

Interpreter-mediated discourse predictability:

A surprisal-based account leveraging large language models

Jessie Jingxian ZENG

Predictability has been empirically associated with processing difficulty in comprehension. In interpreting-mediated encounters, the interpreter and recipients have only one-time immediate access to the utterance, making the predictability effects on comprehension more salient. Nevertheless, Interpreting Studies on prediction focus predominantly on probability prediction in interpreter's processing, with the predictability effects to recipients largely overlooked.

Levelling-out, as one of the widely testified translation and interpreting universals, refers to the tendency of translated/interpreted text to move away from two extremes and gravitate towards the centre of a continuum (Baker, 1996). The tendency toward reduced variability is exemplified as convergence (Laviosa, 2002) in translated and interpreted language. This study hypothesises that convergence will render the interpreted discourse more predictive to recipients compared with the original and non-interpreted comparable discourse. Following surprisal theory (Hale, 2001; Levy, 2008), predictability in this research is quantified as surprisal, the negative log conditional probability given the preceding context, as a post hoc metrics of expectancy. Furthermore, the tendency of levelling-out in interpreted discourse also suggests attenuated information peaks and more constant information rate, conforming more closely to the Uniform Information Density (UID) hypothesis which posits that speakers prefer to maintain a more uniform distribution of information when selecting among linguistic alternatives (e.g., Jaeger, 2010; Jaeger & Levy, 2007).

To testify the hypotheses, this study computes surprisal using autoregressive large language models for aligned Chinese-English interpreted utterances and non-interpreted comparable utterances, and operationalise uniformity of information density as the variance of per-character surprisal. Data are drawn from the BMPRCUS and BMPRCUK subcorpora of the Chinese/English Political Interpreting Corpus (Pan, 2019) comprising bilateral PRC-US and PRC-UK meetings, which ensures robust comparison across the original, interpreted and comparable discourse. This study examines discourse predictability in the highly demanding context of interpreting for recipients on comprehension. The distinctive data selection and LLM-based approach enable computation and comparison of discourse predictability across languages, offering new insights into discourse processing from the distinctive lens of interpreting corpora.

Key words: Predictability, surprisal, interpreting, corpus linguistics, large language models

References

- Baker, M. (1996). Corpus-based translation studies: The challenges that lie ahead. In H. L. Somers (Ed.), *Terminology, LSP and translation: Studies in language engineering in honour of Juan C. Sager* (pp. 175–186). John Benjamins.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics: Language Technologies* (pp. 1–8). Association for Computational Linguistics.

- Jaeger, T. F., & Levy, R. (2007). Speakers optimize information density through syntactic reduction. In B. Schölkopf, J. C. Platt, & T. Hoffman (Eds.), *Proceedings of the 19th International Conference on Neural Information Processing Systems* (pp. 849–856). MIT Press.
- Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1), 23–62.
- Laviosa, S. (2002). *Corpus-based translation studies: Theory, findings, applications*. Rodopi.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177.
- Pan, J. (2019). The Chinese/English Political Interpreting Corpus (CEPIC). Hong Kong Baptist University Library. <https://digital.lib.hkbu.edu.hk/cepic/>