

REVIEW - WHAT A TRANSPARENT ROMANCE LANGUAGE WITH A GERMANIC GENDER-DETERMINER MAPPING TELLS US ABOUT GENDER RETRIEVAL: INSIGHTS FROM EUROPEAN PORTUGUESE.

This paper explore the gender activation during production in Portuguese, a formal gender language with one determiner per gender. In general, the paper is well written and the review of the literature is appropriate for the field. However, I find some problems with the stats, and therefore I find it hard to comment on the discussion. Among other things, the authors claim to our mixed-effects models that are based on linear regressions but they report F values when they do not do ANOVAs. The most important things that I miss are information about the contrasts used, the data transformation, the logic behind selecting the variables, and the inclusion of some factors for controlling random variance. Also, reporting size effects based on raw RT subtraction is quite surprising for me. I consider that a better treatment of the data should be done in order to support the conclusions they claim in the discussion. Below, I present my comments and suggestions following the page order.

Pg 3. “ we can therefore define grammatical gender as an abstract lexico-semantic feature of nouns whose value has no semantic consequences”. I do not consider this sentence is appropriate, such as there exists some research that does find semantic activation triggered by grammatical gender, see Casado, Palma, & Paolieri (2018; 2021).

Pg 8. “the DSM has been extensively criticized”. In my opinion this is too harsh, considering that only two papers argued against it.

Pg 9. “het huis”. Please, include the English translation for this word.

Pg 11. “PWI”, please introduce the initialism, i.e., Picture Word Interference (PWI)

Pg 13. (Experiments 2 and 3) — there is a typo, please include a space

Pg 13. Hypothesis 2: in my opinion, this hypothesis requieres some theoretical background, like introduce in line with which model are these findings.

Pg 14. “EP”, if you are to use initialism to describe the European Portuguese, introduce this before

Pg 19. “PWIP” — PWI paradigm

Pg 21, Table 5. I cannot understand this table at all. Are the columns corresponding to the distractors? What the numbers under the RT information? The subtraction? why do you consider this is informative?.

Pg 22. Please, use more transparent names for the variables in the model (at least for NEC), and introduce the factors, their levels and the contrast selected.

What contrasts were used for each variable? I miss this information, and this is crucial for understanding the model outcomes.

Why did you introduce the masculine/feminine categories in the model? Do you have any specific predictions about it? I could not find anything in the introduction. Do you do this for control based on previous studies?

What does it mean this (NEC means Transparency Congruency)? why do you include this as a random factor in the model?

I think you should include “the list” as a factor to control for possible confounds related with the different items.

Pg 23. Please include a table with the model outcome, including the estimates and the t-values for each factor and interaction.

Why do you report F values if you are performing mixed-effects models? Please, report the t-values from the linear regressions.

Post-hoc comparisons? did you create an additional model? If so, describe it. Otherwise, please use the function `lsmeans` or similar to report the predictions for each level and the marginal means such as you account for other factors variability

“A look at the RTs”, what does it mean? How did you calculate the effect size? what are the stats for that?

“Planned comparisons revealed that the GCE was due mainly to transparency incongruent distractors (when target and distractor were from different categories of transparency, GC = 836 ms, GI = 855 ms; for instance, an opaque target such as “*tapete*” (rug) paired with a transparent distractor such as “*xadrez*”, $p = .0095$; conversely, for targets and distractors of the same transparency category, $p = .90$).” I cannot understand this paragraph.

“the GCE was due mainly to opaque targets” This is not what I understand from the description, I think it should also include opaque distractors. If not, please correct me and clarify the difference in the text.

Please, in the tables report the estimated values, not raw RT values.

Pg 24. “the gender effect sizes” is this really the subtractions? if so, I do not think you should consider this as an effect size measurement. If this is based on some previous literature please cite this information.

Pg 27. Same comments as with Experiment 1 regarding the stats (also for Experiment 3)

Pg 27. “BNs” if you are to use initialism to describe the Bare nouns, introduce this before

Pg 28. Like different sessions? please specify

Pg 29. When and where the participants saw the lists? in the same session? what was the order of the lists? (I saw that later you introduce this information in Pg. 32)

Pg 33. Please include in the model the order of the list for each participant to account for possible confounds.

Pg 34. “The GCE was due to opaque target nouns” — but not in Experiment 3, I do not think you can say this.

Pg 39, footnote 4. You can actually check this including in a model the desinence of the target and the desinence of the distractor. Please report this analysis if you are going to make such conclusion.