9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

# The Interplay between Speech and Breathing across three Brazilian Portuguese Speaking Styles

*Plínio A. Barbosa*[1], *Sandra Madureira*[2]

[1]Instituto de Estudos da Linguagem, University of Campinas, Brazil
[2]LIACC, Catholic University of São Paulo, Brazil

pabarbosa.unicampbr@gmail.com, sandra.madureira.liaac@gmail.com

## Abstract

This paper introduces the first study on the coordination between breathing and speech in Brazilian Portuguese. Four subjects (two males and two females) were recorded by using the RespTrack device, which allows the measurement of the breathing cycle and the simultaneous recording of speech. In order to shed light on the interplay between the two activities, three speaking styles were investigated: reading, narration and commentary of story characters. Five parameters were measured: breath cycle duration, duration and amplitude of inhalation, duration to speech onset from onset of breath cycle or from peak of inhalation. Results showed stylistic and gender differences: coordination differs between reading and the other two styles; breath cycles are shorter by 1 to 1.5 second in reading; commentary does not differ between the two genders for four out of five parameters; differences between males and females are found in four out of five parameters measured, excepted breath cycle duration, and are interpreted as related to differences in thoracic volume.

**Index Terms**: speech, breathing, speaking style

## 1. Introduction

The interplay between pausing for breathing and speech has been a matter of study from several decades since Stetson's seminal work on motor phonetics relating the production of syllables to chest pulses [1]. As far as 1965, Henderson and colleagues [2] claimed that "reading breaths are taken exclusively at grammatical junctures". This suggests that the preplanning of the respiratory activity would be dictated by the linguistic/prosodic organisation of the read utterance. This is later on confirmed by the study of Grosjean and Collins [3], which pointed out that, "at slow and normal rates, speakers accommodate their need to inhale to the preplanned pause patterns."

A recent study by Fuchs and colleagues [4] showed that inhalation is not the whole story by signalling that the "thoracic/abdominal volume change during sentence production is only partly anticipated via inhalation depth" and that some speakers also make use of the expiratory reserve volume for producing longer sentences, suggesting that "there are some limits on how many words/syllables can be anticipated in a preceding pause." This was also pointed out by Denny [5]: "variability in inspired lung volume prior to speech is only partially accounted for by speech-related concerns such as the length and loudness of the planned utterance."

Włodarczak and Heldner [6], in their study of semi-spontaneous speech breathing patterns in short and long utterances in conversational style, point out the relevance of considering communication needs besides lung volume and respiratory constraints interfering with speech respiratory patterns.

Contrasting spontaneous and reading styles, Henderson and colleagues [2] claimed that these two styles differ since breaths are taken in circa 77 % of the inter-phonatory intervals in the latter as against 34 % in the former (the rest filled by hesitations pauses). For the authors, "pauses in reading serve a different function from pauses in spontaneous speech".

Even if Stetson's Motor Phonetics is a defence of the syllable as a basic unit of speech associated to the so-called chest pulses, in chapters 3 as well as in chapters 6 and 7 chapter of his book (see [7] for a recent review), Stetson develops an account of breath groups whose instantiation is associated with the activity of the abdominal muscles. For doing so, he presents a technique that captures muscle activity at the surface of the abdomen besides muscle activity at the thoracic cage surface. These studies are completed with other sensors to measure pressure at the trachea and the intra-oral cavity. His finding that the ripples along the breath group traces are caused by syllable train production indicate that breathing activity and speech production are related.

Our main goal is to compare the interplay between breathing and speech across three speaking styles, reading, narration and commentary, three styles with very different cognitive and expressive loads. As a specific goal, we investigate which parameter of the breath cycle is better correlated to style.

Our hypotheses are: (1) due to diferences in air cavity volumes we expect diferences across gender affecting the way women and men coordinate breathing and speech; (2) since the commentary style has a higher cognitive load, we expect deeper inhalation in commentary; (3) frequency of inhalation is higher in reading due to task planning specificities.

## 2. Methodology

Three tasks were carried out by four Brazilian subjects (two males and two females), simultaneously recorded with a microphone and RespTrack device. Subjects were one undergrad (20 years) and three grads (from 23 to 27 years) from the first author affiliation (University of Campinas). For the reading task (RE), the four subjects read an excerpt of 670 words from the "Disastrous Monk" story on the origin of the Belém pastries. For the narration task (NR) they told the story they had previously read and they expressed their commentaries (CT) about the characters in the story, the prior general and the main character, the monk Manuel. As the prior general is a critical, rigid, unfriendly and demanding person and the monk is a warm-hearted, sensitive, friendly and enthusiastic, negative and positive commentaries were expected to be expressed.

Two combined devices to monitor the four speakers' respiratory movements of the thoracic cage and the abdomen during the aforementioned tasks were used.

Each subject's breathing during reading, narration and com-

 10.21437/SpeechProsody.2018-75

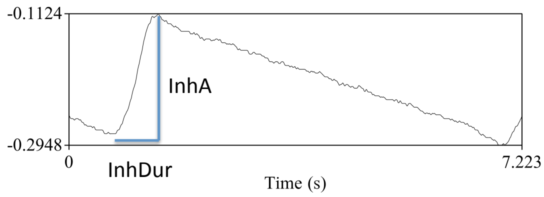


Figure 1: *Breath cycle showing inhalation duration (InhDur) and amplitude (InhA)*

mentary was recorded by the Respiratory Inductance Plethysmography (RIP) apparatus, which measures changes in cross-sectional area of the rib cage and the abdomen by means of two belts placed at the level of the armpits and the navel.

The signal from the belts was recorded and processed by the RespTrack device, designed and built at Stockholm University. The signals from the thorax and the abdomen were summed up to obtain a common trace of the movements of the two cavities.

The simultaneous recording of the speech of each subject was done with a Sony ECM MS907 condenser microphone. The summed-up breathing signals from the four subjects in the three tasks, carried out at a stand position, were submitted to a 5-point moving average smoothing technique to eliminate ripples not associated to any crucial prosodic function. The synchrony button of the RespTrack device was used for ensuring the synchrony between the breath and the speech signals by pressing it some short time before the start of phonation. The synchrony signal and the microphone signal outputs were recorded together by using the two inputs of the Behringer URA 222 sound card. The speech signal and the breath signal were saved as two channels of a stereo WAV audiofile in Praat [8].

The type of breath signal obtained is depicted in Fig. 1, which illustrates a breath cycle from the male speaker FA during the reading task. Inhalation duration (InhDur) and amplitude (InhA) are shown in the figure. The inhalation onset is defined as the minimum of the cycle just before the signal raise (leftmost point of horizontal segment in the figure) and the inhalation peak as the maximum of the cycle (upper point of the vertical segment).

In order to evaluate the five aforementioned hypotheses we measured the following breath signal parameters:

- The breath group duration from two consecutive inhalation onsets DurBG;
- The duration of the inhalation phase DurInhalation;
- The distance of the breath group onset to following speech activity DurSpeechOnset;
- The distance of the inhalation peak and the start of following speech activity DurIntoSP;
- The inhalation amplitude AmpInhalation normalised as a ratio in reference to the maximum amplitude.

Due to the non-normality and heterocedasticity of the original residuals for the five 2-Way ANOVA models with STYLE and GENDER as factors, we used the Scheirer-Ray-Hare (SHR) extension [9, p. 175] to Kruskal-Wallis test instead. For doing so, a formula proposed by Poisot [10] was used. Significance levels both for the general model and for Wilcoxon post hoc tests (used for the STYLE factor, which have three levels) were set to $0.05$. A test for each gender was run by combining data from the two subjects of each gender in order to ensure enough test power. Results are presented in the next section.

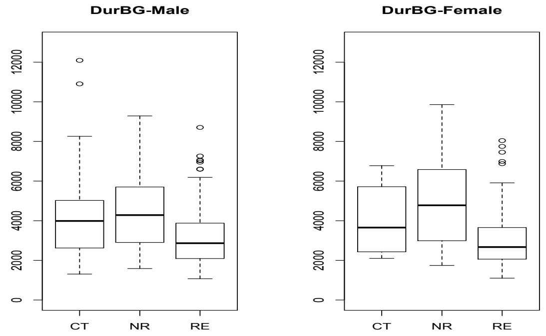


Figure 2: *BoxPlots for DurBG according to style and gender.*

## 3. Results

Table 1 presents the means and standard-deviations in milliseconds (ms) for parameter DurBG for each style according to gender. There is no effect of GENDER, but of STYLE (H $= 54.2$, p $< 10^{-5}$).

Table 1: *Means (standard-deviations) in ms for DurBG for all speakers and styles. Inequalities indicate the higher/lesser significant parameter mean in a style or style grouping (the same holds for the other tables).*

| Gender | Style | Mean (SD) |
|---|---|---|
| | RE | 3169 (1415) |
| Male | NR | 4478 (2044) |
| | CT | 4266 (2341) |
| | RE< (NR=CT) | |
| | RE | 2976 (1306) |
| Female | NR | 5015 (2196) |
| | CT | 4119 (1786) |
| | RE< (NR= CT) | |

The figures reveal that narration and commentary have longer durations in the two styles for both genders. In comparison with reading, the average means for narration and commentary are longer by $1.2$ seconds in males and by $1.6$ seconds in females, as shown in Fig. 2. Because the average breath group duration is the inverse of the inhalation rate, this rate is the same for both genders, because gender is not a significant factor.

Fig. 2 presents the boxplots for the same parameter DurBG.

Table 2 presents the means and standard-deviations in ms for parameter DurInhalation for each style (H $= 22.9$, p $< 10^{-5}$), according to gender (H $= 13.1$, p $< 0.003$). Males do not differ in this parameter across styles, but females do, since analysis of the females' narration revealed a marginal, though significant higher mean than reading. Concerning the reading style, a significant difference for GENDER was found, males' inhalation duration (DurInhaltion) presenting 70 ms more than those of females. For females, commentary does not differ from narration and reading, which is possible due to the higher vari-

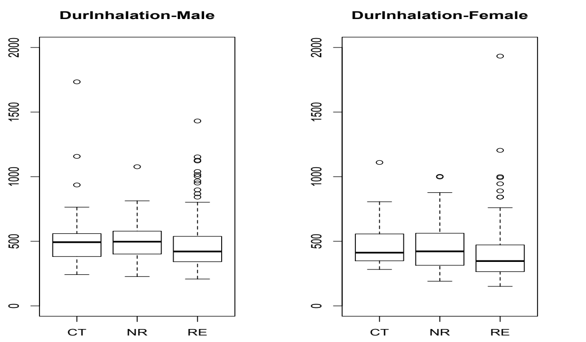


Figure 3: *BoxPlots for DurInhalation according to style and gender.*

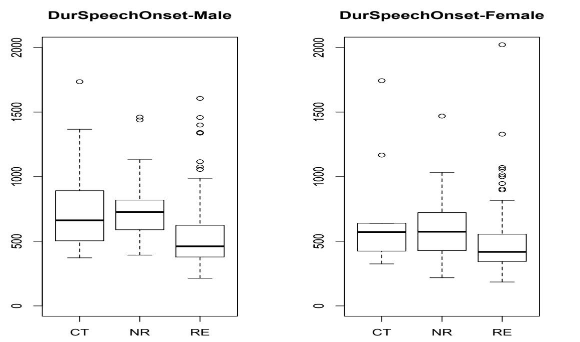


Figure 4: *BoxPlots for DurSpeechOnset according to style and gender.*

ability of commentary.

Table 2: *Means (standard-deviations) in ms for DurInhalation for all speakers and styles.*

| Gender | Style | Mean (SD) |
|---|---|---|
| | RE | 472 (203) |
| Male | NR | 508 (183) |
| | CT | 575 (368) |
| | ns | |
| | RE | 402 (214) |
| Female | NR | 467 (205) |
| | CT | 509 (260) |
| | RE< NR, (CT=NR), (CT=RE) | |

Fig. 3 presents the boxplots for the same parameter DurInhalation.

Table 3 presents the means and standard-deviations in ms for parameter DurSpeechOnset for each style (H $= 55.0$, p $< 10^{-5}$), according to gender (H $= 15.0$, p $< 0.0002$). Both males and females differ across styles with lower values found for reading. There is a marginal significant difference for GENDER, in both reading and narration style, males having higher values in both cases.

Table 3: *Means (standard-deviations) in ms for DurSpeechOnset for all speakers and styles.*

| Gender | Style | Mean (SD) |
|---|---|---|
| | RE | 531 (236) |
| Male | NR | 758 (269) |
| | CT | 807 (451) |
| | RE< (NR=CT) | |
| | RE | 490 (300) |
| Female | NR | 604 (239) |
| | CT | 698 (432) |
| | RE< (NR= CT) | |

Fig. 4 presents the boxplots for the same parameter DurSpeechOnset.

Table 4 presents the means and standard-deviations in ms for parameter DurIntoSP for each style (H $= 98.2$, p $< 10^{-5}$). A GENDER x STYLE interaction is present (H $=$ 10.9, p $<$ 0.005). Both males and females differ across styles having lower values for reading (for females, commentary does not differ from narration and reading, which is possible due to the higher variability of commentary).

Table 4: *Means (standard-deviations) in ms for DurIntoSP for all speakers and styles.*

| Gender | Style | Mean (SD) |
|---|---|---|
| | RE | 60 (96) |
| Male | NR | 250 (142) |
| | CT | 232 (240) |
| | RE< (NR=CT) | |
| | RE | 88 (145) |
| Female | NR | 138 (121) |
| | CT | 189 (186) |
| | RE< (NR= CT), (CT=RE) | |

Fig. 5 presents the boxplots for the same parameter DurIntoSP.

Table 5 presents the means and standard-deviations in normalised units for parameter AmpInhalation for each style (H $= 41.3$, p $< 10^{-5}$) and GENDER (H $= 6.0$, p $< 0.02$). Both males and females differ across styles, significant lower values found for reading. For both genders, commentary does not differ from both narration and reading, which is possible due to the higher variability of commentary. Higher means are found for females.

Fig. 6 presents the boxplots for the same parameter AmpInhalation.

Effect sizes were computed for all variables according to gender. The variable that better explains style for males is DurIntoSP (26 % of effect size) and, for females, is DurBG (22 %).

## 4. Discussion

Genderwise, the following parameters differ: DurInhalation (for reading, 70 ms higher for males), DurSpeechOnset (for reading and narration, respectively 41 and 154 ms higher for males), DurIntoSP (for reading, higher for females by 28 ms

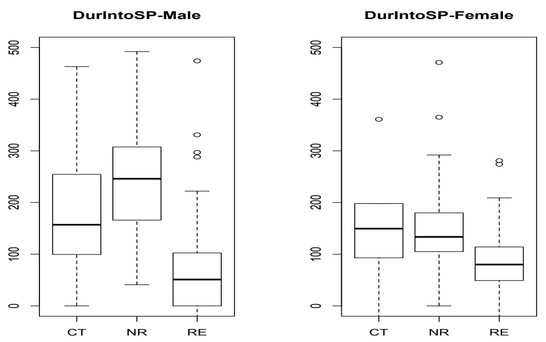


Figure 5: *BoxPlots for DurIntoSP according to style and gender.*

Table 5: *Means (standard-deviations) in normalised units for AmpInhalation for all speakers and styles.*

| Gender | Style | Mean (SD) |
|---|---|---|
| | RE | 0.37 (0.17) |
| Male | NR | 0.47 (0.20) |
| | CT | 0.49 (0.37) |
| | RE< (NR=CT), (CT=RE) | |
| | RE | 0.40 (0.20) |
| Female | NR | 0.66 (0.32) |
| | CT | 0.73 (0.38) |
| | RE< (NR= CT), (CT=RE) | |

and for narration, higher for males by 62 ms), AmpInhalation (circa 0.2 higher for females especially for narration and commentary). This means that, when reading, males have longer inhalations and take a bit more time to start speaking from the start of the breath cycle or from the peak on the inhalation period (in the case of narration). Higher inhalation peaks reached in lesser time were found in narration and commentary for females.

Stylewise, commentary does not differ between the two genders for four out five parameter means (AmpInhalation excepted). Reading presents lesser values in all parameters in comparison with narration and commentary. This means that the coordination between breathing and speech does not differ as far as narrating and commenting are concerned.

Breath cycles are shorter in reading by 1 second for males and 1.5 second for females in comparison to narration and commentary, but both genders do not differ in this variable. Rate of breathing appears not to be relevant to distinguish the two genders, but the other four variables measured here are.

## 5. Conclusions

Our hypotheses concerning the diferences in air cavity volumes across gender and the presence of higher frequency of inhalation in reading and higher amplitude of inhalation in commentary were confirmed. The fact that reading differs from narration and commentary in relation to all of the five breathing parameters measured in this paper can be thought of as related to the kind of cognitive task undertaken by the speaker and to a particular kind of speech style. In reading a text, the speaker is focused in chunking speech into meaningful thought groups, facilitating both his speech production and the listeners decoding task. This is thought to influence breath productivity and distribution as it has been pointed out in the literature on speech and breathing, [2, 3, 6] to mention a few. The constraint on planning and executing a reading task tends to be greater. However, in speaking about what someone has read or experienced or in producing commentaries based on appraisals, the speaker is focused on remembering facts, judging and reporting them to the listener, that is, his focus is on both informing and impressing, or in other words, in communicating. Besides communicative goals and respiratory constraints as pointed out by [6], specific stylistic speaking demands are influential since these are thought to be interwoven in complex ways in speech-breathing mechanisms.

As far as our hypotheses are concerned: (1) differences across gender are found in four out of five parameters and are likely to be related to differences of thoracic volume; (2) commentary pooled with narration and digressed from reading; (3) frequency of inhalation is effectively higher in reading for both genders.

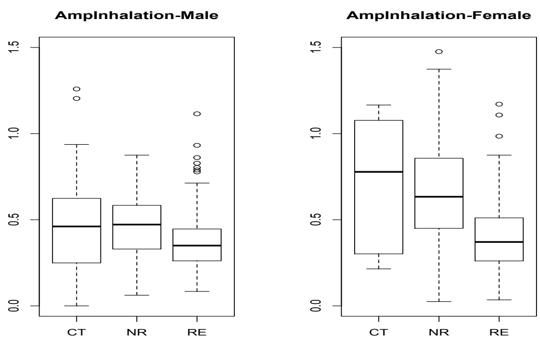


Figure 6: *BoxPlots for AmpInhalation according to style and gender.*

## 6. Acknowledgements

We thank our subjects and the three reviewers who help improve this paper. The first author also thanks the CNPq grant # $302657/2015 - 0$.

## 7. References

[1] R. Stetson, *Motor Phonetics: A Study of Speech Movements in Action*. Amsterdam: North Holland Publishing Co., 1951.

[2] A. A. Henderson, F. Goldman-Eisler, and A. Skarbek, "Temporal patterns of cognitive activity and breath control in speech," *Language and Speech*, vol. 8, no. 4, pp. 236–242, 1965.

[3] F. F. Grosjean and M. Collins, "Breathing, pausing and reading," *Phonetica*, vol. 36, no. 2, pp. 98–114, 1979.

[4] S. Fuchs, P. Hoole, D. Vornwald, A. Gwinner, H. Velkov, and J. Krivokapic, "The control of speech breathing in relation to the upcoming sentence," in *Proceedings of the $8^{th}$ International Seminar on Speech Production*, 2008, pp. 77–80.

[5] M. Denny, "Periodic variation in inspiratory volume characterizes speech as well as quiet breathing," *Journal of Voice*, vol. 14, no. 1, pp. 34–46, 2000.

[6] M. Włodarczak and M. Heldner, "Respiratory constraints in verbal and non-verbal communication," *Frontiers in*

*Psychology*, vol. 8, p. 708, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyg.2017.00708

[7] P. A. Barbosa, "R.H. Stetson, Motor Phonetics: A study of speech movements in action, $2^{nd}$ ed., Amsterdam, North Holland Publishing Co., 1951," *Phonetica*, vol. 74, pp. 255–258, 2017.

[8] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," http://www.praat.org/, 2016.

[9] C. Dytham, *Choosing and Using Statistics: A Biologist's Guide*. Chichester: John Wiley & Sons, 2011.

[10] T. Poisot, "R : test de Scheirer-Ray-Hare'," http://www.cafe-sciences.org/billets/r-test-de-scheirer-ray-hare/, 2008.