

---

## Filler form and dialogue structure

---

Malte Belz  
*Humboldt-Universität zu Berlin*  
[malte.belz@hu-berlin.de](mailto:malte.belz@hu-berlin.de)

Tag  
Datum  
Zeit  
Raum

This study hypothesizes that acoustic filler form is not chosen arbitrarily by the speaker, but is dependent on its context on the level of adjacent segments (or words) as well as on the level of dialogue structure.

**Fillers**, or filled pauses, are a familiar phenomenon of spontaneous speech. The acoustic form of fillers consists of their segmental construction, e.g., vowel quality and quantity, glottalization, fundamental frequency. In German, there are mainly vocalic or vocalic-nasal forms attested [ə ə: əɪm ε ε: εɪm], but sometimes also clicks or glottalized sequences. **Dialogue structure** is understood in terms of chunks that constitute dialogues, such as questions, responses, explanations, etc. These chunks are also called *dialogue moves*. For the annotation of dialogue structure, the dialogue move scheme from Carletta et al. (1997) is adapted, describing various initiating and response moves within dialogues.

A subset of the GECO corpus (Schweitzer & Lewandowski 2013), comprising 8 task-free spontaneous dialogues (25 min each) of 8 German female speakers, is annotated for fillers, filler context and dialogue structure on multiple layers. Different adjacent contexts exhibit different proportions of vocalic/vocalic-nasal, click and glottal fillers. As to dialogue structure, the duration of vocalic and vocalic-nasal fillers is significantly shorter within negative replies to polar questions compared to replies to wh-questions as well as to replies that reflect uncertainty. Filler vowels occurring within narrative sequences are located significantly lower in the vowel space (ca. 50 Hz difference in F1) than filler vowels within replies to wh-question. These findings suggest that filler form varies systematically as a function of dialogue structure. The combination of filler form and dialogue structure thus provides cues for the disambiguation of filler function.

**References:** • Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., Anderson, A. H. (1997): The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1), 13–31. • Schweitzer, A., Lewandowski, N. (2013): Convergence of Articulation Rate in Spontaneous Speech. In: *Proceedings of Interspeech*, 525–529.