

Cross-Lingual Account of Memory and Surprisal for Interpreting

Maria Kunilovskaya, Heike Przybyl, Christina Pollkläsener (University of Saarland),
Ekaterina Lapshinova-Koltunski (University of Hildesheim),
Elke Teich (University of Saarland)
maria.kunilovskaya@uni-saarland.de

Mediated language, a result of translation or interpreting, has been confirmed as identifiably distinct from comparable original production in the target language. The factors that trigger specific linguistic choices in translation and interpreting remain unclear. This computational study investigates the explanatory potential of the information-theoretical account of language processing for parallel simultaneous interpreting data. The data comes from EPIC-UdS (Przybyl et al., 2022), a corpus which contains manual transcriptions of recorded European parliament speeches and their interpreting in both directions for English-German language pair. Our implementation is grounded in memory and surprisal values computed from language-pair-specific MarianMT models, the pre-trained encoder-decoder machine translation Transformer models¹. Interpreting data (transcripts of spoken language) is deemed more suitable than translation to model the memory component in a cross-lingual communicative process as interpreting reflects linear online processing that leaves little room for subsequent correction and editing typical for a translation product. Our aim is to build a model that would approximate available data by varying the amount of context available at the inference time. We expect that the model will return more optimal memory-surprisal trade-off (MST) for liberal translation strategy when more context is available, while more literal and conventionalised choices should achieve comparable MST when the context is limited. In this case, the specificity of production in the situation of cross-lingual mediation can be explained from the rational account of language use, which stipulates that speakers adapt their behaviour to the communicative conditions to keep the processing effort at the necessary minimum.

At the same time, the expected negative correlation of cross-lingual and monolingual surprisal (from a monolingual GPT2) would support the theoretical claim that production effort in interpreting is inversely proportional to the target audience comprehension effort, i.e., the more effort is invested into generating the target, the lower its comprehension cost. The memory component of the model will be represented by the size of the source language context (in sentences) available to generate each segment in interpreting. The document-level context is the premise for various pragmatic and discursive aspects in translation (deixis, ellipsis, lexical cohesion, word order and information flow), and its importance has been emphasized in translation studies and has recently become one of the directions for machine translation improvement (Voita et al., 2019; Läubli et al., 2020). In interpreting studies, the “upstream processing of the previous parts” is hypothesized to make comprehension “easier and faster through gradual construction of a mental model” (Gile, 2008, p. 62). In the context of this study, the translation difficulty of the subsequent sentences in a document is expected to be lower, potentially affecting the allocation of cognitive resources, and hence the outcomes of the mediation process. At the same time, some studies show that production cognitive load is often limited to the sentence boundaries and is not imported to the next sentence (Chmiel et al., 2023; Plevoets & Defrancq, 2020). While cognitive load in interpreting has been a subject of academic scrutiny, we are not aware of any other study where the interpreting data is modelled as directly conditioned by the source language input.

¹ https://huggingface.co/docs/transformers/en/model_doc/marian

References

- Chmiel, A., Janikowski, P., Koržinek, D., Lijewska, A., Kajzer-Wietrzny, M., Jakubowski, D., & Plevoets, K. (2023). Lexical frequency modulates current cognitive load, but triggers no spillover effect in interpreting. *Perspectives*, 32(5), 905–923. <https://doi.org/10.1080/0907676X.2023.2218553>
- Gile, D. (2008). Local cognitive load in simultaneous interpreting and its implications for empirical research. *Forum*, 6(2), 59–77.
- Läubli, S., Castilho, S., Neubig, G., Sennrich, R., Shen, Q., & Toral, A. (2020). A set of recommendations for assessing human-machine parity in language translation. *Journal of Artificial Intelligence Research*, 67, 653–672.
- Plevoets, K., & Defrancq, B. (2020). Imported load in simultaneous interpreting: An assessment. In R. Muñoz Martín & S. L. Halverson (Eds.), *Multilingual mediated communication and cognition* (pp. 18–43). Routledge.
- Przybyl, H., Lapshinova-Koltunski, E., Menzel, K., Fischer, S., & Teich, E. (2022, June). EPIC-UdS-creation and applications of a simultaneous interpreting corpus. In *Proceedings of the 13th Language Resources and Evaluation Conference* (pp. 1193-1200).
- Voita, E., Sennrich, R., & Titov, I. (2019). When a Good Translation is Wrong in Context: Context-Aware Machine Translation Improves on Deixis, Ellipsis, and Lexical Cohesion. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 1198-1212).