Disfluencies in German adult- and infant-directed speech

Charlotte Bellinghausen[1*], Simon Betz[2*], Katharina Zahner[3*], Alina Sasdrich[1], Marin Schröer[2], Bernhard Schröder[1]

[1]*Department of German Studies, University of Duisburg-Essen*
[2]*Faculty of Linguistics and Literary Studies, Bielefeld University*
[3]*Department of Linguistics, University of Konstanz*
shared first authorship (alphabetical order); all corresponding
charlotte.bellinghausen@uni-due.de, *simon.betz@uni-bielefeld.de*,
katharina.zahner@uni-konstanz.de

We investigate the occurrence of disfluencies (lengthening, filled pauses, silent pauses, abandoned utterances, and repairs) in German infant-directed speech (IDS), as compared to German adult-directed speech (ADS). The corpus consists of speech of nine mothers talking to their toddler (IDS condition), or to an adult experimenter (once in a task with low cognitive load, ADS1; once with higher cognitive load, ADS2). In line with previous studies, ADS contains more instances of disfluencies than IDS. Also, with increasing cognitive load in ADS, the number of silent pauses and lengthened instances increases. Overall, our results corroborate earlier findings on IDS in other languages, and on disfluencies and cognitive load. From an explorative perspective, our results also allow to derive hypotheses for future experiments – both for studies regarding IDS and infant speech processing.

INTRODUCTION

Spontaneous speech among adults (adult-directed speech, ADS) contains large amounts of disfluencies and hesitations. Disfluencies occur when speakers are involved in tasks with increased cognitive load, e.g., map tasks [1] and also depend on listeners feedback [2]. The actual frequencies of disfluencies also vary depending on language status ((non)native), as well as psychological or developmental factors like stuttering (cf. [3], Ch. 2).

Infant-directed speech (IDS), on the other hand, is described as highly fluent ([4, 5] on English), with disfluencies becoming slightly more frequent as children grow older (0.58 disfl./100 words in IDS to 7-12 mo-olds vs. 1.03 disfl / 100 words in IDS to 13-24 mo-olds, [6] on Swedish). [5] report that only 5-10% of the prosodic breaks are disfluent in IDS to 6-10 mo-old infants. Similarly, in [6] only 19 filled pauses ("uh"), occurred in a Swedish IDS corpus (35500 words), compared to 249 in ADS (24100 words). Yet, recent processing studies suggest that toddlers discriminate between fluent and disfluent speech ([5, 7] for 22 mo-old children), and also benefit from filled pauses when recognizing novel words ([8] for 24-month-old American children; [9] for 32-mo-old French, English, and English-French children).

In this paper, we compare the occurrence of the disfluencies *lengthening, filled pauses*, *silent pauses*, *abandoned utterances*, and *repairs* in a) German IDS, as compared to ADS and b) in different ADS conditions that differ in cognitive load. Based on the reviewed background, we expect more disfluencies in German ADS compared to IDS, H(ypothesis) 1. We also predict disfluencies to increase with increasing cognitive demands, i.e., increasing task difficulty (H2).

METHODS

The corpus consists of audio-recordings of 9 German-speaking mothers (between 32 and 39 years; currently living in Konstanz and surroundings) in three conditions: one IDS condition, and two ADS conditions (differing in cognitive load). In the IDS condition (86.5 minutes and 6709 words in total), mothers were recorded when unpacking a treasure chest (containing toys) together with their children (all female, between 19 and 24 months). In the ADS1 condition (low cognitive load; 30.0 minutes and 3868 words in total), mothers unpacked the same objects from the treasure chest and talked about them with an adult female experimenter. Finally, in the ADS2 condition (high(er) cognitive load; 9.7 minutes and 1079 words in total), mothers performed a map task (cf. [1]) with the experimenter as the interlocutor. Specifically, the mothers' maps contained a path leading through the depicted objects and their task was to guide the experimenter's way to the destination, see [10] for a more detailed description. These two ADS conditions differ in cognitive load, since ADS1 allows free spontaneous speech on everyday-objects (e.g., a cat, a cup etc.), while ADS2 the speaker needs to find her way through the map in order to guide the interlocutor.

The corpus was annotated following the DUEL guideline, which allows for on-the-fly markup of spontaneous speech elements (DisflUencies, Exclamations and Laughter [11]). Four labelers (authors) received an introduction to the DUEL manual and segmented utterances based on syntactic and pausal criteria in Praat [12]. For each utterance, they marked *lengthening* of words, *filled pauses* (e.g., 'ähm'), *silent pauses*, *abandoned utterances* ('I am leaving - what was the task again?')), and *repairs* (('and the-+ and then) to the left')).

A fifth expert annotator (author) checked the files for compliance with DUEL.

To test the hypotheses outlined in the introduction, we calculated Chi-Square tests to compare two (of the three) conditions in regard to one hypothesis. With respect to H1 (more disfluencies in ADS than in IDS), we chose ADS1 as speech register since in IDS and ADS1 the communicative situation is similar: In IDS the mother talks to the child about object in the treasure chest, in ADS1 she talks to the interviewer about objects in the treasure chest. With respect to H2 (more disfluencies when cognitive load is higher), we compared ADS1 and ADS2 (map task) which had the same register, but the map task required a higher cognitive effort.

RESULTS

Tab. 1 shows the occurrences of the different types of disfluencies per 1000 words in the three conditions. Chi-Square tests assessed the difference in occurrences of disfluencies across conditions. P-values were adjusted using the Benjamini-Hochberg correction [13] to account for the fact that multiple comparisons were run (10 altogether), one for each type of disfluency. As level of significance we chose alpha=0.05. We report the original p-values, along with the adjusted p-values (p_{adjust}).

Tab. 1: Number of occurrences (per 1000 words) for the different types of disfluencies. The values in brackets indicate the absolute numbers of occurrence. N gives the total number of words in each condition.

Condition	Type of disfluencies				
	Length ening	Filled pause	Silent pauses	Aband. utts	Repairs
ADS1 (Treasure Chest) (N=3869 words)	35 (134)	23 (88)	109 (422)	11 (42)	19 (73)
ADS2 (Map Task) (N=1079 words)	117 (126)	27 (29)	151 (163)	15 (16)	14 (15)
IDS (N=6709 words)	56 (378)	2 (16)	67 (447)	11 (77)	8 (52)

H1 (IDS vs. ADS1). As predicted, there were more disfluencies in German ADS than in IDS, Tab. 2a for an overview of results of the Chi-Square Test. Specifically, there were more occurrences of filled pauses (χ^2=16.5, df=1, p<0.0001, p_{adjust}<0.001), and silent pauses (χ^2=10.3, df=1, p=0.001, p_{adjust}=0.005), and more repairs (χ^2=4.6, df=1, p=0.03, p_{adjust}=0.05). Interestingly, instances of lengthening were more frequent in IDS than in ADS1 (χ^2=5.2, df=1, p=0.02, p_{adjust}=0.05). A preliminary analysis of the lengthened instances in IDS reveals that they were not primarily due to hesitations but used for accentuation purposes (for a distinction of different types of lengthening, see [14]).

H2 (ADS1 vs. ASD2). We predicted generally more disfluencies in ADS2 (map task with high cognitive load) than in ADS1 (treasure chest, low cognitive load). This prediction held for the occurrences of *lengthening* (χ^2=25.1, df=1, p<0.0001, p_{adjust}<0.0001) and *silent pauses* (χ^2= 6.8, df=1, p=0.009, p_{adjust}=0.02). For the other

disfluency types there was only a trend, see Tab. 1 for proportions, Tab. 2b for results of the Chi-Square Test.

Tab. 2: Overview of results of Chi-Square Tests for the different types of disfluencies. "Yes" indicates that there is a difference between conditions (direction in brackets); "No" that there is no evidence to assume a difference in the distribution of disfluent occurrences across conditions.

Chi-Square Test	a) Comparison IDS vs. ADS1 (treasure chest)				
	Lengthening	Filled pause	Silent pauses	Aband. utts	Repairs
Difference (yes / no)	**Yes** (IDS > ADS1)	**Yes** (ADS1 > IDS)	**Yes** (ADS1 > IDS)	No	**Yes** (ADS1 > IDS)
	b) Comparison ADS1 (treasure chest) vs. ADS2 (map task)				
	Lengthening	Filled pause	Silent pauses	Aband. utts	Repairs
Difference (yes / no)	**Yes** ADS2 > ASD1	No	**Yes** ADS2 > ASD1	No	No

DISCUSSION

We compared the occurrence of disfluencies in a) different speech registers (IDS vs. ADS) and b) ADS under different amounts of cognitive load (low vs. high). Regarding a) IDS shows only few instances of disfluencies while the ADS condition with low cognitive load shows a higher number of *filled pauses*, *silent pauses* and *repairs*. Hence, our findings on German fit previous studies on fluency in IDS in other languages [4-6]. Yet, there are more instances of *lengthening* in IDS than in ADS, which, based on a preliminary analysis, are not indicative of hesitations but mainly used for highlighting. German IDS has been shown to exhibit many accents [15], more than one would expect in ADS. The difference in lengthening may thus be due to a larger number of accents in IDS than in ADS.

Regarding b), we expected the map task to generally elicit more disfluencies, but this was only the case for *silent pauses* and most strongly for *lengthening*. We see two explanations that may account for the strong effect of lengthening in the map task: First, it might be a strategy of keeping pace with the interlocutor who finds her way through the map. Second, it could be an indication of an iconic relation between the task object and the speech phenomena describing it (path as a continuous line transferred to speech). Yet, it is possible that a more demanding task would further increase the number of disfluencies and hence increase the difference between the ADS conditions differing in cognitive load.

Possibly, disfluencies would also increase in IDS when cognitive load gets higher, e.g., when caretakers are mentally distracted, which provides an interesting hypothesis for future research. As mentioned above, infants distinguish fluent from disfluent speech [5, 7] and use disfluencies for word recognition [8, 9]. In future work, we plan to test whether disfluencies may serve as indicators for different speech registers in German and to what extent children might benefit from this.

References

[1] Belz, M. and Klapi, M., "Pauses following fillers in L1 and L2 German map task dialogues," in *Proceedings of the Disfluency in Spontaneous Speech Workshop (DiSS-2013)*, 2013, pp. 9-12.

[2] Nicholson, H. B. M., Bard, E. G., Anderson, A. H., Flecha-Garcia, M. L., Kenicer, D., Smallwood, L.*, et al.*, "Disfluency under feedback and time-pressure," in *Proceedings of EUROSPEECH-2003*, 2003, pp. 205-208.

[3] Eklund, R., "Disfluency in Swedish human–human and human–machine travel booking dialogues," PhD thesis, Department of Computer and Information Science, Linköping University, Sweden, 2004.

[4] Newport, E. L., Gleitman, H., and Gleitman, L. R., "Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style," in *Talking to children: Language input and acquisition*, Ferguson, S., Ed., Cambridge, UK: Cambridge University Press, 1977, pp. 109-149.

[5] Soderstrom, M. and Morgan, J. L., "Disfluency in speech input to infants? The interaction of mother and child to create error-free speech input for language acquisition," in *Proceedings of the Disfluency in Spontaneous Speech Workshop (DiSS'05)*, Aix-en-Provence, France, 2005, pp. 157-162.

[6] Björkenstam, K. N., Wiren, M., and Eklund, R., "Disfluency in child-directed speech," in *Proceedings of Fonetik 2013, the XXVIth Swedish Phonetics Conference, Studies in Language and Culture*, Linköping, Sweden, 2013.

[7] Soderstrom, M. and Morgan, J. L., "Twenty-two-month-olds discriminate fluent from disfluent adult-directed speech," *Developmental Science,* vol. 10, pp. 641-653, 2007.

[8] Kidd, C., White, K. S., and Aslin, R. N., "Toddlers use speech disfluencies to predict speakers' referential intentions," *Developmental Science,* vol. 14, pp. 925-934, 2011.

[9] Morin-Lessard, E. and Byers-Heinlein, K., "Uh and euh signal novelty for monolinguals and bilinguals: evidence from children and adults," *Journal of Child Language,* vol. 46, pp. 522-545, 2019.

[10] Busse, K., *Are there differences between German IDS and ADS compared to Japanese?*, Bachelor Thesis, Department of Linguistics, University of Konstanz, 2019.

[11] Hough, J., de Ruiter, L., Betz, S., and Schlangen, D., "Disfluency and laughter annotation in a light-weight dialogue mark-up protocol," in *Proceedings of the Disfluency in Spontaneous Speech Workshop (DiSS-2015)*, Edinburgh, UK 2015.

[12] Boersma, P. and Weenink, D. (2016). *Praat: doing phonetics by computer. Version 6.0.23 (version depended on labeller) [Computer program]*.

[13] Benjamini, Y. and Hochberg, Y., "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society,* vol. 57, pp. 289-300, 1995.

[14] Betz, S., Wagner, P., and Voße, J., "Deriving a strategy for synthesizing lengthening disfluencies based on spontaneous conversational speech data," in *Proceedings of the 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum*, München: LMU, 2016, pp. 19-22.

[15] Zahner, K., Schönhuber, M., Grijzenhout, J., and Braun, B., "Konstanz prosodically annotated infant-directed speech corpus (KIDS corpus)," in *Proceedsings of the 8th International Conference on Speech Prosody*, Boston, USA, 2016, pp. 562-566.