Effects of Grammatical Gender on Object Description

Arturs Semenuks[1] (asemenuk@ucsd.edu)
Webb Philips[1] (webb.philips@gmail.com)
Ioana Dalca[2] (ioanadalca@gmail.com)
Cora Kim[3] (cora.kim@fu-berlin.de)
Lera Boroditsky[1] (lera@ucsd.edu)

[1] Department of Cognitive Science, University of California, San Diego
[2] School of Electronic Engineering and Computer Science, Queen Mary University of London
[3] Brain Language Laboratory, Department of Philosophy and Humanities, Freie Universität Berlin

Abstract

Can grammatical gender influence how people conceptualize the referents of nouns? Using an implicit measure, we investigated whether such an effect could be found in a task where neither grammatical nor biological gender is highlighted. In the current study, conducted in English, speakers of French, German and Romanian with knowledge of English were asked to generate adjectives they associate with referents of nouns. Afterwards, the gender valence of the adjectives was measured. The results showed that participants generated more feminine adjectives for nouns with majority feminine translations compared to nouns with majority masculine translations. We found a stronger effect of grammatical gender for some semantic categories than for others. Significant effects of grammatical gender were present starting with the 2nd adjective generated by participants (effects were stronger for adjectives generated 2nd and 3rd by participants, as opposed to the 1st adjective).

Keywords: grammatical gender; conceptualization; categorization; semantic features; representation

Introduction

Can the structure of the language you speak affect how you perceive and conceptualize the world around you? Although the question has a long history (see Koerner (1992) for an overview), the last two decades have seen an especially large increase in the amount of empirical work examining and testing the predictions that the different answers to it imply. This has lead to an accumulation of results showing that there *are* cases where differences in seemingly non-linguistic aspects of cognition between speakers of different languages arise due to specific differences between the structures of those languages (see Wolff & Holmes (2011) for a survey). Nonetheless, the jury is still out on a number of questions aiming to flesh out the details of how, when and why linguistic differences can lead to differences in the operation of cognitive processes.

One question that is not conclusively resolved is what types of effects language could have on cognition. Research supports, for example, the existence of effects of language that play a large role in the operation of some specific cognitive domains. Frank et al. (2012) shows that number words are a cognitive tool that, when suppressed, impairs Western college educated people's ability to encode exact numerical values. Thus number vocabulary plays an important role in numerical cognition. Other work indicates that language also has some early effects on cognition. A body of research has shown that color vocabulary has an effect on categorical color perception (see Regier et al. (2010) for a review). There are also other types of effects and research indicates that cognitive differences between speakers of different languages arise due to the metaphors they use for talking about time, their frequency of use of a particular spatial frame of reference and knowledge of constructions for talking about mental states, among other differences (Wolff & Holmes, 2011).

One unresolved issue is whether language could have broad or pervasive effects. Are effects of language only restricted to specific domains (e.g. specific color distinctions, numbers, orientation in space)? Or can differences in the structure of languages lead to broadly distributed effects across a wide range of contexts? For this question, it is worth examining grammatical features that apply to say all nouns (e.g., gender) or all verbs (e.g., tense). In this paper we consider the role of grammatical gender in shaping the way people think about things that are named by nouns – a very broad potential scope of influence.

Grammatical Gender

Grammatical gender is a system of categorizing nouns into distinct classes, i.e. genders, which manifests itself via morphologically marking what gender the noun belongs to on some of the words that that noun is morphosyntactically related to (Corbett, 1991).

The forms that gender systems take vary. The ones that have garnered the interest in the research on the effects of language on cognition, however, usually share several properties. First, these systems have masculine and feminine genders, i.e. for nouns with human referents the assignment is fully semantic and dependent on biological sex, bar a small number of isolated exceptions in some languages. Second, the majority of nouns with non-human referents fall into the semantic residue, i.e. the category for which the assignment is based on the word's phonological or morphological, but not semantic properties. Third, a number of words in the semantic residue is assigned to either masculine or feminine gender. Languages with such systems provide the opportunity to ask: can the conceptualization of

a noun's referent accrue gendered properties because of the noun's grammatical gender assignment? If so, grammatical gender could exhibit a broad and pervasive effect on concept formation, since, by having a gender system, a language requires all nouns to be assigned to one of the gender categories. This could result in a broad range of concepts acquiring gendered properties that would be absent otherwise.

Why might we expect entities without biological sex to acquire stereotypically masculine or feminine properties as a result of being assigned to a grammatical gender category? As we mentioned, grammatical gender can be considered to be a system of noun categorization. Biological sex, in that case, is a highly diagnostic feature for determining the category assignment of nouns for humans. As the research on categorization indicates, diagnostic features better suited for differentiating between members and non-members of a category draw more attention to themselves (Tversky, 1977). In addition, category features could develop or become accentuated in response to the requirements that tasks involving categorization impose (Schyns & Rodet, 1997; Medin et al., 1993). This could mean that nouns belonging to a grammatical gender strongly associated with humans of a particular sex might acquire gendered properties just by virtue of being in those genders. Intuitions of native speakers of languages with grammatical gender for why nouns get assigned to their respective genders in their language supports this view (Basetti, 2011).

Prior Research and the Goals of the Current Study

A body of prior work supports the idea that grammatical gender influences conceptualization of objects and animals (Basetti & Nicoladis, 2016). For example, Konishi (1993) asked native speakers of German and Spanish to rate nouns on a variety of semantic differential scales in their native language. The words were either feminine in German and masculine in Spanish or vice versa. Subjects tended to give higher ratings on scales associated with potency for words that were masculine in their native language. In another set of experiments by Sera et al. (1994), Spanish and English speaking adults and children were asked to either give human names or assign male or female voices to pictures of objects. The results showed that older Spanish speakers tended to assign names and voices in these tasks in a grammatical gender-congruent fashion, which was not the case for English speakers or Spanish preschoolers. Phillips and Boroditsky (2003) investigated the effects of grammatical gender by teaching native speakers of English[1] an artificial language with grammatical gender and analyzed whether it influenced their performance on a similarity-rating task. The results indicated that participants rated pictures of humans and objects as more similar when they were in the same grammatical gender.

However, other studies suggest that there are limitations

on when and how grammatical gender could have an effect on conceptualization. For example, Vigliocco et al. (2005) found an effect of animal name grammatical gender on similarity in a triad similarity judgment task when the stimuli were names of animals, but not when the stimuli were pictures of animals. Furthermore, the effect was not found for either names or pictures of artifacts. In another experiment, Kousta et al. (2008) asked participants to name pictures that appeared at a random location on the screen, and afterwards analyzed cases where participants' produced an incorrect word. The subjects were either Italian or English monolinguals, or Italian-English bilinguals. The monolinguals responded in their native language, whereas the bilinguals participated in the experiment in both Italian and English. The results showed that the grammatical gender of the Italian noun describing the picture tended to affect the responses of Italian monolinguals - they preserved the grammatical gender of the target word almost twice as often as English monolinguals. The responses of Italian-English bilinguals followed a pattern similar to Italian monolingual pattern when they were tested in Italian. When they were tested in English, however, their pattern of errors was similar to the one exhibited by English monolinguals. Ramos and Roberson (2010) examined how speakers of Portuguese conceptualize inanimate objects in a series of experiments using different methodologies and compared their responses to the responses of an English-speaking group. The authors found that the grammatical gender of an object's name strongly influenced whether the Portuguese speakers assigned a male or a female voice to it, but had a smaller influence on the participant's responses in word similarity rating task, or in a triad similarity judgment task, especially when the stimuli were pictures, and not words.

These findings raise several issues. One alternative interpretation of the previously obtained effects is that they are by-products of task demands or of the particular stimuli that were used in the study. For example, when people are asked to give a name to an object, they might consciously or unconsciously understand that gendered information is important for this task. This seems especially likely when participants come from cultures where sets of male and female names have very little overlap, or when they are asked to give the name to something that can be very easily anthropomorphized, e.g. an animal. This then could lead to a pattern of responses that suggests an effect of grammatical gender on ingrained features of concepts, but that actually arose due to the demands of a particular context. Similar arguments could be made for other methodologies.

Another possibility is that some of the results arise because of the influence of information contained in the experimental stimuli. Given a set of words, the ones that are more similar phonologically or morphologically are more likely to prime each other (Kinoshita & Lupker, 2004), and words in the semantic residue are assigned to a grammatical gender specifically based on their phonology or morphology. When a speaker of Spanish is participating in a semantic differential scale task and reads a masculine noun

---

[1] English does have *pronominal* gender, but not a full-fledged *grammatical* gender system.

ending in "o", this weakly activates other words of similar morphology and phonology. That, supported by the experimental context, could lead to higher activation of frequently used words for human males and their properties. It, in turn, could lead to participants giving higher ratings to masculine words on scales associated with masculinity.

The experiment described below aims to address these issues. Three groups of non-native English speakers whose native languages have grammatical gender systems were asked to generate adjectives for English nouns. The adjectives were then scored on how feminine or masculine they were, and adjectives generated for nouns differing in grammatical gender between the languages were examined. This design addresses the issues discussed above in the following ways. The manipulation is implicit - there is no way for participants to guess that gender is of interest in the study – they are simply given a list of nouns and asked to generate adjectives that describe the referent. The adjectives are generated for English words, which eliminates the possibility of phonetic or morphological similarity between words driving the effect.

Methods

Participants

A total of 273 participants completed the study. Out of them, 99 were Romanian-English bilinguals (69 female; mean age = 24.03, SD = 6.83; mean English proficiency = 4.38/5, SD = 0.62), 90 were German-English bilinguals (53 female; mean age = 24.37, SD = 5.91; mean English proficiency = 4.08/5, SD = 0.80) and 84 were French-English bilinguals (51 female; mean age = 28.96, SD = 11.52; mean English proficiency = 4.38/5, SD = 0.64).

Materials

A total of 225 nouns served as stimuli. Out of them, 200 were the most frequently used nouns in the English language based on the data from the Corpus of Contemporary American English (Davies, 2008). These words were used for several reasons. First, it decreases the possibility of an effect arising due to an unconscious experimenter bias in the choice of items. Second, these words are often encountered in written and spoken language, so their meaning is unlikely to be misunderstood by the participants. Lastly, the procedure allows for an easier selection of items in future experiments using a similar methodology, but examining speakers of languages other than the ones used in this experiment.

We were also interested in examining whether the semantic category of the noun modulates what adjectives are generated for it, as previous work in the field has shown that the semantic category of the noun affects participants' responses (e.g. Vigliocco et al., 2005). Because the list described above lacked nouns with animal referents, 25 such words were added to the stimuli.

Procedure

All of the instructions in the experiment were given in English. At the very beginning participants were asked to verify their native language proficiency by translating an English sentence.

After this, the participants were informed that in the next part of the study they would see a list of nouns one by one, and that they would need to list the properties that they associate with their referents. For each noun the participants had to list three adjectives. Half of the participants were instructed to personify the nouns, i.e. imagine them as humans, and generate adjectives that would best suit those personifications, whereas the other half was not given personification instructions (and were instructed simply to produce adjectives). This was done in order to establish whether explicit invitations to personify (as when assigning names or voices) are necessary to induce effects of grammatical gender.

Next, the participants continued onto the adjective generation portion of the experiment. This part of the study consisted of 225 trials. During each trial the participants saw an English noun from the stimulus set and were asked to generate three adjectives for it. The participants were presented with all of the items shown in a random order.

After this, the participants were asked to translate all of the nouns they had encountered previously into their native language. At the end of the experiment, the participants filled in a questionnaire about their language background, education and residency.

Post-Processing

Grammatical Gender of Translations Some of the nouns were translated by the participants in multiple ways. For example, German participants translated English "difference" both as "*(der) Unterschied*" (masculine) and "*(die) Differenz*" (feminine).

Due to this, we first established the grammatical gender for each of the translations generated by the participants. Afterwards, the number of translations belonging to each of the grammatical genders was calculated for each noun in each language. We call the most common grammatical gender among the noun's translations in a particular language its *most common gender* (MCG) in that language.

Noun Semantic Categories In order to analyze the effect of noun's semantic category, each noun was categorized as being either an abstract noun, a noun denoting an animal, a body part or a concrete object by two coders.

Noun Selection After the most common grammatical genders of the nouns were established, only the nouns that had most common masculine gender in one of the experimental languages, and feminine in one of the others were left for analysis, leaving a total of 68 nouns. In cases where the noun's MCG in the third language was neuter, only adjectives generated by participants speaking the other

two languages were analyzed. 39 denoted abstract concepts, 13 - animals, 4 - body parts and 12 - other concrete objects.

Adjective Gender Valence Our next task was to operationalize how masculine or feminine the generated adjectives were. The participants generated a total of 45972 adjective tokens for the nouns remaining in the sample. Out of them, 36426 were also generated for nouns for humans (i) whose referent was clearly biologically male or female (e.g. "man", "woman") or (ii) that differed in translation depending on the sex of the referent (e.g. English "friend" = French "*ami*" (masc) or "*amie*" (fem)). For each of those adjectives, we calculated its token frequency among adjectives generated for males (a total of 12564) and adjectives generated for females (a total of 3681 tokens). After that, we subtracted the frequency of each adjective among adjectives generated for biological females from its frequency among those generated for biological males. We consider this to reflect adjectives' relative gender valence: relatively more positive values indicate more masculine connotations, relatively more negative values indicate more feminine connotations.

Results

The data were analyzed with linear mixed-effects models using the lme4 and multcomp packages for R (Bates et al., 2015; Hothorn et al., 2016). The most common gender (MCG) of the noun in the participant's native language, the noun's semantic category, personification condition, i.e. whether the participant was asked to personify the nouns, and the order of the adjective were modeled as fixed effects, whereas subject IDs and items were modeled as random effects.

As a manipulation check, we examined whether the biological gender of the referent for nouns denoting humans predicted what adjectives would be generated for it. As can be seen on figure 1, nouns for males on average received more masculine adjectives (M = 0.001) compared to nouns for females (M = -0.005). We compared mixed-effect models with gender valence of the adjective as the dependent variable, random intercept for subject and by-subject random slope for MCG of the noun, and a random intercept for noun. The model with MCG of the noun fit the data significantly better ($\chi 2 = 21.871$, $p < 0.001$).

Following that, we analyzed the data from nouns for non-humans. Only comparisons between adjectives generated for nouns that had either masculine or feminine MCG are reported. Besides the specified fixed effects, all of the models below contain random intercept for subject and by-subject random slope for MCG, as well as random intercept for noun and by-noun random slope for personification.

Adjectives generated by the participants for nouns with majority masculine translations had higher gender valence ratings (M = 0.0007) than adjectives generated for nouns with majority feminine translations (M = 0.0002), as figure 2 shows. To investigate whether the difference is significant, we compared mixed-effect models with

personification, noun's semantic category, order of the adjective, as well as all of their interactions as fixed effects. Comparison of the models with and without MCG as a fixed effect revealed that it significantly improves the fit of the model ($\chi 2 = 14.988$, $p < 0.01$).

The effect, however, could potentially arise due only to the adjectives generated by participants who were asked to personify the noun's referent. To examine that possibility, we compared mixed-effect models with and without an interaction between personification and MCG. The models contained all other possible main effects and interactions between them. Model comparison revealed no significant differences between them ($\chi 2 = 2.3089$, $p = 0.13$).
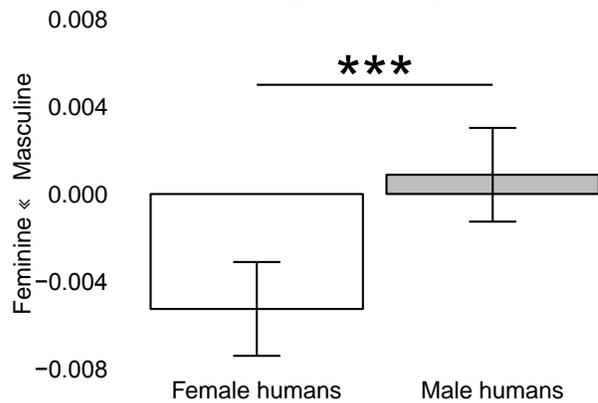

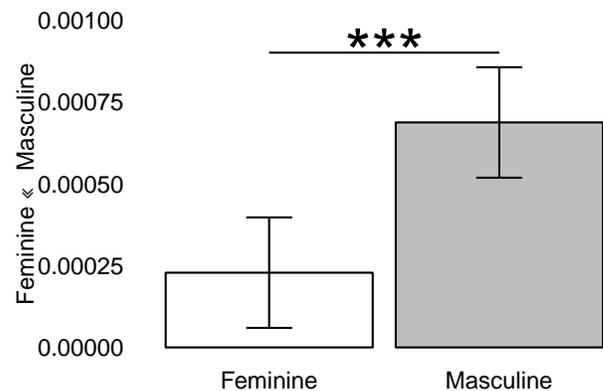Figure 1: Gender valence of adjectives for humans.

Figure 2: Gender valence of adjectives for non-humans.

Comparison of models with and without an interaction between noun category and MCG revealed a significant difference between them ($\chi 2 = 9.1268$, $p < 0.05$). Further multiple comparisons analysis showed a significant difference of the gender valences of adjectives for animal nouns (z-value = 3.713, $p < 0.01$) and a marginally significant difference for abstract nouns (z-value = 2.929, $p = 0.078$) in the expected gender-congruent directions, but no significant differences in gender valence for nouns for body parts or concrete objects (see fig. 3). We note however that the categories of body parts and concrete objects contained very low numbers of items (4 and 12 respectively). Further,

our design does not license us to generalize the results found for these specific items to their respective semantic categories as a whole.
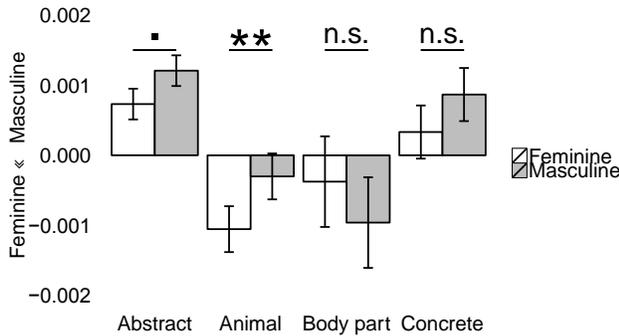

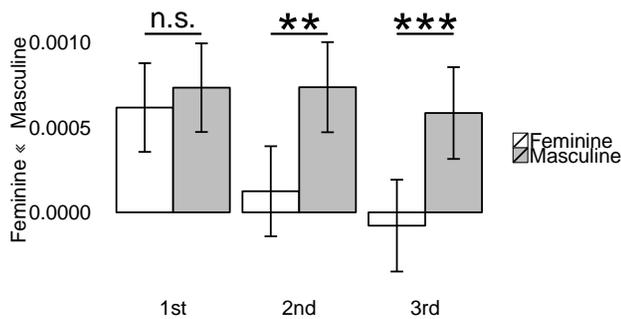Figure 3: Gender valence of adjectives belonging to different semantic categories.

Figure 4: Gender valence of first, second and third adjectives generated by participants for non-human nouns.

Last, we examined when the effects of grammatical gender emerge by looking separately at adjectives generated first, second, and third for each noun. Comparison of models with fixed effects of MCG, personification and noun category, as well as all of their interactions, but differing in having or not having a fixed effect of adjective order, revealed a highly significant difference in their fit ($\chi2 = 23.465$, $p < 0.001$). Additionally, a comparison of models with and without an interaction between MCG and adjective order was conducted. The models included all other possible main effects and interactions between them. The comparison revealed that the model with the interaction fit the data significantly better ($\chi2 = 10.413$, $p < 0.01$). Multiple comparisons analysis showed that adjectives generated first did not differ significantly depending on the MCG of the noun they were generated for (z-value = -0.163, p = 1.00). However, the difference in average gender valence was significant for second (z-value = 3.643, p < 0.01) and third (z-value = 4.054, p < 0.001) adjectives generated for nouns with different MCG.

Discussion

What do the results tell us? We see that the grammatical gender of a noun in a particular language influences what adjectives are generated for it: feminine nouns tend to elicit relatively more feminine adjectives compared to masculine nouns.

The effect appeared even though the task was conducted in English and did not invite participants to think about biological or grammatical gender. This removes two possibilities for why it arose. The first possibility that can be ruled out is that phonological or morphological properties of the nouns used in the experiment made the grammatical gender more salient to the participants, since English nouns do not contain in themselves any information related to grammatical gender. Thus operations situated solely at the lexical level of processing could not explain the effect. Second, gender was not highlighted in the experimental context for the participants who received no instructions to personify. Additionally, the selection procedure for the items minimized the possibility that the stimuli set would implicitly push the participants towards thinking about gender when participating in the experiment. This suggests that the results could be taken as evidence for the existence of effects of grammatical gender on how referents of nouns are conceptualized. The exact mechanism through which this effect takes place is a question for future work, but the current study provides some suggestions. Similarly to some prior work (e.g. Vigliocco et al., 2005) grammatical gender had the most effect on adjectives for nouns denoting animals. The interaction between the noun's semantic category and the gender of the noun found in the current study provides some support for the hypothesis that anthropomorphization of the noun referent is the mechanism through which the effect comes into being.

Cross-linguistically nouns often constitute the most frequently occurring word class in a language (Liang & Liu, 2013), and they are also used in everyday language for reference to humans, animals, objects, relations, categories and other types of entities. Because of this, the effect we found could have a broad and pervasive influence, affecting a wide range of processes relying on how referents of nouns are conceptualized.

The data also indicate that the effect did not emerge solely due to an invitation to personify the nouns that half of the participants received. Both participants who received instructions to personify and those who did not, showed effects of grammatical gender. It appears that an explicit invitation to personify (as when assigning names or voices in prior studies) is not necessary to induce effects of grammatical gender.

Of course, it is possible that participants personified the nouns even without being given the instruction to do so. For those who did not receive this instruction, the experiment did not make any suggestions for the participants to anthropomorphize the nouns' referents. If participants did engage in such unprompted personification, it seems likely that the effect would be observed outside of the context of this experimental task as well.

Last, we found that significant grammatical-gender

effects emerged starting with the second adjective participants generated for a given noun. This suggests that gender information is quite central in people's mental representations. There was however, no significant effect observed at the very first adjective.

This has the possibility of explaining why some studies using tasks where participants need to respond rapidly fail to observe effects of grammatical gender (e.g. Vigliocco et al., 2005; Kousta et al., 2008). The effect of grammatical gender is small in comparison to some other effects, for example the effect of cultural associations (Beller et al., 2015). It has also been suggested that more abstract pictures induce more schematic ways of conceptualizing what they depict and that verbal description of the less abstract pictures induces a similar effect (Holmes & Wolff, 2010). This allows one to hypothesize that the strength of the gender effect in combination with its temporal development could leave it unnoticeable in conditions where the participants need to respond rapidly and where other perceptual or conceptual features of the stimulus are highly activated due to task demands or stimuli properties.

Conclusion

The results obtained in this study support the view that grammatical gender affects object conceptualization. The effect was obtained in absence of any phonological or morphological aspects of the word carrying information about grammatical gender. Additionally, the effect was obtained without participants being invited to think about gender by any experimental instructions or demands (and participants could not have guessed that gender was of interest in the study). Finally, effects of grammatical gender emerged starting with the second adjective participants generated for a given noun. This suggests that gender information is quite central in people's mental representations (but did not emerge on the very first adjective). Furthermore, it is possible that the effect is quite pervasive, as it has the potential to affect anything that could be named by a noun.

References

Bassetti, B. (2011) The grammatical and conceptual gender of animals in second language users. In Cook, V.J. and Bassetti, B. (Eds.) *Language and Bilingual Cognition* (pp. 357-384). Hove, UK: Psychology Press.

Bassetti, B., & Nicoladis, E. (2016). Research on grammatical gender and thought in early and emergent bilinguals. *Int. journal of bilingualism: interdisciplinary studies of multilingual behaviour*, 20(1), 3-16.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Beller, S., Brattebø, K. F., Lavik, K. O., Reigstad, R. D., & Bender, A. (2015). Culture or language: what drives effects of grammatical gender?. *Cognitive Linguistics*, *26*(2), 331-359.

Corbett, G. (1991). *Gender. Cambridge textbooks in linguistics*. Cambridge: Cambridge.

Davies, M. (2008). *The Corpus of Contemporary American English: 425 million words, 1990 – present.* Available online at http://corpus.byu.edu/coca/.

Frank, M. C., Fedorenko, E., Lai, P., Saxe, R., & Gibson, E. (2012). Verbal interference suppresses exact numerical representation. *Cognitive psychology*, 64(1), 74-92.

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., Scheibe, S., & Hothorn, M. T. (2016). Package 'multcomp'. *Simultaneous inference in general parametric models. Project for Statistical Computing, Vienna, Austria.*

Holmes, K. J., & Wolff, P. (2010). Simulation from schematics: Dorsal stream processing and the perception of implied motion. In *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 2704-2709).

Kinoshita, S., & Lupker, S. J. (2004). *Masked priming: The state of the art*. Psychology Press.

Koerner, E. F. (1992). The Sapir-Whorf Hypothesis: A Preliminary History and a Bibliographical Essay. *Journal of Linguistic Anthropology*, 2(2), 173-198.

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of psycholinguistic research*, 22(5), 519-534.

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *JEP: Learning, Memory, and Cognition*, 34(4), 843.

Liang, J., & Liu, H. (2013). Noun Distribution in Natural Language. *Poznań Studies in Contemporary Linguistics* 49(4), 509-529

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for Similarity. *Psych. Review*, 100(2), 254-278.

Phillips, W., & Boroditsky, L. (2003) Can quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the 25th annual meeting of the Cognitive Science Society* (pp. 928-933).

Regier, T., Kay, P., Gilbert, A. L., & Ivry, R. B. (2010). Which Side Are You on, Anyway? In B. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience*. Oxford: Oxford University Press.

Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology Learning Memory and Cognition*, 23, 681-696.

Sera, M. D., Berge, C. A., & del Castillo Pintado, J. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development*, 9(3), 261-292.

Tversky, A. (1977). Features of similarity. *Psychological review*, 84(4), 327.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: implications for language learning and language use. *JEP: General*, 134(4), 501.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 253-265.