Incremental RSA Model Explains Adjective Ordering Preferences by Communicative Efficiency Across Contexts

Hening Wang & Fabian Schlotterbeck & Michael Franke
University of Tübingen

Motivation: Various predictors [1] have been proposed to account for cross-linguistic adjective ordering preferences (e.g. *big brown box* vs. *brown big box*). We focus on two closely related, cognitively oriented predictors, subjectivity [2] and discriminatory strength [3]. They are both grounded in the notion of communicative efficiency, assuming that utterances which are less costly and more informative are more useful. However, depending on the perspective taken (listener vs. speaker) and the assumptions about the underlying structure (hierarchical vs. linear), different implications follow. Assuming a listener who restricts the set of potential referents by interpreting an AAN sequence from right to left, according to its hierarchical compositional structure, more informative adjectives are expected to be integrated earlier to reduce the risk of misidentification. Consequently, a rational speaker would produce them **later** [4, 5]. On the other hand, speakers tend to maximize utility at each incremental step by selecting the most useful utterances and thus place them linearly **earlier** [3]. [6] pitted these two hypotheses against each other in a slider rating experiment showing that, despite the overall subjective-first bias, subjective adjectives (e.g., *big*) are preferred linearly later when their communicative efficiency is effectively reduced through contextual manipulation (see Fig. 1 for an illustration and Fig. 2a for the result). This finding aligns with the predictions of the discriminatory strength hypothesis but contradicts those of the subjectivity hypothesis. [6] also proposed an incremental computational model within the Rational Speech Act framework [7] (in the following RSAaop), qualitatively accounting for their data but only if a general subjective-first bias is assumed.

Methods & Results: We extend the application of the RSAaop [6]. One novelty is the transition from WebPPL to the state-of-the-art machine learning framework NumPyro, built on JAX. This shift substantially improves the efficiency of MCMC inference, particularly when dealing with intermediate marginal distributions over large discrete utterance and state spaces used in recursive computations. By fitting the model to the data, we show that the best-fitted model with an incremental speaker accounts quantitatively for various effects of discriminatory strength (see Fig. 2b). The overall subjective-first bias still requires explicit encoding via utterance cost. Moreover, a non-incremental speaker cannot capture all effects related to discriminatory strength, in particular effects related to in-context informativity of gradable adjectives. One possible explanation why effects derived from the subjectivity hypothesis were limited in our data and model thereof is that the variation in the **number** and **size** of objects in the contexts used may not be large enough to create a substantial risk of misidentification. We address this issue through simulation in a large set of randomly generated contexts with greater variability along these two dimensions. We find that a general subjective-first preference emerges. This holds not only for the non-incremental speaker, in a similar vein but in a different setting from [5], but also for the incremental speaker (see 3a). Furthermore, only an incremental speaker can overcome the preference for the subjective-first ordering under specific parameter settings (see 3b). This may help explain the conventionalization of the subjective-first bias.

Conclusion & Outlook: We showed that RSAaop captures effects of communicative efficiency on adjective ordering, both within experimental contexts and in a large set of simulated communicative contexts. We also applied the model to the free production data and will present preliminary results at the workshop.

References:

[1] Scontras, G. (2023). *Annual Review of Linguistics*, *9*, 357–376. [2] Scontras, G., Degen, J., & Goodman, N. D. (2017). *Open Mind*, *1*(1), 53–66. [3] Fukumura, K. (2018). *Journal of Memory and Language*, *101*, 37–50. [4] Scontras, G., Degen, J., & Goodman, N. D. (2019). *Semantics and Pragmatics*, *12*, 7–1. [5] Franke, M., Scontras, G., & Simonic, M. (2019). *CogSci*, 344–350. [6] Schlotterbeck, F., & Wang, H. (2023). *Proceedings of the Society for Computation in Linguistics*, *6*(1), 121–132. [7] Goodman, N. D., & Frank, M. C. (2016). *Trends in cognitive sciences*, *20*(11), 818–829.

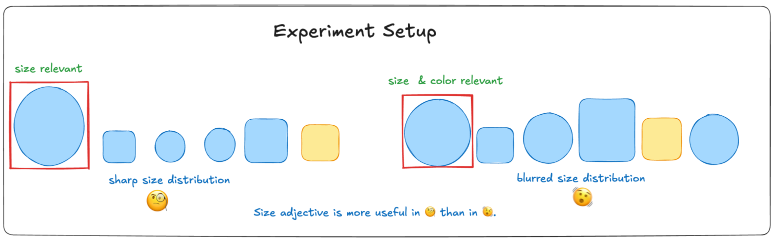
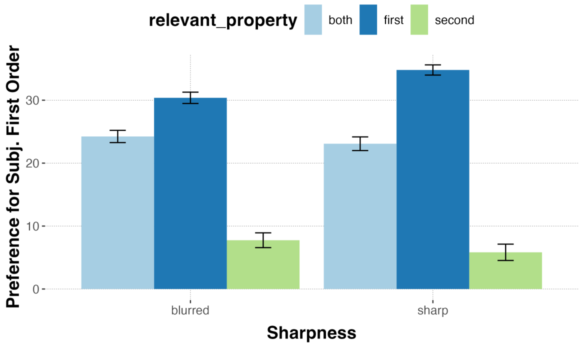
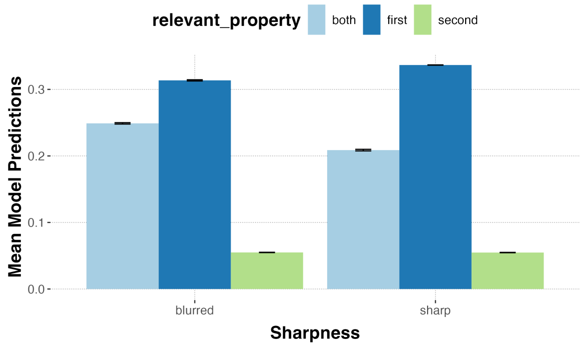


Figure 1: Item setup for the experiment in a mixed design. Left: size-relevant context in the high-contrast (sharp) condition; right: size- and color-relevant context in the low-contrast (blurred) condition. The relevant property, as a within-subject factor, determines the contextual informativeness of all adjectives, while sharpness, as a between-subject factor, modulates the in-context informativeness of the gradable adjectives.

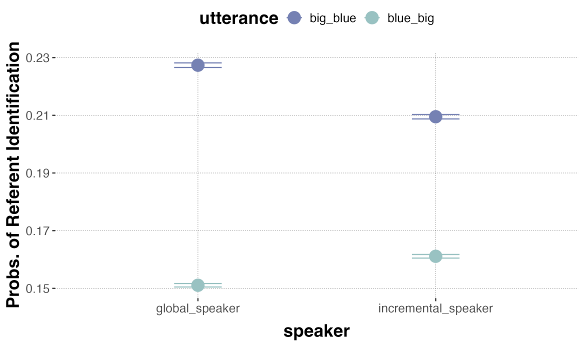


(a) Experiment results from the slider rating task [6]. The y-axis indicates the preference rating for the subjective-first ordering, modulated by contextual informativeness of relevant property. This preference decreases when the subjective adjective is less informative—specifically, in blurred compared to sharp conditions when the subjective adjective is the relevant property.

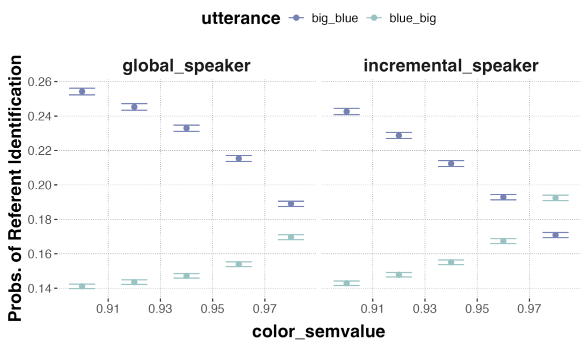


(b) Posterior predictions of mean preference ratings from the best-fitting model. The model successfully captures the mean ratings across conditions but does not fully account for the variance. This discrepancy may be due to individual differences in the data, as the analysis aggregates over all participants.

Figure 2: **Left:** experiment result and **right:** posterior predictive from the best-fitting model



(a) Mean probabilities of identifying the target referent aggregated over all parameter settings: For both non- and incremental speaker models, the subjective-first ordering yielded a higher probability than the reverse order.



(b) The x-axis represents the values of the parameter *color semvalue*, which determines how inherently informative the color adjective is—the higher the value, the more informative it becomes. When color is sufficiently informative, only the incremental speaker model is able to reverse the otherwise preferred subjective-first ordering.

Figure 3: Results of the simulations. The y-axis indicates the probability of successful referent identification for each utterance. 10,000 samples per parameter combination.