



HAL
open science

Adaptation of respiratory patterns in collaborative reading

Gérard Bailly, Amélie Rochet-Capellan, Coriandre Emmanuel Vilain

► **To cite this version:**

Gérard Bailly, Amélie Rochet-Capellan, Coriandre Emmanuel Vilain. Adaptation of respiratory patterns in collaborative reading. Interspeech 2013 - 14th Annual Conference of the International Speech Communication Association, Aug 2013, Lyon, France. pp.1653-1657. hal-00851890

HAL Id: hal-00851890

<https://hal.science/hal-00851890>

Submitted on 19 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptation of respiratory patterns in collaborative reading

G rard Bailly, Am lie Rochet-Capellan & Coriandre Vilain

GIPSA-lab, UMR 5216 CNRS/Universit  de Grenoble -France

{firstname}.{lastname}@gipsa-lab.grenoble-inp.fr

Abstract

Speech and variation of respiratory chest circumferences of eight French dyads were monitored while reading texts with increasing constraints on mutual synchrony. In line with previous research, we find that speakers mutually adapt their respiratory patterns. However a significant alignment is observed only when speakers need to perform together, i.e. when reading in alternation or synchronously. From quiet breathing to listening, to speech reading, we didn't find the gradual asymmetric shaping of respiratory cycles generally assumed in literature (e.g. from symmetric inhalation and exhalation phases towards short inhalation and long exhalation). In contrast, the control of breathing seems to switch abruptly between two systems: vital vs. speech production. We also find that the syllabic and the respiratory cycles are strongly phased at speech onsets. This phenomenon is in agreement with the quantal nature of speech rhythm beyond the utterance, previously observed via pause durations.

Index Terms: respiration, coordination of speech and respiration, rhythm, pauses

1. Introduction

During speech production, the physiological need of air inhalation interrelates with the need of air exhalation to produce speech sounds. The interrelation of breathing and speech production depends on higher-factors such as the syntactic and prosodic structure of the discourse. Breathing is also involved in joined actions during interactions and more specifically, verbal communication, notably the coordination of turn-taking.

Pulmonary ventilation is controlled by two systems: the ventral and dorsal breathing centers in the medulla and the pneumotaxic and apneustic centers of the pons in the brain. The central pattern generators of the medulla seem to be responsible for the automatic and permanent generation of rhythmical respiratory patterns. This generation is largely below the level of consciousness. The frequency, the amplitude and the shape of the ventilation automatically adapts to physiological requirements (temperature, emotion, cardiovascular state, vigilance, etc). Forced inhalations and exhalations can also be monitored by the pons that overrides the automatic regulation.

During quiet breathing, inhalation and exhalation are similar in amplitude and duration (symmetrical). The patterns of respirations when speaking are very different from the rather symmetric patterns in quiet: slow expirations follow rapid or even partial inhalations [1, 2] and a greater ventilation is generally needed for speech than for quiet breathing [3]. The content and structure of the discourse shape the mutual patterning of speech and respiration. During speech production, inhalation pauses mainly occur at phrase or sentence boundaries [4, 5]. These pauses combine with linguistic cues (e.g. syntactic structure or lexical choices) and prosodic cues (e.g. prepausal lengthening, boundary tones) to

signal links between previous speech chunks with the upcoming flow of speech. Breathing pauses are also involved in speech perception [6, 7] and participate to turn-taking coordination in the listener-speaker interaction [1, 8].

McFarland [1] found that respiratory movements may provide a valuable marker of conversational exchange and coordination between participants. Changes in breathing patterns were indeed observed during listening and significant synchronicity was observed during periods of simultaneous speech and laughter as well as at turn-taking boundaries. Most authors agree on a continuum of respiratory control according to the demand of the task [1, 2, 9]. The present work aims at analyzing this hypothetical gradual adaptation of respiratory patterns to increased synchronization constraints.

2. Experiment

In our experiments, speakers were asked to read the same 13 texts under four various conditions. Each text was a short paragraph (each 10 ± 1.5 sentences; 131 ± 15 syllables). The conditions varied the coordinative constraint with the reading partner.

2.1. Conditions

The four experimental conditions were the following:

1. Reading aloud alone
2. Reading aloud paragraphs in alternation
3. Reading aloud sentences in alternation
4. Reading aloud texts in synchrony/chorus [10]

Six minutes of quiet breathing (condition 0) were recorded at the beginning of the session.

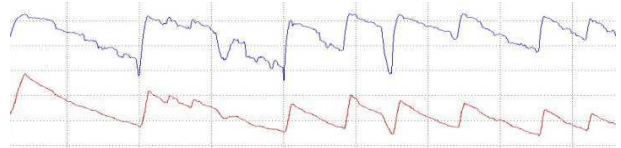


Figure 1. Comparing the variation of the thoracic circumference estimated by the VISURESP[®] vest (bottom) and the BIOPAC[®] belt (top) during speech production.

2.2. Experimental settings and procedure

Reading partners sit in two different rooms and wear headset microphones. The text is presented as a PowerPoint presentation with one paragraph per page. The experimenter toggles the page once the paragraph has been read. Screens of both speakers are thus refreshed simultaneously in condition 1. Prior to reading, speakers go through the 13 texts with the experimenter in order to agree on the way to pronounce unusual proper nouns. They then went through conditions 0 to 4. The feedback from the partner is switched off for conditions 0 and 1. Any text produced with disfluencies or laughters (that occur sometimes drastically during condition 4) was read again. A typical recording session lasts about one hour.

Variations of respiratory chest circumferences are monitored by two BIOPAC[®]SS5LB belts. Following McFarland [1],

each belt is placed around the subject's thorax with strain transducers placed aside just below the right axilla. The best position for the belt was determined from the results of a calibration session where we compared movements recorded by the SS5LB belts and the Resptrace® vest (see Figure 1). We found that this positioning of the BIOPAC®SS5LB belt trustfully captures thoracic movements while minimizing the impact of mechanical coupling of the rib cage with heart and head movements.

2.3. Subjects

In order to further study inter-speaker adaptation, one male (gb) and one female (arc) speakers each interacted with four different reading partners of the same gender and similar social status. All pairs knew each other well.

2.4. Processing of data

Speech data is first automatically forced-aligned with phonetic transcriptions of text fragments thanks to speaker-independent HMM. Inhalation and exhalation respiratory phases are detected using the zero-crossings of the acceleration of the respiratory signal in the bandwidth [0.05Hz-10Hz]. All segmentations are then checked by hand with Praat [11].

3. Results

3.1. Respiratory patterns

Figure 2 displays the elementary ventilation cycles collected during one session for one female dyad. For each speaker, we distinguish production vs. perception cycles respectively concomitant with the other speaker's speech production and perception. This was done for condition with or without mutual acoustic feedback. Our data reproduce the low skewness of the profiles during quiet breathing that contrast with the high skewness of profiles collected during speech production [2, 12]. But contrary to the hypothetical "[...] continuum of respiratory control, with perhaps quiet breathing and speech at the extremes [...]" [1, p.130, see also 9], the observed control strategies are more in line with an abrupt switch between two modes: (a) a regular patterning of respiration with symmetrical profiles vs. (b) an irregular one at the service of speech production. The two types of respiratory cycles will be terms *vital* vs. *speech* cycles in the following. We will argue that speech perception does not involve the active recruitment of vital cycle. Figure 3 shows two examples characteristic of these rapid switches that occur even in condition 3. Note that two of our interlocutors often favored apnea when leaving the floor: one is a trombone player and the other proved to be an apnea swimmer!

This rapid switching militates for the takeover of a vital function by a cognitive process. Such quick blocking of regular breathing has also been evidenced for deglutition [13] where the apnea is able to interrupt the respiratory cycle at anytime.

We computed three characteristics of each respiratory cycle: (1) its period; (2) its inhalation quotient (duration of the inhalation over the period); (3) its phase with reference to the closest cycle of the interlocutor. The general structure of phasing relations between cycles produced by interlocutors is illustrated in Figure 4. We do not obtain regular phasing relations between close respiratory patterns in condition 2 - where the interlocutor is listening to the speaker reading a whole paragraph. Condition 3 is characterized by rapid

switching and more apnea than condition 2. In this condition, partners' cycles tend to be phase locked. With no surprise Condition 4 compels speakers to strongly synchronize their inhalations.

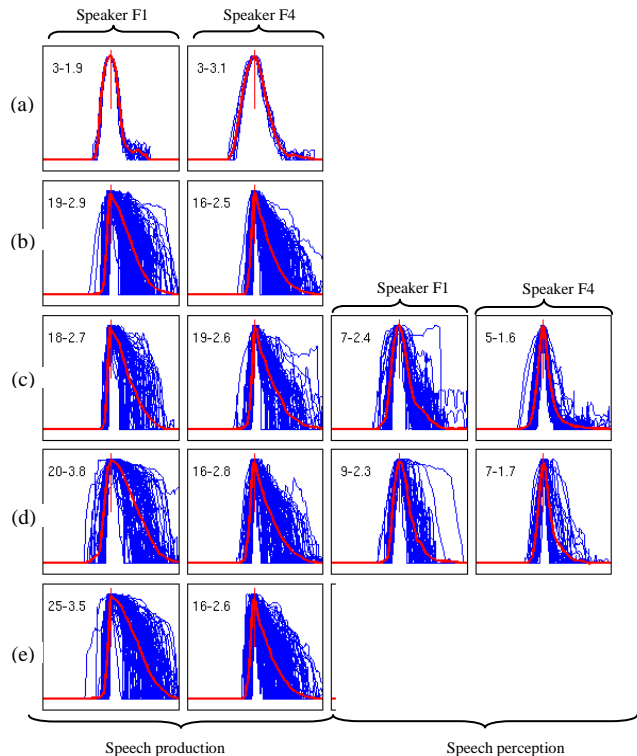


Figure 2. Ventilation profiles (variation of rib cage circumference against time) for the female dyad #4 (speakers F1 and F4). Conditions for top to bottom: (a) quiet breathing, (b) reading alone, then with mutual acoustic feedback, reading in alternation (c) paragraphs and (d) sentences, (e) in chorus (speech production & perception profiles equal here). Note the switch between a low skewness (first number ranges 3-10 in each caption) for breathing and hearing condition in comparison with the high asymmetry (skewness close to 20) for speech production. The second number gives the average period in s.

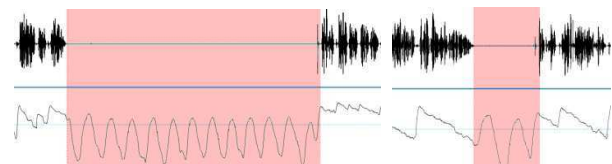


Figure 3. Abrupt switch between speech and quiet breathing patterns. The variation of the breast section is displayed under the acoustic signal. Respiratory patterns of speaker F2 reading aloud alternatively paragraphs (condition 2 displayed on the left) vs. sentences (condition 3 displayed on the right). Note that speech cycles seem to "interrupt" vital cycles asynchronously.

These observations are coherent with the switching hypothesis between control structures evidenced previously. If the cognitive monitoring of respiratory patterns overrides the automatic regulation with no negotiation, the patterns generated by the two systems have no reason to be related. The phase locking in condition 3 a priori between two different types of cycle may be resulting from three phenomena (see Figure 5): (a) the use of apnea; (b) the use of vital cycles for speaking; (c) the absence of unlocking because of a fast turn taking rate.

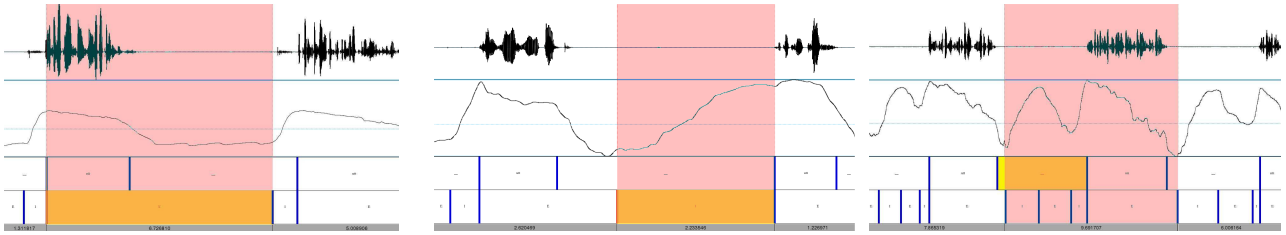


Figure 5: Handling turns in condition 2. From left to right: apnea (speaker M1); using/prolonging a vital cycle (speaker M3); approximate phase locking by alternating with vital cycles (speaker M3).

Table 1. Mean asymmetry coefficient of the respiratory cycles according to conditions. For each speaker and listener, we gray the conditions for which distributions of these coefficients are statistically different from condition 1 ($p < .05$). Note that most listening periods fulfill this hypothesis.

Dyads	Speaker	Speaker						Interlocutor							
		Speak			Listen			Speak			Listen				
		0	1	2	3	4	2	3	0	1	2	3	4	2	3
1	F1	29	19	17	21	13	30	36	31	21	21	23	23	23	35
2	ar F2	36	17	14	20	19	39	37	20	25	17	16	25	18	32
3	c F3	28	18	25	20	9	28	30	39	28	24	26	13	40	33
4	F4	27	14	12	16	14	31	36	34	20	17	20	17	27	36
5	M1	36	34	-	24	26	-	32	26	21	-	14	23	-	38
6	gb M2	36	34	31	26	31	31	33	37	26	29	25	23	33	48
7	M3	38	32	29	25	24	31	37	35	21	19	25	16	33	52
8	M4	37	25	23	21	20	32	38	29	15	15	18	14	30	37

3.2. Inter-speaker adaptation

Figure 6 displays the evolution of the mean ventilation period and the mean syllabic rate of dyads according to conditions. If speakers adapt their speaking rate to increased coordination, they do not seem to be compelled to adapt their respiratory strategy: male dyads maintain notably different ventilation periods in condition 4 despite a perfect adjustment of their speaking rate.

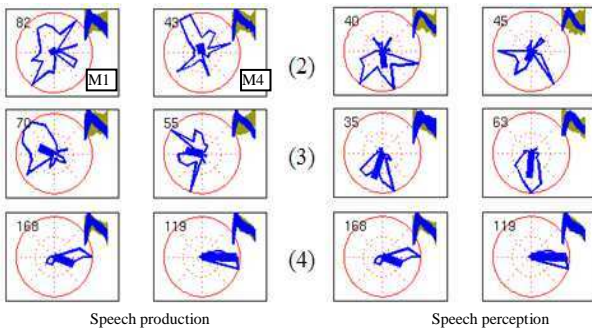


Figure 4: Phase relations between respiratory cycles of the interlocutors of the male dyad #8. No consistent phase relations are observed in condition 2 (first row). There is a tendency for cycles to be phase locked in condition 3 (π for production and $-\pi/2$ for perception), while being strongly locked in phase in condition 4 (second and third row). Normalized speaker's average respiratory cycles are superposed with the corresponding interlocutor's average cycle for each condition (top right). Note the more symmetric profiles for the listening conditions 2 and 3.

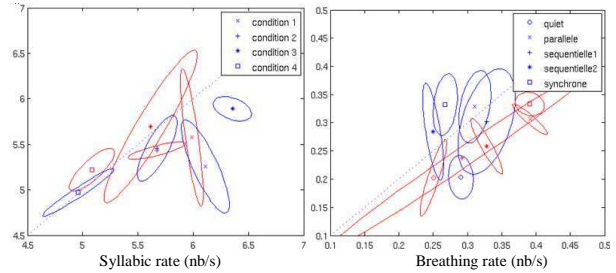


Figure 6. Inter-speakers adaptation. Interlocutor's mean values (ordinate) vs. speaker's mean values (abscissa). The proximity to the dotted line means convergence of behavior. If speakers mutually adapt their speaking rate to increase coordination, they do not always adapt their respiratory strategy (see notably males figured in blue).

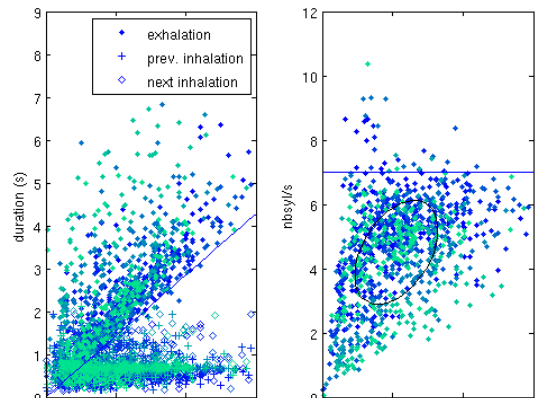


Figure 7. Durations of the inhalations and exhalations as a function of the number of syllables in the breath group to be produced. Our data show an absolute limitation of 7 syllables per expired second and of about 1s for the inhalation.

3.3. Breath groups

Phonotactic and linguistic factors such as the structural organization of the text exert a strong influence on the size of the breath group and thus on the duration and amplitude of the respiratory cycle that will carry the utterance. The duration of the exhalation is known to correlate with the number of syllables in the breath group [14]. The influence of the previous and next groups on the intermediate pause is controversial [15] and the relationship between pause and inhalation durations is rather weak since silent as well as audible exhalation or inhalation can be used to encode text and discourse structure [16].

Figure 7 shows that for the French read corpus studied here, the breathing cycle of speech production adapts to the breath group: while the inspiratory duration is rapid and limited to less than 1s, the expiratory phase rarely goes below 140ms per syllable... As displayed on Figure 8, the expiratory phase

extends largely the phonation that be released in the interval 20-80% of the available duration of expiration.

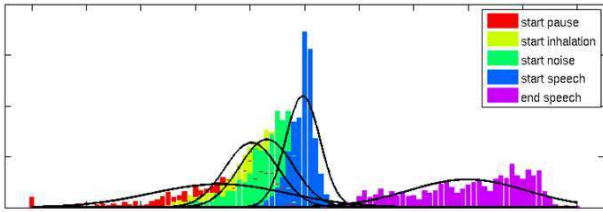


Figure 8. Distributions of respiratory and acoustic events within two normalized expiratory duration. Inhalation lasts approximately 20% of the expiratory duration. When it is made audible, the breathing noise lasts 50% of that duration. Speech starts almost in synchrony with the exhalation and ends between 20% and 80% of the expiratory duration.

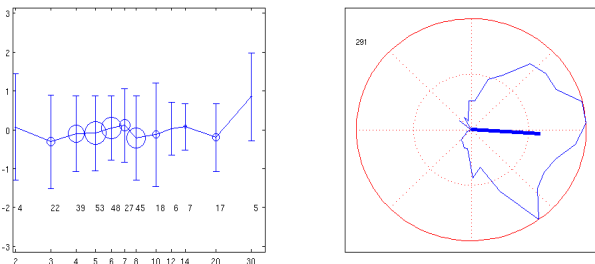


Figure 9: Synchronization between syllabic clocks across two adjacent respiratory cycles. Left: Mean and standard deviations of the phase according to estimated silent cycles. Right: General phase diagram.

3.4. Coordination between speech and respiration

Several authors have shown that the modal distribution of silent pauses [4, 16, 17] and that these modes are harmonic of each other. We questioned our data for a possible coordination of the planned respiratory cycles with the rhythmic organization of speech. Following our previous work [18], we choose to characterize speech rhythm by the successions of vowel onsets. We compute an underlying regular syllabic clock for the breath group preceding the pause as the clock with a constant frequency equal to the mean syllabic duration of that breath group that minimizes phase shifts with the vocalic onsets. We then considered the mean phase shift between this virtual clock and the first three vocalic onsets of the next breath group. The results are display in Figure 9: surprisingly, the virtual syllabic clock of each breath group seems to go over the silent pause for a rather large number of interleaved beats (1 to 20) and start smoothly in phase with the actual vocalic onsets of the next breath group.

This apparent coordination of respiratory and syllabic clocks may explain the quantal effect observed in the distribution of pause durations: the pause interval that encompasses silent expiration, the eventual apnea and next inspiration is paced by an underlying smooth time-keeper that spans long stretches of speech production. Note that these distributions have been computed within paragraphs uttered during conditions 1, 2 and 4. We did not observed such synchronization regularities across speakers for condition 3.

4. Discussion

We argue here that the double neuro-anatomical control of the breathing should be mirrored somehow in the organization of the breathing cycles. The switching hypothesis that we put forward here does not prevent any kind of mutual influence between these controls. As put forward in the introduction, the central pattern generators of the medulla automatically adapts the frequency, amplitude and shape of the ventilation for the sake of maintaining a stable state of physiological resources. In the same way of emotional arousal or increase of ambient temperature, adaptation and regulation of ventilation is required after speaking, and even more between speeches. No surprise that the average vital breathing characteristics changes between and during speeches to prevent and anticipate irregular oxygen deficits.

We however do not observe strong switching patterns for some dyads. Their distributions of shape parameters in these cases were often bimodal. This may be due to the fact that the text was available to the listeners and that they were focusing on smooth turn taking in the absence of a competing cognitive task. It would be interesting to suppress textual information and impair mutual attention with a secondary task focusing on linguistic content such as phoneme or syllable monitoring. On the reverse, rapid recovery of the vital control could be favored by increasing the oxygen debt during speech production (speaking louder, quicker or with no breath).

As in previous works, we showed that respiratory patterns may change very abruptly and that the respiratory cycle can be shaped and timed in a flexible way to fulfill cognitive demands [19]. Within the limit of our experimental paradigm, we found little evidence of coordination between respiratory patterns of speakers and listeners until they are really involved in a joint action.

5. Conclusions and perspectives

In our study, respiratory strategies and changes in breathing patterns are a complex by-product of reading, speech planning and coordinative demands. The passive submission of the respiratory system to the cognitive demand is unrealistic: if speakers are able to manage large oxygen deficits, the long-term recovery of the debt may result in out of breath or hyperventilation strategies. We should now put more effort in understanding how speakers regulate this oxygen debt. The use of respiratory belts combined with plethysmographs and pneumotachographs may provide a more comprehensive understanding of the short- and long-term control of air volume.

Dyads were in separate rooms and were instructed not to move their body. However, co-presence and gesturing may be relevant factor to understand the coordinative respiratory control. We also choose people that new each other quiet well. Together with the training of F1 and M1 to the task, this may have an effect on the coordinative profiles we observed. It will be now interesting to study the impact of social relationship and affect on respiratory patterning when speaking and inter-speakers coordination.

6. Acknowledgements

Part of this work was financed by the ANR ORTHOLEARN.

7. References

- [1] McFarland, D.H., *Respiratory markers of conversational interaction*. Journal of Speech and Hearing Research, 2001. **44**: p. 128-143.
- [2] Conrad, B. and P. Schönle, *Speech and respiration*. Archiv für Psychiatrie und Nervenkrankheiten, 1979. **226**: p. 251-268.
- [3] Hoit, J.D. and H.L. Lohmeier, *Influence of continuous speaking on ventilation*. Journal of Speech and Hearing Research, 2000. **43**: p. 1240-1251.
- [4] Campione, E. and J. Véronis. *A large-scale multilingual study of silent pause duration*. in *Speech Prosody*. 2002. Aix-en-Provence, France.
- [5] Winkworth, A.L., et al., *Variability and consistency in speech breathing during reading: Lung Volumes, speech Intensity, and linguistic factors*. Journal of Speech and Hearing Research, 1994. **37**: p. 535-556.
- [6] Whalen, D.H., C.E. Hoequist, and S.M. Sheffert, *The effects of breath sounds on the perception of synthetic speech*. The Journal of the Acoustical Society of America, 1995. **97**(5): p. 3147-3153.
- [7] Lieberman, P., *Intonation, Perception and Language*. 1967, Cambridge MA: MIT Press.
- [8] Guaitella, I., *Etude expérimentale de la respiration en dialogue spontané*. Folia phoniatrica, 1993. **45**: p. 273-279.
- [9] Smith, A. and M. Denny, *High-frequency oscillations as indicators of neural control mechanisms in human respiration, mastication, and speech*. Journal of Neurophysiology, 1990. **63**: p. 745-758.
- [10] Cummins, F., *Practice and performance in speech produced synchronously*. Journal of Phonetics, 2003. **31**(2): p. 139-148.
- [11] Boersma, P. and D. Weenink, *Praat, a System for doing Phonetics by Computer, version 3.4*, in *Institute of Phonetic Sciences of the University of Amsterdam, Report 132*. 182 pages. 1996.
- [12] Hoit, J.D., et al., *Speech breathing in women*. Journal of Speech and Hearing Research, 1989. **32**: p. 353-365.
- [13] Martin-Harris, B., et al., *Breathing and swallowing dynamics across the adult lifespan*. Archives of Otolaryngology - Head and Neck Surgery, 2005. **131**: p. 762-770.
- [14] Whalen, D.H. and J.M. Kinsella-Shaw, *Exploring the relationship of inspiration duration to utterance duration*. Phonetica, 1997. **54**: p. 138-152.
- [15] Zvonik, E. and F. Cummins. *The effect of surrounding phrase lengths on pause duration*. in *EuroSpeech*. 2003. Geneva, CH.
- [16] Bailly, G. and C. Gouvernayre. *Pauses and respiratory markers of the structure of book reading*. in *Interspeech*. 2012. Portland, OR.
- [17] Goldman, J.-P., et al. *Étude statistique de la durée pausale dans différents styles de parole*. in *Journées d'Etude sur la Parole (JEP)*. 2010. Mons, Belgique.
- [18] Barbosa, P. and G. Bailly, *Characterisation of rhythmic patterns for text-to-speech synthesis*. Speech Communication, 1994. **15**: p. 127-137.
- [19] McFarland, D.H. and A. Smith, *Effects of vocal task and respiratory phase on prephonatory chest wall movements*. Journal of Speech and Hearing Research, 1992. **35**: p. 971-982.