

## Supplementary materials

### Post-hoc power analyses

Post-hoc power analyses were conducted on the first 10 “pilot” participants to determine appropriate sample sizes. The `mixedpower` function from the `mixedpower` package was used for the analysis (Kumle et al., 2021). The effects of a `clmm` and an `lmer` model on both datasets were similar, so we could continue fitting an `lmer` model on our data. The code for these analyses are publicly available at OSF (see [https://osf.io/r8hck/?view\\_only=c07896d7a6214acead46b5d3880d75b8](https://osf.io/r8hck/?view_only=c07896d7a6214acead46b5d3880d75b8)).

For the first analysis, an `lmer` model was fitted on the valence ratings with accent, type of valence and their interaction as fixed effects. Participants and items were added as random effects and a random slope of accent by participants was also included. Power was estimated over two different sample sizes: 40 (our current sample size) and 80 (as requested by an anonymous reviewer). A  $t$ -value of 2 was determined as the significance threshold. The power analysis showed no differences between the two sample sizes. Forty or 80 participants would allow for 2% or 3% power (contrast 1) and 99% or 100% power (contrast 2) to measure a plausibly-sized interaction between accent and valence.

For the second analysis, we fitted an `lmer` model on our arousal ratings with accent, type of arousal, and their interaction as fixed effects. Participants and items were included as random intercept and a random slope of accent by participants was also added. Power was estimated over two different sample sizes (40, 80). A  $t$ -value of 2 was determined as the significance threshold. The analysis showed that 40 or 80 participants would allow for 57% or 89% power to measure a plausibly-sized effect between accent and arousal. This shows an increase in

power with a larger sample size, however, our results already showed a significant interaction with the smaller sample size.

Kumle, L., Vø, M. L. H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavioral Research*, 53, 2528-2543. <https://doi.org/10.3758/s13428-021-01546-0>

\*\*\*\*\*

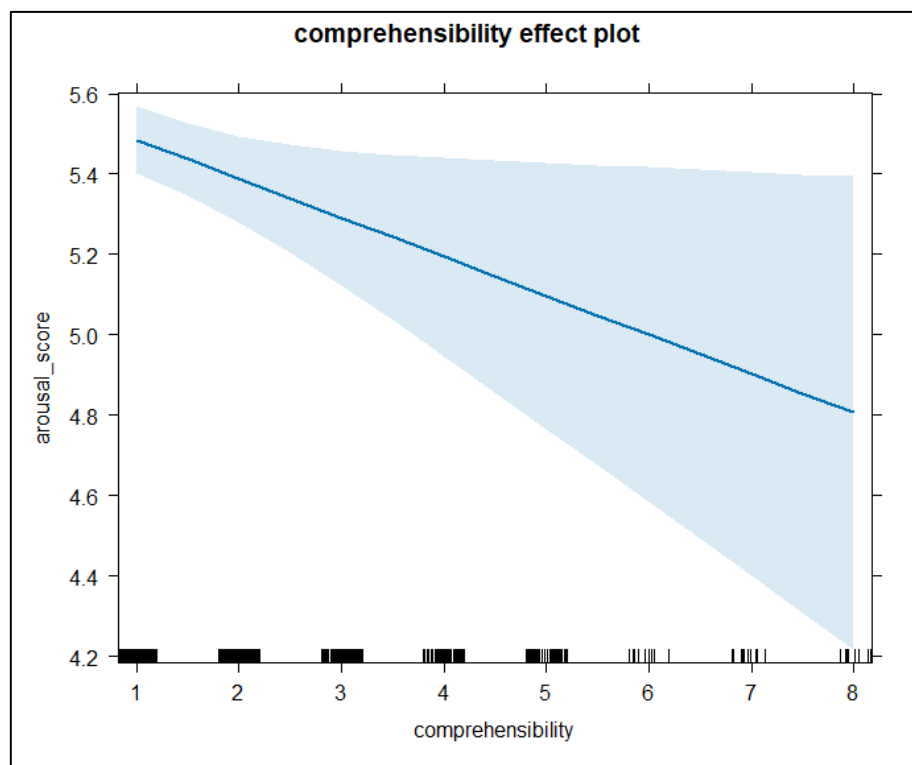


Figure S1. Results of the post-hoc analysis testing the interaction between comprehensibility and accent on valence ratings.

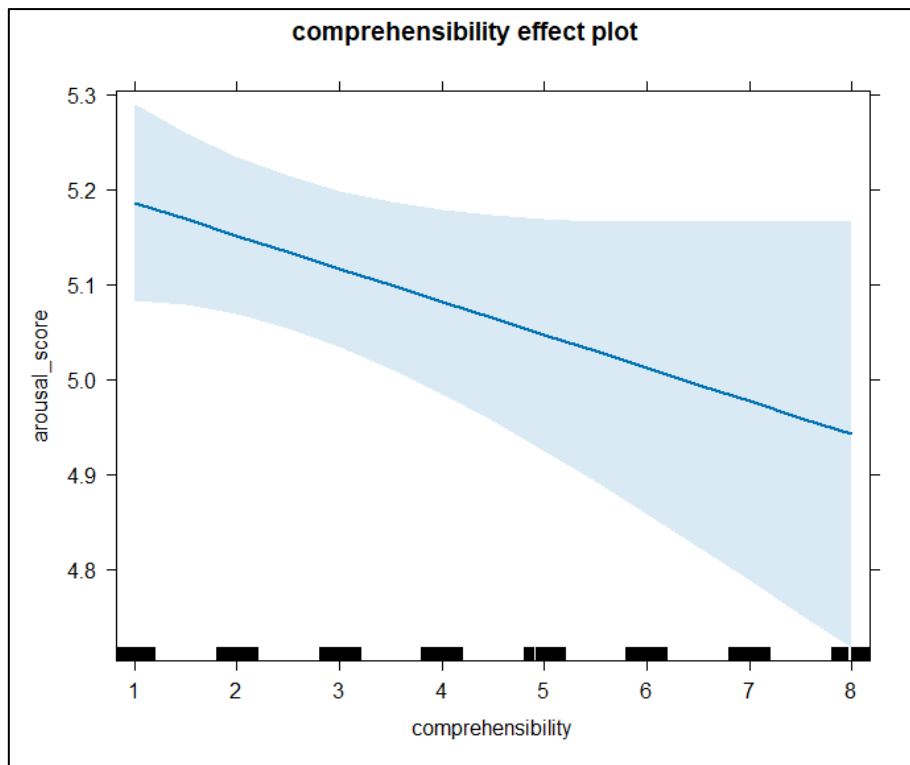


Figure S2. Results of the post-hoc analysis testing the interaction between comprehensibility and accent on arousal ratings.