Are longer um… filler particles associated uh with higher surprisal?

Christina Pollkläsener

Filler particles (FPs), i.e. hesitation markers like *uh* or *um*, are often linked to problems in language planning of upcoming material. More recent research has refined this perspective by linking the occurrence of FPs to probabilistic properties of language: speakers are more likely to produce FPs before words that are unexpected or less predictable in context, as quantified by surprisal (e.g. Dammalapati et al., 2021; Zámečník, 2019). Importantly, this has also been reported for simultaneous interpreting (Pollkläsener et al., 2025).

Beyond their mere presence, the phonetic form and duration of FPs may provide further insight into the extent of the underlying production difficulty. Germanic languages like German and English usually distinguish between two major types of FPs: vowel (uh, äh) and vowel-nasal combinations (um, ähm). Clark and Fox Tree (2002) proposed that speakers choose between these FP forms to announce the initiation of either a minor or a major production delay. They suggested that minor delays were associated with vocalic FPs and major delays with vowel-nasal combinations.

We will explore the following two research questions: Do form and duration of FPs vary as a function of upcoming processing difficulty, operationalised as surprisal in the subsequent linguistic material? Does this effect differ between original, non-mediated speech and simultaneous interpreting?

To investigate this, we use the spoken German and English components of the EPIC-Europarl-UdS Corpus (Kunilovskaya & Pollkläsener, 2026), which contains audio recordings and transcripts of European Parliament speeches and their simultaneous interpretations. Manual transcriptions are time-aligned with the audio using WebMAUS (Schiel, 2015), after which we correct alignments for FPs and annotate their pause context following an adapted annotation scheme based on Muhlack et al. (2022). Annotations are carried out in Praat (Boersma & Weenink, 2025).

We estimate surprisal in our corpus with language models (GPT-2, English: Radford et al., 2019; German: Schweter, 2020) and annotate each word with surprisal. Surprisal is an information-theoretic measure that estimates the level of information conveyed by a word (w_i) in its context. For surprisal estimation, FPs are taken out of the corpus so that they are excluded from the context of words following FPs. They are reinserted at a later stage. In the talk, we will present distributions of the different FP types and their duration in the corpus, show results on the link between surprisal and FP type/duration, and elaborate on challenges during annotation.

References

Boersma, Paul & Weenink, David (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.48, retrieved 9 December 2025 from https://praat.org

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3

Dammalapati, S., Rajkumar, R., & Agarwal, S. (2021). Effects of duration, locality, and surprisal in speech disfluency prediction in English spontaneous speech. In *Proceedings of the Society for Computation in Linguistics 2021,* 91-101. https://aclanthology.org/2021.scil-1.9

Kunilovskaya, M., & Pollkläsener, C. (2026). EPIC-EuroParl-UdS: Information-Theoretic Perspectives on Translation and Interpreting. In *Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026)* (pp. 6998–7013). European Language Resources Association (ELRA). https://doi.org/10.63317/3txs6tgs4wsu.

Muhlack, B., Trouvain, J., & Jessen, M. (2022). Acoustic characteristics of filler particles in German. *Phonetik und Phonologie im deutschsprachigen Raum*.

Pollkläsener, C., Kunilovskaya, M., & Teich, E. (2025). Surprisal explains the occurrence of filler particles in simultaneous interpreting. *SKASE Journal of Translation and Interpretation*, 18(2). https://doi.org/10.33542/JTI2025-S-4

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog, 1*(8), 9.

Schiel, F. (2015). A Statistical Model for Predicting Pronunciation. In: *Proc. of the ICPhS 2015*, paper 195.

Schweter, S. (2020). *German GPT-2 model* [Data set]. Zenodo.

Zámečník, J. (2019). *Disfluency prediction in natural spoken language* (Doctoral dissertation, Dissertation, Universität Freiburg, 2019).