealed significant stiffness differences between the opening and the closing gestures of the velum, at both rates, which means that, on the one hand, the velum closes more rapidly than it opens [11, 12] (a difference in frequency response of closing and opening gestures has also been observed for jaw and tongue dorsum movements) and, on the other hand, that the opening movement is longer than the closing movement as shown by our results. The stiffness of the closing an opening movements increases significantly at fast speech. Thus, the velocity of the gestures also increases, while the duration decreases.

It has been proposed by [4] that the opening gesture of velum starts after the tongue dorsum gesture and overlaps the following oral gesture. In fact, the SRL results indicate that, at the normal rate, the onset of velum opening follows the consonant release by about 20-30 ms. It seems to us that this pattern could be caused by simple constraints of the executing motor system: stops and fricatives require small velopharyngeal openings. So in that contexts, velum cannot open earlier and the beginning of the nasal vowel will be acoustically and perceptually similar to an oral vowel.

Furthermore, the overlap patterns observed in our study reveal that the emergence of a consonantal segment (N) is caused by an overlap among adjacent gestures, namely the velum closing gesture, the oral gesture and the glottal gesture. At the higher speaking rate, there was greater overlap among these gestures and consequently a decrease of N duration. In some tokens, the relative timing shifted and gestures were produced essentially simultaneously ("bimbo" < "bimo"), an effect previously reported by [6].

7. Conclusion

This papers addressed the question of whether and to what extent speech rate affects the production of Portuguese nasal vowels. The hypothesis that velum duration, velum stiffness and inter-gestural timing would show the influence of rate was tested in an articulatory (EMMA) study of nasal vowels, produced at different speaking rates and word position conditions. The results suggest that speech rate acts to reduce the duration of velum gestures and to increase the stiffness and the inter-gestural overlap. This information is crucial to realistically represent the movement of articulatory gestures in our articulatory based text-to-speech system [13]. Nevertheless, further investigations are needed, since other factors, such as dialectal and individual factors can affect inter-gestural cohesion.

8. Acknowledgments

This work was supported by FCT (Fundação para a Ciência e a Tecnologia) grant SFRH/BD/18058/2004 and HERON project. Many thanks are due to GIPSA-Lab, Solange Rossato, for help with articulatory design, and Christophe Savariaux, for help with the EMMA recordings.

9. References

[1] 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.

[2] F. N. Gregio, "Configuração do trato vocal supraglótico na produção das vogais do português brasileiro: dados de imagens de ressonância magnética," Tese de Mestrado, PUC/SP, 2006.

[3] L. Lovatto, A. Amelot, L. Crevier-Buchman, P. Basset, and J. Vaissière, "A fiberscopic analysis of nasal vowels in Brazilian Portuguese," in *ICPhS' 2007*, Saarbrücken, 2007, pp. 549–552.

[4] E. Albano, "O português brasileiro e as controvérsias da fonética atual: pelo aperfeiçoamento da fonologia articulatória," *DELTA*, vol. 15, pp. 23–51, 1999.

[5] C. Oliveira and A. Teixeira, "On gestures timing in European Portuguese nasals," in *ICPhS' 2007*, Saarbrücken, 2007.

[6] L. Oliveira and S. Marin, "Patterns of velum coordination in Brazilian Portuguese," in *PAPI*, Barcelona, 2005.

[7] K. Roon, A. Gafos, P. Hoole, and C. Zeroual, "Influence of articulator and manner on stiffness," in *ICPhS' 2007*, Saarbrücken, 2007, pp. 409–412.

[8] A. Amelot, "Étude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français," Thèse de doctorat, Université Paris III, 2004.

[9] P. Basset, A. Amelot, and L. Crevier-Buchman, "Etude multi-paramétrique des consonnes nasales du français: Prise de données simultanées aérodynamiques et fibroscopiques," *Revue Parole*, vol. 39-40, pp. 113–135, 2006.

[10] R. Krakow, "Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate," in *Nasals, Nasalization, and the Velum*, M. K. Huffman and R. A. Krakow, Eds. Academic Press, 1993, pp. 87–116.

[11] A.-P. Benguerel, H. Hirose, M. Sawashima, and T. Ushijima, "Velar coarticulation in French: fiberscopic study," *Journal of Phonetics*, vol. 5, pp. 149–158, 1977.

[12] M.-J. Solé, "Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization," *Language and Speech*, vol. 38, no. 1, pp. 1–23, 1995.

[13] A. Teixeira, C. Oliveira, and P. Barbosa, "European portuguese articulatory based text-to-speech: First results," in *Computational Processing of the Portuguese Language, PROPOR 2008*. Springer, 2008.