

The role of context and age in the use of French causal connectives by humans and LLMs

Joanna Blochowiak, Ru-Yi Zeng, Marie-Catherine de Marneffe & Liesbeth Degand
(University of Louvain, Belgium)

Keywords: causal connectives, human annotations, Large Language Models annotations, minimal context, extended context

Large Language Models (LLMs) have achieved remarkable success across a wide range of Natural Language Processing tasks. Yet their ability to handle fine-grained pragmatic distinctions, especially those relying on subtle contextual and subjective cues, remains an open question (Gao et al. 2021). This paper investigates this issue through a classic but unresolved problem in French linguistics: the choice between the causal connectives *parce que* and *car* ('because'). While traditionally distinguished along semantic and pragmatic lines – *parce que* being associated with objective causality and *car* with subjective reasoning (Degand & Pander Maat 2003, Pit 2003) – recent research challenges this dichotomy, suggesting greater overlap and ongoing change in usage, particularly among younger speakers (Nazarenko 2000, Zufferey et al. 2018, Blochowiak & Grisot, 2022).

Building on recent pragmatic accounts that conceptualize subjectivity as an emergent property derived from multiple cues distributed across discourse (Blochowiak et al. 2020, 2024), we test the hypothesis that connective choice depends on contextual information extending beyond the sentence level. We address this hypothesis through two complementary studies: a *Minimal Context Study* and an *Extended Context Study*, comparing human judgments with LLM behavior.

In the *Minimal Context Study*, we used as a test set a corpus of 420 French excerpts drawn from journalistic texts (*Le Monde*) and SMS messages (Blochowiak et al. 2020, *SMS4science* Fairon et al. 2006). In each excerpt, the original connective (*parce que* or *car*) was removed and systematically replaced with a <MASK> token. Three human participants and two LLMs were asked to predict the missing connective. Results from three experiments show that human annotators perform poorly on isolated sentences, with low accuracy and agreement, especially in SMS data. GPT-3.5, tested via few-shot prompting, exhibits inconsistent and overall weaker performance than humans, though it performs unexpectedly better on SMS than on newswire texts. As training corpus, we used a dataset of 10,000 sentences (5,000 containing *car* and 5,000 containing *parce que*) sourced from 4,200 *SMS4science* text messages and 5,800 Belgian news articles. A fine-tuned CamemBERT model trained on 10,000 examples from the training corpus outperforms both humans and GPT-3.5, achieving an accuracy of 66.7%, with a strong effect of text genre.

The *Extended Context Study*, currently ongoing, investigates whether access to broader discourse context and speaker age modulate connective choice. Using the same corpus enriched with preceding and following sentences, we compare younger (20–30) and older (39–59) speakers across genres. The same task is then applied to LLMs, including GPT-4.5 and a fine-tuned CamemBERT model, to assess whether models benefit from extended context and whether their behavior aligns with specific age profiles.

In sum, the results from the *Minimal Context Study* had shown that CamemBERT outperforms LLMs and humans in low context environments. The ongoing *Extended Context Study*, enriched context and speaker variation, should inform us whether LLMs align more closely with specific age profiles and

whether they capture variation in connective use, contributing to ongoing debates on the pragmatic competence of LLMs (Gao et al. 2021).

References

- Blochowiak J. & Grisot C. (2022). The mapping between the subjective-objective divide and French causal connectives: Is it about the relation, the connectives or the speakers? In Maria Josep Cuenca & Liesbeth Degand (eds.), *Discourse Markers in Interaction: From Production to Comprehension*, 45–68 Berlin: De Gruyter.
- Blochowiak J., Grisot C. & Degand L. (2024) « Subjectivité, langage et pragmatique : que nous dit l'usage argumentative des connecteurs causaux au sujet de la subjectivité ? » in *Nouvelles approches linguistiques de la littérature et de la fiction*, Michele Costagliola d'Abele & Jana Altmanova (eds.) University Press de l'Université de Naples L'Orientale.
- Blochowiak, J., Grisot, C., & Degand, L. (2020). What Type of Subjectivity Lies Behind French Causal Connectives? A Corpus-based Comparative Investigation of 'car' and 'parce que'. *Glossa: A journal of general linguistics*, 5(1).
- Degand, L., & Pander Maat, H. (2003). A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. In A. Verhagen & J. M. van de Weijer (Eds.), *Usage-based approaches to Dutch* (pp. 175–199). Lot.
- Escoufflaire L., Blochowiak J., Degand L., & de Marneffe, M-C. (2024). "Which connective fits best: 'car' or 'parce que'? A challenge for both humans and LLMs." In: *JADT 2024 Proceedings: 17th International Conference on Statistical Analysis of Textual Data*, JADT: Brussels 2024.
- Fairon, C., Klein, J., Paumier, S. (2006). Le corpus SMS pour la science. *Base de données de 30.000 SMS et logiciels de consultation*. Louvain-la-Neuve, Presses universitaires de Louvain.
- Gao, S., Alawad, M., Young, M. T., Gounley, J., Schaefferkoetter, N., Yoon, H. J. et al. (2021). Limitations of transformers on clinical text classification. *IEEE journal of biomedical and health informatics*, 25(9), 3596-3607.
- Nazarenko, A. (2000). *La cause et son expression en français* (p. 169). Ophrys.
- Pit, M. (2003). *How to express yourself with a causal connective: subjectivity and causal connectives in Dutch, German and French*. Amsterdam: Rodopi.
- Zufferey, S., Mak, W., Verbrugge, S., & Sanders, T. (2018). Usage and Processing of the French Causal Connectives 'car' and 'parce que'. *Journal of French Language Studies*, 28(1), 85-112.