Speech–breathing interactions under physical load

Heather Weston

Leibniz-Centre General Linguistics (ZAS) and Humboldt-Universität zu Berlin, Germany
weston@leibniz-zas.de

Abstract

People frequently speak during physical activities, and this physical load has been found to have perceptible effects on the voice. Competition for the lungs is thought to play a key role in explaining why load affects speech, yet few studies have investigated the breath cycle itself. This dissertation aims to explore interactions between speech-breathing parameters and acoustic features of the speech signal during physical activity. To do this, acoustic and respiratory data were recorded from 48 female German speakers exercising at different intensities while performing different speech tasks. The range of data makes possible diverse analyses; some studies were planned pre-experiment but working with the data has prompted reconsideration of the original plan. The main challenge at this stage is combining the possible studies into a coherent whole.

Index Terms: human speech production, physical task stress, speech breathing, automatic stress recognition

1. Motivation and research aims

1.1. Speech under stress

This work fits within the research area of speech under stress, with stress defined as "the non specific response of the body to any demand" [1]. Cognitive and emotional demands, or *load*, have been found to have perceptible effects on the voice, including changes in pitch, loudness and speech rate [2]. Physical load has also been found to affect the voice, though it has been less widely researched, as seen in a recent review [3]. The most widely obtained finding is an increase in mean fundamental frequency. Some studies have investigated further parameters, including voice quality measures [4], jitter and shimmer [5], and articulation rate [6], but the description of speech under physical load is far from complete. It is also difficult to compare results between studies: methodologies are diverse in terms of speech tasks and physical load, and substantial interspeaker variability has been reported. Most importantly, although breathing is thought to play a key role in explaining how load affects spoken language, few studies have investigated speech–breathing interactions.

1.2. Speech breathing

Speech breathing is sensitive to linguistic structure. In read speech, almost all breath pauses occur at syntactic/grammatical breaks [7], and the volume of air inspired is correlated with the length of the upcoming utterance [8]. Results are similar for spontaneous speech [9].

Speech breaths are characterized by short, fast inhalations and long, controlled exhalations, during which speech is produced. Under physical load, however, breaths are taken more frequently and a greater volume of air is inspired. These changes to general breathing patterns are predicted to have different "knock-on effects" on spoken language. First, changes in airflow/volume may affect vocal fold behavior in diverse ways, resulting in changes in the acoustic signal. And second, an increase in breath frequency may reduce the time available for speech in between breaths. In other words, fewer words can be produced per breath. This is predicted to lead to shorter speech chunks and/or to disrupt the coordination of breath pauses with grammatical breaks, as has been found in [10].

1.3. Aims of the dissertation project

The aims of this doctoral research are: 1) to investigate how changes in the breath cycle affect speech under physical load, and 2) to publish a corpus of speech and respiratory data to enable further research on speech under physical load. Prior to data collection, several studies were planned using the theoretical, physiologically based hypothesis that increased respiratory drive affects the three components that generate speech: the lungs, the vocal folds and the articulators.

Specific questions are:

- **Lungs**: How do speech breaths change under physical load in terms of frequency and volume of air inspired? Do more frequent breath pauses disrupt coordination of pauses and syntactic structure?
- **Vocal folds**: Changes in volume inspired may affect airflow and in turn the behavior of the vocal folds. How does fundamental frequency change under physical load? What measures of voice quality are most affected?
- **Resonance/articulators**: Changes in vocal fold behavior and/or airflow may disrupt timing relations with the articulators. Is voicing affected in fricatives and plosives? Are there changes in the vowel space?

2. Dissertation project

Data were recorded in August and September 2020 and will be published as an open corpus in 2022. The corpus was designed to allow multiple analyses. Some studies were planned pre-experiment but working with the data has revealed difficulties with the methods and also new possibilities. Consequently, I am now reconsidering which studies to combine for a coherent dissertation and which to pursue as individual publications.

2.1. Project corpus

Speech and respiratory data were collected from 48 female German native speakers performing speech tasks in three conditions: 1) sitting still; cycling at 2) light intensity; and 3) moderate intensity. Respiratory data were recorded using respiratory inductance plethysmography, which uses elasticated bands at rib cage and abdomen to record movements of the chest wall. Speech tasks and labelling progress are as follows:

1. **Sustained vowels**: /a, i, u/; 3 repetitions each per condition (1,296 tokens); fully labeled and delimited

2. **Read passage**: 126-word passage taken from spontaneous speech; 3 readings per condition (≈ 4.5 hours); fully transcribed and segmented into speech chunks and pauses (>130 ms)

3. **Prompted monologues**: speakers responded to 9 "small talk" prompts; 3 two-minute trials per condition (≈ 12 hours); fully transcribed; speech/pause segmentation is in progress.

2.2. Studies using the corpus

2.2.1. Changes in speech breathing under physical load

This confirmatory study tests how physical load affects speech breathing; it is the basis for connecting changes in breathing with changes in speech. Under increasing load, it is predicted that the (i) breath cycle shortens; (ii) volume of air inspired increases; (iii) breath frequency increases; and (iv) proportion of inhalation to total cycle duration increases. The breath signal is labeled at inhalation onset and peak using a MATLAB/Praat pipeline. Labeling is in progress.

2.2.2. Changes in speech chunking under physical load

This exploratory study asks: do more frequent breaths affect the coordination of breath pauses with grammatical structure? If so, there should be more ungrammatical pauses under physical load. This study tests for possible effects at the utterance planning level (i.e., chunking of speech). Pauses in the reading task and monologues will be manually annotated as breath/silent and grammatical/non-grammatical. Originally, an analysis of pause duration was planned, but this is being reconsidered after difficulties separating utterance-final fricatives and aspirated stops from breath noise.

2.2.3. Fundamental frequency (f0)

This confirmatory study investigates mean f0 in all three speech tasks, asking a) whether f0 increases even with low-intensity activity, and b) if results are conditioned by speech task. Results from the vowel [11] and reading tasks [12] show a significant increase in f0 with light and moderate physical activity, with a comparable absolute increase per speaker between tasks. Analysis of the monologues is in progress. Other f0 statistics will also be investigated (SD, range, min./max.) to explore variability under load. A perception study is also planned to test whether significant results have real-world relevance.

2.2.4. f0 at utterance onset and offset

This is a possible follow-up or extension of 2.2.3, arising from observations in the reading data that there are perceptible peaks in f0 at utterance onset and sometimes offset (speaker-specific) after inhalations. Could this be driving the increase in mean? The main work for this study is devising an appropriate method.

2.2.5. Voice quality

Results will be presented at INTERSPEECH 2021 [11]. Strength of excitation, H1 and HNR were found to significantly change under load, though there were speaker differences. Currently, the other point vowels are being analyzed. Additionally, the spontaneous speech data contain some "naturally sustained vowels" (prolongations). It is thus possible to corroborate these results with a more ecologically valid task.

2.2.6. Articulator precision

This exploratory study is being planned. The original plan was to test for changes in the vowel space by comparing F1/F2 across conditions. A reduced vowel space would indicate target undershoot. Preliminary work on the reading data shows greater reduction/coarticulation under load, making it difficult to select "pure" vowels. Annotation has also revealed sublexical errors; here it could be interesting to investigate whether place or manner of articulation is more affected under load. There is little work on this area, so such an analysis would extend current knowledge but it seems less connected to the other studies.

2.2.7. Disfluencies

Testing whether physical load affects fluency, positively or negatively, is an interest of the greater project in which this dissertation is embedded. The read and spontaneous speech contain many disfluencies, and full annotation may not be possible in the timeframe. One option that is theoretically motivated in terms of breathing is to test whether filled pauses (*uh*/*um*) are affected by load. The hypothesis, inspired by [13] is that if speakers subconsciously conserve airflow under load, they should tend to use closed syllable *um* rather than open *uh*.

3. Future work

This dissertation began as exploratory research based on theoretical hypotheses and diverse literature. After looking at the data collected last fall, it has become clear that some methods may not be feasible (e.g., reliably measuring pause duration under load) while other studies show potential for expansion, such as an in-depth exploration of effects on f0. The main difficulty is combining the possible analyses into a coherent thesis. The current structure is organized along the three physiological components required to generate the speech signal: the lungs, vocal folds, and articulators. Here, the perturbation of the respiratory system is at the center of the story: does increased respiratory drive affect speech production at different scales, from the segmental level to coordination of breath-taking with syntactic structure. However, taken together, these studies may span too broad an area.

4. Main contributions of the research

The main contribution of this research is a better understanding of how respiration and spoken language affect each other during physical activity. This is of theoretical interest for current models of speech production, as it explores whether the physiological state of the body can affect the way utterances are structured (i.e., into smaller chunks) and conveyed (i.e., perceptible vocal effects). The corpus is also of interest, enabling further research and comparison of results between different speech tasks.

At the same time, I believe that the work may also have practical relevance, for example in communications technologies for emergency services or incorporating breath-pause patterns in synthesis of spontaneous speech or for automatic stress recognition. At INTERSPEECH 2021, I hope to be able to explore these ideas further.

5. Acknowledgements

This work is part of the SALAMMBO project, jointly funded by the French National Research Agency and the German Research Foundation. The author thanks her supervisors, Dr. Susanne Fuchs and Prof. Christine Mooshammer, and Dr. Amélie Rochet-Capellan for their guidance and support.

6. References

[1] H. Selye H, *Stress without Distress*. Philadelphia, PA: Lippincott, 1974, p. 14.

[2] C. Kirchhübel, D. M. Howard, and A. W. Stedmon, "Acoustic correlates of speech when under stress: research, methods and future directions," *International Journal of Speech, Language and the Law*, vol. 18, no. 1, pp. 75–98, 2011.

[3] M. Van Puyvelde, X. Neyt, F. McGlone, N. Pattyn, "Voice stress analysis: a new framework for voice and effort in human performance," *Frontiers in Psychology*, vol. 9, art. 1994, 2018.

[4] K. W. Godin and J. H. L. Hansen, "Physical task stress and speaker variability in voice quality," *EURASIP Journal on Audio, Speech, and Music Processing,* vol. 2015, art. 29, 2015.

[5] A. Primov-Fever, R. Lidor, Y. Meckel, and O. Amir, "The effect of physical effort on voice characteristics," *Folia Phoniatrica et Logopaedica*, vol. 65, no. 6, pp. 288–293, 2013.

[6] J. Trouvain and K. P. Truong, "Prosodic characteristics of read speech before and after treadmill running," In *Proceedings of INTERSPEECH 2015 – 16th Annual Conference of the International Speech Communication Association*, Dresden, Germany, Sep. 2015, pp. 3700–3704.

[7] A. L. Winkworth, P. J. Davis, E. Ellis, and R. D. Adams, "Variability and consistency in speech breathing during reading: lung volumes, speech intensity, and linguistic factors," *Journal of Speech, Language, and Hearing Research* vol. 37, no. 3, pp. 535–556, 1994.

[8] D. H. Whalen and J. M. Kinsella-Shaw, J. M., "Exploring the relationship of inspiration duration to utterance duration," *Phonetica* vol. 54, no. 3-4, pp. 138–152, 1997.

[9] A. Rochet-Capellan and S. Fuchs, S., "The interplay of linguistic structure and breathing in German spontaneous speech," In *Proceedings of INTERSPEECH 2013 – 14th Annual Conference of the International Speech Communication Association,* Lyon, France, Aug. 2013, pp. 2014–2018

[10] S. E. Baker, J. Hipp, and H. Alessio, "Ventilation and speech characteristics during submaximal aerobic exercise," *Journal of Speech, Language, and Hearing Research,* vol. 51, pp. 1203–1214, Oct. 2008.

[11] H. Weston, L. L. Koenig, and S. Fuchs, "Changes in glottal source parameter values with light to moderate physical load," in *Proceedings of INTERSPEECH 2021 – 22nd Annual Conference of the International Speech Communication Association,* Brno, Czech Republic, Aug./Sep. 2021 (accepted).

[12] H. Weston, S. Fuchs, and A. Rochet-Capellan, "Speech during light and moderate physical activity: effect on f0 and vocal intensity," in *Proceedings of ISSP 2020,* Providence, RI, USA, Dec. 2020 (accepted).

[13] Z. Zhang, "Respiratory-laryngeal coordination in airflow conservation and reduction of respiratory effort of phonation," *Journal of Voice*, vol. 30, no. 6, pp. 760.e7-760.e13, 2016.