

Aleksandra Ćwiek, Marcin Włodarczak, Mattias Heldner, Petra Wagner

Acoustics and discourse function of two types of breathing signals

Abstract Breathing is fundamental for living and speech, and it has been a subject of linguistic research for years. Recently, there has been a renewed interest in tackling the question of possible communicative functions of breathing (e.g. Rochet-Capellan & Fuchs, 2014; Aare, Włodarczak & Heldner, 2014; Włodarczak & Heldner, 2015; Włodarczak, Heldner, & Edlund, 2015). The present study set out to determine acoustic markedness and communicative functions of pauses accompanied and non-accompanied by breathing. We hypothesised that an articulatory reset occurring in breathing pauses and an articulatory freeze in non-breathing pauses differentiates between the two types. A production experiment was conducted and some evidence in favour of such a phenomenon was found. Namely, in case of non-breathing pauses, we observed more coarticulation evidenced by a more frequent omission of plosive releases. Our findings thus give some evidence in favour of the communicative function of breathing.

1. Introduction

Respiration is a key aspect of speech and voice production. The air stream generated by the respiratory mechanism is the very first step to producing sound and the driving force for a majority of the speech sounds in the world's languages (e.g. Catford, 1988; Flohr & Müller, 2009). It has also been claimed that respiration is closely related to utterance planning (Hird & Kirsner, 2002) in that inhalation amplitude correlates both with sentence length (Fuchs et al., 2013) and with speech volume (Hoole & Ziegler, 1997; Ohala, 1990: 30 and 39).

In recent years, the question has been addressed whether respiratory behaviour could have a communicative function (e.g. Schegloff, 1996; Aare et al., 2014; Rochet-Capellan & Fuchs, 2014; Włodarczak & Heldner, 2015; Włodarczak et al., 2015; Heldner & Włodarczak, 2016). This work aims at providing evidence for a possible co-occurrence of certain articulatory and respiratory behaviours in pauses in speech.

As pointed out by Grosjean et al. (1979) and Grosjean and Collins (1979), pauses can be classified as breathing or non-breathing. As the terms suggest, breathing pauses include respiratory (and in particular exhalatory) noise whereas the non-breathing pauses does not. Local and Kelly (1986) examined silent pauses in spontaneous speech and suggested a

distinction between two types of silent pauses referred to as ‘holding’ silences and ‘trail-off’ silences (Local & Kelly, 1986: 195) where the ‘holding’ silences included breath holding accomplished by a “glottal hold”. The authors furthermore posited that these two types of silent pauses possessed different communicative functions: while ‘holding’ silences implied that the speaker wanted to maintain the turn, ‘trail-off’ silences signalled a willingness to yield the turn. In other words, they observed that subjects used a “glottal hold” to demonstrate their intention to continue speaking after the pause and exhalation noise to release the turn (Local & Kelly, 1986: 201ff). Further evidence for audibility of this phenomenon has not yet been provided. However, other studies have also indicated that various aspects of breathing behaviour are relevant for turn-taking (e.g. Rochet-Capellan & Fuchs, 2014; Heldner & Włodarczak, 2016).

Supported by the findings presented above, we hypothesise that a non-breathing pause is characterized by articulatory suspension. That is, the articulators “freeze” in their position until the articulation is continued after the pause. In addition, we hypothesize that a breathing pause (in addition to the presence of exhalation) is characterized by an articulatory reset during the pause whereby the articulators move back to their neutral position. We furthermore hypothesise that presence or absence of such an articulatory reset can be exploited for communicative purposes. The latter is a theoretical assumption resulting from the differentiation between two types of “silent” pauses. To verify this assumption, a reading experiment has been conducted, followed by acoustic analyses.

2. Method

The present study examines the possible co-occurrence of articulatory resets and breathing pauses by examining spectral similarity between segments preceding and following a pause. The spectral similarity measure used for the acoustic analysis was originally developed by Lewandowski (2012) to test the phenomenon of phonetic adaptation between speakers. A Matlab script written by Lewandowski (2012: 270) uses the amplitude envelope analysis and compares amplitude envelopes of two different signals. Using this method, a degree of similarity between two segments can be obtained on a scale from 0 (no similarity) to 1 (meaning maximal similarity, i.e. identical signals). The original script was modified to raise the sampling rate of the comparisons to ensure better results by shorter audio files. This was done because the originally compared files were of longer duration. In addition, we increased the scale from 0 to 10.

By experimentally controlling respiration and pause environment, we could compare similarity in different respiratory conditions. The hypothesised articulatory reset in breathing

pauses led us to expect a lower similarity in the breathing pause condition than in the non-breathing pause condition. The highest similarity values were thus expected to occur in cases where breathing was not controlled (a quasi-natural breathing and pause condition). The distinction between the two different pause environments of the silent pause allowed us to test, whether satisfactory similarity values of different phonemes preceding and following a pause could be obtained. The contrast between presence and absence of the intonation phrase boundary gives a possibility to investigate whether breathing pauses are more likely to occur at main syntactical boundaries, as stated by Grosjean and Collins (1979: 98).

2.1. Experimental Design

A reading experiment was conducted to test the hypotheses. Three variables were manipulated:

1. **Respiratory behaviour** was manipulated on three levels – (1) “silent” pause with respiration, potentially accompanied by noise, (2) silent pause without respiration and (3) natural (or uncontrolled) respiratory and pause behaviour. The presence of a silent pause does not entail the absence of underlying articulatory movements.
2. **The phonetic environment** preceding and following the pause was constructed to only consist of two variations – /p/ # /p/ or /a/ # /p/ where # stands for a pause and /a/ or /p/ segments represent the phonemes before and after the pause: an open front unrounded vowel and a voiceless bilabial stop.
3. **Intonation phrase (IP) boundary** was also manipulated on two levels: (1) a pause placed on a natural IP boundary and (2) a pause placed within a phrase, i.e. not on a natural IP boundary.

Prior to the final experiment, two pilot studies with two German speakers each were carried out, to test whether the task was not too demanding.

For the reading task a total of 40 sentences was designed, ten for each condition presented in Table 1. The sentences were read in three blocks to enable the best possible control of the breathing behaviour. During the first reading, participants were asked to make a pause at the specified place (marked by #) and take a breath, during the second reading, to make a pause and hold their breath and finally, during the third reading, to read the sentences in a naturally spontaneous manner. This order was kept constant throughout all participants. Admittedly, a different production order (i.e. beginning with the natural productions) may have resulted in an overall higher level of naturalness in the data recordings.

Table 1: Sentence distribution across the conditions with an example sentence.

Environment	IP boundary	Example sentence
/p/ # /p/	+	Es war knapp, # passte dann aber doch noch.
/p/ # /p/	-	Er starb # plötzlich, am 23. Juni 1877.
/a/ # /p/	+	Sonja mag Prosa, # primär liest sie aber Goethes Gedichte.
/a/ # /p/	-	Mehr als eine Pizza # pro Person ist etwas übertrieben.

2.2. Experimental Setting

Six female and three male speakers without known speech or hearing disorders participated in the experiment. Three participants reported having pulmonary disorders (in all cases asthma). However, as this is unlikely to influence the aspects of breathing under investigation, this condition did not disqualify them from participating. All speakers were native speakers of German. Additionally, two participants reported being bilingual. The age of the speakers at the time of the experiment ranged from 21 to 40 years (mean = 27,33, median = 27) and they were all students of Bielefeld University. The purpose of the study was not disclosed to the participants.

All recordings took place in a soundproof studio at Bielefeld University. Acoustic and respiratory signals were recorded and monitored simultaneously. The respiratory signal was registered using respiratory inductance plethysmography (RIP). For this procedure, two wired elastic bands are placed around the abdomen and around the chest. The belts are sensitive to stretching and can thus capture respiratory movements. A sample RIP signal is presented in Figure 1. Apart from the RIP measurements, an additional technique of recording participants' breathing was undertaken. A 3M Littmann Classic II S.E. stethoscope with a lavalier microphone inside the stethoscope tube was used to record audio signal of the respiratory sounds. The chest piece of the stethoscope was fixed on the larynx with a cloth strap. The RIP signal allowed to inspect whether the condition of breathing or sustaining a breath in a pause, respectively, was fulfilled by participants. The belts are however sensitive to any movement, therefore in ambiguous cases a stethoscope recording would help identify if the respiratory behaviour was in fact as requested.

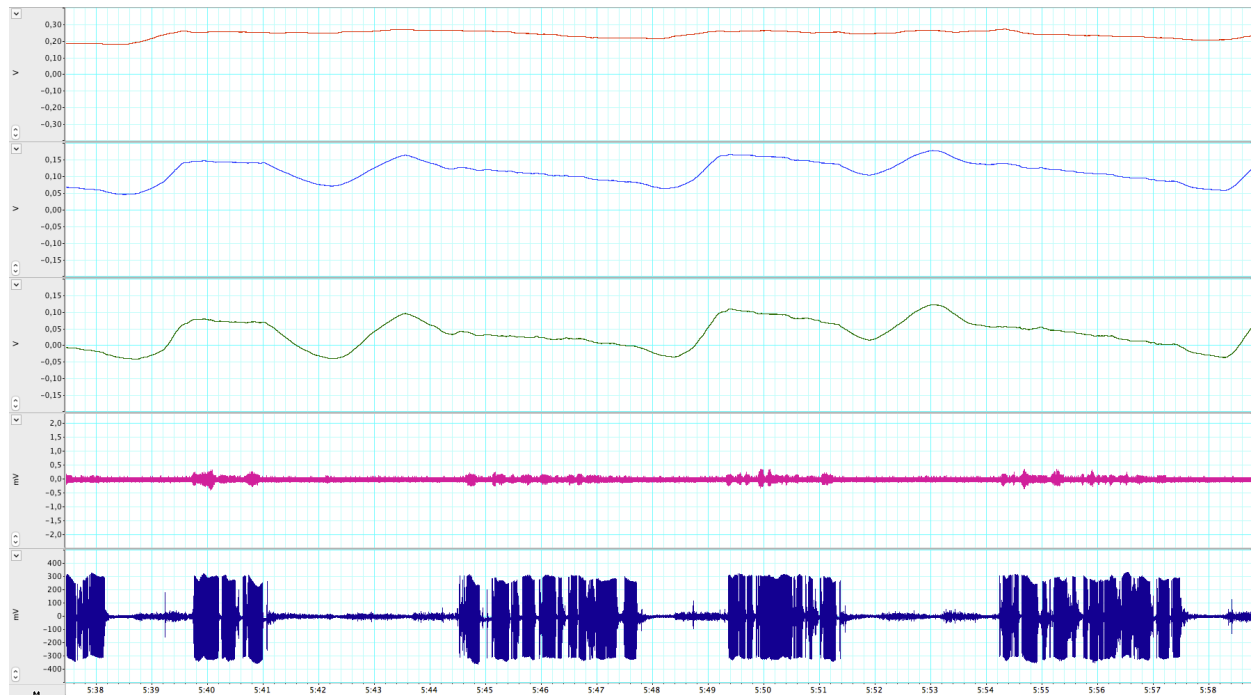


Figure 1: An example of a signal in LabChart. First three channels represent the RIP measurements, from top to bottom: abdomen, sum of both and chest signal. The fourth channel is a recording from a microphone, which was used for synchronisation with the audio signal of higher quality recorded with a separate programme. The last channel is the audio signal from the stethoscope.

3. Results

In order to test whether breathing and non-breathing pauses interact with articulatory behaviour before and after the pauses, a two-factorial ANOVA was carried out, with IP-boundary and respiratory condition as factors, and acoustic similarity before and after the pause as dependent variable. As this analysis yielded no statistical effect or interaction, the IP-boundary condition was left out of the further analyses. In addition, due to an excessive burst omission of the first segment /p/ in the pause environment /p/ # /p/, a further analysis of this condition was impossible as the respective token number decreased substantially.

Table 2: Means and standard deviations of the similarity measure in three breathing conditions for the pause environment /a/ # /p/.

Condition	Mean	SD
Breathing	4.65	1.01
Non-breathing	4.67	1.06
Natural	5.16	1.25

A one-way ANOVA was conducted to compare the effect of respiratory behaviour on amplitude envelope similarity between (1) breathing pause, (2) non-breathing pause, and (3) natural conditions for the pause environment /a/ # /p/, with a total number of 540 observations (i.e. N = 180 per condition). Means and standard deviations for the three conditions are shown in Table 2. There was a significant main effect of the respiratory

behaviour on the similarity measure [$F(2, 16) = 7.50, p = .005^{**}, \eta^2 = 0.29$].

A post hoc comparison (Tukey) revealed significant differences between the natural and both breathing and non-breathing conditions, as presented in Table 3.

Table 3: The results of post hoc Tukey test for the pause environment /a/ # /p/. The comparisons between the natural and both breathing and non-breathing conditions are highly significant.

	Estimate	Std. Error	z-value	Pr(> z)
non-br. – br.	0.01512	0.14943	0.101	1.000
nat. – br.	0.50860	0.14943	3.404	0.002**
nat. – non-br.	0.49348	0.14943	3.302	0.003**

Table 4: Burst omission frequencies in the first /p/ segment in condition /p/ # /p/.

Subject	Breathing	Non-breathing	Natural
1	0	0	11
2	2	14	8
3	1	8	12
4	3	9	9
5	1	0	9
6	5	13	13
7	0	0	9
8	0	0	6
9	2	0	11
Omissions	14	44	88
Non-omission	166	136	92

As mentioned above, burst omissions in the environment /p/ # /p/ were common. This lead to an extensive reduction in the amount of material in this condition, which made further analyses of the similarity data impossible. A chi-squared test of independence was performed to examine the relation between the burst omission rate and the three breathing conditions. The relation between these variables was highly significant, $\chi^2(2, N = 146) = 56.93, p < .001$. Another χ^2 test for the first two breathing conditions, that is breathing and non-breathing pauses, was also highly significant at $\chi^2(1, N = 58) = 15.52, p < .001$. Hence, there was not only a difference between all of the conditions, but most importantly between the first two – breathing and non-breathing. In other words, the speakers were more likely to omit the release of the plosive when they were planning to hold their breath during the following pause. The values used to perform the χ^2 tests are presented in Table 4. At the same time, a comparison of the values of each subject revealed large inter-speaker variation.

4. Discussion and Conclusion

The initially conducted two-factorial ANOVA revealed that the IP boundary had no effect on

the amplitude envelope similarity values between the three pause and breathing conditions. The one-way ANOVA results for the pause environment /a/ # /p/ showed a significant effect between the amplitude envelope similarity values of the three breathing and pause conditions. However, a post hoc Tukey test showed that only the natural condition was significantly different from the other conditions. Thus, the spectral similarity measure could not support the hypothesis of presence of an articulatory reset in the breathing pause condition.

Performing an ANOVA for the pause environment /p/ # /p/ was not possible due to lack of tokens caused by the frequent burst omissions in the segment preceding the pause. The χ^2 test run on the burst omission rates in the first /p/ segment in the pause environment /p/ # /p/ showed a highly significant difference not only between all three breathing and pause conditions, but also between the breathing and non-breathing pauses. This indicates that non-breathing pauses lean toward natural breathing and pause behaviour and the articulatory behaviour in this condition differs significantly from the one in a breathing pause.

Expectedly, the highest burst omission frequency occurred in the natural breathing and pausing condition. More interestingly, stronger anticipatory coarticulation was observed in the non-breathing condition – based on burst omission rates in the plosive segment preceding a silent pause – as opposed to significantly less anticipatory coarticulation in the breathing condition. This finding is in line with our initial hypothesis, namely that breathing pauses can be characterized as “articulatory resets” and non-breathing pauses, or breath holds, can be characterized as “articulatory holds”.

The difference found between breathing and non-breathing silent pauses supports our initial assumption that breathing behaviour in silent pauses could be exploited for communicative purposes. A distinction between different pause types based on their phonetic detail is not uncommon, and has a long tradition in disfluency research, where different pause types are distinguished based on their different phonetic form (filled and unfilled) and function (Shriberg, 1994). Our results show that breathing may be an additional phonetic cue that may be functionally exploited in the distinction of different pause types. Usage of respiratory cues could be profited from as conversational strategies in turn-management, coupled with other cues like intonation or lengthening.

Certainly, our findings are based on productions gathered under laboratory conditions, where speakers were specifically asked to breathe or to hold their breaths. It is unclear how these findings generalise to spontaneous interactions. In order to assess our hypothesis of a potential communicative function of different types of silent pauses, further investigation of conversational data is therefore necessary (Wagner, Trouvain & Zimmerer, 2015).

5. Acknowledgements

This work was funded by the Swedish Research Council project 2014-1072 *Andning i samtal* (*Breathing in conversation*) to the second and third author.

6. References

- Aare, K., Włodarczak, M., Heldner, M. 2014. Backchannels and breathing. *Proceedings of FONETIK 2014*. Stockholm, Sweden. 47–52.
- Catford, J. C. (1988). A practical introduction to phonetics (p. 183). Oxford: Clarendon Press.
- Flohr, H., & Müller, H. M. (2009). Grundbegriffe der Phonetik. In H. M. Müller (Ed.), *Arbeitsbuch Linguistik: Eine Einführung in die Sprachwissenschaft* (2nd ed., pp. 47-76). Paderborn: UTB - Schöningh.
- Fuchs, S., Petrone, C., Krivokapic, J., and Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics*, 41(1):29–47.
- Grosjean, F. and Collins, M. (1979). Breathing, pausing and reading. *Phonetica*, 36(2):98–114.
- Grosjean, F., Grosjean, L., and Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive psychology*, 11(1):58–81.
- Heldner, M. and Włodarczak, M. (2016). Is breathing silence? *Proceedings of FONETIK 2016*. Stockholm, Sweden.
- Hird, K. and Kirsner, K. (2002). The relationship between prosody and breathing in spontaneous discourse. *Brain and Language*, 80(3):536–555.
- Hoole, P. and Ziegler, W. (1997). A comparison of normals' and aphasics' ability to plan respiratory activity in overt and covert speech. *Speech production: Motor control, brain research and fluency disorders*, pages 205–212.
- Lewandowski, N. (2011). *Talent in nonnative phonetic convergence*. Doctoral dissertation, Universität Stuttgart. [Online]. Available: <http://elib.unistuttgart.de/opus/volltexte/2012/7402/pdf/Lewandowski.pdf>
- Local, J. and Kelly, J. (1986). Projection and 'silences': Notes on phonetic and conversational structure. *Human studies*, 9(2):185–204.
- Ohala, J. J. (1990). Respiratory activity in speech. In *Speech production and speech modelling*, pages 23–53. Springer.
- Rochet-Capellan, A. and Fuchs, S. (2014). Take a breath and take the turn: how breathing meets turns in spontaneous dialogue. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1658):20130399.
- Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E.

- Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Shriberg, E. E. (1994). *Preliminaries to a theory of speech disfluencies*. PhD thesis, Citeseer.
- Matlab (2009). *version 7.8.0 (R2009a)*. The MathWorks Inc., Natick, Massachusetts.
- Wagner, P., Trouvain, J., & Zimmerer, F. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics*, 48, 1-12.
- Włodarczak, M. and Heldner, M. (2015). Respiratory properties of backchannels in spontaneous multiparty conversation. In *Proceedings of ICPhS 2015*.
- Włodarczak, M., Heldner, M., and Edlund, J. (2015). Communicative needs and respiratory constraints. In *Proceedings of Interspeech 2015*, Dresden, Germany. ISCA.