



## RHYTHMIC REORGANIZATIONS OF SPEECH BASED ON JAW VERTICAL MOVEMENTS

## REORGANIZAÇÕES RÍTMICAS DA FALA EM FUNÇÃO DAS MOVIMENTAÇÕES VERTICAIS DA MANDÍBULA

Alexsandro Rodrigues Meireles<sup>1</sup>

**Resumo:** Este artigo investiga como o aumento da taxa de elocução atua para mudar articulatoriamente os movimentos verticais da mandíbula. Duas mulheres (uma de Minas Gerais e outra de São Paulo) na faixa etária de 30-40 anos de idade foram gravadas no Laboratório de Fonética do Universidade Federal do Espírito Santo com o uso de um magnetômetro NDI Wave com taxa de amostragem de 100 Hz. Os resultados mostram que o aumento da taxa de elocução alterou os parâmetros articulatórios no eixo vertical da seguinte forma: a) diminuição da duração da aceleração; b) diminuição do ponto vertical máximo (y-extremo); c) diminuição do deslocamento da constrição; d) diminuição em módulo da velocidade do vale/pico; e) diminuição da duração gestual; e f) tempo-para-velocidade-de-pico/vale proporcional constante. Além disso, os resultados mostram que a taxa de elocução tende a afetar todos os gestos independentemente de sua posição frasal. No entanto, há evidências que alguns parâmetros articulatórios, se devidamente controlados, podem fornecer indícios para reestruturações rítmicas da fala.

**Palavras-chave:** taxa de elocução, EMA, ritmo de fala.

**Resumen:** Este artículo investiga cómo el aumento de la tasa de elocución actúa para cambiar articulatoriamente los movimientos verticales de la mandíbula. Dos mujeres (una de Minas Gerais y otra de São Paulo) en el grupo de edad de 30-40 años fueron grabadas en el Laboratorio de Fonética de la Universidad del Espíritu Santo con el uso de un magnetómetro NDI Wave con una frecuencia de muestreo de 100 Hz. Los resultados muestran que el aumento de la tasa de elocución alteró los parámetros articulatorios en el eje vertical de la siguiente manera: a) disminución de la duración de la aceleración; b) disminución del punto vertical máximo (y-extremo); c) disminución del desplazamiento de la constricción; d) disminución en módulo de la velocidad del valle/pico; e) disminución de la duración de los gestos articulatorios; y f) tiempo-para-velocidad-de-pico/valle proporcional constante. Además, los resultados muestran que la tasa de elocución tiende a afectar todos los gestos independientemente de su posición frasal. Sin embargo, hay evidencias que algunos parámetros articulatorios, si debidamente controlados, pueden proporcionar indicios para reestructuraciones rítmicas del habla.

**Palabras clave:** tasa de elocución, EMA, ritmo del habla.

---

<sup>1</sup> Federal University of Espírito Santo, Ufes, Vitória, Brasil. meirelesalex@gmail.com  
Orcid: <https://orcid.org/0000-0003-1901-9329>

## 1. INTRODUCTION

The research of Meireles and Barbosa has shown some rhythmic reorganizations (or restructurings) due to speech rate variation in French (Barbosa, 1994, p. 121-123), and in Brazilian Portuguese (Barbosa, 2006, p. 249-254; see also Barbosa, 2007; Meireles, 2009; Meireles and Barbosa, 2008; Meireles, Tozetti and Borges, 2010; Silva and Meireles, 2011; Meireles and Gambarini, 2012; Meireles and Barbosa, 2014).

Rhythmic restructuring is considered as the reorganization of stress groups (henceforth SG) along the utterance when speech rate changes, which tends to yield a smaller number of SG at fast rates. According to Barbosa and Meireles' studies, some minor prosodic boundaries are deleted due to speech rate increase, since it is not possible to decelerate at some phrasal edges to pinpoint phrasal stress, and, at the same time, to maintain a speeded pace imposed by the rate along the whole utterance (Barbosa, 2006, p. 248-249).

Barbosa (2007) and Meireles (2009) have shown that intrinsic segmental duration is normalized through the use of abstract vowel-to-vowel (henceforth VV) duration. Entrained, abstract VV duration is used as the control parameter to generate prosodic variation (cf. Barbosa, 2007; Meireles, 2009; Silva and Meireles, 2011). The maxima of VV duration delimitate produced SGs as rhythmic units in Brazilian Portuguese (cf. Barbosa, 2007; Meireles, 2009; Silva and Meireles, 2011).

The speech rate influence on speech rhythmic reorganizations can be explained by the Dynamical Speech Rhythm model (Barbosa, 2007; Barbosa, 2006, henceforth DSR). This model accounts for prosodic structuring using a dynamical systems approach to language data (cf. Kelso, 1995; Kelso et al., 1986). In this model, speech rhythm is seen as a "consequence of the variation of perceived duration along the entire utterance" (Barbosa 2002, p. 71), in which the main control parameter is the duration of a syllable-sized unit. Entrained, abstract vowel-to-vowel (henceforth VV) duration is used as the control parameter to generate prosodic variation in the DSR model (see Marcus 1976, Morton et al. 1976, Lehiste 1970, Dauer 1983, Pompino-Marshall 1991, for the relevance of the vowel onset or CV transition). The maxima of VV duration delimitate produced SG as rhythmic units in Brazilian Portuguese (henceforth BP) (for a review of the SG as a prosodic unit, see Lehiste 1970, 1978, Frawley 1974, Woodrow 1951, among others). In the same model, BP syllable-sized duration is the primary parameter for explaining lexical stress realization as well as phrase stress, which determines linguistic rhythm (Mattoso Câmara 1970, Massini 1991, Massini-Cagliari 1992).

## 2. DYNAMICAL RHYTHMIC RESTRUCTURING DUE TO SPEECH RATE VARIATION: ACOUSTIC STUDIES

The following acoustic experiments show how speech rate variation acts to change speech rhythm in Brazilian Portuguese.

### 2.1. Experiment 1

By eliciting different speech rates as a perturbational device, which is a classical approach in dynamical systems theory for discovering the underlying parameters of a system, Meireles (2009) carried out a study in order to evaluate the utterance's rhythmic restructurings due to speech rate increase, having as a theoretical framework the DSR model. The corpus (a two-paragraph text) is phonetically balanced in agreement with BP

phone frequency in the NURC corpus (cf. Albano & Moreira 1996) (Spearman coeff. > 0.94). Four male subjects were recorded at three speech rates.

Statistical analyses (t-test) revealed a significant difference ( $\alpha = 0.05$ ) between the slow and fast rates for all speakers, analyzed by the distribution of VV durations in the text, which confirms that the extreme rates are significantly different from the slow rates.

Phrasal prominence placement for this experiment was extremely variable according to the subjects. Therefore, we decided to run another experiment in which variability should be lesser than before. In order to do so, the following experiment was run.

## **2.2 Experiment 2**

A corpus of eleven sentences was read by one female subject (age 20-30) at three speech rates (cf. Meireles, 2009). These sentences were made with different syntactic structures. The results of this experiment have corroborated the hypotheses derived from experiment 1, as follows: (i) VV number/SG proportionally increases with speech acceleration; (ii) SG duration remains constant across the speech rates; (iii) strong syntactic boundaries inhibit rhythmic restructurings; (iv) SG duration standard deviation is, in general, smaller at fast rates, which could contribute to a greater isochrony sensation.

## **2.3. Experiment 3**

Meireles and Barbosa (2008) results have shown, in case of rhythmic restructurings, that (i) the standard deviation of VV unit duration and stress group duration is smaller at faster rates; (ii) the stress group duration tends to be constant with speech rate increase; (iii) the number of VV units (vowel-to-vowel) per stress group proportionally increases with speech rate increase; and (iv) As standard deviation is consistently smaller at fast rates, speech rate increase exacerbates the mixture character of Brazilian Portuguese rhythm, i.e., tendencies to syllable as to stress-timed rhythm.

## **2.4. Experiment 4**

In a different study, Meireles, Tozetti and Borges (2010) corroborated our previous results on the question of rhythmic restructurings with speech rate increase, as follows: (i) standard deviation decrease of durational units with speech rate increase; (ii) the standard deviation of VV duration and/or SG duration is smaller at fast rates; (iii) the stress group duration and/or the VV duration tend to be constant with speech rate increase; (iv) the decrease of standard deviation of SG and/or VV duration is influenced by many factors such as dialect, gender, and sentence structure; and (v) Brazilian Portuguese rhythm varies between stress-timing and syllable-timing according to different speech styles.

## **2.5. Experiment 5**

Silva and Meireles (2011) sheds new light on the phenomenon of rhythmic restructuring since it correlates phonetic variables (mean and standard deviation of VV and SG durations) with social variables (gender and age). The main results of this paper

show that social variables variation can contribute to the reorganization of speech rhythm and modify some phonetic parameters in the following way: (i) the standard deviation of vowel-to-vowel (VV) duration and stress group duration is smaller for the male gender and advanced age group; (ii) stress group duration tends to be vary less according to male gender and the advanced age (rhythmic restructurings make VV units smaller, but with a greater number of VV units per stress group, which results in a statistically constant stress group duration for these variables); (iii) the number of VV units per stress group proportionally increases from the young age group to the advanced age group due to rhythmic restructurings; and (iv) the male gender and the advanced age group emphasize even more strongly the mixed character of Brazilian Portuguese rhythm, i.e., the tendency to mix syllable-timed with stress-timed rhythm.

## 2.6. Experiment 6

Meireles and Gambarini (2012) analyzed the speech rhythm of three Brazilian Portuguese dialects under the light of the DSR model to rhythm typology. The data showed a scale from a less stress-timed to a more stressed-timed dialect as follows: Minas Gerais state (coupling strength = 1.83, speech rate = 6.3) > São Paulo state (1.66, 6.6) > Bahia state (1.26, 5.9) > Espírito Santo state (0.95, 4.8). Also, the results show that stress-timing seems to be associated to faster rates.

## 3. DYNAMICAL RHYTHMIC RESTRUCTURING DUE TO SPEECH RATE VARIATION: AN ARTICULATORY STUDY

Meireles and Barbosa (2014) presented some results regarding rhythm restructuring with speech rate increase. According to the authors, speech rate increase tends to strengthen the right-headedness characteristic of BP, i.e., the greatest phrasal prominences occur to the right of the sentence at fast rates. Another important result of this paper is that a statistical comparison of the sentences' vowels as a function of rate indicated a significant general tendency for smaller displacements (y-extremum and constriction displacement) with speech rate increase. High and mid-high vowels tend to be less high, and low vowels tend to be less low from slow to fast rates.

## 4. METHODOLOGY<sup>2</sup>

Two female native speakers of BP (age 30-40) were recorded acoustically (sampling rate: 20.05 kHz) and articulatorily at the UFES Phonetics Laboratory. Subject MG is from Minas Gerais state, Brazil, and subject SP is from São Paulo state, Brazil. The subjects signed an approved informed consent form explaining the purpose of the experiment. An NDI Wave<sup>3</sup> electromagnetic articulograph speech research system (cf. Perkell et al. 1992, EMMA magnetometer system) was used for tracking jaw movement. The movement data was sampled at 100 Hz, head-corrected, rotated to the occlusal plane, and low-pass filtered at 25 Hz. Pellets were attached to the following articulators: tongue (close to the palatal region), lower lips, jaw (at the lower incisors). Two other pellets were

---

2 The methodology here is exactly the same as in Meireles and Barbosa (2014), since we are replicating the experiment with more subjects and additional results regarding rhythmic restructuring.

3 [https://support.ndigital.com/downloads/documents/guides/IL-1070187\\_rev\\_007.pdf](https://support.ndigital.com/downloads/documents/guides/IL-1070187_rev_007.pdf)

used as reference for the signal acquisition system: one at the nose bridge, and one at the center anterior surface of the maxillary incisors. Only the y-movement of the jaw was measured. As in Erickson (2004), jaw opening was measured “in terms of the lowest vertical position of the mandibular pellet in the syllable from the maxillary occlusal plane” (p. c-134).

The recorded sentences are displayed in table 1. These sentences were designed in such a way as to have an opening-closing alternating jaw pattern throughout the utterance. That is why we have used a sequence of high (or mid-high)-low vowels all the way up to the end of the sentence. For the same reason we have used bilabial consonants at the predicted phrase edges. Ten repetitions of each sentence (randomized within blocks) at three speech rates were recorded. This results in a total of 240 utterances for analysis (4 sentences x 10 repetitions x 3 speech rates x 2 subjects).

To obtain three distinct speech rates, the subjects were asked to read the sentences according to the following instructions and order: (1) normal (henceforth N): speak in a comfortable way; (2) slow (henceforth S): speak as slow as you can preserving the sentence’s meaning and without introducing pauses between words; (3) fast (henceforth F): speak as fast as you can without introducing distortions in speech. Some utterances were also given as prompts to the subject exemplifying each one of the intended speech rates.

Table 1. Sentences used in the experiment with their respective translation (TR) and phonetic transcription (PT). Bolded words represent where phrasal prominence is expected to fall (no such markings appeared in the stimuli for reading).

Sentence 1	Ela diz mão de <b>máfia</b> no carro da moça do <b>papai</b> .
PT	[ɛl.ɯdʒ.izm.ãʊdʒ.im.af.ɾɯŋ.ʊk.aʁ.ʊd.am.os.ɯd.ʊp.ap.aɪ]
TR	She says mafia’s hand in the car of my father’s girl.
Sentence 2	Foi <b>bão</b> gostar demais de <b>papai</b> , mas de <b>máfia</b> !?
PT	[f.oɪb.ɯʊg.ost.aʁdʒ.im.aɪzdʒ.ɪp.ap.aɪm.azdʒ.im.af.ɾɯ]
TR	It were good to love too much my dad, but not mafia... [“were” instead of “was” implies a very informal style]
Sentence 3	Mamãe não quer mais qu’eu <b>babe</b> , mas qu’eu <b>pape</b> .
PT	[m.ãm.ãn.ãʊk.ɛʁm.aɪsk.ɪeʊb.ab.im.ask.ɪeʊp.ap.ɪ]
TR	My mother doesn’t want that I dribble anymore, but that I eat [child language].
Sentence 4	Vou lá levar o <b>pavê</b> pra filha do <b>papai</b> .
PT	[v.ʊʊl.al.ev.ar.ʊp.av.epr.af.ɨ.ɯd.ʊp.ap.aɪ]
TR	I am going there to deliver the “pavê [kind of dessert]” to dad’s daughter.

For transcription and labeling, MAVIS software (Tiede et al. 1999), modified at the University of Southern California, USA, was employed to measure the jaw sensor movement in the vertical dimension (y-axis) (cf. figure 1). The following articulatory variables were measured: (i) jaw maximum extension: measured at y-velocity zero-crossing at maximum opening; (ii) jaw constriction displacement: measured as the

difference between the zero crossings at constriction onset and extremum; (iii) jaw gesture duration: measured between y-velocity zero-crossings at constriction onset and maximum; (iv) constriction jaw gesture peak/valley velocity (y); (v) jaw acceleration duration: measured from the time of zero-crossing at constriction onset to the time of constriction peak/valley velocity; (vi) proportional time-to-peak/valley-velocity: measured from the ratio of y acceleration duration to total constriction formation duration.

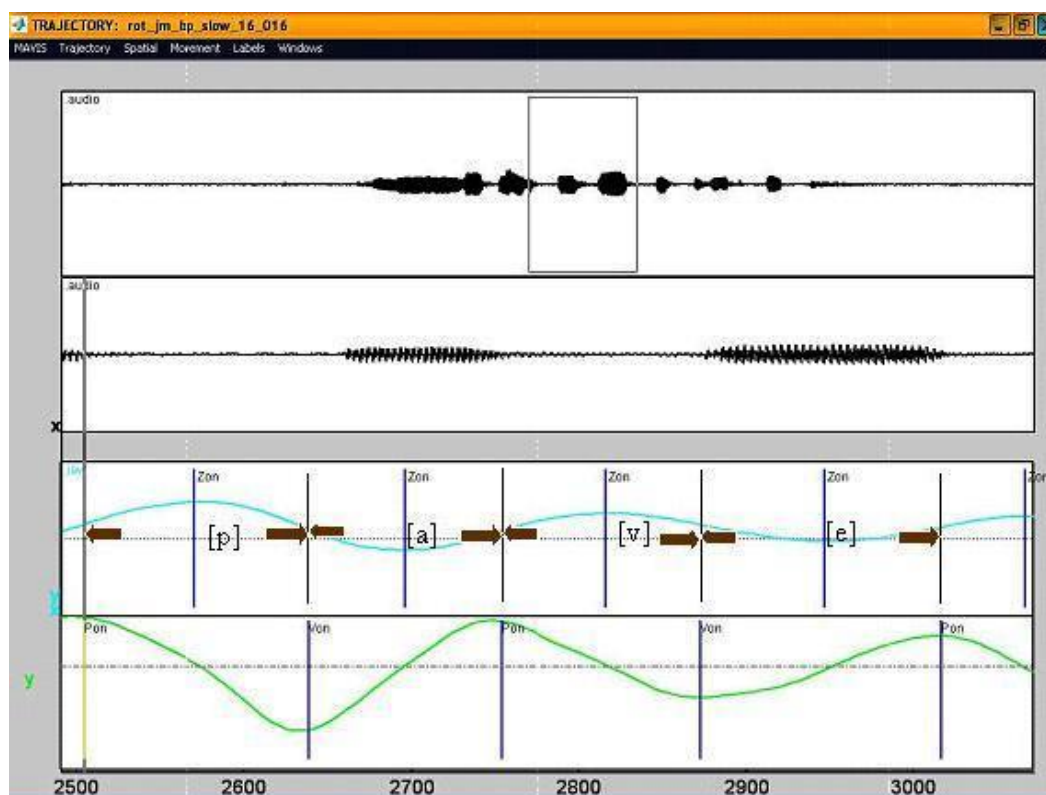


Figure 1. Display of jaw measurements within MAVIS software. Vertical blue lines represent the zero crossings at the points of maximum closing (consonants) and maximum opening (vowels). Vertical black lines delimit the jaw gesture duration (in-between peak and valley velocities), as indicated by the brown arrows (Meireles, 2009, p. 135)

The hypothesis to be investigated here is the diminishment of the number of SG with speech rate increase and the non-occurrence of jaw movement reset after prosodic boundaries at fast rates. According to this pattern, a minor prosodic boundary is immersed in the dynamic realization of a SG finishing with a major prosodic boundary and is reorganized as a function of the restriction of speech rate increase. Eg.: “|| Subiu a tribuna | de um poleiro de ouro ||” is reorganized as “|| Subiu a tribuna de um poleiro de ouro ||” at fast rates.

Before analyzing the data, we checked whether the articulatory parameters passed the normality test, which is mandatory for using the Analysis of Variance (ANOVA). First we made histograms and quartile-quartile (QQ) plots of the data, in order to visualize the data, and then, ran a Shapiro-Wilk normality test to check whether the data is statistically equal to a normal distribution. If the sentences did not pass the normality test, we used the Kruskal-Wallis test followed by a Pairwise Wilcoxon test with Bonferroni as the adjustment method. The Wilcoxon test was used to observe the differences among rates. In case the data passed the normality test, we used an Analysis of Variance (ANOVA) followed by a Tukey post-hoc test to observe the differences among rates.

The articulatory analyses also confirmed that the speakers produced at least two distinct speech rates for each sentence with just one exception. To do so, a Kruskal-Wallis test with articulatory VV duration as the dependent variable and rate as a factor was run. Statistical results have shown a consistent difference among speech rates for all sentences (SENTENCE 1: SP:  $S \neq F$ , MG:  $S \neq N \neq F$ ; SENTENCE 2: SP: no differences among rates; MG:  $(S = N) \neq F$ ; SENTENCE 3: SP:  $S \neq F$ , MG:  $(S = N) \neq F$ ; SENTENCE 4: SP:  $(S = N) \neq F$ , MG:  $(S = N) \neq F$ ).

As we found in previous works, there was a decrease of standard deviation with speech rate increase as can be seen in tables 2 and 3. By now, we are very confident that the decrease of standard deviation is one of the reorganizations of speech data caused by rate influence. In all cases, the smallest standard deviation was always found at the fast rates.

Table 2: Statistical results for the standard deviation of the articulatory VV duration with speech rate increase. Bold data stands for subject SP, and the other data refers to subject MG; S = slow rate, N = normal rate; F = fast rate; n.s. = non significant.

Sentence	ANOVA or Kruskal-Wallis	p <	Tukey or Wilcox
1	<b>n.s.</b> F (2,9) = 28.28	<b>n.s.</b> 0.000133	<b>S = N = F</b> S $\neq$ N $\neq$ F
2	<b>F (2,18) = 45.82</b> F (2,9) = 18.68	<b>8.10<sup>-8</sup></b> 0.00063	<b>S <math>\neq</math> N <math>\neq</math> F</b> (S = N) $\neq$ F
3	<b>K-W = 6.7048</b> F (2,20) = 14.39	<b>0.036</b> 0.00014	<b>S <math>\neq</math> (N = F)</b> S $\neq$ (N = F)
4	<b>F (2,9) = 6.372</b> K-W = 8.7411	<b>0.02</b> 0.01265	<b>S <math>\neq</math> N <math>\neq</math> F</b> S $\neq$ (N = F)
All	<b>K-W = 19.936</b> K-W = 13.331	<b>0.00047</b> 0.00127	<b>(S = (N) <math>\neq</math> F</b> S $\neq$ (N = F)

Table 3: Averages of the standard deviation of the articulatory VV duration with speech rate increase. Bold data stands for subject SP. The other data is for subject MG.

Sentence	Slow	Normal	Fast
1	<b>114.6</b> 219.2	<b>102.3</b> 145.5	<b>79.0</b> 74.1
2	<b>161.2</b> 62.3	<b>131.4</b> 53.9	<b>103.7</b> 29.0
3	<b>89.0</b> 59.6	<b>59.2</b> 49.5	<b>57.8</b> 32.9
4	<b>152.1</b> 121.4	<b>185.9</b> 69.3	<b>108.1</b> 67.2
All	<b>137.5</b> 115.0	<b>124.6</b> 70.8	<b>83.8</b> 50.5

## 5. OVERALL EFFECTS OF RATE ON DURATION AND KINEMATICS

The present section analyzes the effects of speech rate increase on the articulatory duration and kinematics. The aim of this section is to find articulatory patterns which correspond to speech rhythm restructurings.

As we have shown in Meireles and Barbosa (2014), our results suggest that speech rate increase tends to strengthen the right-headedness characteristic of BP (see tables 4 and 5). For sentence 1, with duration as a function of the articulatory VVs [afi] and [ai], the greatest phrasal prominence occurred at [afi] for slow (both subjects) and normal rates (both subjects), and there was no statistical difference between the VVs at [ai] for fast

rate (only for subject MG). For sentence 2, with duration as a function of the articulatory VVs [aim] and [afi], the greatest phrasal prominence occurred at [aim] for slow (both subjects) and normal rates (only for subject MG. Subject SP had a prominence at [afi] for this rate), and at [afi] for fast rate (significant difference for subject SP and n.s. for subject MG). For sentence 3, with duration as a function of the articulatory VVs [ab] and [ap], the greatest prominence at the slow rate occurred at [ap] for subject SP and at [ab] for subject MG. For the normal and normal rates, the greatest phrasal prominence occurred at [ap] for both subjects. For sentence 4, with duration as a function of the articulatory VVs [epR] and [aI], there was no statistical difference among rates for subject MG but with decreasing value of F with speech rate increase. The greatest phrasal prominence for subject SP in this sentence for all sentences was at [epR].

Table 4: Statistical results (ANOVA or Kruskal-Wallis) of the duration as a function of the articulatory VVs shown in table 5. Bold data stands for subject SP, and the other data refers to subject MG. n.s. = non significant.

Sent	Slow	Normal	Fast
1	<b>K-W = 8.3077, p &lt; 0.004</b> F(1,4) = 15.86, p < 0.017	<b>K-W = 8.3077, p &lt; 0.004</b> F(1,10) = 72.68, p < 0.00007	<b>K-W = 9, p &lt; 0.0028</b> n.s.
2	n.s. F(1,8)=85.69, p<0.000016	<b>F (1,4) = 53.02, p &lt; 0.0019</b> F(1,6) = 22.61, p < 0.0032	<b>K-W = 12.789, p &lt; 0.00035</b> n.s.
3	<b>F (1,8) = 12.6, p &lt; 0.008</b> F (1,6) = 5.396, p < 0.0593	<b>K-W = 3.8571, p &lt; 0.0496</b> F (1,6) = 6.032, p < 0.0494	n.s. F(1,12) = 15.04, p < 0.0022
4	<b>F(1,4) = 55.02, p &lt; 0.0018</b> n.s.	<b>F(1,4) = 22.19, p &lt; 0.010</b> n.s.	<b>F (1.14) = 18.9, p &lt; 0.00074</b> n.s.

Table 5: Mean and standard deviation of the articulatory VVs shown below. Bold data stands for subject SP. The other data is for subject MG. Underlined data means the VV where the phrasal prominence occurs.

Sent	Slow	Normal	Fast
1	<b>afl(460,73), aI (127, 9)</b> afl(335,25), aI(220,43)	<b>afl(383,58), aI (123,9)</b> afl(294,26), aI(178,20)	<b>afl(301,73), aI (111,10)</b> afl(186,11), aI(178,17)
2	<b>aIm(452,45), afl(430,83)</b> aIm(732,79), afl(385,29)	<b>aIm(296,35), afl(448,7)</b> aIm(544,78), afl(355,14)	<b>aIm(251,24), afl(398,27)</b> aIm(285,92), afl(272,15)
3	<b>ab (347,54), ap (548,115)</b> ab (379,25), ap (339,24)	<b>ab (333,19), ap (578,16)</b> ab (307,21), ap (344,21)	<b>ab (283,26), ap (363,93)</b> ab (208,19), ap (262, 32)
4	<b>epR (270,33), aI (115,15)</b> epR (237,54), aI (227,39)	<b>epR (213,16), aI(127,27)</b> epR (182,15), aI (175,25)	<b>epR (134,12), aI (109,11)</b> epR (138,9), aI (147,7)

## 6. ARTICULATORY-RHYTHMIC REORGANIZATIONS

So far we have presented general articulatory changes due to speech rate increase, but no relation of these changes to rhythmic restructurings of speech. Therefore, we intend to investigate in this section whether there is any articulatory parameter related to the predicted phrasal prominences due to strong syntactic boundaries (cf. table 1). In order to do so, six parameters will be individually observed for sentences 1 and 2 (only for subject MG): 1) acceleration duration; 2) y-extremum; 3) constriction displacement; 4) peak/valley velocity; 5) Articulatory gestural (henceforth AT) duration; 6) proportional time-to-peak/valley-velocity.

The analysis of the evolution of VV duration across rates did not show any rate differences at the phrasal prominences, i.e., there was no rhythmic restructuring for sentences 3 and 4 for any of the subjects. It was also not possible to analyze the rhythmic restructurings for subject SP, since she did not produce the same VVs across rates for sentences 1 and 2.



We observed in sentence 1 (fig. 2) a rhythmic restructuring of the utterance at fast rates. At slow and normal rates the VV duration contour was almost the same. Nevertheless, at fast rates, according to table 1, there is no peak at the articulatory VV [os].

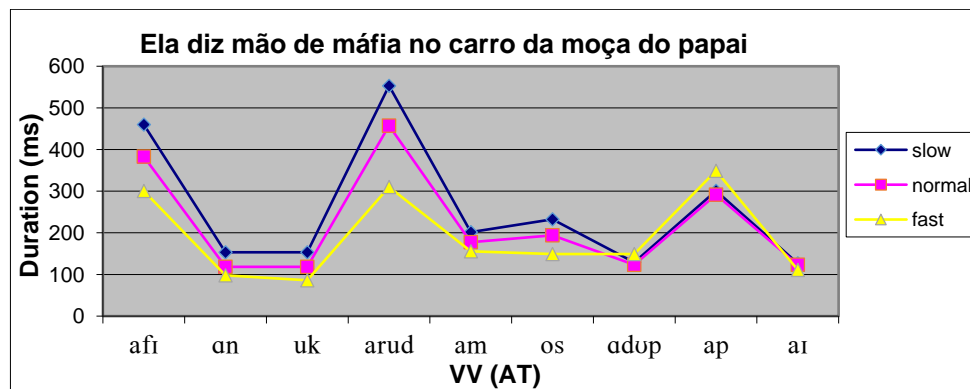


Figure 2. Rhythmic restructuring for the sentence “Ela diz mão de máfia no carro da moça do papai”, in which the accentual prominence at the VV [os] at slow and normal rates is reorganized at fast rates. Subject MG.

Focusing on the gestures in the VV [os] (sentence 1), we have tested whether the proposed articulatory hypotheses proposed in Meireles (2009) could explain such rhythmic restructuring. A one-way ANOVA or Kruskal-Wallis test with articulatory parameters as a function of rate was then run and its results are presented in table 6.

Table 6. One-way ANOVA or Kruskal-Wallis test for the articulatory dependent variables related to the reorganized accentual prominence in [os] of the sentence “Ela diz mão de máfia no carro da moça do papai”.

Sentence 1				
Dependent variables	[o]		[s]	
	ANOVA or Kruskal-Wallis	p <	ANOVA or Kruskal-Wallis	p <
1 - Acceleration Duration	n.s.	n.s.	K-W = 6.2692	0.044
2 - Y-extremum	n.s.	n.s.	n.s.	n.s.
3 - Constriction Displacement	n.s.	n.s.	F (2,9) = 13.543	0.002
4 - Velocity (modulus)	n.s.	n.s.	F (2,9) = 6.604	0.02
5 - AT Duration	n.s.	n.s.	K-W = 7.6154	0.023
6 - Proportional T-P/V-Velocity	n.s.	n.s.	n.s.	n.s.

Table 6 shows that the hypothesis of articulatory reduction with speech rate increase was not found for the dependent variables of the gesture [o]. For this gesture, the only confirmed hypothesis was the constant proportional time-to-peak-velocity. On the other side, for the consonantal gesture [s], the hypotheses were confirmed, except the diminishment of the y-extremum. These results suggest that, in case of speech rhythmic restructurings, articulatory reorganizations related to speech rate increase follow the hypotheses proposed in Meireles (2009). Yet, they may be not manifested on all gestures of the respective VVs.

For sentence 2 (fig. 3), a speech rhythmic restructuring occurred from slow to normal rate for the articulatory unit [aIzdZI<sub>p</sub>], though not manifested at fast rates. Here, the durational contour does not help to visualize this restructuring, since neither z-score nor smoothed z-score was employed in the articulatory analysis.

Based on the articulatory unit [aIzdZI<sub>p</sub>], we have analyzed, as for sentence 1, if the proposed articulatory hypotheses could explain such rhythmic restructuring. A one-way ANOVA or Kruskal-Wallis test with dependent articulatory variables as a function of rate was run and its results are in table 7.

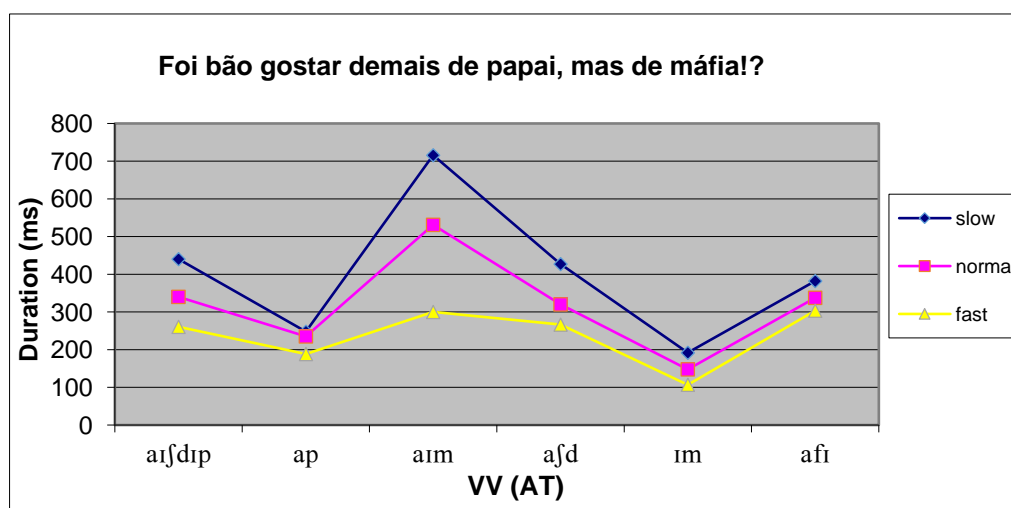


Figure 3. Rhythmic restructuring for the sentence “Foi b3o gostar demais de papai, mas de m3fia”, in which the accentual prominence at the articulatory unit [aIzdZI<sub>p</sub>] at slow and normal rates is reorganized at fast rates.

Table 7. One-way ANOVA or Kruskal-Wallis test for the articulatory dependent variables related to the reorganized accentual prominence in [aIzdZI<sub>p</sub>] of the sentence “Foi b3o gostar demais de papai, mas de m3fia”.

Sentence 2				
Dependent variables	[aIzdZI]		[p]	
	ANOVA or Kruskal-Wallis	p <	ANOVA or Kruskal-Wallis	p <
1 - Acceleration Duration	K-W = 6.7846	0.034	K-W = 6.40	0.04
2 - Y-extremum	n.s.	n.s.	n.s.	n.s.
3 - Constriction Displacement	F (2,9) = 5.6169	0.027	F (2,9) = 4.2517	0.05
4 - Velocity (modulus)	F (2,9) = 5.5608	0.027	n.s.	n.s.
5 - AT Duration	n.s.	n.s.	F (2,9) = 40.683	0.00004
6 - Proportional T-P-Velocity	n.s.	n.s.	n.s.	n.s.

Table 7 shows that, differently from sentence 1, there was confirmation of the articulatory reduction hypothesis for variables 1, 3 and 4 and the constant proportional time-to-peak-velocity (variable 6) for the jaw gesture of [aIzdZI<sub>p</sub>]. For the [p] gesture, on the other hand, hypotheses were confirmed for factors 1,3, 5 and 6.

These results confirm that all gestures composing an articulatory jaw unit need at least some articulatory parameters to be changed in the VV to be restructured.

After investigating whether the proposed articulatory hypotheses work in the VVs restructured by the speech rate increase, we inquired whether there should be significant reductions in the studied articulatory parameters only for the cases in which the VVs were restructured. To verify this question we ran one-way ANOVAs for the dependent articulatory variables as a function of rate for all sentences.

Regarding the acceleration duration, almost all sentences showed a acceleration duration reduction for all gestures, independently of its phrasal position (SUBJECT MG: sentence 1:  $K-W = 27.341$ ,  $p < 1.156 \cdot 10^{-6}$ ; sentence 2:  $K-W = 32.694$ ,  $p < 7.955 \cdot 10^{-8}$ ; sentence 3:  $K-W = 49.581$ ,  $p < 1.713 \cdot 10^{-11}$ ; sentence 4:  $K-W = 67.467$ ,  $p < 2.237 \cdot 10^{-15}$ ; SUBJECT SP: sentence 1:  $K-W = 16.679$ ,  $p < 0.00024$ ; sentence 2:  $K-W = 8.2856$ ,  $p < 0.016$ ; sentence 3 = n.s.; sentence 4 = n.s.). Nevertheless, a fact called our attention: the acceleration duration contour for the gestures [o] and [s] in sentence 1 (subject MG, ellipsed in figure 4(a)) and the articulatory units [aIzdZI] and [p] in sentence 2 (subject MG, ellipsed in figure 4(b)) was different from the other gestures. These durational contours are in agreement with the phrasal boundaries predicted in table 1. Therefore, it is expected that some derivations in the articulatory signal could reveal greater modifications in the gestures at or close to a phrasal boundary.

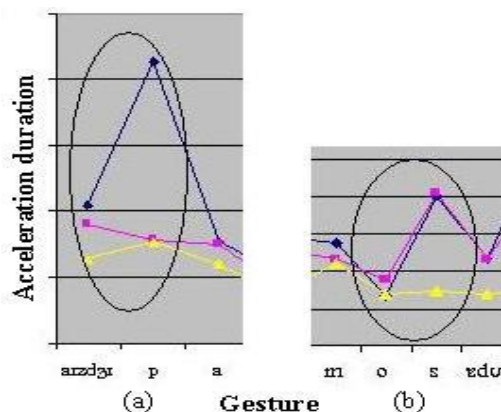


Figure 4. Gestures in sentences 1 (b) and 2 (a) from subject MG representing different acceleration duration patterns with speech rate increase. Scales are different for the two sentences.

Regarding the articulatory parameters y-extremum, peak/valley velocity, and proportional time-to-peak/valley-velocity, no special pattern was found for the gestures at or close to a phrasal boundary for both subjects.

Similarly to the acceleration duration, the constriction displacement parameter showed no clear influence of the phrasal boundary in the articulatory modifications of the gestures. For all gestures and subjects there was a consistent smaller constriction displacement with speech rate increase, though not statistically significant. Yet, the gestures in the VV [os] (sentence 1, subject MG) showed a different pattern than the other gestures. For the gesture [s] there was no peak at the fast rate (figure 5) for the constriction displacement. It is worth to remind that it is exactly at this place where a speech rhythmic restructuring from normal to fast rate occurred. So, also the constriction displacement parameter, if properly normalized, may work as a cue for identifying phrasal prominences along the utterance.

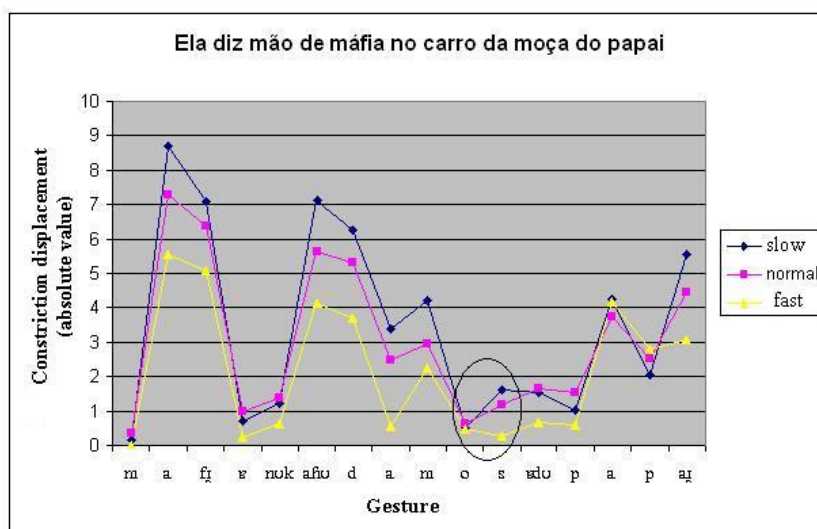


Figure 5. Constriction displacement for sentence 1 (subject MG) showing a different pattern for the gesture [s] (see elipsis).

Similarly to the constriction displacement, the articulatory VV duration was significantly different for all gestures in the sentences (SUBJECT MG: sentence 1: K-W = 20.288,  $p < 3.931 \cdot 10^{-5}$ ; sentence 2: K-W = 13.04,  $p < 0.0015$ ; sentence 3: K-W = 46.572,  $p < 7.711 \cdot 10^{-11}$ ; sentence 4: = K-W = 43.548,  $p < 3.496 \cdot 10^{-10}$ ; SUBJECT SP: sentence 1: K-W = 15.968,  $p < 0.00035$ ; sentence 2: K-W = 7.6405,  $p < 0.022$ ; sentence 3: K-W = 6.6559,  $p < 0.036$ ; sentence 4: n.s.). However, a distinct pattern was found for the gestures [o] and [s] for the MG data in sentence 1 (fig. 6, see the ellipsis), in which at fast rates there was an articulatory pattern change. Again, this is the place where the rhythmic restructuring was found. Because of this we suggest that an analysis of the articulatory VV duration using z-score and smoothing, as in our acoustic studies, could also provide articulatory evidence of the position of phrasal prominences.

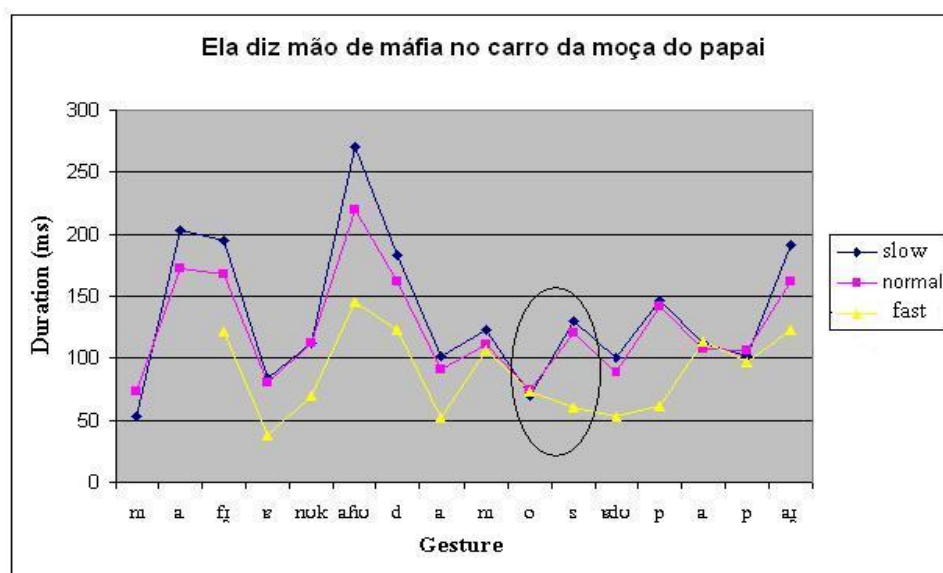


Figure 6. VV duration of the gestures in sentence 1 (subject MG) showing a different pattern for the gestures in [os] (see ellipsis).

Summing up, the articulatory analyses presented in this section have shown that speech rate affects uniformly all gestures in the utterances disregarding their phrasal stress

position. Yet, there was evidence that the factors acceleration duration (stiffness of the spring in the task dynamic model), constriction displacement and AT duration can, if normalized with, for instance, z-score and smoothing, provide articulatory evidence for the speech rhythmic restructurings. It is worth to recall that in the DSR model these restructurings are analysed through methods applied to the acoustic duration of VV units.

## 7. CONCLUSION

The results of this articulatory study suggest that speech rate increase affects all gestures in a sentence disregarding their phrasal position. Nevertheless, future work is needed since acceleration duration, constriction displacement and AT duration seem, though not conclusive, to reflect the rhythmic restructurings found on the articulatory side. Finally, main results have shown that rhythmic structure variation is modified gradually with speech rate increase, i.e., quantitative aspects of speech are acting to modify speech rhythm, showing how a dynamical systems approach to language is perfectly adequate to linguistic descriptions.

## Acknowledgements

This work was supported by the Espirito Santo Research Foundation (FAPES, grant 221/2017). Many thanks are due to USC Phonetics Laboratory, Haskins Laboratories, Dani Byrd and Sungbok Lee, for help with the articulatory design.

## REFERENCES

- ALBANO, E. C. & MOREIRA, A. A. (1996). Archisegment-based letter-to-phone conversion for concatenative speech synthesis in Portuguese. *Proceedings ICSLP'96*, v.3, 1708-1711.
- BARBOSA, P. A. (1994). *Caractérisation et génération automatique de la structuration rythmique du français*, PhD Thesis, ICP/INP in Grenoble, France.
- BARBOSA, P. A. (2002). Explaining Cross-Linguistic Rhythmic Variability via a Coupled-Oscillator Model of Rhythm Production. In BEL, B. & MARLIEN, (eds.). *Proceedings of the Speech Prosody 2002 Conference*, 11-13 April, Aix-en-Provence: Laboratoire Parole et Langage, 163-166.
- BARBOSA, P. A. (2006). *Incursões em torno do ritmo da fala*. Campinas, São Paulo: Pontes Editores, Fapesp.
- BARBOSA, P. A. (2007). From syntax to acoustic duration: a dynamical model of speech rhythm production. *Speech Communication*. 49 (1-2), 725 - 742.
- DAUER, R. M. (1983). Stress-timing and syllable-timing re-analysed, *Journal of Phonetics* 11, 51-62.
- FRAISSE, P. (1974). *La psychologie du rythme*, Paris: Presses Universitaires de France.
- KELSO, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge: MIT Press. [Paperback edition, 1997].
- KELSO, J. A. S., SALTZMAN, E. L., & TULLER, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics* 14, 29-59.
- LEHISTE, I. (1970). *Suprasegmentals*. MIT Press, Cambridge.
- LEHISTE, I. (1978). Temporal organization and prosody: perceptual aspects. In Joint Meeting of A.S.A. and A.S.J.1, Honolulu, 1-17.
- MARCUS, S. M (1976). *Perceptual-centres*. Unpublished PhD Thesis, Cambridge University.
- MASSINI, G. (1991). *A duração no estudo do acento e do ritmo do português*. Campinas: Iel/Unicamp, MD.
- MASSINI-CAGLIARI, G. (1992). *Acento e Ritmo*. São Paulo: Contexto.
- MATTOSO CÂMARA Jr., J. (1970). *Estrutura da língua portuguesa*. Petrópolis: Vozes.
- MEIRELES, A. R. (2009). *Self-organizing rhythms in Brazilian Portuguese: speech rate as a system perturbation*. Germany: VDM Verlag, 2009.

- MEIRELES, A. R. & BARBOSA, P. A. (2008). Speech rate effects on speech rhythm. In: *Speech Prosody 2008 Conference, 2008, Campinas. Proceedings of the Speech Prosody 2008 Conference*. Campinas: RG. v.1. p.327 – 330, 2008.
- MEIRELES, A. R. & BARBOSA, P. A. (2014). Articulatory Reorganizations of Speech Rhythm due to Speech Rate Increase in Brazilian Portuguese. In: *Speech Prosody 2014, 2014, Dublin, Irlanda. Proceedings of Speech Prosody 2014, 2014*.
- MEIRELES, A. R. & GAMBARINI, V. P. (2012). Rhythm Typology of Brazilian Portuguese dialects. In: *Speech Prosody 2012, 6h International Conference on Speech Prosody, 2012, Shanghai. Proceedings of Speech Prosody 2012, 2012.v. 1*.
- MEIRELES, A. R., TOZETTI, J. P. & BORGES, R. R. (2010). Speech rate and rhythmic variation in Brazilian Portuguese. In: *Speech Prosody 2010 Conference, 2010, Chicago. Proceedings of the Speech Prosody 2010 Conference*. Chicago: RG. v.1. p.1-4, 2010.
- MORTON, J., MARCUS, S. & FRANKISH, C. (1976) Perceptual-centers (P-centers), *Psychological Revue* 83 (5), 405-408.
- PERKELL, J., COHEN, M., SVIRSKY, M., MATTHIES, M., GARABIETA, I. & JACKSON, M. (1992). Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *J. Acoustical Soc. America* 92, 3078-3096.
- POMPINO-MARSCHALL, B. (1991). *The syllable as a prosodic unit and the so-called P-centre effect*. Munique, Alemanha.
- SILVA, J. de A. & MEIRELES, A. R. (2011). Sociophonetic study of the Capixaba speech rhythm. *Journal of Speech Sciences* 1(1):3-13. 2011. Available at: <http://www.journalofspeechsciences.org>.
- TIEDE, M. K., VATIKIOTIS-BATESON, E., HOOLE, P. & YEHIA, H. (1999). Magnetometer data acquisition and analysis software for speech production research. *ATR Technical Report TR-H 1999*. Kyoto, Japan: ATR Human Information Processing Labs.
- WOODROW, H. (1951). Time Perception. In STEVENS, S. S., editor, *Handbook of Experimental Psychology*, 1224-1236. Wiley, New York.

Recebido: 18/6/2019

Aceito: 10/2/2021

Publicado: 15/2/2021