Conceptual and grammatical gender at scale: Normative estimates across Czech and English

James Brand[1], Mikuláš Preininger[1], Adam Kříž[1] & Markéta Ceháková[1]
[1]Charles University, james.brand.ac@gmail.com

The extent to which grammatical gender 'rubs off' on conceptual gender has provoked much discussion in the linguistic relatively literature (Samuel, Cole & Eacott, 2019), with mixed results depending on task and contextual constraints. At the core of most studies that investigate whether grammatical gender has an effect on the conceptual representation of gender are carefully controlled and selected stimuli, where word forms in a grammatically gendered language typically have either masculine, feminine or neuter forms, which can be compared to translation equivalent word forms from a control language that does not have grammatical gender. In our study, we take an exploratory approach to better understand the importance of stimuli for any studies aiming to test hypotheses related to linguistic relativity and grammatical gender by creating a large database of images and words that have been normed for conceptual gender association, which can be used to explore where grammatical gender does and does not influence the associated gender across different participant groups.

Our approach focuses on Czech - a language with grammatical gender, and English - a language without grammatical gender. We conducted a large-scale norming study, where native Czech (n = 731), native English (n = 299) and L1Cz-L2Eng (n = 521) speakers rated over 500 concepts taken from the Multilingual Picture database (Duñabeitia et al., 2022), which were presented both as words and as visual images in colour and in grayscale. Participants were presented with lists of words or images, presented in Czech for Czech participants, and in English for the English and L1czL2eng participants, and were asked to rate how they associated the meaning of the word/image in terms of how feminine-neutral-masculine it was, using a 7-point Likert scale (see Vankrunkelsven et al., 2022). All concepts were coded for grammatical gender in Czech (FEMININE/NEUTER/MASCULINE), animacy (ANIMATE/INANIMATE) and semantic category (ANIMAL/BODY/CLOTHING/FOOD/HOUSEHOLD/MUSIC/NATURE /OTHER/PEOPLE/PLACES/SPORTS/TOOLS/TRANSPORT).

We modelled the participant responses using cumulative link mixed-effects models (see Taylor et al., 2021), where the data was coded as an ordinal response, with a random intercept for participant and a random slope of condition (Czech/English/L1CzL2eng) on the item intercept. Three separate models were run – colour images, grayscale images and words, with each model providing an estimate of the latent gender association, and also with an estimate of the difference between the ratings across the three participant conditions, which we can then analyse to see which concepts, if any, differ based on the grammatical gender, animacy and semantic category. Unsurprisingly, our results show that for animates, the Czech and the L1czL2eng groups differ from the English condition, whereby the ratings are more congruent to the grammatical gender in Czech for both images and words, but this trend is not observed for inanimates. However, when we look at these patterns by the semantic categories, we can get a more nuanced picture, with ANIMALS, FOOD and PEOPLE all showing a trend towards grammatical gender congruency for images and words, but not for the other categories.

Whilst this is an explicit assessment of gender conceptualisation, we plan to use these stimuli to further ascertain whether grammatical gender effects can be observed in more implicit designs, such as in Sato and Athanasopoulos (2018).

References

Duñabeitia, J.A., Baciero, A., Antoniou, K., Antoniou, M., Ataman, E., Baus, C., Ben-Shachar, M., Çağlar, O.C., Chromý, J., Comesaña, M. and Filip, M., 2022. The multilingual picture database. *Scientific data*, 9(1), p.431.

Sato, S. and Athanasopoulos, P., 2018. Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition*, 176, pp.220-231.

Samuel, S., Cole, G. and Eacott, M.J., 2019. Grammatical gender and linguistic relativity: A systematic review. *Psychonomic bulletin & review*, 26, pp.1767-1786.

Taylor, J.E., Rousselet, G.A., Scheepers, C. and Sereno, S.C., 2022. Rating norms should be calculated from cumulative link mixed effects models. *Behavior Research Methods*, pp.1-22.

Vankrunkelsven, H., Yang, Y., Brysbaert, M., De Deyne, S. and Storms, G., 2022. Semantic gender: Norms for 24,000 Dutch words and its role in word meaning. *Behavior Research Methods*, pp.1-13.