Pauses following fillers in L1 and L2 German Map Task dialogues

Malte Belz & Myriam Klapi

Humboldt-Universität zu Berlin, Germany

Abstract

Fillers and pauses in spoken language indicate hesitations. Filler type (*uh* vs. *um*) is believed to signal a minor or major following speech delay in L1. We examined whether advanced speakers of L2 German use pauses following filler type (*äh* vs. *ähm*) in the same way as native speakers do. Two Map Task corpora of L1 and L2 were contrasted with respect to speaker role, filler type and the exact time interval of fillers and pauses. Speaker role influenced the disfluency patterns in L1 and L2 in the same way. Filler type had no impact on the length of the following pause, but the time interval patterns differed significantly. Longer filler intervals are followed by longer pauses in L2 and by shorter pauses in L1. These results suggest that filler type in German is not used to indicate the length of the following delay. Advanced learners seem to have adopted this pattern of use, but cannot overcome their hesitations as fast as native speakers, probably due to their less automatised speech production.

Index Terms: Fillers, Pauses, Spontaneous speech, L1, L2, Map Task, German, Disfluencies, Contrastive Analysis

1. Introduction

In this paper we examine whether German native (L1) and non-native speakers (L2) use the fillers *äh* and *ähm* and their following pauses in the same way.

Fluency in spontaneous speech is not constantly achieved. Disfluencies in spontaneous speech are frequent and have been analysed both for L1 ([1–6], inter alia) and L2 speakers [7, 8]. Among the most commonly studied disfluency categories are fillers and pauses. Fillers, like the English *uh* and *um*, or the German *äh* and *ähm* – sometimes also called filled pauses or hesitations [5, 9] – are well-described, although many studies propose different vantage points [10]. Together with pauses (also known as silent pauses or unfilled pauses [5]) fillers are central to the research of hesitation phenomena in L1 and L2 speech production [11–13]. Taken together, they constitute about 78% of the overall occurrence of disfluencies in spontaneous speech [5]. When examined separately, the use of fillers and pauses is often related to hesitation and repair phenomena. Consequently, a combined use raises the question of whether a more serious problem in speech production has been encountered. As learners have to deal with non-native speech processing, they may experience a working memory capacity overload [14]. Therefore, deviations in delay behaviour can be expected, and interesting implications can be drawn from this phenomenon when applying contrastive analysis.

Prior research has made evident that native speakers often use fillers and pauses in order to find time for processing decisions [15]. According to [7] and [14], limitations in L2 proficiency cause patterns of error and repair which are different from those of native speakers. Hence, one would expect that, because learners of a foreign language have to process a higher cognitive load, they will differ from native speakers in their filler and pause patterns. The question which arises at this point is whether this is true regarding advanced L2 speakers. Do they use fillers with following pauses as L1 speakers do? In order to evaluate hypotheses of this kind, it is essential to conduct contrastive research, thus enabling a comparison to native speakers. Differences in the use of the patterns described above may relate to less automatised speech production processes and monitoring in L2 [16–18].

Clark & Fox Tree [9] examined the English fillers *uh* and *um* in combination with following pauses. Their results suggest that filler types affect the length of their respective following pauses. *Uh*s preceding pauses signal a minor delay, whereas *um*s preceding pauses signal a major delay. Example 1 illustrates their use:

(1) ich sags dir 0.6 s ähm 1.7 s also du musst äh 0.4 s nach äh re/ rechts hoch

I'm telling you 0.6 s um 1.7 s well you have to uh 0.4 s go uh upwards to the right

(BeMaTaC_L1_2013-01/2012-01-19-A, 04:32)

The L1 speaker in the example above inserts *ähm* as well as a 1.7 s pause before giving directions. Before specifying the exact direction, the speaker inserts *äh* together with a shorter pause (0.4 s). From the example above, it may be assumed that *uh* and *um* in Map Task dialogues behave differently with respect to their following pause. In genres like interviews, however, the differences observed for post-filler pauses may not be perceived quite as clearly [19].

No quantitative study of combined fillers preceding pauses has yet been made in a German L1/L2 Map Task setting. The present approach attempts to bridge this gap by demonstrating a contrastive corpus analysis of two recently constructed corpora. Our hypotheses are the following:

1) It is often implicitly assumed that fillers in different languages show similar patterns in use, when their form seems to be identical (Engl. *uh/um* vs. Germ. *äh/ähm*). As a first step, it is crucial to examine whether the length of the following pause in German is influenced by the filler type (*äh* vs. *ähm*), as it is in English [9]. Our contrastive analysis suggests that the two filler types deviate in relation to the length of their following pauses. This prediction is relevant for both L1 and L2 speaker categories and should be observable in both groups.

2) According to the given experimental Map Task design, the speaker role (i.e. instructor vs. instructee) is expected to affect the length of pauses. Instructors take up the highest amount of speaking time (see section 2.1). We expect this effect to be observable for both L1 and L2.

3) If type of filler turns out to be the only influence on the following pause length, then the filler length is not anticipated to affect the length of the pause. Though no evidence for a similar use of filler categorisation preceding pauses has been found for German yet, we expect fillers to behave in the English way, as stated by [9]. As the proficiency of learners in our data exceeds intermediate levels, we expect them to adopt the native-like pattern.

2. Method

2.1. Corpora

The Berlin Map Task Corpus (BeMaTaC) [20, 21] and the Hamburg Map Task corpus (HAMATAC) [22] both use a Map Task design [23], where one speaker (= instructor) instructs another speaker (= instructee) to reproduce a route on a map with landmarks. This design is suitable for multilevel linguistic research, as it enables spontaneous dialogues elicited in a controlled context [21]. BeMaTaC has been inspired by HAMATAC and follows the same experimental design, enabling comparable and contrastive studies of native and non-native German.

BeMaTaC (version 2013-01) consists of 12 dialogues, 16 native German speakers and 11192 tokens in the relAnnis format. In order to conduct the present research we accessed BeMaTaC via ANNIS [24], an open-source browser-based search and visualisation tool for deeply annotated corpora.

HAMATAC (version 0.2 [2011-09-30]) consists of 24 dialogues (21433 words) by 24 advanced learners of German. For lack of a standardised L2 proficiency test, we rely on the meta-data, consisting of learners with an advanced proficiency level (20 out of 24). Participants' native languages covered a wide range (Romance, Slavic, Persian and Non-Indo-European languages). We extended the corpus with further annotation layers. The corpus was converted with the SaltNPepper converter [25] to the relAnnis format.

2.2. Data

All fillers preceding pauses were extracted from these two corpora (Table 1), every instance linked to its metadata role (instructor vs. instructee) and subject ID. The exact filler interval time as given in the transcriptions were extracted and calculated in L1 and L2, as well as for pauses in L1. Pauses annotated in HAMATAC were extracted from the vocal transcription tier as given in deciseconds. Zero-length pauses were not considered as relevant instances and therefore not taken into account.

Table 1: *Frequencies of fillers preceding pauses*

Fillers preceding pauses	BeMaTaC äh	BeMaTaC ähm	HAMATAC äh	HAMATAC ähm
Actual numbers	34	42	108	142
In %	44.7	55.0	43.2	56.8
In % of overall words or tokens	0.67		1.16	
In % of overall fillers	27.74		28.09	

The L1 data were extracted via ANNIS using the token boundaries we obtained from a PRAAT transcription [26]. The L2 data could not be extracted quite as easily. Therefore, we exported the PRAAT voice transcription tier and calculated the exact filler interval times.

Describing a path through a Map Task is rather challenging since instructor and instructee cannot see each other. A high working memory load can be expected, especially if subjects have to conceptualise their message in a foreign language. As we expected higher hesitation levels, we did neither apply a cut-off of length for fillers nor for pauses.

2.3. Model

Since there are individual differences in disfluency length distribution, we applied a linear mixed-effects model to the data, which allows us to treat subject IDs as random effects while looking for significant patterns between fixed effects. We started with a full model, including as many fixed effects and their interactions as technically possible. Then we reduced the complexity of the model stepwise by comparing their AIC and performing log-likelihood tests. The remaining fixed effects which seem to predict pause length are language type (L1 vs. L2), role (instructor vs. instructee) and interaction of language with filler length (see Table 2).

3. Results

The findings are summarised in Table 1. L2 speakers use the described phenomenon nearly twice as much as L1 speakers (1.16% vs. 0.67%). We observe that in both groups approximately every fourth instance is followed by a pause (27.74% vs. 28.09%). We see that pause length differs with respect to the preceding filler type (Figure 1). Pauses following *äh* exhibit a large variance in L1 ($\bar{t} = 0.57\,s$, $\hat{\sigma} = 0.41\,s$, $median = 0.43\,s$) and in L2 ($\bar{t} = 0.95\,s$, $\hat{\sigma} = 1.33\,s$, $median = 0.6\,s$), whereas pauses following *ähm* have a more narrow variance, both in L1 ($\bar{t} = 0.63\,s$, $\hat{\sigma} = 0.43\,s$, $median = 0.59\,s$) and L2 ($\bar{t} = 0.91\,s$, $\hat{\sigma} = 0.73\,s$, $median = 0.7\,s$).

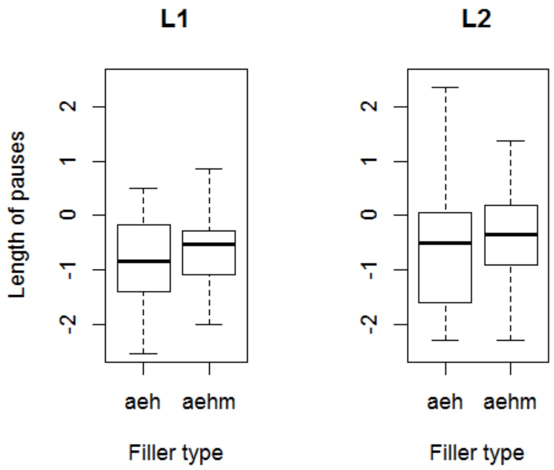

Figure 1: *Filler type and variance of pause length in L1 and L2 (logarithmic scale).*

Differences can be seen regarding the overall length variance of pauses depending on filler types. Nevertheless, no significant results were found regarding the interaction between filler type and pause length, either for L1 or for L2, as will be shown below.

As far as we can tell, filler type (*ähm* vs. *äh*) has no significant impact on the length of the following pause ($Df = 282$, $p < 0.96$), nor does interaction between language and filler type ($Df = 282$, $p < 0.28$). These effects were therefore excluded from the model, among others.

As expected, the instructor role exhibited a significant effect ($Df = 288$, $p < 0.043$) compared to the instructee role for both language types. The main effect of filler length does not appear to be interpretable in a way that makes sense to us due to its strong interaction with language type (L1 vs. L2). However, the interaction between L2 and filler length was significant ($Df = 288$, $p < 0.0016$). This indicates that filler length had a different effect on the duration of the unfilled pause for L2 compared to L1. More specifically, L2 speakers tend to produce longer pauses following respectively longer filler intervals. Thus, the longer a filler stretches in time, the longer the following pause seems expected to be. This effect does not depend on filler type, as shown in Figure 2.

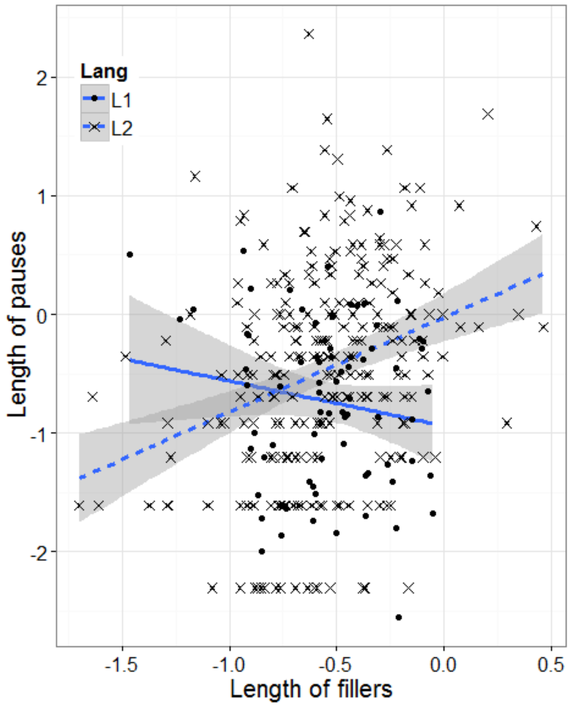

Figure 1: *Interaction of pauses preceding fillers in L1 and L2 (logarithmic scale).*

Random slopes were ruled out after a χ^2-test calculation, which proved not to be significant. To avoid spurious correlations of fixed effects, the logarithms of filled pauses were centered. Model coefficients are illustrated in Table 2.

Table 2: *Results of the linear model.*

	Value	Std. Error	DF	t	p
(Intercept)	–1.15	0.23	288	–5.09	0.000
Lang	0.82	0.22	33	3.77	0.000
Role	0.32	0.16	288	2.04	0.042
log(FPlength)	–0.24	0.26	288	–0.90	0.370
Lang:log(FPlength)	0.98	0.31	288	3.20	0.001

4. Discussion

First, advanced learners of German insert pauses after fillers with a similar frequency in Map Task dialogues when compared to native speakers (27.74% in L1, 28.09% in L2). These similarities between native speakers and learners are expected when stipulating an advanced L2 level. They herewith verify the average competence level of L2 speakers in HAMATAC.

Second, we expected pauses in L1 to differ systematically in duration when used after *äh* and *ähm* respectively. Nonetheless, the results failed to reach a significant level, as has been calculated above. This is confirmed for both groups. This implies that there is no difference between *äh* and *ähm* regarding their following delays, being inconsistent with the findings of [9] for English.

Figure 1 might suggest differences in pause length for both filler types. However, these narrow discrepancies proved not to be significant. This was found by applying the best linear mixed-effects model as described above. It excludes the interaction of filler type with language as a fixed effect. Therefore, no significant difference of how advanced learners use pauses following *äh* and *ähm* compared to native speakers was detected, suggesting that advanced learners of German do not behave differently from German native speakers regarding the combination of filler types with following pauses.

A significant effect is found for the dependence of length of pauses on the speaker role, hence verifying hypothesis 2. Filled pauses are significantly longer when speaking as instructor than when speaking as instructee. This finding is consistent with the results of Bortfeld et al. [1], who also showed an influence of the factor role on the number of disfluencies produced. This holds for both L1 and L2 and might reflect higher cognitive demands that instructors have to deal with.

As to the third hypothesis, our anticipation regarding the length of *äh* and *ähm* predicted no effect on the length of the following pause. This prediction was surprisingly falsified. The model exhibits that pause length is influenced significantly by the presence of the interaction between L2 speakers and filler length. This finding suggests that pauses may indeed be dependent on the time it takes a speaker to articulate a filler. The duration of *äh* and *ähm* may therefore be interpreted as a signal of an upcoming planning pause. Hence, the implication which becomes evident is that when L2 speakers of German take longer to utter a filler, they somehow signal that they need a longer planning phase and tend to insert a longer silent pause. Since the finding for native speakers of German is opposite regarding the dependent pause (the longer the filler, the shorter the pause), it is suggested that L2 speakers show a deviating pausing behaviour with respect to fillers. This finding might also suggest that speech production in Map Task descriptions is hard to process for learners, despite their high level of competence in German.

5. Summary and Conclusion

The implications of this pilot study are two-fold. The current results indicate that the role of participants (i.e. instructor vs. instructee) within the Map Task significantly influences their disfluency patterns, confirming the findings of Bortfeld et al. [1]. This holds both for L1 and L2, providing us with a more comparable and thus more reliable environment for contrastive research.

We did not find a correlation between filler type and length of the following pause between L1 and L2, what one would have expected for learner speech production, namely a non-nativelike use of fillers with pauses. Since there is no such evidence, our finding suggests either that learners have adopted the use of these patterns at an earlier stage, or that there is no difference in the distinctive filler types. We argue for the latter, thus implying that filler type seems not to affect German learners concerning the process of planning in speech production. Our results imply an observable difference in the use of delays when compared to English. Even though no direct comparison has been made in this study, it is possible that the use of delays combined with fillers follows a language-specific pattern.

Our findings suggest that L1 and L2 speakers have different pausing behaviours depending on the time spent for uttering a filler, regardless of filler type. These results show that German learners deviate significantly from German native speakers in using this specific disfluency pattern, which might be related to less automatised speech processing and monitoring in non-native speech production, as described by Levelt [17] and Declerck & Kormos [18]. With the objective of identifying differences in learner speech disfluencies and L2 acquisition, a more fine-grained stratification of proficiency control may result in the emergence of a new measure for automatisation in L2 speech production.

6. Acknowledgements

We are grateful to Anke Lüdeling, Felix Golcher and Amir Zeldes for their continuous support. We thank Robert Eklund for his encouragement in contributing to this workshop.

7. References

[1] H. Bortfeld, S. D. Leon, J. E. Bloom, M. F. Schober, and S. E. Brennan, "Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender", *Language and Speech*, vol. 44, no. 2, pp. 123–147, 2001.

[2] J. E. Arnold, M. Fagnano, and M. K. Tanenhaus, "Disfluencies Signal Theee, Um, New Information", *Journal of Psycholinguistic Research*, vol. 32, no. 1, pp. 25–36, 2003.

[3] E. E. Shriberg, "Preliminaries to a Theory of Speech Disfluencies," Unpublished Dissertation, University of California, Berkeley, 1994.

[4] E. E. Shriberg, "To 'errrr' is human: ecology and acoustics of speech disfluencies", *JIPA*, vol. 31, no. 1, pp. 153–169, 2001.

[5] R. Eklund, "Disfluency in Swedish human–human and human–machine travel booking dialogues", Dissertation, Department of Computer and Information Science, Linköpings Universitet, Linköping, Sweden, 2004.

[6] F. Ferreira and K. G. D. Bailey, "Disfluencies and human language comprehension", *Trends in Cognitive Sciences*, vol. 8, no. 5, pp. 231–237, 2004.

[7] J. Kormos, "Monitoring and Self-Repair in L2", *Language Learning*, vol. 49, no. 2, pp. 303–342, 1999.

[8] C. L. Rieger, "Disfluencies and hesitation strategies in oral L2 tests". In: *Proceedings of DiSS'03*, Disfluency in Spontaneous Speech Workshop, pp. 41–44, 2003.

[9] H. H. Clark and J. E. Fox Tree, "Using uh and um in spontaneous speaking", *Cognition*, vol. 84, no. 1, pp. 73–111, 2002.

[10] M. Corley and O. W. Stewart, "Hesitation Disfluencies in Spontaneous Speech: The Meaning of um", *Language and Linguistics Compass*, vol. 2, no. 4, pp. 589–602, 2008.

[11] S. S. Reich, "Significance of pauses for speech perception", *Journal Psycholinguist. Research*, vol. 9, no. 4, pp. 379–389, 1980.

[12] R. Griffiths, "Pausological Research in an L2 Context: A Rationale, and Review of Selected Studies", *Applied Linguistics*, vol. 12, no. 4, pp. 345–364, 1991.

[13] P. Trofimovich and W. Baker, "Learning Second Language Suprasegmentals: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech", *Stud. Sec. Lang. Acq.*, vol. 28, no. 1, pp. 1–30, 2006.

[14] L. Temple, "Second language learner speech production," *Studia Linguistica*, vol. 54, no. 2, pp. 288–297, 2000.

[15] H. H. Clark and T. Wasow, "Repeating Words in Spontaneous Speech", *Cognitive Psychology*, vol. 37, pp. 201–242, 1998.

[16] W. J. M. Levelt, "Monitoring and self-repair in speech", *Cognition*, vol. 14, no. 1, pp. 41–104, 1983.

[17] W. J. M. Levelt, *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press, 1989.

[18] M. Declerck and J. Kormos, "The effect of dual task demands and proficiency on second language speech production", *Bilingualism*, vol. 15, no. 4, pp. 782–796, 2012.

[19] D. C. O'Connell and S. Kowal, "Uh and Um Revisited: Are They Interjections for Signaling Delay?", *J. Psycholinguist. Res.*, vol. 34, no. 6, pp. 555–576, 2005.

[20] L. Giesel, M. Klapi, D. Krüger, I. Nunberger, O. Rasskazova, and S. Sauer, "A deeply annotated multimodal map-task corpus of spoken learner and native German", *Proc. DGfS*, Potsdam, 2013.

[21] S. Sauer and A. Lüdeling, "BeMaTaC: A Flexible Multilayer Spoken Dialogue Corpus for Contrastive SLA Analyses". In: *ICAME* 34, 2013.

[22] T. Schmidt, H. Hedeland, T. Lehmberg, and K. Wörner, HAMATAC – The Hamburg MapTask Corpus. http://www.exmaralda.org/files/HAMATAC.pdf

[23] A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. S. Thompson, and R. Weinert, "The HCRC Map Task Corpus", *Language and Speech*, vol. 34, no. 4, pp. 351–366, 1991.

[24] A. Zeldes, J. Ritz, A. Lüdeling, and C. Chiarcos, "ANNIS: A Search Tool for Multi-Layer Annotated Corpora", in *Proceedings of Corpus Linguistics*, 2009.

[25] F. Zipser and L. Romary, "A model oriented approach to the mapping of annotation formats using standards". In: Proceedings of the Workshop on Language Resource and Language Technology Standards, *LREC 2010*, Malta, 2010.

[26] P. Boersma, "Praat, a system for doing phonetics by computer", *Glot International*, vol. 5, no. 9, pp. 341–345, 2001.

8. URLs

BeMaTaC
http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/bematac

HAMATAC
http://vs.corpora.uni-hamburg.de/corpora/z2-hamatac/public/index.html

Proceedings of

DiSS 2013

The 6th Workshop on Disfluency in Spontaneous Speech

KTH Royal Institute of Technology
Stockholm, Sweden
21–23 August 2013

TMH-QPSR
Volume 54(1)

Edited by
Robert Eklund

Conference website: http://www.diss2013.org
Proceedings also available at: http://roberteklund.info/conferences/diss2013

Cover design by Robert Eklund
Front cover photo by Jens Edlund and Joakim Gustafson
Back cover photos by Robert Eklund

Proceedings of DiSS 2013, The 6th Workshop of Disfluency in Spontaneous Speech
held at the Royal Institute of Technology (KTH), Stockholm, Sweden, 21–23 August 2013
TMH-QPSR volume 54(1)
Editor: Robert Eklund
Department of Speech, Music and Hearing
Royal Institute of Technology (KTH)
Lindstedtsvägen 24
SE-100 44 Stockholm, Sweden

ISBN 978-91-981276-0-7
eISBN 978-91-981276-1-4
ISSN 1104-5787
ISRN KTH/CSC/TMH--13/01-SE
TRITA TMH 2013:1

Printed by Universitetsservice US-AB, Stockholm, Sweden, 2013