Is breathing sensitive to the communication partner?

Amélie Rochet-Capellan[1], Gérard Bailly[1] and Susanne Fuchs[2]
[1] GIPSA-lab, UMR 5216 CNRS/Université de Grenoble –France
[2] Centre for General Linguistics, Berlin – Germany
{amelie.rochet-capellan,gerard.bailly}@gipsa-lab.grenoble-inp.fr, fuchs@zas.gwz-berlin.de

Abstract

This paper investigates breathing profiles in eleven female speakers (subjects) when talking successively with the same two females (partners). Breathing kinematics of the two interlocutors was recorded synchronously by means of two Inductance Plethysmographs. In order to understand the implication of breathing in dialogue, we analyzed changes in breathing pauses according to the main dialogue events (listening, backchannels, turns start and turns continuation). Breathing and syllable rates were also compared among partners and subjects. The duration of inhalations and related pauses was reduced before a turn continuation in comparison to a turn start. The delay between speech offset in a breathing cycle and the onset of the next inhalation increased when a speaker and a listener swap roles as compared to a speaker who continued the turn. This was observed for both partners and subjects. The partners differed in their breathing and articulation rates but the two rates were not clearly correlated. In agreement with previous works, the current study shows that breathing kinematics is strongly linked to dialogue events. However, it doesn't show any clear effect of partner on speaker's breathing. This last result is discussed relative to methodological aspects.

Index Terms: Breathing, Respiration, Spontaneous dialogue, Interpersonal adaptation, Breathing rate, Syllable rate, Inhalation pauses

1 Introduction

Breathing is actively involved in speech production as it provides the airflow required to generate speech sounds. The vital need of air is also a constraint that organizes the discourse into inhalation pauses and speech intervals. Several studies have analyzed breathing in reading tasks and to a lesser extent in spontaneous speech. These studies consistently showed that speech production is achieved by a specific control of breathing, visible in the clear reduction of the inhalation duration relative to the exhalation phase, and as compared with quiet breathing [1-3].

In text reading, the inhalation pauses consistently occur at syntactic boundaries [4]. The duration and amplitude of inhalation are mainly related to the syntactic constituents of the text (e.g. more air is inhaled before a new paragraph than before a sentence inside a paragraph) [5-7]. The properties of inhalation are also related to the length of the upcoming utterance [2, 8-10]. Similar behavior can be found in spontaneous speech, with greater inhalation before a main than an embedded clause, but more inhalation pauses occurring at non-syntactic locations as compared to read speech [11-13].

Analyses of breathing noises showed that they may play a role to indicate continuity between two related speech groups in text reading [14], and could improve the quality of synthetic speech [15]. This indicates that breathing noises are useful for speech perception. Furthermore, when listening to read speech, the listener breathing changes according to the properties of the speech signal, suggesting some influences of the speaker on the listener breathing and interaction between the control of breathing and perceptual processes [16-18].

Few studies have analyzed interpersonal influences of breathing in verbal collaborative tasks. During collaborative reading, readers coordinate their breathing. They breathe in-phase when reading synchronously, and in anti-phase when reading in alternation [19]. During choir singing, singers synchronize their breathing, especially when singing in unison [20].

In dialogue, breathing pauses have to be coordinated with interpersonal constraints and in particular with turn taking events. Despite hypotheses about the implication of breathing in conversation [21], few behavioral studies have analyzed the breathing profiles in spontaneous dialogue. Preliminary analyses of inhalation pauses in spontaneous dialogue were based on short recordings and single dyads [22-23]. They showed adaptation of breathing to dialogue constraints. A more systematic investigation of breathing in scripted and spontaneous dialogue was provided in [24]. This study found evidence for interpersonal alignment of breathing related to turn taking and laughter. Yet, no study addressed changes in breathing profiles during dialogue according to the conversational partner, nor the potential role of breathing in communication.

Several studies found interpersonal adaptations during verbal interactions. These adaptations may occur at different levels. For example, interlocutors involved in a dialogue may converge in speech rate or intensity, f0 or formants values [25-27]. The same speaker may thus behave differently when talking to two different partners. As breathing is specifically involved in speech production and verbal interaction, it could also be involved in interpersonal adaptation. We tested potential changes in speakers' breathing according to their interlocutor by analyzing breathing profiles during spontaneous dialogue. In addition, we tested how breathing pauses durations were linked to the main communicative events of the dialogues (listening, backchannels, turns continuation vs. turns start).

2 Methods

2.1. Subjects

The participants were eleven subjects (S01-S11, age: 31 years (mean) ± 7 (standard deviation), body mass index 21.3 ±1.5) and two partners (P01, age 42, BMI: 20.4, and P02 age 28, BMI: 20.8). Subjects and partners were all native female speakers of German, students or academics with a university degree. The subjects were naïve to the purpose of the study while the partners were not.

2.2. Procedure and data acquisition

Partner and subject were sitting and facing each other with a distance of ~1.5 m. They were instructed to keep their hands on their legs in order to avoid torso and arm movements that could strongly affect the recording of breathing. Subject and

10.21437/SpeechProsody.2014-110

partner’s task was to talk with each other about a topic chosen in agreement with one another (holidays, sports, cooking, or movies). In total, each subject had five short conversations with each partner (2.5 min, for each trial), starting with P01 or P02. The topic of the conversation could change or remain the same over the five trials.

The acoustic signals were recorded using two directional microphones (Sennheiser HKH50 P48). The rib cage and the abdominal kinematics were recorded by means of two Inductance Plethysmographs (RespitraceTM). One band was positioned at the level of the axilla (rib cage) and the other band at the level of the umbilicus (abdomen). The acoustic and the breathing signals were recorded synchronously for the two speakers by means of a six channels voltage data acquisition system. All signals were sampled at 11030 Hz.

2.3. Post-processing and labeling

After the recording, the breathing data were sub-sampled at 100 Hz and pass-band filtered (1-40Hz). The onset and offset of inhalation movements were detected automatically from the sum of the rib cage and the abdomen displacements, and manually corrected when required. The breathing cycle was defined from the onset of an inhalation to the onset of the next inhalation (*I+PI* on Figure 1).

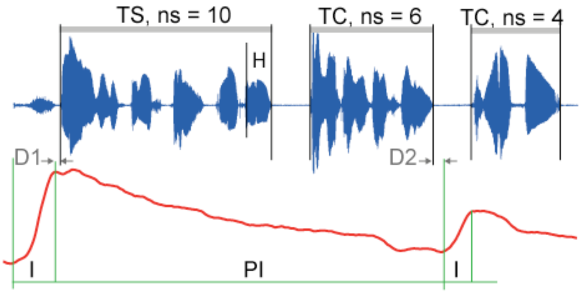

Figure 1. Top: The acoustic signal and the corresponding labeling of speech groups (*TS*: turn start, *TC*: turn continuation, *ns*: number of syllables, *H*: hesitation). Bottom: corresponding breathing kinematics and labeling of the breathing cycle (*I*: inhalation + *PI*: post inhalation). D1: *delayOnset* and D2: *delayOffset* are delays between the main breathing and speech events (see text for details).

Speech productions were labeled in Praat [28] by a trained phonetician. The boundaries of the inter-pausal units (IPUs) were detected. The vocalized hesitations (*uh, uhm mmm*,...) and the non-verbal communicative noises (laughter, mouths noises) were also delimited. Hesitations were distinguished from backchannels (*mhm*) easily as the first ones occurred during speaking and the second ones during listening phases. The spoken productions were transcribed for each IPU. On the basis of this transcription, the number of syllables was derived automatically from the output of the BALLOON toolkit [29]. The vocalized hesitations were considered as one (*uh, uhm*) or two syllables (*mmm*). IPUs were also labeled according to their main function in the dialogue as: turn start, turn continuation or backchannel (Figure 1). A turn start (*TS*) was defined when one interlocutor became the speaker and her interlocutor became the listener. This initial turn could be followed by one or several continuation (*TC*) separated by silent pauses. When the turn holder held the floor, the listener could produce short utterances like “mhm”, “okay”, “yes”, “aha” that did not intend to take the floor, but rather signal to the speaker to continue talking. These utterances were labeled as backchannels (*BC*) [30].

2.4. Data selection and analyses

Each breathing cycle (*I+PI* on Figure 1) was characterized by its total duration (*durC*) and the duration of inhalation (*durI*). For each trial, we computed the breathing rate (as the number of labeled breathing cycles divided by the sum of the durations of these breathing cycles) and the articulation rate (total number of syllables in the speech groups divided by the sum of their durations). Vocalized hesitations were not considered in the computation of articulation rate.

The breathing cycles were classified according to their function in the conversation (*cycle_type*): listening cycle (*LI*, the subject is not speaking while her interlocutor is); backchannel cycle (the cycle included only one or more *BC* units); turn start (the cycle started with a *TS* IPU), and turn continuation (the cycle started with *TC* IPU).

The inhalation noises could be used to signal the continuation of a theme or major thematic breaks during text reading [14] : their amplitudes, durations and phasing relations with speech contribute to encode thematic structure. We investigated if similar strategies could also be found in unconstrained face-to face dialogues to signal turn start or continuation by analyzing pauses related to inhalation. For *BC, TS* and *TC* cycles, we computed the delay between inhalation offset and the onset of the first IPU (*delayOnset*) and between the end of the last speech unit and the onset of the next inhalation (*delayOffset*).

The changes in breathing and syllable rate according to the partner were tested using ANOVAs. The dataset was separated for subjects and partners, with partner as a within subject factor. This allows testing the effect of the partner on subjects’ behaviors, and differences between the two partners. The results were considered as significant when p<.01.

The effect of partner and task (*TS*, *TC*, *BC*, *LI*) on the duration of inhalation was tested using Linear Mixed Models (lme4 & languageR packages in R version 2.14). The dataset was split between the partners and subjects’ data. The subjects and dialogue trials were taken as random factors. The inhalation duration was log-transformed to obtain normally distributed residuals. The results were considered significant with p<.01. Additionally we run lmer for the *delayOnset* and *delayOffset*, but the residuals were nonlinear, even when transformed. We will therefore only provide descriptive statistics.

3 Results

3.1. Global description of the dialogue

We first characterized the dialogue at a global level. On average, speech intervals represented 37% of the total duration for P01, 40% for P2, 39% for subjects talking with P01 and 43% for subjects talking with P02. The effect of partner on the duration of subjects’ speech failed to reach significance (F(1, 10)=3.9, p=.08), due to variability in subjects’ behaviors. For example, some subjects were talking more than the partners while the reverse was observed for other subjects.

3.2. Relationships between breathing and syllable rates

The breathing and syllable rates were globally smaller for P01 as compared with P02 (F(1, 10)=18.4, p<.01 and F(1, 10)=67, p<.0001), see Figure 2. By contrast, subjects’ breathing and syllable rates were not different when either talking with P01 or P02. Positive correlations were observed between the average syllable and breathing rates for subjects (r=.61 when talking with P01, .43 when talking with P02) and for partners (r~.3

for P01 and P02). For subjects, the correlation was significant only when talking with P01 (p<.05).

No clear correlation was observed between subjects and partners' breathing and syllable rates (Figure 3).

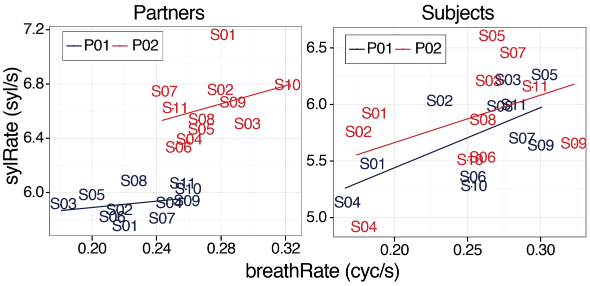

Figure 2. Average syllable rate according to average breathing rate for partners (left) and subjects (right).

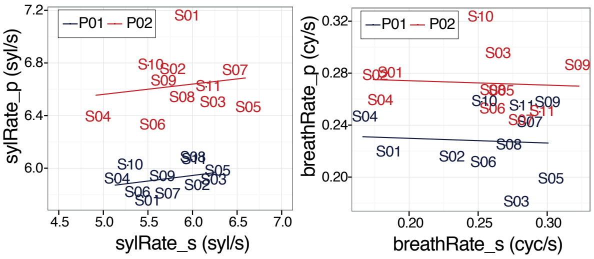

Figure 3. Partner's vs. subject's parameters, left: average syllable rate, right: average breathing rate

3.3. Duration of inhalation

Globally, inhalations were shorter for P02 than P01 (here we used lmer since our dataset was unbalanced, t=-7.3, pMCMC=.0001), while subjects did not show significant changes in *durI* according to the partner (t=-2.3). Inhalations were significantly shorter before turn continuations than before turn starts, backchannels and listening cycles (all |t|>2.7, pMCMC<.004). This was observed for both partners and subjects (Figure 4). The difference between *BC* and *LI* was significant for partners (t=2.7, pMCMC=.004) but not subjects (t=2.3).

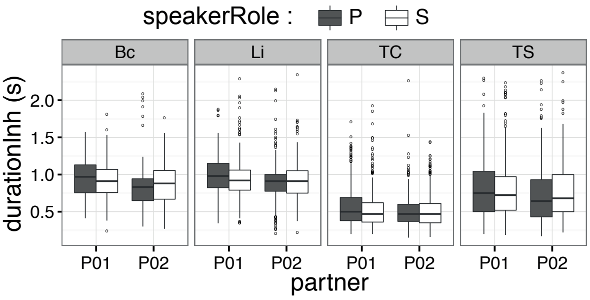

Figure 4. Duration of inhalation (*durI*) according to the function of the first speech group on the breathing cycle (*BC*, *TC*, *TS*) and for listening cycle (*LI*), for dyad involving P01 and P02 and for partners (*P*) and subjects (*S*).

3.4. Delay between inhalation offset and speech onset

The duration of pauses between inhalation offset and speech onset (*delayOnset*) were shorter for turn continuations than turn starts and backchannel cycles (see Figure). P02 may also initiate backchannels earlier on the breathing cycle as compared with P01.

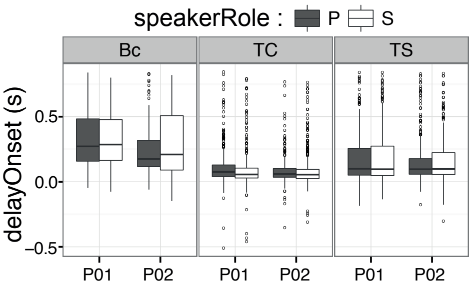

Figure 5. Delay between the offset of inhalation and the beginning of speech according to the dialogue event (*BC, TC, TS*) for the dyads involving P01 and P02 and for the partners (*P*) and the subjects (*S*).

3.5. Delay between speech offset and next inhalation

When speakers talked and then inhaled again, the delay between speech offset and the inhalation onset was shorter when the speaker continued the turn than when she started a new turn. This was observed for both subjects and partners (Figure). *delayOffset* was generally longer when the speaker switched her role and became the listener. This was particularly the case for P02 as compared with P01, while no clear change in subjects' behaviors was observed according to the partner.

The distribution of *delayOffset* is given in Figure 7 when the next breath group started with a turn continuation, a turn start (black and gray bars) or a turn start only (blue bars). The duration of *delayOffset* was in most of the cases shorter than 200 ms. The probability of the next speech group to be a turn start increased with the increase of the duration of *delayOffset*. This was particularly evident for the two partners. Additionally, P02 produced longer *delayOffset* than P01. The distribution was comparable between P01 and subjects who talked to her, but an asymmetry was observed between P02 and subjects talking to her, with a larger number of observations with shorter *delayOffset* for subjects than for P02.

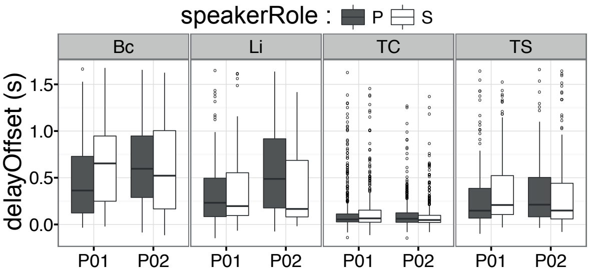

Figure 6. Delay between the offset of the last speech group on the breathing cycle and the onset of the next inhalation according to the dialogue event (*BC, LI, TC, TS*) for the dyads involving P01 and P02 and for the partners (*P*) and the subjects (*S*).

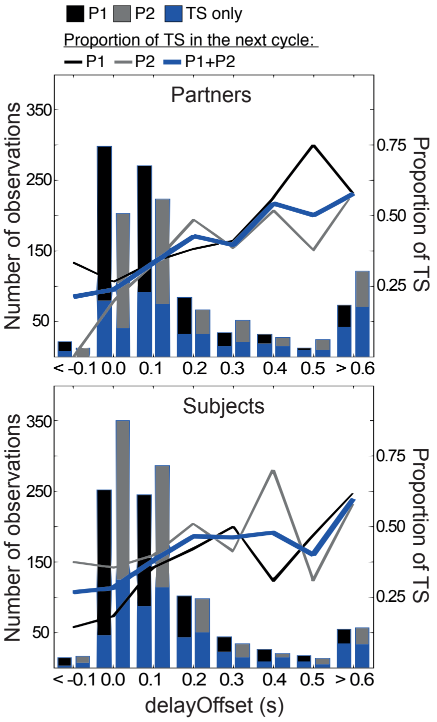

Figure 7 Distribution of the breathing cycles according to the delay between speech offset on the breathing cycle and the onset of the next inhalation. The results are given for the partners (*top*) and the subjects (bottom). The superimposed curves represent the proportion of turn start (*TS*) in the next breathing cycle (*TS*/(*TS*+*TC*)) for each lag.

4 Discussion and conclusion

The aim of the current study was to investigate changes in breathing profiles with respect to the conversation partner during spontaneous dialogue. We observed differences in the two partners' behaviors. P02 was breathing and speaking faster than P01. If breathing would determine dialogue rhythms [21] and if interlocutors were systematically adapting their physiological rhythms to each other through the verbal exchange (as observed before in more controlled tasks [19-20]), we could expect an increase in subjects' syllable and breathing rates when talking with P02 as compared to P01. This was yet not the case: the current analyses of breathing and syllable rate did not provide clear evidence of subjects' adaptation to partners' behaviors. This could be explained by the fact that the recordings were relatively short (2.5 mins * 5 recordings for each dyad) and that longer and continuous interpersonal interactions may be required to generate adaptations in physiological rhythms. Moreover, familiarity with the interlocutor may play a role in conversations. For example, romantic partners or members of a same family may display stronger adaptation of physiological rhythms [31] than persons who did not know each other very well as in the current study.

Our experimental paradigm involved two non-naïve partners. This choice allows testing the same partners for all subjects, similarly as in our previous work on breathing adaptation during listening [18]. However, for dialogue, this choice introduced an asymmetry in the interlocutors' roles: partners were aware they should maintain the dialogue, ending in conversations globally balanced between subjects and partners. Some variability was yet observed according to the subject: some subjects were talking more than the partners, some others less, with some interaction between subjects and partners. This suggests that the partners were not fully controlling the dialogue and that a better understanding of interpersonal adaptation of breathing may require a better description of socio-cultural and human factors [32]. It is also possible that partners were adapting more to the subjects than the reverse, as they were non-naïve, which could have preserved the subjects' rhythms.

The main difference in subjects' behaviors according to the partner was observed in the delay between speech offset and the onset of the next inhalation. Even if the effect was not statistically tested, it seems that subjects reduced this delay when talking with P02 as compared with P01. By contrast this delay was longer when talking with P02. Detailed analyses of turn types in the current dataset [33] showed that P02 was interrupted more often than P01, which could explain the longer delay at the end of the breathing cycle. The reverse could also be true: P02 could be interrupted more often due to longer pauses before inhalation. Despite the variability in subjects and partners' behaviors, and complex profiles due to inter-individual interactions, inhalation pauses during dialogue also showed strong and consistent patterns related to the dialogue events. The duration of inhalation was shorter when inhalation occurred inside a turn than before the start of a new turn. This "compression" of the breathing gap was also observed in the delay before the inhalation onset and after the inhalation offset. It suggests that partners and subjects avoid gaps inside turn by strongly reducing inhalation. This cue could signal to their interlocutor that they want to keep the floor. Such a strategy may limit the chance of the interlocutor to take the turn. It may also have physiological consequences such as hyper- or hypo-ventilation [3].

As readers used their breathing to indicate thematic changes or continuations [14], interlocutors of a dialogue may control the characteristics of their breathing noises – timing, amplitude and duration – to take or maintain their turn. Because the position of the microphone relative to the speakers' mouth was not precisely controlled in the current study, breathing noises could not be reliably analyzed. Faster inhalation may yet be related with louder noises, indicating to the interlocutor that the speaker wanted to keep or take the turn. By contrast, one could expect silent inhalations during listening, as inhalation noises produced by the listener may interfere with the perception of her interlocutor's speech.

Breathing during dialogue is a complex stream that alternates between different control levels of breathing, partially explained by the switching between speaking and listening [24]. Previous work found a tendency of interlocutors to breathe in-phase or in anti-phase at turn taking [24]. A similar analysis of interpersonal coordination in breathing has been carried out using the current dataset. Results showed no clear coordination of breathing profiles at a global level, but specific coordination patterns according to the type of turn [33]. Together with the current analyses, this seems to confirm the idea that during dialogue, breathing is strongly shaped by the communicative constraints [23]. However more studies are now required to understand the participation of breathing to these communicative constraints, the relationship between breathing changes and mutual accommodation in general [34] and the consequences of dialogue constraints on ventilation.

5 Acknowledgements

This work was funded by a grant from the BMBF (01UG0711) and the French-German University to the PILIOS project. The authors want to thanks Caroline Magister, Jörg Dreyer, Anna Sapronova, and Uwe Reichel for their help with data collection and labelling.

6 References

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

[2] McFarland, D. H. & Smith, A. (1992). Effects of vocal task and respiratory phase on prephonatory chest wall movements. *J Speech Hear Res, 35*, 971-982.

[3] Hoit, J. D., & Lohmeier, H. L. (2000). Influence of continuous speaking on ventilation. *J Speech Hear Res, 43*(5), 1240-1251.

[4] Grosjean, F. & Collins, M. (1979). Breathing, pausing and reading. *Phonetica*, 36(2), 98-114.

[5] Cutler, A., D. Dahan and W. Van Donselaar (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40, 141-201.

[6] Conrad, B., Thalacker, S., & Schönle, P. (1983). Speech respiration as an indicator of integrative contextual processing. *Folia Phoniatr, 35*, 220-225.

[7] Winkworth, A. L., Davis, P. J., Ellis, E., & Adams, R. D. (1994). Variability and consistency in speech breathing during reading: lung volumes, speech intensity, and linguistic factors. *J Speech Hear Res, 37,* 535-556.

[8] Whalen, D. H. & Kinsella-Shaw, J. M. (1997). Exploring the relationship of inspiration duration to utterance duration. *Phonetica, 54*, 138-152.

[9] Huber, J. E. (2008). Effects of utterance length and vocal loudness on speech breathing in older adults. *Respir Physiol Neurobiol*, 164, 323-330.

[10] Fuchs, S., Petrone, C. Krivokapic, J. & Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *J of Phonetics 41*, 29-47.

[11] Winkworth, A. L., Davis, P. J., Adams, R. D., & Ellis, E. (1995). Breathing patterns during spontaneous speech. *J Speech Hear Res, 38*, 124-144.

[12] Wang, Y. T., Green, J. R., Nip, I. S., Kent, R. D., & Kent, J. F. (2010). Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatr Logop, 62*, 297-302.

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

[14] Bailly, G. & Gouvernayre, C. (2012). Pauses and respiratory markers of the structure of book reading. *Proceeding of Interspeech*, 2012, Portland, OR. Paper?

[15] Whalen, D. H., Hoequist, C. E., & Sheffert, S. M. (1995). The effects of breath sounds on the perception of synthetic speech. *J Acoust Soc Am*, *97*, 3147-X.

[16] Ainsworth, S. (1939). Empathic breathing of auditors while listening to stuttering speech. *JSHD* IV, 139–156.

[17] Brown, C.T. (1962). Introductory study of breathing as an index of listening. *Speech Monographs*, 29(2), 79–83.

[18] Rochet-Capellan, A., & Fuchs, S. (2013). Changes in breathing while listening to read speech: the effect of reader and speech mode. *Frontiers in Psychology*, *4*.

[19] Bailly, G., Rochet-Capellan, A., & Vilain, C. (2013). Adaptation of respiratory patterns in collaborative reading. *Proceedings of Interspeech 2013, Lyon*, paper 54.

[20] Müller, V. & Lindenberger, U. (2011). Cardiac and respiratory patterns synchonize between persons during choir singing. *PLoS ONE* 6(9), e24893.

[21] Warner, R.M., Waggener, T.B. & Kronauer, R.E. (1983). Synchronized cycles in ventilation and vocal activity during spontaneous conversational speech. *J Appl Physiol Respir Environ Exerc Physiol*. 54(5), 1324–34.

[22] Autesserre, D., Nishinuma, Y. & Guitella, I. (1989). Breathing, pausing, and speaking in dialogue. *Proceedings of Europeech Paris,* 2433–2436.

[23] Guaïtella, I. (1993). Experimental study of respiration in spontaneous dialogue. *Folia Phoniatrica* 45, 273–279.

[24] McFarland, D. H. (2001). Respiratory markers of conversational interaction. *J. Speech Lang. Hear. Res.,* 44, 128-X.

[25] Babel, M. & Bulatov, D. (2011). The role of fundamental frequency in phonetic accommodation. *Speech and Language* 55(2) 231–248.

[26] Manson, J.H., Bryant, G.A., Gervais, M.M. & Kline, M.A. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behaviour* 34(6), 419-426.

[27] Levitan R., Hirschberg J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *Proceedings of Interspeech 2011*, Florence, 3081–3084.

[28] Boersma, P., & Weenik, D. (2011). Praat: doing phonetics by computer [Computer program]. Version 5.3, retrieved from http://www.praat.org/

[29] Reichel, U.D. (2012). PermA and Balloon: Tools for string alignment and text processing. *Proceedings of Interspeech 2012, Portland,* paper 346.

[30] Beňuš, S., Gavano, A. & Hirschberg, J. (2011) Pragmatic aspects of temporal accommodation in turn-taking. *J of Pragmatics* 43, 3001–3027.

[31] Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, *12*(4), 748-762.

[32] Giles, H., Coupland, N., & Coupland, J. (1991) Accomodation Theory: Communication, context, and consequence. *Contexts of Accomodation: Developments in Applied Sociolinguistics*, 1–68.

[33] Rochet-Capellan, A. & Fuchs, S. (under review).

[34] Lee, C.-C., A. Katsamanis, M. P. Black, B. Baucom, A. Christensen, P. G. Georgiou and S. S. Narayanan (2014). Computing vocal entrainment: A signal-derived PCA-based quantification scheme with application to affect analysis in married couple interactions. *Computer, Speech, and Language* 28(2), 518-539.