The Control of Speech Breathing in Relation to the Upcoming Sentence

Susanne Fuchs[1], Phil Hoole[2], Dominique Vornwald[2], Anne Gwinner[3], Hristo Velkov[1,] Jelena Krivokapić[4]

1: ZAS Berlin, 2: IPSK Munich, 3: University Constance, 4: Yale University New Haven
E-mail: fuchs@zas.gwz-berlin.de

Abstract

This study investigates the extent to which the length of a pause and the inhalation depth reflect the planning of an upcoming sentence. Thoracic and abdominal volume changes during the pause and the upcoming sentence were recorded for 24 German speakers using Respitrace. The target sentences varied between long (24 syllables) or short (14 syllables) and simple or grammatically complex. Results provide evidence that thoracic / abdominal volume changes during sentence production are only partly anticipated via inhalation depth and duration. Speakers also make use of the expiratory reserve volume in the realization of longer sentences.

1 Introduction

We investigate the relationship between inhalation depth/duration and the length and grammatical complexity of the upcoming utterance. Our work was inspired by Whalen and Kinsella-Shaw (1997), who discuss this relation from a speech planning perspective. The hypothesis they investigated was that the longer the upcoming sentence, the deeper the inhalation and/or the longer the inhalation duration, assuming there is an anticipation of the upcoming sentence length during the breathing pause. The assumption in their study was that the amount of air inhaled at the onset of speech is sufficient for the whole upcoming utterance. However, Wilder (1983) reported that the inhalation is relatively independent of sentence length and complexity, but that the air consumption during sentence production depends to a great extent on sentence length and complexity.

Whalen and Kinsella-Shaw (1997) found that there is a significant correlation between utterance length and inhalation duration, and that complexity of the sentence did not affect inhalation duration. Our study extends the investigation of the relation between inhalation depth / duration and the length and complexity of an upcoming utterance by using a greater sample (24 speakers compared to three speakers in Whalen & Kinsella-Shaw 1997) and a more natural experimental protocol (see the methods section for more details).

2 Aims of this study

The general aims of our study are to investigate the relation between sentence length, grammatical complexity and breathing pauses under controlled conditions in order to: 1) examine the planning of the upcoming sentence occurring during the pause at utterance onset and 2) investigate speaker-specific mechanisms of abdominal and thoracic breathing control during read speech production.

3 Methods

3.1 Experimental design and speech material

Whalen and Kinsella-Shaw (1997) mention a methodological problem occurring in most previous studies where authors did not control for the length and complexity of the preceding sentence, which is problematic, since pause duration can be a result of both, the preceding and upcoming sentence. In order to control for this, Whalen and Kinsella-Shaw's subjects uttered each time only the target sentence of varied length and grammatical complexity after doing some breathing maneuvers. Before we started our recording sessions we did an acoustic pre-test with 10 German speakers, adapting the methods described in

Whalen and Kinsella-Shaw (1997). Since many subjects found the task to do breathing maneuvers before speaking difficult, we chose a different procedure. In our study, the utterance of interest was always preceded by a sentence with a fixed length and ending with a colon. We decided to adjust the length of the preceding sentence for each subject so that it would naturally induce the subjects to breathe in, without causing them to be completely out of breath. As opposed to Whalen and Kinsella-Shaw (1997) we did not instruct our subjects to read the target sentence within one breath, but varied the length of a target sentence within a range that allowed the sentence to be spoken within one breath.

The corpus consisted of short and long sentences (14 vs. 24 syllables) with 2 degrees of grammatical complexity, complex and simple (determined by the presence vs. absence of a finite subordinate clause for short sentences and the presence vs. absence of a restrictive vs. non-restrictive subordinate clause for long sentences). For example:

short, simple

Sonja Wunderlich besuchte die Komische Oper. (Sonja Wunderlich visited the Komische Oper.)

short, complex

Sonja Wunderlich, die Tanz liebt, besuchte die Oper.
(Sonja Wunderlich, who loves dance, visited the Komische Oper.)

long, simple

Sonja Wunderlich bestaunte in einer warmen Sommernacht im Monat August die Oper.
(Sonja Wunderlich marveled at the opera during a warm summer night in August.)

long, complex

Sonja Wunderlich sagte einem Freund, der uns morgens anrief, sie begeistert sich für Oper.
(Sonja Wunderlich told a friend who phoned us in the morning that she is crazy about opera.)

All sentences started either with <Sonja> /zɔnja/ or <Tonja> /tɔnja/ in order to investigate effects of breathing control on initial obstruents. Sentences always ended with the same noun <Oper> /o:pɐ/ in order to investigate the effect of sentence length and the amount of thoracic/abdominal volume changes during the production of obstruents in final utterance position. All sentences were presented in a randomized order and repeated 10 times, giving a total of 2160 sentences (90 sentences * 24 speakers). To date abdominal and thoracic volume changes have been analyzed separately, since there is a large amount of inter-speaker variability between these two measures.

Thoracic and abdominal expansion during speech and quiet breathing has been recorded using Respitrace. Simultaneously, intraoral pressure (IOP) has been monitored with a small piezo-resistive intraoral pressure sensor (Endeveco) that was glued to the posterior end of the subject’s hard palate. These data, together with the acoustic recordings, were saved using a multi-channel DAT recorder and were subsequently imported and analyzed using Matlab, and PRAAT for the acoustic analyses. 24 native speakers of German served as subjects (9 males and 15 females). We report the results of the 14 speakers analyzed to date.

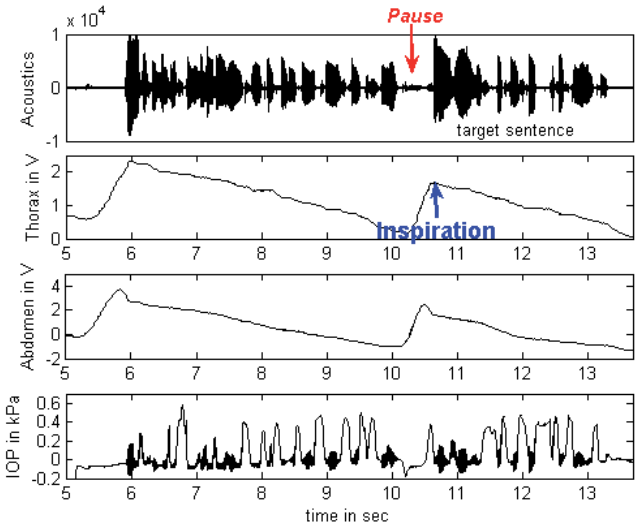

Figure 1: *Data obtained synchronously (1st track acoustics, 2nd track thoracic changes, 3rd track abdominal changes, 4th track intraoral pressure*

3.2 Labeling procedures

Based on the acoustic signal the beginning and end of the pause, the target sentence, the first word, and the first obstruent were labeled. All other measures were carried out semi-automatically. Thoracic and abdominal inhalation depth was calculated as the difference between the maximal and the minimal expansion from the beginning of the pause till the end of the first word. Volume changes during sentence production were calculated as the difference between maximal thoracic / abdominal expansion and their corresponding values at the end of the sentence.

4 Results

All tokens where speakers realized more than one breathing pause (or any other kind of pause) were excluded. Only significant correlations with p<0.05 are reported. Since our corpus involved only 2 different sentence lengths (l=long and s=short) and 2 different kinds of grammatical complexity (n=not complex, c=complex) the data are in many cases bi-modally distributed. Thus, for all the correlations we used the Spearman's rho correlation coefficient, which does not assume a normal distribution. In the majority of the analyses run, sentence length had an effect on the dependent variable, but grammatical complexity had not. We report only the most general results.

4.1 Acoustic pause duration and inhalation depth

Acoustic pause duration correlates positively with thoracic inhalation depth (R = 0.45-0.81, except for 2 speakers who do not show an effect) and to a smaller degree with abdominal inhalation depth (R = 0.25-0.66, except for 2 speakers who do not show an effect).

4.2 Acoustic sentence duration and thoracic/ abdominal volume changes

Sentence duration correlated highly with thoracic volume changes during the whole sentence (R = 0.71-0.93) and to a lesser degree and a higher inter-speaker variability with abdominal volume change (R = 0.3-0.88). The longer and the more complex a sentence, the larger the volume change, in particular for the thorax.

4.3 Relationship between breathing pauses and upcoming sentences

Results for 6 out of 14 speakers showed a weak positive correlation between acoustic pause and sentence duration (R = 0.22-0.48). Figure 2 provides an overview for all data when the 14 speakers were pooled together. It shows that sentence duration varies from short / not complex to long / complex sentences (2-4.5 sec), but that the acoustic pause duration is relatively stable in comparison (400-800 ms).

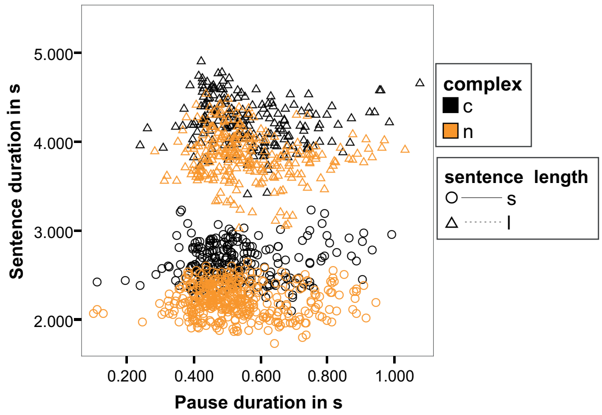

Figure 2: *Scatterplot with pause duration (x-axis) and sentence duration (y-axis) for all the 14 speakers. Black markers: complex sentences, orange markers: not complex sentences; triangles: long sentences, dots: short sentences.*

When taking the physiological measures (thoracic/ abdominal) inhalation depth and volume change during sentence production into account, the correlations between the two variables become slightly higher. For 10 out of 14 speakers a weak to moderate positive correlation was found between thoracic inhalation depth and thoracic volume changes during an upcoming sentence (R = 0.27-0.59) and for 11 speakers we observed a correlation between abdominal inhalation depth and volume changes during an upcoming sentence (R = 0.25-0.72).

This means that most speakers take a deeper breath for longer or more complex sentences. However, the thoracic volume differences during inhalation before long and short sentences are often marginal and therefore not significant. Figure 3 displays the mean values of thoracic changes during inhalation for each subject. Different markers correspond to the different conditions (black filled dots = long complex sentences, empty dots = long, not complex sentences, filled squares = short complex sentences, empty squares = short not complex sentences).

In contrast, differences in thoracic changes during sentence production are in most cases significant between the different conditions as can be seen in Figure 4. Long and complex sentences coincide with larger thoracic changes in comparison to shorter and not complex sentences.

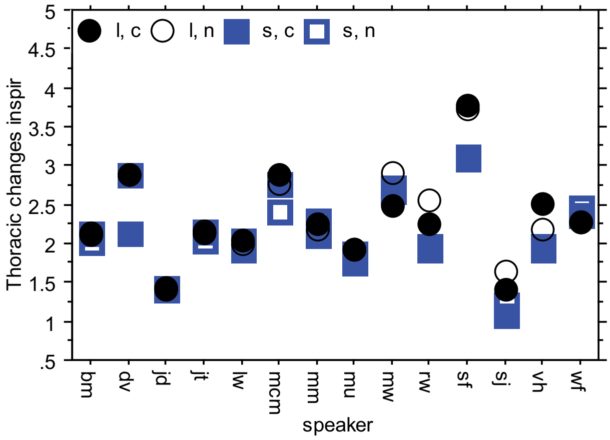

Figure 3: *Scatterplot with means of thoracic changes during inhalation in V, different markers correspond to the different sentence conditions: l=long, s=short, c=complex, n=not complex*

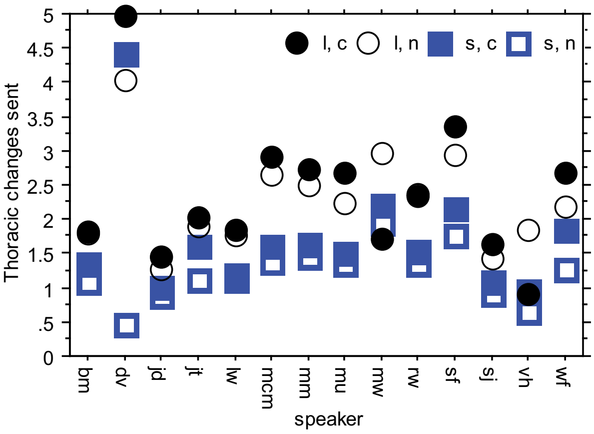

Figure 4: *Same as figure 3, but for means of thoracic changes during sentence production*

That means that the relation between the two variables is not a one-to-one. Our preliminary analysis provides evidence that inhalation depth would increase only of about 30% (or even less) for a doubling in the amount of thoracic volume change.

5 Discussion

Is there evidence for planning during breathing pauses, i.e., do speakers' breathing patterns show anticipation of the upcoming sentence? The correlations found so far are rather weak and provide only partial evidence for an anticipation of the upcoming utterance. The differences between Whalen and Kinsella-Shaw's findings and ours might on the one hand be related to the fact that they used a much wider variety of sentence length (48 sentences ranging from 5 to 82 syllables compared to 2 degrees of length in this study, namely 14 and 24 syllables) and grammatical complexity (with 1 to 7 embedded clauses in Whalen and Kinsella-Shaw in comparison to 2 degrees of complexity in this study) and they were therefore able to provide a more global picture on the relation between inhalation depth and sentence duration. On the other hand Whalen and Kinsella-Shaw's findings may also be different since they instructed their subjects to read the sentences in a single breath. This task may have caused speakers to inhale deeper for longer and more complex sentences which may not be the case in a more natural setting where speakers can also increase breathing frequency or rely on the expiratory reserve volume if necessary.

6 Conclusion

We conclude that thoracic/abdominal volume change during sentence production is only partly anticipated via inhalation depth. Speakers also make use of the expiratory reserve volume in the realisation of longer sentences. Based on our results we also suggest that pause and sentence duration are only linearly correlated within a certain range, i.e. there are some limits on how many words/syllables can be anticipated in a preceding pause. However, further work is clearly needed to investigate this issue.

Acknowledgements

This work was supported by a grant from the Federal Ministry of Education and Research to the PB1 project at ZAS and from the DFH Saarbrücken to the German-French project PILIOS.

References

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

[2] C.N. Wilder. Chest wall preparation for phonation in female speakers. In Bless, Abbs. *Vocal fold physiology: Contemporary research and clinical issues.* 109-123. College Hill Press: San Diego, 1983.