SPEECH BREATHING DURING PHYSICAL ACTIVITY IS OPPORTUNISTIC

Heather Weston

Humboldt-Universität zu Berlin, Germany
westonhe@hu-berlin.de

ABSTRACT

The ability to speak during diverse physical activities benefits individuals in many ways. What is currently unclear, however, is how individuals modify their respiratory behavior to accommodate both the slow, controlled breaths used for speech and the rapid, deeper breaths required by exercise. While the phonetics literature characterizes the execution of these two tasks as a *competition* between phonatory and respiratory goals, scholarship on neural control of respiration discusses the relationship in terms of *coordination*. To better understand speech breath support during physical activity, the current study presents empirical evidence from respiratory and acoustic recordings of read speech from 48 female speakers at rest and during exercise. Modifications of both the inspiratory and expiratory phases of respiration were observed that suggest that speech breathing during physical activity is opportunistic – linguistic structure is exploited to insert quick inspirations and expiratory airflow is increased to meet the metabolic needs of exercise.

Keywords: speech production, respiratory pattern, talking while walking, dyspnea, human performance

1. INTRODUCTION

This empirical study investigates how speech breath support is affected by simultaneous physical activity. The motivation rests on the following observation: speech and exercise modify the respiratory cycle in opposing ways. Breathing is slowed by speech to accommodate utterances but accelerated by exercise to meet increased metabolic needs. Although speaking during physical activity is part of daily life and crucial for diverse occupations, it is currently unclear *how* speakers negotiate between conflicting linguistic and exercise respiratory patterns. A better understanding of speech breathing during physical activity has both theoretical and practical relevance. Describing how speech sounds are produced under different speaker circumstances lies at the very heart of phonetics, while experimental data can provide insight into real-life speech-breathing situations that pose challenges for diverse groups of speakers, including those with certain respiratory pathologies, occupational voice users, athletes, and users of automatic speech recognition technology.

Breath-taking is commonly quantified with temporal and volumetric measures – the duration/ frequency of each breath and the volume of air inspired. A small body of work has investigated such changes in speech breathing during exercise. For temporal measures, respiratory rate (breaths/minute) has been found to significantly increase [1] by up to 25% [2], suggesting speakers accommodate speech during exercise by producing shorter chunks (i.e., pausing more). But work assessing volumetric changes suggests that this account is incomplete. Two studies [3, 4] measuring the volume of air inspired per minute – the product of breath rate and breath volume – found a significant increase during exercise. But it is unclear whether rate, volume, or both increased; based on also finding fewer syllables per phrase during exercise, [4] concluded that mainly breath rate increased, but stressed the need to pursue this hypothesis by measuring breath rate and volume separately. Work assessing volumetric measures has reported inconsistent results, finding no change in breath volume during exercise [5, 6], a decrease [7], or an increase [2]. These differences may stem from methodological choices: two studies [2, 7] did not measure breath volume directly but inferred changes based on breath noise intensity in the acoustic signal.

One commonality among this scholarship is the conceptualization of speech during exercise as a “competition” between phonatory and non-speech respiration. This is an interesting characterization of two *volitional* activities – speaking and exercising – and one that is absent in the literature on neural control of respiration, which seeks to understand how different respiratory patterns are *coordinated.* While this area is still incompletely understood and different models have been proposed [8], three strands of research add useful perspectives to investigate speech breathing during exercise. First, there is consensus that mammalian breathing comprises three phases: an inspiratory phase, a post-inspiratory phase, and an active expiratory phase incorporated to meet increased metabolic needs [9]. Second, these phases are proposed to be independent oscillators controlled by separate areas [10] of the brain’s respiratory rhythm generator, the preBötzinger complex [11]. Third, the “pacemaker” neurons that dictate phase rhythms are highly flexible and can respond to situational needs on a breath-by-breath basis [12, 13]. This account explains how breathing can be continuously adapted to meet dynamic exercise and linguistic (e.g., utterance length) goals. Related to this is evidence from respiratory physiology showing that breath rate and depth are independent of one another during exercise [14] and other types of stress [15].

Taken together, this research highlights a need for empirical evidence on the timing and depth of speech-breaths during inspiratory and expiratory phases. The present study thus extends previous phonetic work by assessing breath rate, duration, and volume in each phase to take a closer look at the "movable parts" of the breath cycle to see which are adjusted and how.

2. METHODS

2.1 Participants, materials, and equipment

A total of 48 healthy, non-smoker female native German speakers participated in the experiment. One participant was excluded due to a recording error. The average age was 23.6 years (*SD:* 3.8).

Participants read a 126-word passage designed to approximate spontaneous speech: a female German speaker was asked *What do you need to host a party?* Her monologue was transcribed, the disfluencies removed, and punctuation added. Participants were told the text stemmed from spoken speech and read it aloud once for familiarization prior to the experiment.

Exercise was performed on a low-noise stationary bicycle (daum electronic, Germany). Breathing was recorded via respiratory inductance plethysmography (Ambulatory Monitoring, USA), a non-invasive method that captures the expansion/contraction of the rib cage and abdomen as changes in raw voltage using elasticated belts with inductance wires. Thoraco-abdominal displacement is thus used as a proxy for breath volume. Speech was sampled at 22,050 Hz and recorded with a head-mounted microphone (beyerdynamic, Germany) placed 4 cm from the corner of the mouth at a 90° angle.

2.2. Experimental design, conditions, and procedure

The study used a within-participant design with three physical workload conditions: rest (sitting still), low workload, and moderate workload. Condition order was fixed (rest > low > moderate) because the cardiovascular systems may take several hours to recover after exercise [16] and responses are highly individual [17]; if condition order were randomized, it would thus be unclear if the rest condition actually captured baseline respiration values.

The physical workload conditions were defined following guidance by health organizations [e.g., 18] as a percentage of maximal heart rate (HR_{max}): low = 35% and moderate = 65%. These relative values were translated into participant-specific target heart rates (beats per minute) using a standard method in sports science, the Karvonen formula, given in (1).

$$\text{target HR} = [(HR_{max} - HR_{rest}) \times \%\ \text{intensity}] + HR_{rest} \tag{1}$$

Following [19], HR_{max} was predicted as: 208 – (0.7 × age). HR_{rest} was estimated by having each participant recline for 10 minutes and taking the average HR of minute 11 using a wrist-worn heart rate monitor worn throughout the experiment. The use of age and resting pulse accounts for physiological factors that affect respiratory function and thus makes the physical effort required for each workload comparable across participants, regardless of their age or level of fitness.

To validate the calculated target HRs for each condition, participants were asked to rate their level of perceived exertion using the Borg scale [20]. The average rating for low workload was 9.7 ("very light") and for moderate 13.8 ("somewhat hard").

The procedure was as follows: the text passage was presented on a monitor at eye level and read aloud three times per condition. The rest condition was followed by 4-minute cycling warm-ups to reach the target HR for low and moderate workload, respectively. Heart rate was monitored in real time using a tablet readout and was maintained by adjusting resistance on the bike between trials.

2.3. Preparation and analysis of respiratory data

To smooth motion noise, the respiratory signals were downsampled from 22,050 Hz to 100 Hz and filtered (passband: 1–10 Hz) using a custom MATLAB [21] protocol. The rib cage and abdomen signals were summed using a 2:1 correction [22] to estimate total change in lung volume during speech breathing.

To obtain duration and displacement measures for each breath, inspiratory peaks and expiratory troughs in the summed respiratory signal were automatically detected using MATLAB's findpeaks function [23] and manually corrected. Inspiratory time/ displacement was calculated by subtracting the timestamp/voltage of the preceding trough from a given peak. (For expiratory measures: trough – peak).

To aid interpretation of displacement measures, the raw signal in volts was converted to a physiologically functional unit, percentage of maximal displacement (≈ % of vital capacity), using an inspiratory vital capacity maneuver [24]. To assess respiratory differences between rest and exercise conditions, statistical analyses were run in R [25; Version 4.2.2]. Linear mixed-effects models were estimated using the packages lme4 [26], lmerTest [27] and emmeans [28], with condition as the independent variable (three levels: rest, low, moderate), the outcome for each respiratory variable as the dependent variable, and random slopes and intercepts for speakers (lmer syntax: variable ~ workload + (1 + workload | speaker), data). Prior to analysis the data were log-transformed to account for upper outliers. To account for different numbers of observations (i.e., number of breaths per trial/ participant), differences between conditions were assessed using estimated marginal means (EMMs) – the means taken from the statistical model for each response variable at each level of the predictor variable. The differences between conditions were tested for statistical significance using contrast analysis to assess pairwise differences. The Kenward-Roger method was used to obtain the degrees of freedom needed to compute *p*-values.

3. RESULTS

3.1. Statistical models

EMMs for breath duration and thoracoabdominal displacement are plotted in Figure 1. The original model and the log-transformed model (retransformed to original scale) are plotted together for comparison. Inspiratory and expiratory displacement, and inspiratory time, were found to increase significantly for all workload comparisons (rest vs. low, rest vs. moderate, and low vs. moderate; *ps* < .001) in both models. Expiratory time decreased significantly from low to moderate and from rest to moderate (*ps* < .001), but no significant change was found from rest to low (*p* = .151; log model: *p* = .854). Respiratory rate showed no significant change between rest and low but increased significantly from rest to moderate and from low to moderate (*ps* < .001). Table 1 gives the group means and the log model statistics.

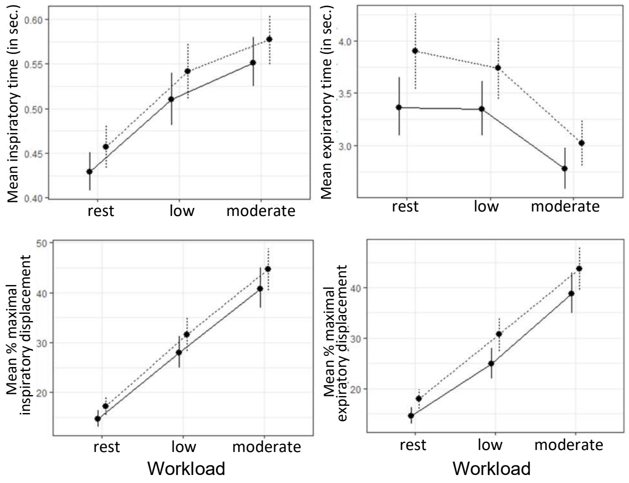

Figure 1: EMMs for log model (*solid line*) and original model (*dashed line*) for breath duration (*top*) and displacement (*bottom*) with increasing physical workload. Respiratory rate shown in Figure 2 (*inset*).

Table 1: Inspiratory/expiratory times and thoraco-abdominal displacement (in % max. displacement) across work conditions: group means with standard deviations; (log) model statistics in gray.

Var.	**Cond.**	**Mean (*SD*)**	**EMM**	***SE***	**Conf. level**
Insp. time (sec.)	rest	0.44 (0.17)	0.43	0.01	[0.41, 0.45]
	low	0.52 (0.19)	0.51	0.01	[0.48, 0.54]
	mod.	0.56 (0.17)	0.55	0.01	[0.52, 0.58]
Exp. time (sec.)	rest	3.59 (1.90)	3.36	0.14	[3.10, 3.65]
	low	3.51 (1.63)	3.35	0.13	[3.10, 3.62]
	mod.	2.86 (1.19)	2.78	0.10	[2.59, 2.98]
Insp. displ. (%)	rest	16.84 (9.97)	14.69	0.79	[13.2, 16.4]
	low	30.39 (15.25)	27.93	1.54	[25.0, 31.2]
	mod.	43.99 (17.89)	40.73	1.99	[36.9, 44.9]
Exp. displ. (%)	rest	17.48 (11.30)	14.58	0.80	[13.1, 16.3]
	low	29.50 (18.13)	24.83	1.49	[22.0, 28.3]
	mod.	43.07 (18.99)	38.72	2.01	[34.9, 43.0]
Breaths /min.	rest	13.55 (4.0)	13.5	0.55	[12.4, 14.6]
	low	13.87 (3.6)	13.9	0.50	[12.9, 14.9]
	mod.	16.74 (4.0)	16.7	0.54	[15.6, 17.8]

3.2. Further exploration

The statistical analysis showed that speakers adjust breath length, rate *and* depth during exercise, but the measures are affected differently by workload level: at low workload, resting breath rate and expiratory duration can be maintained, and only breath depth is increased; moderate workload shows changes in all measures. These findings are explored here to glean more insight into time–volume relations.

Of note in the inspiratory phase is the significant increase in breath depth with no change in breath rate until moderate workload. However, the histogram in Figure 2 and the higher *SD*s during exercise suggest that breath depth does not increase across the board but rather varies, showing upper outliers but also shallower inspirations (e.g., below mean at rest).

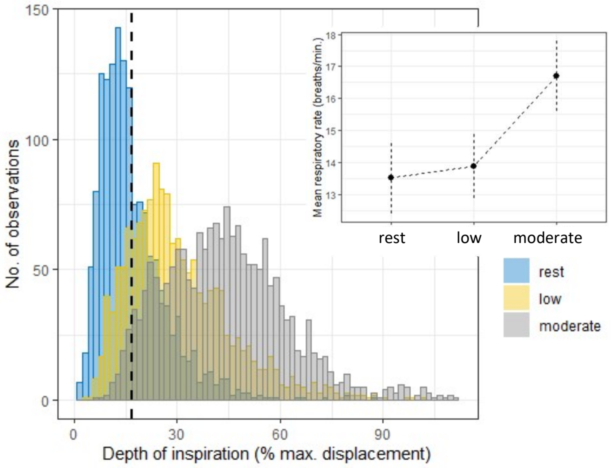

Figure 2: Histogram showing frequency of inspiratory depth; the black dashed line indicates mean depth at rest (16.8%). Inset shows EMMs for mean respiratory rate.

To test this, the non-parametric Fligner-Killeen test was used to evaluate the homogeneity of group variances. The test was significant, X^2 (2, 4860) = 459.4, $p < .001$, indicating that depth of inspiration varied more widely in some conditions than others.

Moving on to the expiratory phase, a striking observation was that speakers seemed to expel more air in a shorter time during exercise. Spearman correlations calculated for each condition showed good correlations between expiratory time and depth at rest ($R = .65$) and low workload ($R = .61$) but weak ($R = .32$) for moderate workload, suggesting that time–volume relations are not dictated by physical activity per se but rather by its intensity.

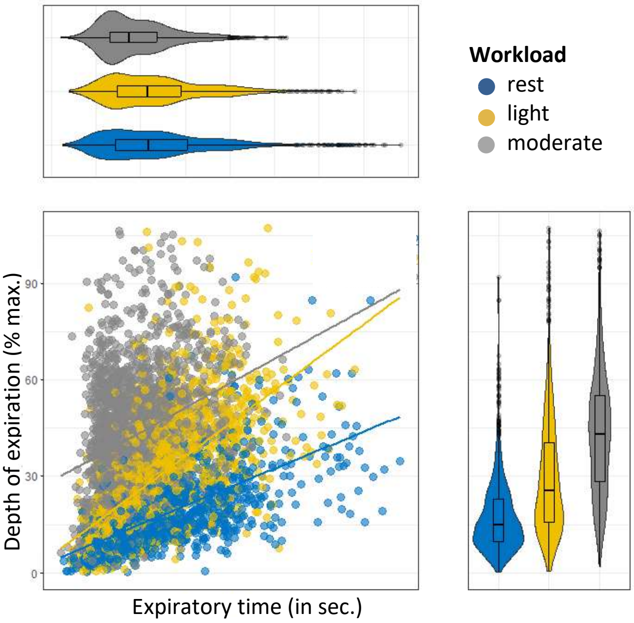

Figure 3: Scatter plot showing relationship of expiratory time to expiratory depth (% max. displacement) with regression lines; wrapped boxplots show the distribution of the expiratory time (*top*) and depth (*right*) data.

4. DISCUSSION

This study found that individuals adjust the timing *and* volume of their breaths to speak during physical activity, but the changes demonstrate more flexibility than previously though. Analyzing breath rate and depth separately revealed two independent – but interrelated – "movable parts" that can be adjusted in different ways to adapt to exercise level. At low workloads, speakers can accommodate their preferred utterance length by increasing breath depth alone. But greater workloads require a further increase in depth and an increase in rate, which may break up speech into smaller chunks. This flexibility may not have been apparent in previous studies because most assessed moderate or heavy workloads. In the present study, the low workload condition is equivalent to brisk walking and the moderate condition to jogging.

Looking at the inspiratory and expiratory phases separately also provided further insights. While inspiratory depth significantly increased with workload, it appeared that variability was also greater. It was rather surprising that small-amplitude breaths were seen at moderate workload – it suggests that breath volumes do not increase across the board. One possible explanation is that speakers become "opportunistic" and insert inspirations where possible. For example, during data labeling, some speakers took small-amplitude in-breaths at phrase breaks that had none in the rest condition (e.g., after the utterance-initial adverb "Additionally,"). Longer syntactic breaks – between sentences – were used for deeper breaths. These observations would fit with neural-control accounts of breathing rhythms being determined on a cycle-to-cycle basis [12, 13].

In the expiratory phase, it was striking that speakers expelled significantly more air compared to rest in the same amount of time (low workload) or less time (moderate workload). Again, timing and volume were adjusted separately to meet the respiratory demands for a given physical workload. But there may be different acoustic consequences for different time–volume relations, such as an increase in airflow during speech or the appearance of audible out-breaths (or both). The former may result in breathier speech, while the latter creates additional noise. Indeed, a widely obtained finding for speech during exercise is the presence of non-phonated expirations during moderate/vigorous activity in [2, 5, 7], with the latter two studies observing that expiratory volume was distributed over both phonated and non-phonated stretches of the expiratory phase. This could explain the weaker time–volume correlation observed in the present study under moderate workload. A next step is thus to further analyze the expiratory phase – most relevant for speech production – by assessing airflow changes using the three-phase breath cycle described in the work on respiratory neural control.

To come back to the question of *how* speakers negotiate between the opposing respiratory patterns required by simultaneous speech and exercise, it seems that situation-responsive, cycle-to-cycle variability is key. Breath-taking during exercise is interpreted here as being *opportunistic*: breaths may be deepened to preserve pause placement, and linguistic structure may be exploited to insert different-sized inspirations. Expiratory airflow may be regulated between phonated and non-phonated sections. These hypotheses must now be tested with further analyses, but the data presented here highlight the respiratory system's flexibility to meet different metabolic and behavioral goals. The guiding question is perhaps not *Which respiratory pattern "wins"?* but *Which time–volume relations occur and when?*

5. ACKNOWLEDGEMENTS

Parts of this work were jointly funded by the French Agence Nationale de la Recherche and the Deutsche Forschungsgemeinschaft through the SALAMMBO project. The author wishes to thank Susanne Fuchs and Amélie Rochet-Capellan for their guidance and support and Theres Weißgerber for her assistance throughout the project. The author also thanks two anonymous reviewers for their valuable suggestions.

6. REFERENCES

[1] Fuchs, S., Reichel, U. D., Rochet-Capellan, A. 2015. Changes in speech and breathing rate while speaking and biking. *Proc. 18th ICPhS Glasgow*.

[2] Trouvain, J., Truong, K. P. 2015. Prosodic characteristics of read speech before and after treadmill running. *Proc. Interspeech 2015 Dresden*, 3700-3704.

[3] Baker, S. E., Hipp, J., Alessio, H. 2008. Ventilation and speech characteristics during submaximal aerobic exercise. *J. Speech Lang. Hear. Res.*, 51(5), 1203–1214.

[4] Ziegler, A., Van Swearingen, J., Jakicic, J. M., Verdolini Abbott, K. 2019. Phonation demonstrates goal dependence under unique vocal intensity and aerobic workload conditions. *J. Speech Lang. Hear. Res.*, 62(8), 2584–2600.

[5] Doust, J. H., Patrick, J. M. 1981. The limitation of exercise ventilation during speech. *Resp. Physiol.*, 46(2), 137–147.

[6] Meckel, Y., Rotstein, A., Inbar, O. 2002. The effects of speech production on physiologic responses during submaximal exercise. *Med. Sci. Sports Exerc.*, 34(8), 1337–1343.

[7] Patil, S., Sangwan, A., Hansen, J. H. L. 2010. Speech under physical stress: A production-based framework. *Proc. 2010 IEEE Cape Town*, 5146–5149.

[8] Del Negro, C. A., Funk, G. D., Feldman, J. L. 2018 Breathing matters. *Nat. Rev. Neurosci.* 19, 351–367.

[9] Richter D. W. & Smith J. C. 2014 Respiratory rhythm generation in vivo. *Physiology* 29(58–71).

[10] Anderson, T. M., Ramirez, J. M. 2017. Respiratory rhythm generation: triple oscillator hypothesis. *F1000Res.*, 6, 139.

[11] Smith, J. C., Ellenberger, H. H., Ballanyi, K., Richter, D. W., Feldman, J. L.(1991. Pre-Bötzinger complex: a brainstem region that may generate respiratory rhythm in mammals. *Science*, 254(5032), 726–729.

[12] Carroll, M. S., & Ramirez, J. M. (2013). Cycle-by-cycle assembly of respiratory network activity is dynamic and stochastic. *J. Neurophysiol.*, 109(2), 296–305.

[13] Ramirez, J. M., & Baertsch, N. A. (2018). The dynamic basis of respiratory rhythm generation: one breath at a time. *Ann. Rev. Neurosci.*, 41, 475–499.

[14] Nicolò, A., Girardi, M., Bazzucchi, I., Felici, F., Sacchetti, M. 2018. Respiratory frequency and tidal volume during exercise: differential control and unbalanced interdependence. *Physiol. Rep.*, 6(21), e13908.

[15] Tipton, M. J., Harper, A., Paton, J. F. R., Costello, J. T. 2017. The human ventilatory response to stress: rate or depth?. *J. Physiol.*, 595(17), 5729–5752.

[16] Romero, S. A., Minson, C. T., Halliwill, J. R. 2017. The cardiovascular system after exercise. *J. Appl. Physiol. (1985)*, 122(4), 925–932.

[17] Travers, G., Kippelen, P., Trangmar, S. J., González-Alonso, J. 2022. Physiological function during exercise and environmental stress in humans—An integrative view of body systems and homeostasis. *Cells*, 11(3), 383.

[18] Physical Activity Guidelines Advisory Committee. 2008. *Physical Activity Guidelines Advisory Committee Report, 2008*. U.S. Department of Health and Human Services, p. D-3. Retrieved from https://health.gov/sites/default/files/2019-10/CommitteeReport_7.pdf, 10 Aug. 2020.

[19] Tanaka, H., K. D. Monahan, K. D., Seals, D. R. 2001. Age-predicted maximal heart rate revisited *J. Am. Coll. Cardiol.*, 37(1), 153–156.

[20] Borg, G. A. 1982. Psychophysical bases of perceived exertion. *Med. Sci. Sports Exerc.*, 14(5), 377–381.

[21] MATLAB. 2010. version 7.10.0 (R2017b). Natick, Massachusetts: The MathWorks Inc.

[22] Banzett, R. B., Mahan, S. T., Garner, D. M., Brughera, A., Loring, S. H. 1995. A simple and reliable method to calibrate respiratory magnetometers and Respitrace. *J. Appl. Physiol.*, 79(6), 2169–2176.

[23] findpeaks. (2007-2016). Natick, Massachusetts: The MathWorks Inc. https://de.mathworks.com/help/signal/ref/findpeaks.html.

[24] Wanger, J., Clausen, J. L., Coates, A., Pedersen, O. F., Brusasco, V., Burgos, F., Casaburi, R., Crapo, R., Enright, P., van der Grinten, C. P. M., Gustafsson, P., Hankinson, J., Jensen, R., Johnson, D., MacIntyre, N., McKay, R., Miller, M. R., Navajas, D., Pellegrino, R., Viegi, G. 2005. Standardisation of the measurement of lung volumes. Eur. Respir. J. 26 (3), 511–522.

[25] R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[26] Bates, D., Maechler, M., Bolker, B., Walker, S. 2022. lme4: Linear Mixed-Effects Models Using Eigen and S4. https://github.com/lme4/lme4/.

[27] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.*, 82(13), 1–26.

[28] Lenth, R.V., Love, J., Hervé, M. 2022. Package emmeans: Estimated marginal means, aka least-squares means. https://github.com/rvlenth/emmeans.