Breath and Non-breath Pauses in Fluent and Disfluent Phases of German and French L1 and L2 Read Speech

Jürgen Trouvain[1], *Camille Fauth*[2], *Bernd Möbius*[1]

[1] Computational Linguistics and Phonetics, Saarland University, Saarbrücken, Germany
[2] Phonetics, Université de Strasbourg, France
trouvain@coli.uni-saarland.de, cfauth@unistra.fr, moebius@coli.uni-saarland.de

Abstract

In this study we examined the read speech of native and non-native speakers with respect to pausing details of audible breathing, particularly in disfluent phases. 20 German and 20 French native speakers read the same narrative text in their native (L1) and in their non-native language (L2). Some expected results were confirmed: more frequent pauses and more frequent disfluencies in L2, as well as longer duration of pauses filled with breath noise than those without. However, the analysis also reveals that in fluent phases the vast majority of pauses contains audible inhalation - which requires a re-interpretation of the terms "unfilled" and "silent" pauses. Most disfluent phases are marked by genuinely silent pauses (i.e. without breathing noises), which are also shorter than those in fluent phases. So-called "filled pauses" are virtually not present. Surprisingly, French speakers use more but shorter pauses than the Germans as an L2 pausing strategy. The results suggest that the widely assumed concept of pauses in phonetics, prosody and fluency research should be renewed and enriched with phonetic detail that goes beyond "silent" vs. "filled" pauses in order to get a better understanding of the prosodic make-up of fluent and less fluent phases in speech.

Index Terms: pauses, inhalation noises, fluency, non-native speech

1. Introduction

Pauses play a central role for fluency rating and speech rate measurement. Besides articulation rate, pauses are an important indicator of utterance fluency and are thus essential for the assessment of how well a learner masters a non-native language (L2). Non-native speakers usually produce more pauses than speakers in their native language (L1), which leads to a slower speaking rate, more inter-pause stretches and often to a reduced temporal fluency [1, 2, 3, 4]. The literature on disfluencies distinguishes between "filled" and "unfilled" pauses whereby the first type includes hesitation particles like [ə:] or [ə:m] in many languages, and the latter type is also named "silent" pause [2, 3, 4, 5]. It is the assumption that in comparison to L1 speech in L2 the number of hesitations increases, which enforces the effect of slowing down and reduced fluency. However, it has been observed that "filled pauses" rarely occur in read speech [2].

A further problem comes with the label "silent" in "silent" pauses. "Silence" in a narrow acoustic sense is probably seldom observable in data used in phonetics and linguistics. From a phonetic point of view "silence" would be the absence of phonetic activity which excludes the acoustic correlates of inhalation (and sometimes exhalation) in speech pauses. Speech respiration activity in pauses is in fact an infrequently studied topic but we know for instance that breath noises in pauses can serve various functions, e.g. signalling the length of the upcoming phrase or marking a higher-level break (e.g. a paragraph) [6, 7, 8, 9]. The perception of breathing noises can also help improve memorising sentences [10]. In addition, a change of the respiratory setup, e.g. by inducing physical stress, may have a substantial impact on the prosodic phrase structure [11].

In this study we investigate the pauses of French and German speakers in their L2 *and* in their L1, producing quasi-identical texts in both languages. A particular interest lies in the differences between breath vs. non-breath pauses, and pauses in fluent vs. disfluent phases. Most studies on L2 fluency investigated spontaneous speech. In contrast to this we deal with read speech, which has the advantage of comparing different speakers, different languages, different states of language proficiency, and different time points of fluency testing. With respect to L1-L2 differences of the same speaker/s it is interesting to learn more about the assumed increase of the number of breath and non-breath pauses and their durational characteristics in general, and the pauses in disfluent phases in particular. We can expect differences in pausing strategies on an individual level (e.g. producing a high number of pauses, maybe combined with a short average pause duration) but we would not expect differences between the two examined languages.

2. Method

2.1. Data

The analysed data are a subset of the IFCASL bilingual phonetic learner corpus [12]. 20 German and 20 French native speakers read the same narrative text ("The three little pigs") in versions of their L1 and in their L2. Both language versions contained 13 sentences. In both language groups 10 speakers were at a beginning level of L2 (BEG) and 10 at an advanced level (ADV). In total 80 recordings were annotated and analysed. Recording durations of the text readings were between 50 and 200 seconds.

2.2. Annotation

The annotation took place with the help of the speech editor Praat in the following steps (with one annotation tier for each step). First, the entire recording was segmented in *articulation phases* and *pauses*. Pauses before the first syllable and after the last syllable of the text were discarded. Silent parts in the acoustic signal can also occur as a consequence of articulatory activity in closure phases of plosives and are often

 doi: 10.21437/SpeechProsody.2016-7

indistinguishable from pause silence. For this study we decided to use a fixed duration of 50 ms as articulatory closure time when a plosive follows a pause. There is no generally agreed cut-off point for a pause, it varies e.g. from 100 ms [13] to 200 ms [1, 2] and 400 ms [3]. In this study we had no fixed threshold but we defined a pause when there was a perceived pause plus a silence (excluding the closure phases of plosives), which also included pauses shorter than 100 ms.

In step 2, all pauses were annotated for whether they contained one or more of the following categories: *silence*, audible *inhalation*, clearly identified audible *exhalation*, and *clicks*, which occur rather frequently (cf. [14]) but were not considered for further analysis here. In addition there was the category of *unclear* cases. See Figures 1 and 2 for example illustrations.

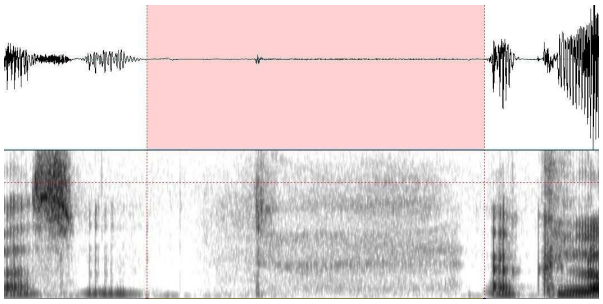

Figure 1: *Waveform and spectrogram (0-8kHz) of a section of 1.4 sec with a pause with a click at the beginning of the inhalation.*

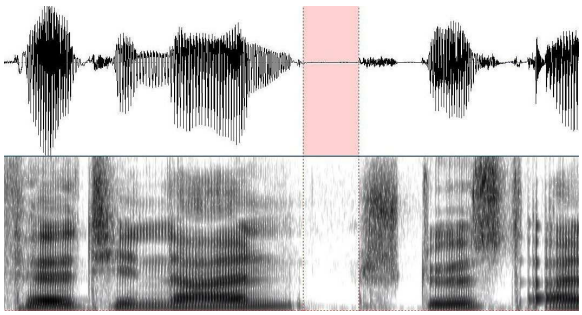

Figure 2: *Waveform and spectrogram (0-8kHz) of a section of 1.4 sec with a pause without inhalation noise.*

Central for the annotation of the data is the identification of *breath sounds* and *silences* in speech pauses, a task which has not received much attention in phonetic and prosodic research. Acoustically similar events to breath noise are on the one hand the aspiration noise of a fortis stop (e.g. in English or in German) and a voiceless /h/-realisation. An informal inspection of these sounds showed for aspiration noises a rather high intensity and a very short duration while voiceless productions of /h/ are often longer than aspiration noises but slightly less intense. In contrast to these, inhalation noises are by far less intense and much longer (usually exceeding 200 ms).

In step 3, the auditorily identified *disfluent phases* were annotated for the following phases (in line with [4, 15]):

- *reparandum:* the section before the repair which was corrected later in the repair,
- *repair:* the section which have been corrected by repetition,
- *disfluent pause:* a pause (no matter whether with or without inhalation), often between reparandum and repair, or solo in an ungrammatical location, leaving the impression of disfluency,
- *disfluent articulation*: unusual prolongations of sounds and sometimes the production of non-phonemic sounds.

In this study only disfluent pauses were considered.

After the annotation process, pauses were regarded either as *non-breath pauses* (containing neither inhalation nor exhalation but possibly clicks) or as *breath pauses* (containing either inhalation or exhalation). Thus, a breath pause may or may not contain silent segments. Pauses were additionally distinguished in whether they were produced *in fluent phases* or *in disfluent phases.*

3. Results

3.1. Frequency of occurrence of all pauses

Table 1 shows that there are only small differences regarding the total number of pauses between German (DE) and French (FR) speakers, and between beginners and advanced speakers when speaking in their L1. The number of pauses generally increases for L2, with a higher number for both groups of BEG. The tendency for an increase for L2 is for both language groups, but German speakers do not show values as extreme as French speakers.

Table 1. *Mean number of pauses for each subject group.*

		L1	L2
FR	BEG	25.7	50.2
	ADV	26.3	44.6
	all	26.0	47.4
DE	BEG	27.0	40.9
	ADV	26.0	31.7
	all	26.5	36.3

3.2. Duration of all pauses

Table 2 shows the average values for pause duration. French speakers pause longer in their L1 than German speakers. However, in their L2 the French speakers decrease their pause durations whereas the German speakers show just the opposite trend by lengthening their pauses. Both tendencies can be observed for all 40 speakers.

Table 2. *Mean pause durations (in ms) for each subject group.*

		L1	L2
FR	BEG	638	523
	ADV	630	492
	all	634	507
DE	BEG	533	637
	ADV	520	573
	all	526	605

3.3. Frequency of occurrence of breath vs. non-breath pauses

When dividing all pauses in either breath or non-breath pauses, the breath pauses clearly dominate for all sub-groups for L1 as well as for L2 (Table 3). Only three of the 40 speakers showed more pauses without than with audible breathing.

In L1, French speakers produced fewer breath pauses than the Germans – in contrast to L2, where the French subjects roughly doubled the number of breath pauses, whereas the German speakers kept the number of breath pauses constant. French and German speakers increased the number of non-breath pauses in L2, but the French only to a smaller extent.

Table 3. *Mean number of breath and non-breath pauses for each subject group.*

		L1		L2	
		breath	non-breath	breath	non-breath
FR	BEG	17.0	8.7	34.1	16.1
	ADV	15.4	10.9	28.9	15.7
	all	16.2	9.8	31.5	15.9
DE	BEG	20.9	6.1	21.9	19.0
	ADV	20.7	5.3	18.8	12.9
	all	20.8	5.7	20.4	16.0

3.4. Duration of breath vs. non-breath pauses

As could be expected from the durational characteristics of pauses in general (Table 2), breath pauses and non-breath pauses in native French are longer than in native German. In L2, this picture is reversed, with French speakers shortening their breath and non-breath pauses, and Germans lengthening both types of pauses.

Table 4. *Mean duration of breath and non-breath pauses (in ms) for each subject group.*

		L1		L2	
		breath	non-breath	breath	non-breath
FR	BEG	771	292	585	250
	ADV	714	345	592	302
	all	743	318	589	277
DE	BEG	564	231	685	366
	ADV	580	248	682	280
	all	572	239	683	323

3.5. Frequency of occurrence of pauses in disfluent phases

From 357 pauses in disfluent phases in total only two also contained hesitation particles. Table 5 shows the number of pauses in disfluent phases which increases from L1 to L2. In L1, the numbers are extremely low, though not zero as would be expected from a perfectly fluent reader. This level of fluency was achieved by 15 French speakers but only by 9 German speakers.

Table 5. *Mean number of pauses in disfluent phases for each subject group. Mean number of breath pauses in parentheses.*

		L1	L2
FR	BEG	0.7	8.5 (2.4)
	ADV	0.2	6.9 (1.1)
	all	0.5	7.7 (1.8)
DE	BEG	1.0	13.0 (3.2)
	ADV	1.2	4.2 (0.4)
	all	1.1	8.6 (1.8)

The differences in disfluent pausing between individuals were even larger for L2 (Figures 1 and 2). French beginners range from 0 to 37, French advanced learners from 0 to 28. For the German speakers, the beginners range from 5 to 26, and from 0 to 13 at the advanced level.

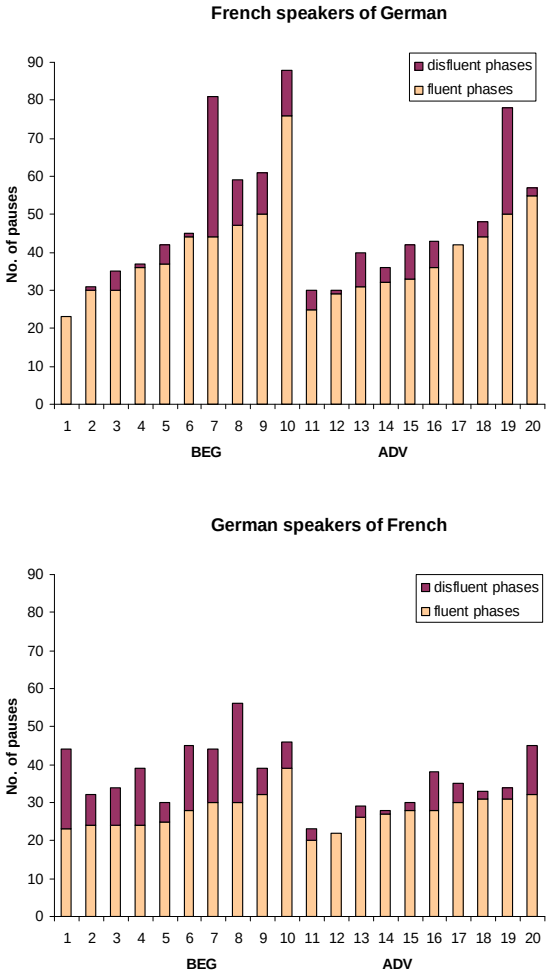

Figure 3: *Number of pauses (fluent and disfluent pauses accumulated in a bar) per individual speaker. Top: L2 German. Bottom: L2 French. Beginners: 1-10, advanced: 11-20.*

3.6. Duration of pauses in disfluent phases

With the temporal differences between breath and non-breath pauses in mind (Table 4) the duration of disfluent pauses (Table 6) reflects the fact that the predominant form of pause in disfluent phases of the L2 performance is the non-breath pause (Table 5). The differences to the general pause duration (Table 2) are substantial (more than 200 ms on average).

Table 6. *Mean duration of all pauses (in ms) and of the subset of pauses in disfluent phases in L2 for each subject group.*

		L2		
		all pauses	disfl. pauses	diff.
FR	BEG	523	321	234
	ADV	492	316	208
	all	507	319	221
DE	BEG	637	360	277
	ADV	573	361	205
	all	605	361	243

4. Discussion and conclusions

As expected there were no differences in how often French and German speakers produce pauses in their L1. It was also expected that in L2 all speakers produce more pauses. It appears as a surprise that in L1 French speakers use longer pauses than their German peers but shorter pauses in L2.

Pauses filled with breath noise were longer than those without, as expected, and the analysis also revealed that in fluent phases the vast majority of pauses contains audible inhalation – which requires a reinterpretation of the terms "unfilled" and "silent" pauses that are often mentioned in the literature [2, 3, 4, 5]. Future studies have to show further details of the acoustic content of "silent" pauses, particularly the production of tongue clicks [14].

A second desideratum is the research on the physiology of inhalation and exhalation in speech pauses (e.g. kinematics) [9] and its acoustic and perceptual correlates. So far, we investigated *audible* inhalation noises and we were missing possible *inaudible* inhalation. Thus it is unclear which physiological effort leads to which acoustic result.

It was expected that in L2 the number of pauses in general would increase, particularly the number of pauses in disfluent phases. It must be stressed how strongly individuals vary between each other [5], and this to a considerable degree across levels of proficiency (Figure 3). Speakers at an advanced level of general language proficiency do no generally perform among the best regarding their disfluent pauses, andconversely, various speakers with a low level of L2 mastery can be found with a performance at an advanced level regarding pausing. It must thus be a concern at all proficiency levels to offer training for utterance fluency.

For fluency research it might also be interesting to see that so-called "filled pauses", i.e. pauses containing hesitation particles, are virtually not present in our read material, confirming the evidence from other studies [2]. In addition, most disfluent phases are marked by genuinely silent pauses (i.e. without breathing noises), which are also shorter than those in fluent phases. Obviously, these two features are stable markers of utterance fluency in read texts – in contrast to spontaneous speech where often a great number of hesitation particles can be found. In read speech important planning steps are missing or reduced, such as the conceptual preparation of the formulation, the morpho-syntactic encoding, the selection of words and aspects of social interactivity [16, 17]. Comparisons with non-read forms of speech, e.g. spontaneous dialogues, would allow a better understanding of the utterance fluency of speech production.

Likewise, a more detailed knowledge of individual patterns of disfluency can be very important for automatic (and human) fluency assessment. It would be a general benefit if we would know more about the aspects on which the *perceived* fluency are based in the speech production of various types of speech situations, e.g. to correlate the utterance fluency of read speech with that of spontaneous speech. Furthermore, it would be a benefit for an individual L2 learner to hear whether s/he has improved on utterance fluency and to use visualisations and other forms of feedback, e.g. in a computer-assisted pronunciation and fluency training.

The general pattern of pausing in native speech is that longer pauses (which usually contain inhalation) are used for marking higher syntactic-prosodic breaks, and that shorter pauses (most of them non-breath pauses) are used for within-sentence breaks [6]. This pattern has also been found in this study across both languages. It was expected that there is more than one strategy in L2 read speech (compared to L1), e.g. using more but shorter pauses. However, it was unexpected to find that two different main strategies could be allocated to the two language groups, with the French using extremely more pauses and shortening them in L2, while the German speakers producing slightly more pauses but lengthening them. A language-dependent pausing pattern, here in texts read aloud, would be a new finding and should be investigated in more detail. This would also include a detailed analysis of articulation rate and a possible compensation of low speed of articulation and short pause duration, and vice versa.

To summarise, the results suggest that the widely assumed concept of pauses in phonetics, prosody and fluency research should be renewed and enriched with phonetic detail that goes beyond "silent" vs. "filled" pauses in order to get a better understanding of the prosodic make-up of fluent and less fluent phases in speech.

5. Acknowledgements

This research was supported by DFG and ANR for the project IFCASL (www.ifcasl.org). The authors appreciate very much the support of various kinds by Sascha Schäfer, Jeanin Jügler and Zofia Malisz.

6. References

[1] Lennon, P. 1990. Investigating fluency in EFL: A quantitative approach. Language Learning 40, 387–417.

[2] Cucchiarini, C., Strik, H., Boves, L. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. Journal of the Acoustical Society of America 111/6, 2862-2873.

[3] Tavakoli, P. (2011). Pausing patterns: differences between L2 learners and native speakers. ELT Journal, 65(1), 71-79.

[4] Lickley, R.J. 2015. Fluency and disfluency. In: Redford, M.A. (ed) The Handbook of Speech Production. Hoboken, NJ: John Wiley, pp. 445-474.

[5] De Jong, N.H., Groenhout, R., Schoonen, R. & Hulstijn, J.H. 2015. Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. Applied Psycholinguistics (36), 223-243.

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

[7] Gee, J.P. & Grosjean, F. 1983. Performance structures: a psycholinguistic and linguistic appraisal. Cognitive Psychology 15, pp. 411-458.

[8] Strangert, E. 1991. Pausing in texts read aloud. Proc. ICPhS Aix-en-Provence, Vol. 4.

[9] Fuchs, S., Petrone, C., Krivokapić, J. & Hoole, Ph. 2013. Acoustic and respiratory evidence for utterance planning in German. Journal of Phonetics 41, pp. 29-47.

[10] Whalen, D.H., Hoequist, Ch.E. & Sheffert, S. 1995. The effects of breath sounds on the perception of synthetic speech. Journal of the Acoustical Society of America 97, pp. 3147-3153.

[11] Trouvain, J. & Truong 2015. Prosodic characteristics of read speech before and after treadmill running. Proc. Interspeech, Dresden, pp. 3700-3704.

[12] Fauth, C., Bonneau, A., Zimmerer, F., Trouvain, J., Andreeva, B., Colotte, V., Fohr, D., Jouvet, D., Jügler, J., Laprie, Y., Mella, O. & Möbius, B. 2014. Designing a bilingual speech corpus for French and German language learners: a two-step process. Proc. 9th Language Resources and Evaluation Conference (LREC), Reykjavik, pp. 1477-1482.

[13] Trouvain, J. 2004. Tempo Variation in Speech Production. Implications for Speech Synthesis. PhD thesis, Saarland University. Phonus 8, Phonetics, Saarbrücken.

[14] Trouvain, J. 2014. Laughing, breathing clicking -- The prosody of nonverbal vocalisations. Proc. *Speech Prosody* (SP7), Dublin, pp. 598-602.

[15] Shriberg, E. 2001. To 'errrr' is human: ecology and acoustics of speech disfluencies. Journal of the International Phonetic Association 31, 153–169.

[16] Levelt, W. 1989. Speaking. Cambridge, Massachusetts: MIT Press.

[17] Segalowitz, N. 2010. Cognitive Bases of Second Language Fluency. Routledge.