

The evolution of silence: the role of inter-turn speech pauses in the co-evolution of language and cooperation

Theresa Matzinger^{*1,2}, Michael Pleyer³, Elizabeth Qing Zhang⁴, and Przemysław Żywicznyński³

*Corresponding Author: theresa.matzinger@univie.ac.at

¹Department of English and American Studies, University of Vienna, Vienna, Austria

²Cognitive Science Hub, University of Vienna, Vienna, Austria

³Center for Language Evolution Studies, Nicolaus Copernicus University in Toruń, Toruń, Poland

⁴School of Linguistic Sciences and Arts, Jiangsu Normal University, Xuzhou, China

Language and cooperation are closely intertwined aspects of human social interaction, with language serving as an important tool in the assessment of potential cooperation partners (Fox Tree, 2002; Brennan & Williams, 1995). Since pauses between turns of a conversation perform important social-communicative functions, we argue that pauses and their perception may have played a significant role in the co-evolution of language and cooperation.

We support this with data from an experiment testing how listeners use the duration of inter-turn speech pauses to assess others' knowledge, confidence and willingness to grant requests – all of which are highly relevant when evaluating others as potential cooperation partners (Authors, 2023; Authors, *in prep*). We hypothesized that, in general, speakers making long pauses would be regarded as less apt and willing (Roberts & Francis, 2013) but that listeners would be more tolerant towards long pauses in non-native speakers. This is because in non-native speakers, long pauses may result from prolonged cognitive processing when answering in a non-native language (Cenoz, 2000; Guyer et al., 2019) rather than from a lack of knowledge or willingness. Crucially, since evaluating others' cooperativeness is important across cultures, and pause production and perception are similar cross-linguistically (Matzinger & Fitch, 2021), we predicted similar effects across languages.

In our experiment, 100 native Polish-speaking raters listened to short staged conversations, during which a speaker asked questions or made requests that were answered or granted by either native speakers of Polish or native Chinese-speaking non-native speakers of Polish. The pauses before the answers were

manipulated to be either short (200 ms) or long (1200 ms; cf. Roberts & Francis, 2013; Dingemanse & Liesenfeld, 2022). After listening, the raters judged each respondent on their knowledge, confidence and willingness. To test for cross-linguistic similarities, we replicated the experiment with Chinese raters and the two languages reversed.

Our results suggest similarities across languages and cultures: as predicted, Polish and Chinese raters perceived native speakers making long pauses as less knowledgeable, confident and willing (Fig. 1; see non-overlapping 95% confidence intervals). Also, for both rater groups, linear mixed effects models revealed a mediating effect of non-native accent on perceived willingness (interaction: *pause duration* * *accent*: $p < 0.001$ each), but not on knowledge and confidence. A potential reason for the difference between the findings on willingness versus knowledge and confidence is that requests may be more socially engaging and more directly relevant for interpersonal cooperative interactions than knowledge that reflects on partners' competence but not cooperativeness. The raters may therefore have evaluated willingness more carefully.

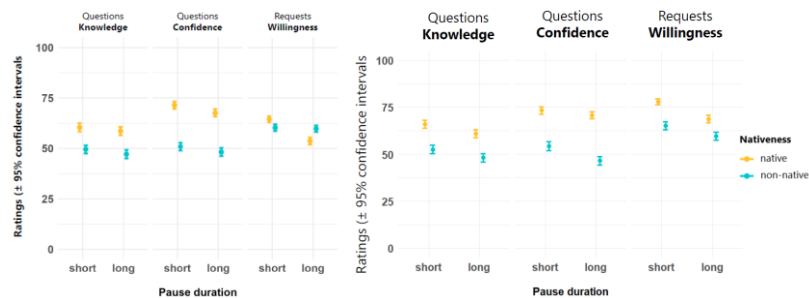


Figure 1. Mean ratings (\pm 95%-confidence intervals) of perceived knowledge, confidence and willingness when native vs. non-native speakers made short vs. long inter-turn pauses. Left: Polish raters, right: Chinese raters.

To further support the point that pauses may have been used as communicative signals in early language evolutionary stages when fully-developed shared linguistic communication systems have not yet been established, we will also show results from a follow-up study testing whether similar relationships hold if participants don't know the language spoken in the conversations. We predict that understanding the content of the answers should not play a pivotal role.

Finally, since pauses in turn-taking occur across species (Levinson, 2016), we will also discuss inter-turn pauses as promising candidates to investigate links between communication and cooperativeness across species.

Acknowledgements

Michael Pleyer was supported by project No. 2021/43/P/HS2/02729 co-funded by the National Science Centre and the European Union Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie grant agreement No. 945339.

References

- Brennan, S. E. & Williams. M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34, 383-398.
- Cenoz, J. (2000). Pauses and hesitation phenomena in second language production. *ITL: Review of Applied Linguistics*, 127, 53-69.
- Dingemanse, Mark & Andreas Liesenfeld. (2022). From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics 1 (Long Papers)* (pp.5614-5633). Dublin: Association for Computational Linguistics.
- Fox Tree, J. E. (2002). Interpreting pauses and ums at turn exchanges. *Discourse Processes*, 34, 37-55.
- Guyer, J. J., Fabrigar, L. R. & Vaughan-Johnston, T. I. (2019). Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion. *Personality and Social Psychology Bulletin*, 45, 389-405.
- Levinson, S. C. (2016). Turn-taking in human communication—origins and implications for language processing. *Trends in Cognitive Sciences*, 20(1), 6-14.
- Matzinger, T., & Fitch, W. T. (2021). Voice modulatory cues to structure across languages and species. *Philosophical Transactions of the Royal Society B*, 376(1840), 20200393.
- Roberts, F. & Francis, A. L. (2013). Identifying a temporal threshold of tolerance for silent gaps after requests. *The Journal of the Acoustical Society of America*, 133, EL471-EL477.