

Bayesian models in psycholinguistics

Bruno Nicenboim (Department für Linguistik, Universität Potsdam)

Statisticians have been arguing for decades about the relative merits of the frequentist vs. Bayesian statistical methods. For a long time, the difficulty and the effort required in applying Bayesian methods made their merits virtually irrelevant for many researchers. The last years, however, have seen several rapid revolutions in probabilistic programming languages that allow researchers to fit Bayesian models in a relatively straightforward way.

In the first part of the talk I will discuss briefly the main advantages of using Bayesian methods for data analysis in psycholinguistics (and cognitive science in general):

(i) The possibility of directly answer the question we are interested in by quantifying our uncertainty about the parameters of interest, instead of focusing on rejecting a strawman null hypothesis.

(ii) The flexibility in defining hierarchical models (also known as mixed effects or multilevel models) that are highly relevant for the repeated-measures designs used in linguistics and psycholinguistics (see also Nicenboim & Vasishth 2016).

In the second part of the talk, I will discuss my research on computational modeling of sentence processing phenomena. I will use Nicenboim & Vasishth (under review in JML) as a case study that takes advantage of Bayesian methods.

The focus of Nicenboim & Vasishth (under review in JML) is on disentangling the predictions of two influential models of sentence processing: the *activation-based model* (Lewis & Vasishth 2005) and the *direct access model* (McElree 2000). These two models assume that constraints on working memory influence the speed and accuracy of dependency resolution processes. The models give virtually identical predictions when comparing only mean reading times (or accuracy). However, the use of Bayesian methods to fit reading times and accuracy simultaneously reveals interesting differences between the models: it shows that the direct access model provides a better fit to the data, and, in addition, where the activation-based model fails and how it should be improved.

References

Lewis, R. L. & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29 (3), 375–419. doi:10.1207/s15516709cog0000_25

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29 (2), 111–123. doi:10.1023/A:1005184709695

Nicenboim, B. & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas - Part II. *Language and Linguistics Compass*, 10 (11), 591–613. doi:10.1111/lnc3.12207. eprint: <https://arxiv.org/abs/1602.00245>

Nicenboim, B. & Vasishth, S. (Under review in JML). Models of retrieval in sentence comprehension: A computational evaluation using Bayesian hierarchical modeling. eprint: <https://arxiv.org/abs/1612.04174>