

## Is there an Effect of Gender-Fair Formulations in the German Language?

Johannes Beller<sup>1</sup>, Juella Kazazi

*Institute of Psychology, Technische Universität Braunschweig, Braunschweig, Germany*

Received 21.08.2012, accepted 14.12.2012, published online 31.12.2012

### 1 Why is this question important? 2 Why has this question not been fully answered yet?

In German the generic masculine refers to a generalizing denotation which is grammatically masculine. For example, in the sentence “Wissenschaftler führen Studien durch” (engl. “scientists conduct studies”) “scientists” is meant to be a generic masculine, because usually one implicitly refers to both “Wissenschaftler” (“male scientists”) and “Wissenschaftlerinnen” (“female scientists”) but uses only the masculine form “Wissenschaftler”. [1] Since the 1970s, the use of a generic masculine language, as a sexist one, has been highly debated and alternatives like a gender-fair language have been suggested, with the term gender-fair language referring to the use of formulations, which imply an equal linguistic treatment of men and women – such as in the sentence “Wissenschaftler und Wissenschaftlerinnen führen Studien durch”.

Although language use has somewhat changed over time, gender-fair language is not yet generally accepted and its use is far from being the norm. [2] The frequent use of the generic masculine orientation is said to reinforce the idea of “male as norm”, an idea that for example Hellinger and Bußmann (2001) believe has widespread negative effects: The selection of male expressions as the norm results in female invisibility. Take for example the following German sentence “Wissenschaftler führen Studien durch”. It is true that male scientists conduct studies; female scientists conduct studies as well but are however not explicitly named and thereby linguistically “invisible”. [3] Empirically this frequent use of masculine generics has been shown to cause a cognitive over-representation of males. Hellinger and Bußmann (2003), for example, found that, in German, “masculine terms automatically trigger expectations as to a most suitable representative of the noun, which in frequent cases is a male.” [4] When reading masculine forms, readers tend to associate them predominantly with men [5], which leads to a number of disadvantages for females. However, especially in administrative texts, laws and official forms, equal linguistic treatment is frequently awkward and complicated. Following this line of thought it is often argued that the generic masculine is preferable due to its brevity, clarity and comprehensibility. Thus it seems important to scientifically investigate the possible positive and negative effects of generic masculine versus gender-fair language: Is there an effect of gender-fair formulations on the subjective evaluation and cognitive processing of a text?

<sup>1</sup>johannesbeller@gmail.com

Some empirical studies explored the potential positive and negative effects of gender-fair formulations in the German language (for an overview see e.g. refs [6–9]). These studies, however, mostly used subjective data and are therefore susceptible to multiple methodological errors. [10] One notable study using objective data is the study of Braun et al. (2007). [11] They examined whether the use of gender-fair language has a negative effect on the cognitive processing and “quality” – defined as the subjective evaluation of comprehensibility, syntax and readability by participants in their paper – of a package leaflet. Comparing one generic masculine and two gender-fair text versions, the results of their experiment showed among other things that both female and male participants recalled similar amount of information. The authors conclude that all texts are processed similarly successfully and consequently that an argument of gender-fair language having negative impact on the subjective evaluation and cognitive processing of a text is without substance.

Although this study has many strong points – for example the large sample size, an experimental approach and the additional use of objective data – it should be mentioned that it shows some methodological problems and could be improved upon. For example, one inconsistency can be found in the description of the sample of participants. Furthermore there seem to be errors in the statistical results because the reported degrees of freedom sometimes exceed the reported sample size. Lastly some of the assumptions of their statistical analyses (MANOVA and ANOVA) seem violated, e.g. the homogeneity of variance assumption. [12] Considering these limitations the questions about the possible positive and negative effects of a gender-fair language use remain as yet to be answered. Thus there is a need to replicate and improve upon the methodologies of the former studies. In the following we will outline some approaches how this could be accomplished.

### 3 How could this question be answered in future studies?

Future studies should continue to explore whether the use of gender-fair formulations has an impact on subjective and objective variables. One important new line of research

could be initiated by the use of methods, where the movements of the reader's eye are being tracked. One could for example analyze the reading style of participants (i.e. cumulative gaze duration and total number of glances). Taken the aforementioned limitations into account, future studies could (a) use an objective eye tracking method in addition to subjective ratings and objective information recall, (b) measure information recall using an open question format and thus gaining more power to detect between-group differences and (c) implement parameter-free modern statistical methods to circumvent possible violations of statistical assumptions.[13] To provide a good example we conducted a small exemplary study, which we will describe in the following.

### Example study

The experiment was conducted on 16 participants (average of 20.9 years old, SD = 2.8, 14 female psychology students and 2 male psychology students). Participants read two different versions of a text (generic masculine vs. gender-fair formulations; eight participants in each version) of a fictitious package leaflet and were asked to answer questions regarding demographic information, subjective text quality (i.e. the subjective evaluation of syntax, linguistic quality, content, readability and clarity) and preference as well as an

open-format questionnaire concerning the recall of contents in the leaflet. In addition to the above mentioned measures the cumulative gaze duration, i.e. the reading length, and the total number of gazes were captured. This use of eye-tracking methods follows the rationale that the possible positive and negative effects of gender-fair language should be most obvious in eye-tracking data because eye movements capture the underlying cognitive processes very well. If, for example, gender-fair formulations make the text harder to process, participants will need to spend more time looking at the specific formulations, thus increasing the cumulative gaze duration. In general we followed the protocol of.[11] Subjective ratings, recall performance, preference and eye tracking data were used as dependent variables. We calculated a Permutation-MANOVA, which has minimal statistical assumptions.[14,15] Furthermore the psytabs package was used.[16] All analyses were conducted in R.[17] Further informations regarding our experiment might be obtained from the website of the first author.<sup>2</sup>

### Results

The descriptive statistics are depicted in table 1. Regarding subjective text quality (i.e. the subjective evaluation of syntax, linguistic quality, content, readability and clarity), information recall and preference larger values are better while for the eye tracking data smaller values are better.

Table 1: Means and standard deviations of the dependent variables for the two text versions. Standard deviations are depicted in parentheses.

	Dependent variable	Generic masculine	Gender-fair
<i>Subjective text quality</i>	Syntax	9.75 (1.83)	10.88 (3.00)
	Linguistic quality	9.38 (1.92)	10.25 (2.71)
	Content	10.12 (4.55)	10.88 (1.73)
	Readability	10.75 (2.82)	10.38 (2.26)
	Clarity	11.12 (2.70)	11 (1.93)
<i>Eye tracking</i>	Reading time per paragraph	13.36 (4.11)	13.89 (5.00)
	Reading time per formulation	0.59 (0.66)	1.51 (1.34)
	Gaze frequency per formulation	1.46 (1.15)	2.65 (1.60)
<i>Information recall</i>	Correct answers	3.62 (1.30)	2.25 (1.39)
<i>Preference</i>	Preference votes	14	1

Regarding subjective text quality the Permutation-MANOVA showed no significant group differences,  $F(1, 15) = 0.50, p = .731, f^2 = 0.03$ . There were also no significant differences for the eye tracking data for the whole text regarding reading length,  $F(1, 15) = 0.29, p = .628, f^2 = 0.02$ . If only the data for the gender-fair formulations are being considered, there were significant and large effects for the cumulative gaze duration,  $F(1,7) = 6.40, p = .018, f^2 = 1.07$ , as well as for the gaze frequency  $F(1,7) = 2.91, p = .036, f^2 = 0.49$ . [18] The recall performance approached significance, Wilcoxon test:  $p = .057$ . The participants favored the use of the generic masculine over gender-fair formulations, multinomial test:  $p = .018$ . The preference is being reasoned by the distracting impact of gender-fair

formulations.

### Discussion

Our sample study asked whether the use of gender-fair language could negatively or positively impact the subjective evaluation and cognitive processing of a text. Our analyses showed that the subjective ratings did not differ significantly between text versions. This seconds the results of [11]. Contrary to [11] we found a nearly significant difference between text versions in case of information recall, thereby not corroborating [11]. Participants recalled less information in the gender-fair condition. As the result only approached significance and our sample size was relatively small one should not over interpret the results. However,

<sup>2</sup><https://sites.google.com/site/johannesbeller>

it seems possible that because of the open-answer format our questions were harder to answer as opposed to the more easy format of [11] (multiple choice answers with four options only one being correct). This high difficulty level might have allowed a better differentiation of groups. To put it another way, it might not matter if the text is gender-fair if only easy questions are asked, but it matters when the questions are difficult.

## 4 General conclusions

We asked whether there is an effect of gender-fair formulations in the German language. We argued that although studies exist which tackle this, the question remains as yet to be definitely answered because the methodologies of most studies could be improved. Going beyond the current literature we suggested to (a) use objective eye tracking methods to determine whether the use of gender-fair formulations changes the way a text is read (b) use an open-answer format when measuring information recall and (c) calculate modern non-parametric statistical tests like the Permutation-MANOVA.

Additionally we conducted an exemplary study, which, albeit with a small sample size, incorporated our previously made suggestions. Discrepant findings were found. On the one hand, the differences regarding the whole text between text versions were – even descriptively – only marginal let alone statistically significant. On the other hand, if one focused on the formulations themselves, which are changed between texts, both number of gazes and the reading time differ strongly with large effect sizes in favor of the generic masculine text. Thus one could speculate that gender-fair formulations influence the reading speed and the number of gazes needed to process the information locally. In our sample the effect is yet not large enough to significantly impair the text globally. So the ratio of gender-fair formulations to “normal” text might be an important aspect. Future studies should pursue this new arc of questioning. Another line of argument could be that participants focus more attention on gender-fair formulations as compared to generic masculine formulations but have approximately the same gaze duration and gaze frequency on the whole text. They spend less time on the relevant information (e.g. the number of tablets to take) and more time on “unimportant information” like the gender-fair formulations. This could be an explanation for the poor recall performance of participants in the gender-fair text version.[19] This is also in line with the observation that the participants justified their preference over the generic masculine text version by reasoning that the gender-fair text version would have a distracting effect.

So do gender-fair formulations negatively or positively impact the subjective evaluation and cognitive processing of a text after all? This and other questions remain to be answered:

- Is the negative effect of gender-fair formulations on information recall using an open-answer format sub-

stantive? I.e. can it be replicated?

- Is the ratio of gender-fair formulations vs. other text important in diminishing or enhancing the possible negative effects of gender-fair formulations?
- Do participants really spend more time on gender-fair formulations and neglect the important information, e.g. how many tablets to take?

Incorporating our suggestions for future research and answering the above questions the possible positive and negative effects of gender-fair formulations could finally be investigated more holistically.

## 5 Acknowledgments

We thank Elke Muhrer, Klaus Reinprecht and Lukas Weißmeyer for their thoughtful comments on this paper.

- [1] Duden. *Die Grammatik*. (2009). Mannheim: Dudenverlag.
- [2] Mucchi-Faina, A. (2005). Visible or influential? Language reforms and gender (in)equality. *Social Science Information*, 44, 189–215.
- [3] Hellinger, M., & Bußmann, H. (2001). *Gender Across Languages: The Linguistic Representation of Women and Men (Volume 1)*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- [4] Hellinger, M., & Bußmann, H. (2003). *Gender Across Languages: The Linguistic Representation of Women and Men (Volume 3)*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- [5] Gygas, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23, 464–485.
- [6] Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name your favorite musician. Effects of masculine generics and of their alternatives in German. *Journal of Language and Social Psychology*, 20, 464–469.
- [7] Rothmund, J., & Christmann, U. (2002). Auf der Suche nach einem geschlechtergerechten Sprachgebrauch: Führt die Ersetzung des ‘generischen Maskulinums’ zu einer Beeinträchtigung von Textqualitäten? *Muttersprache*, 2, 115–136.
- [8] Verweken, D., & Hannover, B. (2012). Ambassadors of Gender Equality? How Use of Pair Forms versus Masculines as Generics Impacts Perception of the Speaker. *European Journal of Social Psychology*, 42, 754–762.
- [9] Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. In K. Fiedler (Ed.), *Social communication* (pp. 163–187). New York: Psychology Press.
- [10] Chan, D. (2009). So why ask me? Are self-report data really that bad? In C. E. Lance and R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 309–336). New York: Psychology Press.
- [11] Braun, F., Oelkers, S., Rogalski, K., Bosak, J., & Sczesny, S. (2007). “Aus Gründen der Verständlichkeit ...”: Der Einfluss generisch maskuliner und alternativer Personenbezeichnungen auf die kognitive Verarbeitung von Texten. *Psychologische Rundschau*, 58, 183–189.

- [12] Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics*. Boston: Pearson.
- [13] Beller, J. & Baier, D. (in press). Are the Effects Studied by Psychologists Really Linear and Homogeneous? *Europe's Journal of Psychology*.
- [14] Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26, 32–46.
- [15] Zapala, M. A., & N. J. Schork. (2006). Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. *Proceedings of the National Academy of Sciences*, 103, 19430–19435.
- [16] Beller, J., & Kliem, S. (2012). psytabs: Produce well-formatted tables for psychological research (Version 0.1) [Computer software]. Hannover, Germany. Retrieved from <http://cran.R-project.org/web/packages/psytabs/>
- [17] R Core Team (2012). R: A Language and Environment for Statistical Computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- [18] Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- [19] Baddeley, A. D. (1992). Working memory. *Science*, 255, 556–559.