Modeling latent processes during garden-pathing with data from a large benchmark study

Dario Paape[1], Tal Linzen[2] & Shravan Vasishth[1]
[1]University of Potsdam, [2]New York University

Multinomial processing trees (MPTs) are a class of model that has seen much empirical success in cognitive psychology [1]. MPTs decompose cognitive phenomena into multiple latent processes that occur serially and probabilistically. Paape & Vasishth [2] applied an MPT model to garden-pathing in temporarily ambiguous English sentences (e.g., *When the little girl attacked the lamb remained calm*). They broke garden-pathing down into a mixture of several probabilistically occurring subprocesses: paying attention to the stimulus (or not), initially adopting the incorrect analysis (or not), and actually engaging in syntactic reanalysis (or not; "triage" [3]). The results showed, inter alia, that readers do not always carry out reanalysis, but when they do, it is much more costly than had previously been assumed based on simple comparisons of condition means.

In the current work, we apply the Bayesian MPT model of Paape & Vasishth [2] to a large benchmark data set on garden-pathing in English recently published by Huang et al. [4], called SAP. The SAP data contain GPT-2 surprisal values for the disambiguating word for every sentence. This allows us to integrate the surprisal measure into the MPT model, and to ask whether surprisal alone is sufficient to explain differences in processing between conditions. The SAP experiments also differ from the experiments modeled by Paape & Vasishth (P&V) in two important aspects:

- The P&V experiments allowed rereading during self-paced reading while SAP did not, meaning any reanalysis must take place "in situ".
- P&V used grammaticality judgments while SAP used comprehension questions.

P&V assumed that readers reject garden-path sentences as ungrammatical if reanalysis is not carried out ("triage") or fails. However, when asked *Did the girl attack the lamb?*, readers may respond "yes" even if reanalysis succeeded, due to a pragmatic inference. Our updated MPT model takes such inferences into account by adding another subprocess, as shown in Figure 1. Furthermore, the updated MPT models spillover effects by assuming that reanalysis, if it takes place, can be postponed to the post-critical region, so that its associated processing cost is paid there.

The updated MPT model was fitted to a subset of the SAP data (NP/Z and MV/RR garden paths, see Table 1) in Stan [5]. The probabilities of each subprocess being triggered, as well as the cost of reanalysis, are assumed to be influenced by ambiguity, garden-path type, and surprisal, as well as by individual differences between readers and sentences, which are included as random effects. Model comparisons were also carried out using approximate leave-one-out cross-validation [6] to compare the predictive performance of different model versions on unseen data.

Model comparisons indicated that the MPT model with its mixture of processes has significantly better predictive performance than a simple multivariate model fitted to reading times and question responses that only accounts for condition means. Importantly, an MPT model containing ambiguity and garden-path type as categorical predictors outperforms a model that only contains surprisal as a predictor of the process probabilities. However, surprisal does affect the probability of garden-pathing in the MPT, with an increase of one SD of surprisal increasing the probability by about 3% (see Figure 2).

Overall, the results show that the Paape & Vasishth's MPT can account for data from different paradigms and tasks, as well as effects of surprisal on garden-pathing. The results are also in line with previous work showing that surprisal alone is insufficient to explain garden-path effects [4, 7]. In future work, we plan to fit the MPT to combined data from different tasks to model shared latent processes across experiments.

NP/Z AMBIGUITY: Comma (unambiguous) versus no-comma (ambiguous)

(1) When the little girl { attacked, / attacked } the lamb remained relatively calm . . .

MV/RR AMBIGUITY: Unreduced RC (unambiguous) versus reduced RC (ambiguous)

(2) The little girl { who was fed / fed } the lamb remained relatively calm . . .

Table 1. Example stimuli used in the SAP benchmark.

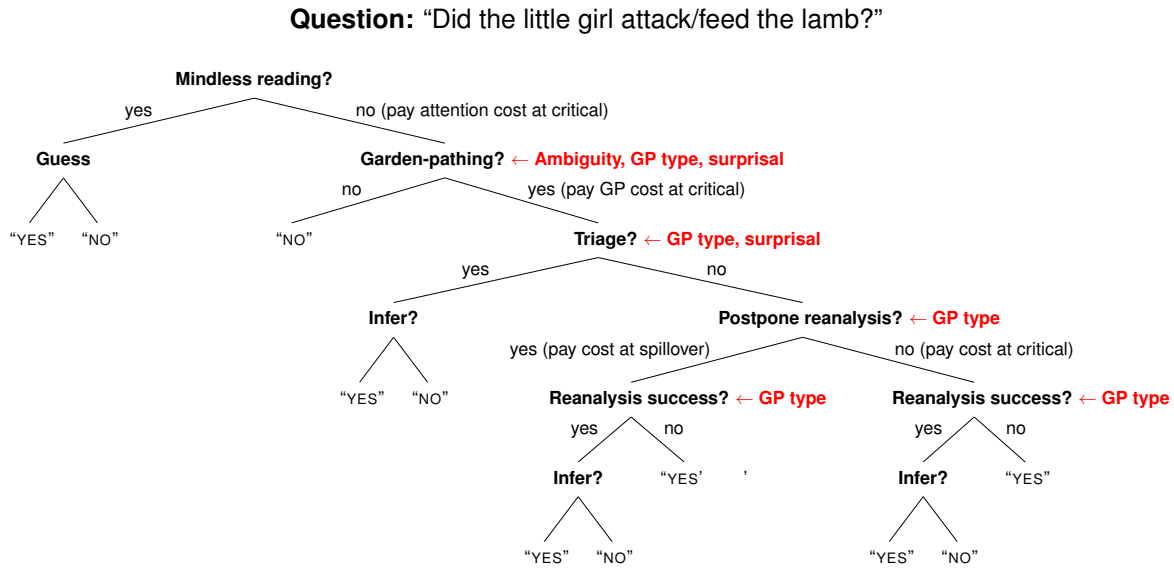

Figure 1. Extended multinomial processing tree model of garden-pathing and reanalysis.

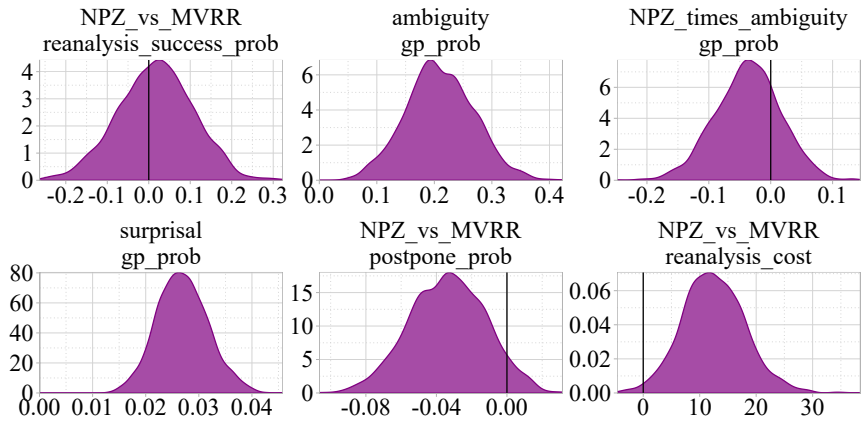

Figure 2. Effect estimates for different predictors on a subset of the MPT parameters.

References. [1] Singmann et al., 2024, *Psychol Bull* [2] Paape & Vasishth, 2002, *Cognitive Sci* [3] Fodor & Inoue, 2000, *Lang Speech*. [4] Huang et al., 2024, *J Mem Lang* [5] Stan Development Team, *mc-stan.org* [6] Vehtari et al., 2017, *Stat Comput*. [7] Van Schijndel & Linzen, 2021, *Cognitive Sci*.