CHANGES IN SPEECH AND BREATHING RATE WHILE SPEAKING AND BIKING

Susanne Fuchs, Uwe Reichel, Amélie Rochet-Capellan

ZAS Berlin (Germany), LMU Munich (Germany), *Univ. Grenoble Alpes, GIPSA-Lab,* CNRS Grenoble (France)
fuchs@zas.gwz-berlin.de, reichelu@phonetik.uni-muenchen.de, amelie.rochet-capellan@gipsa-lab.grenoble-inp.fr

ABSTRACT

Speech communication is embedded in many daily activities. In this paper we investigate the effect of biking on respiratory and speech parameters. Breathing and speech production were recorded in eleven subjects while speaking alone and while speaking and biking with different rates. Breathing frequency, speaking rate, speech and pause intervals, overall intensity and f0 were analyzed for the different tasks. It was hypothesized that cyclical motion increases breathing frequency, which leads to a restructuring of speech and pause intervals or an increase in speech rate. Our results generally confirm these predictions and are of relevance for applied sciences.

Keywords: dual task, breathing, pauses, sequencing, speech rate

1. INTRODUCTION

In our daily life, speech communication occurs during all kinds of activities, for instance while walking to a shop, to school or to work: people are talking while walking, running, biking... However, in how far rhythmical motor actions such as walking or biking can influence the way we talk, is not well understood. Most of the literature on these combined activities is concerned with the reduced attention caused by the dual process and with the speed of cognitive processing. For example, talking on a mobile phone changes the way we walk [1], and mathematical operations are slower when people additionally walk on a treadmill than when they just focus on the operations (for a literature review see [2]).

In order to fill this gap, we propose to investigate if and how speech production changes when people move. Our working hypothesis is that speech production, specifically the temporal sequencing between speech and silent pause intervals, is influenced by continuous rhythmical movements. This hypothesis is grounded on the following rational.

It has been shown for mammals that locomotion and breathing are synchronized [3]. With an increase in the speed of locomotion, respiratory demands for oxygen supply change and lead in general to a higher breathing frequency. In the motor control literature a synergy between respiration and body motion or even posture has often been reported [4,5]. Since locomotion increases breathing frequency, breathing cycles will be shorter while moving. On the other hand, speech production involves a specific control of breathing: inhalations are shortened and faster than in vegetative breathing to reduce pauses, and speech sounds are produced during longer and slower exhalation periods [6]. Inhalations also have to be synchronized with the linguistic content to preserve the discourse flow [7]. Consequently, changes in breathing rate due to physical effort may have strong consequence on speech rhythm.

Facing the necessity to adapt breathing both to the physical effort and to speech production, speakers may adopt different strategies when talking while biking. They could first increase their speaking rate with raised physical effort to produce a similar amount of speech in a single breathing cycle as when speaking without biking. Alternatively, they could preserve a similar speech rate, but chunk their speech differently, introducing more pauses in the speech flow. Furthermore it is likely that certain global phonetic parameters, like f0 and intensity may be affected by the physical effort. In particular, if muscular tension during biking is transmitted to the vocal apparatus, we suppose that mean f0 will be higher during the dual task and mean intensity could increase as well.

The study reported in this paper was designed to address these hypotheses by analyzing global parameters of speech production when speaking with different physical efforts.

2. METHODOLOGY

2.1. Experimental set-up

The data presented here are part of a larger project designed to study the relation between locomotion, breathing and speaking. In this paper we focus on the respiratory and acoustic data, but locomotion was also recorded is synchrony with breathing and speech using 12 VICON MX3+ cameras. Respiratory kinematics was recorded by means of Inductance Plethysmography (formerly Respitrace) simultaneously with the acoustic speech signal by means of a head mounted microphone (Beyer-dynamic Opus 54.16/3). For the biking task a low noise ergometer was used (ergo_bike, Daum electronic). The participants fulfilled the following tasks in successive order:

- Quiet breathing (ca. 3min) – **(Qon)**
- Speaking only (ca. 8 min) – **(S)**
- Biking only (ca. 5 min, 70 W) **(B)**
- Biking only with effort (ca. 2 min, 140 W) **(Be)**
- Biking & Speaking (ca. 8 min, 70 W) **(SB)**
- Biking with effort & Speaking (ca. 2 min, 140 W **(SBe)**

The experiment took about two hours, and participants were paid for their attendance. Spontaneous speaking was elicited in condition S, SB, SBe by presenting the subjects ten different items in written form displayed on a wall. These items consisted of a variety of tools, survival kits, food, and drinks that could be taken for a trip to a lonely island. Participants were instructed to answer the following questions: "Imagine you have to go to a lonely island and could bring with you only five of these ten items. Which ones would you chose? Please chose five items, bring them in an order, and motivate your choice. Why wouldn't you take the other five items?" None of the ten items was ever repeated to avoid learning effects. Each condition except of Qon, Be, SBe involved two or three trials.

For the biking task, subjects were instructed to bike in a comfortable way, and defined their own speed according to the resistance of the bike. The resistance was set to 70 W in the normal biking condition and 140 W in the biking with effort condition.

2.2. Speakers and stimuli

Eleven native speaker of German (10 females) were recorded in this study. All of them had a normal Body Mass Index. Their age ranged from 18-37 and they had no reported history of voice, breath, language, hearing, or motor disorders.

2.3. Data acquisition and processing

Acoustical and breathing data were sampled synchronously at 11030 Hz.

For the breathing recording, the Respitrace gains were the same for the rib cage and abdomen belt. After the recording, signals were resampled to 100 Hz and low-pass filtered at 40 Hz. In the current study we analyzed breathing using the sum of two thorax' and one abdomen signals [8] to capture changes in overall lung volume.

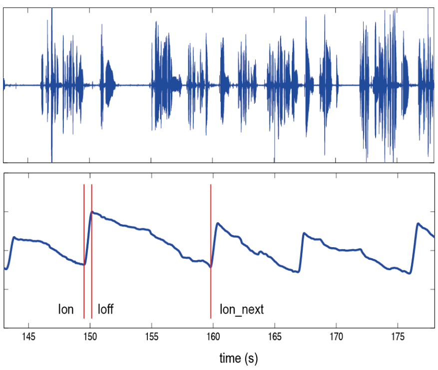
In a first step speech intervals were labeled manually using Praat (version 5.3.53 [9]). Speech was written in the canonical form to allow later matching with a German lexicon and to carry out an automatic syllable count. We obtained the number of syllables from the syllabified canonical transcription of the utterances, which was automatically derived by the BALLOON toolkit [12].

Onset and offset of inhalation phases of breathing signals were detected automatically using zero crossing of velocity profiles and peak velocity threshold. Boundaries were then checked manually and corrected when required. Noisy intervals or interval with artifacts were ignored. The breathing cycle was defined from the onset of an inhalation to the onset of the next inhalation (see Figure 1), since these landmarks could be detected with a greater reliability than the offset of exhalation.

2.4. Acoustic and respiratory analyses

Praat functions were used to compute the intensity (minimum pitch: 100 Hz, window: 8 ms, time step: 2 ms) and the f0 (autocorrelation, f0 floor 75 Hz, ceiling 600 Hz, time step: 10 ms) of the signals. We characterized each speech unit by: its duration, the number of syllables, the average intensity, and the average f0. For every trial involving speech production (S, SB, SBe), the percentage of speech and silent pause intervals were calculated by dividing the sum of the duration of the speech intervals and the sum of the duration of the pauses intervals by the total duration. We also computed the global speech rate as the total number of syllables in the speech intervals divided by the sum of the duration of these intervals. Furthermore, we computed the duration of each breathing cycle. This measure was used to determine the breathing frequency in each trial as the number of breathing cycles in the trial divided by the sum of their duration.

Figure 1: Sample of acoustic signal (top) and associated breathing signal (bottom). Vertical lines on the breathing signal indicate onset and offset of inhalation (In, Ioff) and offset of the cycle, taken as the onset of the next inhalation (Ion_next)



2.5. Statistical analysis

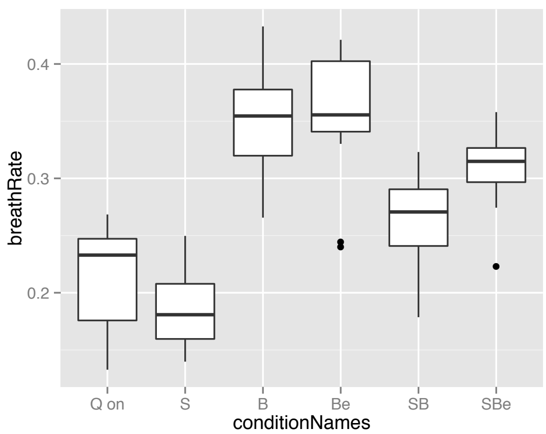
For the analyses of the breathing rate we first ran a within subjects ANOVA to test the effect of the six conditions. We then tested the effect of biking on speech breathing rate by comparing S with SB and Be conditions using post-hoc comparisons with Bonferroni correction. To better investigate the effect of effort and its interaction with speech, we ran a second ANOVA restricted to the B, Be, SB and SBe conditions and two within subjects factors: effort (strong vs. normal) and speech (with vs. without). ANOVAs were also used for the analyses of speech parameters with the condition (S, SB and SBe) as a within subject factor.

3. RESULTS

3.1. Breathing rate

As expected, breathing rate varied according to the condition (see Figure 2). ANOVA revealed a significant effect of the condition ($F(5,50)=36.7$, $p<0.001$). Considering only the conditions involving speech, breathing rate was significantly increased in the dual task and with increased effort (S<SB<SBe). The analysis of the biking conditions showed that the global effect of effort was almost significant ($F(1, 10)=4.8$, $p=.0535$). The presence of speech during biking clearly decreases breathing rate in comparison to biking only ($F(1, 10)=24.0$, $p<.001$), especially for the low effort condition (interaction effort x speech: $F(1, 10)=11.0$, $p<.01$).

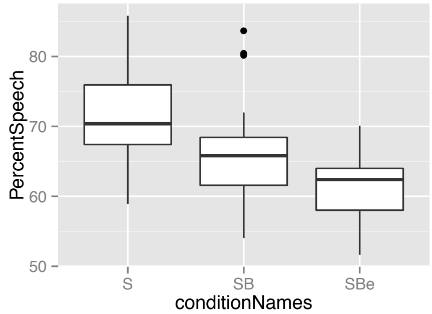
Figure 2: Average breathing rate (cycles/s) in the different conditions (see text for details).



3.2. Speech intervals and pauses

Figure 3 displays the percentage of speech relative to the overall duration. Biking clearly induced a decrease in the percentage of speech and an increase in percentage of pauses, especially in biking with greater effort ($F(2, 20)=18.6$, $p<.0001$).

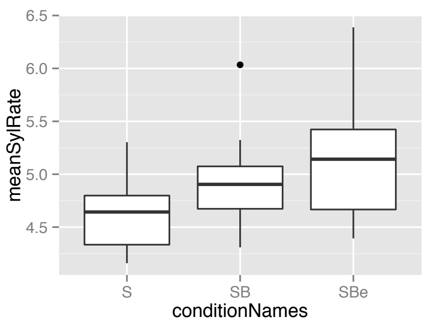
Figure 3: Average percentage of speech relative to the overall duration in the different speaking conditions.



3.3. Speech rate, intensity and f0

Figure 4 displays the average speech rate in the different conditions. Speech rate rises from an average of 4.6 syll/s in the single speaking task to 5.0 syll/s in the dual task, and reaches the highest values (mean=5.2 syll/s) when subjects bike with increased effort ($F(2, 20) = 15.6$, $p<.0001$).

Figure 4: Average speaking rate (syll/s) according to the speaking condition.



Similarly to speech rate, average intensity and average f0 show an effect in the dual task. Mean intensity increases in speaking and biking as compared with speaking alone ($F(2, 20)=45.96$, $p<.0001$). A similar effect was observed for f0 ($F(2, 20)= 10.2$, $p<.001$).

Figure 5: Average intensity (dB) according to the speaking condition

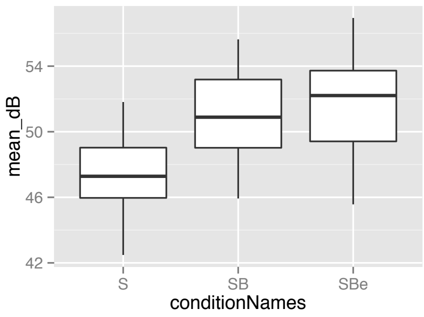
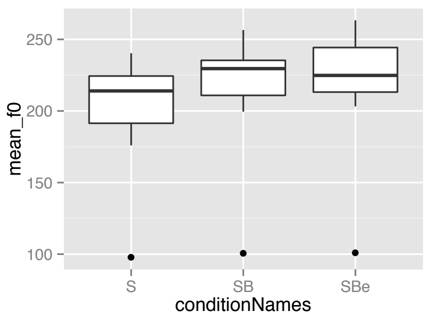


Figure 6: Average f0 (Hz) according to the speaking condition (the lower value refer to the male speaker data).



6. SUMMARY AND DISCUSSION

This study was designed to evaluate how speech parameters and breathing change in a single and dual task and with different efforts. In accordance with our expectations, we observed an increase in breathing rate while biking in comparison to quiet breathing. Breathing rate in biking also tended to rise with increased biking effort. In the dual task, we clearly observed an effect of speech and physical effort on breathing rate. When speaking and biking in parallel, subjects were not able to ventilate as fast as when they biked alone: speech production required to slow down the breathing rate. However, as required by physical activity, speaking and biking with a larger effort involved faster breathing than biking with less effort. These first results on breathing rate suggest a compromise of ventilation between the speaking and the biking tasks.

Biking also increased the percentage of pauses within the speech flow as compared with speaking alone. This result could be explained due to the necessity to breathe more often. The shortening of speech intervals with increased physical effort is accompanied by an increased speaking rate. Faster speech could indicate that the speakers try to maintain a certain level of information density within speech chunks. Speakers may also be aware of the higher breathing rate and the potential to get out of air and compensate for it by increasing speech rate. Further linguistic analyses are now required to better quantify where speech pauses occurred (e.g. Do they only occur in the inhalation phase?).

We also observed an increase in average intensity and f0 while biking. The higher intensity might be related to the noise induced by biking: subjects are breathing faster, produce louder breathing noise and although we used a low noise ergometer, the bike itself also produces some noise. These factors could create a kind of Lombard effect [14]. It is also possible that the increased muscular effort propagates among modalities, i.e. from limb to speech muscles.

In the introduction we proposed two possible strategies how speakers perform speech under increased breathing rate constraints due to motion: a) speakers may increase their speaking rate due to shorter breathing cycles in order to include enough content in one breath group and b) speakers may restructure the speech and silent pause intervals, but keep the speech rate constant. Our results suggest that speakers adopt an intermediate strategy with changes in both. It could now be interesting to contrast strategies between different populations, especially between sportive vs. non-sportive populations. Rhythmical physical activity such as biking may also be a path to further explore for speech pathology. For example, if physical effort transfers to speech production, it could be a helpful strategy for patients suffering from Parkinson's disease.

Speech and pause intervals are crucial elements for speech recognition systems [10], and may also provide a window investigating speech planning [11]. Planning speech while moving is a common daily situation that may induce differences in the way speech and pauses are chunked together. Moreover, mobile devices for oral communication provide situations where people are speaking while moving. Hence, our findings might be also relevant for speech recognition technology to adapt to the speaker's physical effort. This adaption refers to phonetic variance induced amongst others by effort-dependent speaking rates.

ACKNOWLEDGMENTS

This work was supported by a grant (01UG0711) from the BMBF to the LabPhon group at ZAS. Thanks to the University of Applied Science for their support, to Jörg Dreyer (ZAS) for excellent technical support, and to our participants.

7. REFERENCES

[1] Lamberg, E.M., Muratori, L.M. 2012. Cell phones change the way we walk. *Gait and Posture* 35, 688-690.

[2] Al-Yahyaa, E., Dawesa, H., Smith, L., Dennisa, A., Howells, K., Cockburn, J. 2011. Cognitive motor interference while walking: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews* 35, 715–728.

[3] Bramble, D.M., Carrier, D.R. 1983. Running and breathing in mammals. *Science* 219(4582), 251-256.

[4] Kuznetsov, N.A., Shockley, K.D., Richardson, M.J., Riley M.A. 2011. Effect or precision aiming on respiration and the postural-respiratory synergy. *Neuroscience Letters* 502(1), 13-17.

[5] Larsson, M. 2014. Self-generated sounds of loco-motion and ventilation and the evolution of human rhythmic abilities. *Animal Cognition* 17, 1–14.

[6] Conrad, B. & Schönle, P. (1979). Speech and respiration. *Arch. Psychiatr. Nervenkr.* 226, 251–268.

[7] Rochet-Capellan, A. & Fuchs, S. 2013 The interplay of linguistic structure and breathing in German spontaneous speech. *Proceedings of Interspeech, Lyon*, paper 1228.

[8] 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, 2169-2176.

[9] Boersma, P. Weenink, D. 2013. Praat: doing phonetics by computer [Computer program]. Version 5.3.53, retrieved April 2013 from http://www.praat.org/

[10] Zellner, B. 1994. Pauses and the temporal structure of speech, in E. Keller (Ed.) *Fundamentals of Speech Synthesis and Speech Recognition*. (pp. 41-62). Chichester: John Wiley.

[11] Krivokapic, J. 2007. *The planning, production, and perception of prosodic structure*. PhD thesis at the University of Southern California.

[12] Reichel, U.D. 2012. PermA and Balloon: Tools for string alignment and text processing. *Proceedings of Interspeech*, Portland, Oregon, paper no. 346.

[14] Étienne Lombard 1911. Le signe de l'élévation de la voix. In: Annales des maladies de l'oreille, du larynx, du nez et du pharynx. 37, 101–119.