**Appendix: Retrospective power analysis for Experiment 3**

We carried out a power analysis for Experiment 3 according to the procedure used by Vasishth, Mertzen, Jäger & Gelman (2019). We used a range of possible effect sizes, based on the estimate of the effect size extracted from the Bayesian analysis and its credible interval boundaries. We generated repeated data simulations to determine the proportion of cases where an effect of the relevant size is detected. We defined the discovery rate as the proportion of cases where the absolute t-value of the effect of interest exceeds 2. The code and the associated data are available at <https://osf.io/9XB2Y/> as supplementary martials.

Table A1 summarizes the results of the power analyses for the two discussed effects: (i) the contrast between wh-questions and the *regarding*-phrase condition at the critical region; and (ii) the contrast between *regarding* and *with* conditions at the spillover region.

|  |  |
| --- | --- |
|  | Effect size |
| Contrast | Low boundary of CI | Estimate | High boundary of CI |
| Critical word: *wh* vs. *regarding* | 29 ms | 60 ms | 92 ms |
| 54% | 96% | 100% |
| Spillover region: *regarding* vs. *with*  | 6 ms | 31 ms | 57 ms |
| 13% | 60% | 98% |

Table A1. Retrospective power analysis of Experiment 3. We report a range of effect sizes derived from the effect size estimate and its 95% credible intervals, and the corresponding power estimates (the probability of detecting the effect under repeated sampling, expressed as a percentage).