

Connective Frequency, Exposure to print and Discourse Expectations: A Bayesian Account of Individual Differences in Coherence Judgments

Antoine Cruchon¹, Sayaka Sato¹, Pascal Gyga² and Sandrine Zufferey¹

¹University of Bern, Institute of French Language and Literature

²University of Fribourg, Department of Psychology

Adult readers are often treated as relatively homogeneous once basic literacy has been acquired, yet recent work suggests substantial variability in discourse-level competence, especially in the comprehension and evaluation of discourse connectives from the written mode (Crible & Cuenca, 2017; Scholman Demberg, & Sanders, 2020; Zufferey et al., 2022). This study examines this variability through the lens of discourse expectations. During comprehension, readers anticipate how incoming segments relate to previous discourse, and connectives constrain these expectations by signaling the coherence relation (Sanders, Spooren, & Noordman, 1992). However, not all connectives provide equally strong cues: frequent connectives are highly understood, whereas rare connectives are likely to be encountered mainly by readers with greater exposure to print (Zufferey & Gyga, 2020).

In this present study, we argue that connective frequency is a distributional correlate of expectation strength. Repeated exposure strengthens lexical representations and stabilizes form–function mappings (Bybee, 2006; Perfetti, 2007). Frequent connectives should therefore make coherent uses easier to recognize and accept, whereas rare connectives should lead to lower or less categorical coherence judgments, even when used correctly. In order to empirically assess this claim, our design compared frequent and rare monofunctional French connectives within the same discourse relations, allowing frequency effects to be tested while holding relational meaning constant.

In addition, discourse relation type may further modulate these expectations. The study focused on Cause, Concession, Addition, and Condition, which differ in inferential properties, notably basic operation and polarity (Sanders et al., 1992; 2018). Causal relations may benefit from a privileged interpretative status: according to the causality-by-default hypothesis, readers tend to construct causal interpretations when possible (Sanders, 2005). Causal coherence may therefore remain relatively robust even when signaled by a rare connective. By contrast, concessive and conditional relations may depend more strongly on explicit connective marking.

Individual experience provides a third layer of expectation. Exposure to written language has been shown to affect discourse-level competence and connective comprehension (Zufferey & Gyga, 2020; Scholman et al., 2020; Zufferey et al., 2022). Readers with greater exposure to print are more likely to have encountered low-frequency written connectives and to have internalised their discourse functions. Exposure to print should therefore attenuate the processing cost for rare connectives, especially for relations that depend strongly on explicit marking.

Native French-speaking adults completed a coherence judgment task involving sentences consisting of frequent and rare monofunctional French connectives expressing Cause, Concession, Addition, or Condition. Participants also completed measures of Exposure to print (ART-F). Coherence ratings were analyzed using a Bayesian zero–one-inflated beta mixed model, suitable for bounded judgments with genuine boundary responses (Ospina & Ferrari, 2010). This approach distinguished mean coherence ratings from categorical response tendencies, interpreted here as an index of expectation stability: categorical responses reflect stronger expectations, whereas intermediate responses may indicate uncertainty in accepting the intended relation.

Results showed that rare connectives received lower coherence ratings overall, supporting the view that lexical frequency constrains connective-based expectations. However, this rarity cost was relation-dependent. Causal relations were comparatively robust to rarity, consistent with the idea that causal expectations can often be generated from discourse content itself. Concessive and conditional relations were more strongly penalized, suggesting that when a relation is less expected or more dependent on explicit marking, connective availability becomes crucial. Exposure to print increased coherence ratings and attenuated rarity costs, especially for rare connectives and more demanding relations.

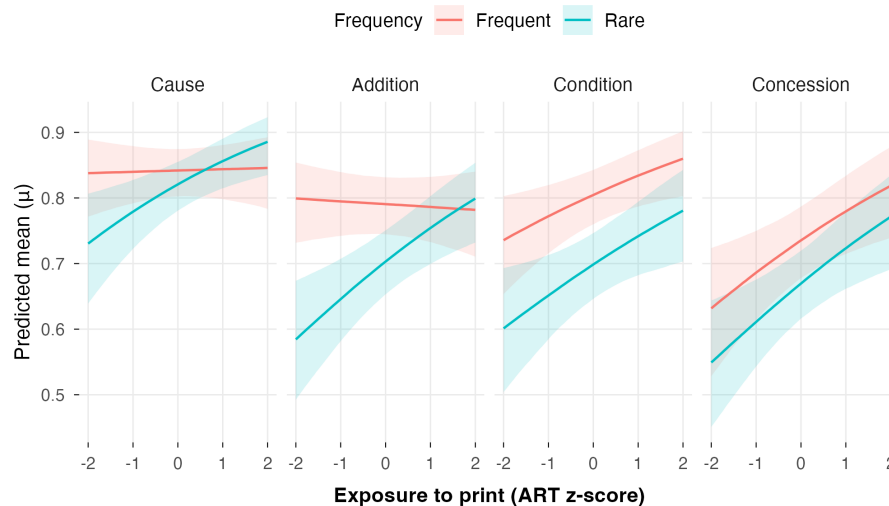


Figure 1. Predicted mean coherence ratings (μ) by ART-F-R scores. Predicted mean coherence ratings (μ) are presented as a function of ART-F-R scores (z-score) across discourse relations.

These findings suggest that adult variability in connective comprehension is affected by differences in learned discourse expectations. Connective frequency reflects the distributional strength of form–function mappings; relation type determines how much interpretation relies on explicit marking; and print exposure modulates the individual availability of these mappings. More broadly, the study contributes to expectation-based and probabilistic accounts of discourse comprehension by showing that coherence judgments are shaped by both language-wide distributional regularities and individual written experience (Asr & Demberg, 2020; Scholman et al., 2020).

References

- Asr, F. T., & Demberg, V. (2020). Interpretation of discourse connectives is probabilistic: Evidence from the study of *but* and *although*. *Discourse Processes*, 57(4), 376–399. <https://doi.org/10.1080/0163853X.2019.1700760>
- Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733.
- Crible, L., & Cuenca, M. J. (2017). Discourse markers in speech: Characteristics and challenges for corpus annotation. *Dialogue and Discourse*, 8(2).
- Ospina, R., & Ferrari, S. L. P. (2010). Inflated beta distributions. *Statistical Papers*, 51(1), 111–126. <https://doi.org/10.1007/s00362-008-0125-4>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Sanders, T. (2005). Coherence, causality and cognitive complexity in discourse. *Proceedings/Actes SEM-05, First International Symposium on the Exploration and Modelling of Meaning*, 105–114.
- Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15(1), 1–35. <https://doi.org/10.1080/01638539209544800>

Scholman, M. C. J., Demberg, V., & Sanders, T. J. M. (2020). Individual differences in expecting coherence relations: Exploring the variability in sensitivity to contextual signals in discourse. *Discourse Processes*, 57(10), 844–861. <https://doi.org/10.1080/0163853X.2020.1813492>

Zufferey, S., & Gygas, P. (2020). “Roger broke his tooth. However, he went to the dentist”: Why some readers struggle to evaluate wrong and right uses of connectives. *Discourse Processes*, 57(2), 184–200. <https://doi.org/10.1080/0163853X.2019.1607446>

Zufferey, S., Tskhovrebova, E., Wetzel, M., & Gygas, P. (2022). Individual differences in the ability to master connectives: The importance of exposure to print. In M. Cuenca & L. Degand (Eds.), *Discourse Markers in Interaction: From Production to Comprehension* (pp. 69–88). De Gruyter Mouton. <https://doi.org/10.1515/9783110790351-004>