Distribution of gender-inclusive orthographies in German press texts

Carolin Müller-Spitzer[1], Samira Ochs[1], Jan Oliver Rüdiger[1]

[1] *Leibniz-Institut für Deutsche Sprache Mannheim*

Gender-inclusive orthographies are highly debated in German-speaking society and scholarship. Gender symbols, and the gender star in particular, give rise to heated discussions (Müller-Spitzer, 2022; Rummel & Hetjens, 2021). Psycholinguistic investigations of gender-inclusive forms have become more established recently (e.g. Friedrich et al., 2021; Körner et al., 2022; Zacharski & Ferstl, 2023). However, studies on the actual use and distribution of gender symbols remain scarce (we are aware of only three such studies: Krome, 2021, 2022; Sökefeld, 2021) and rarely go beyond theoretical considerations (e.g. Kotthoff, 2017). Therefore, this paper presents a short-term diachronic corpus study of gender-inclusive orthographies in German-language press texts.

The study is based on a sub-corpus of the DeReKo (German Reference Corpus; Kupietz et al., 2010, 2018) which was specifically compiled for the purpose of this study. It comprises 16 German-language press publications from Germany, Austria and Switzerland. The sources were chosen because they a) are all fully available from 2015 to 2021, b) cover different formats (magazines, newspapers, online news) and genres/audiences, and c) are issued by different publishers. We used two search queries:

> (1) **Lexicon-based search**: From a comprehensive list of role nouns with systematic gender differentiation provided to us by the *Duden* (of the type *Manager/in, Besucher/in* etc.), we annotated the 142 most frequent role nouns for masculine and feminine inflectional forms and forms with gender symbols. These were then extracted from the sub-corpus.

> (2) **Regex search**: We extracted tokens with gender symbols (asterisk, capital I, colon, brackets, slash, underscore) via a full text search based on Regular Expression. With this search, we aimed at finding role nouns that were not included in the lexicon-based search. Masculine and feminine forms were not retrieved with this search. Furthermore, the data had to be cleaned manually because it is prone to contamination with false positives (e.g. missing spaces between tokens and syntactic colons).

The lexicon-based search allows for a comparison of masculine, feminine, and gender-inclusive forms. Masculine and gender-inclusive forms are strongly negatively correlated ($r(5)$ = -0.81, p = 0.02), indicating that gender-inclusive orthographies might gradually replace masculine generics. To test this, further annotations and analyses are necessary, as generic and specific uses of masculine forms cannot be distinguished automatically. We will give insights into the process of manually annotating for genericity from other parts of our project.

Both searches show that the gender star has been constantly on the rise since 2015, surpassing all other forms (cf. Figure 1 for the Regex search). It mostly occurs in plural forms (89%). Besides the asterisk, only the colon plays a noticeable role in the dataset. The capital I is on the decline; all other forms are negligible. An increase in the use of gender star forms can be seen for almost all sources. However, with the Regex search, we identified two main drivers of this development: the newspaper *taz* and the magazine *Brigitte*. Both show an extreme rise in gender star forms and considerably higher relative frequencies compared to other media. Besides that, *taz* is one of the first newspapers to start using the gender star already in 2015. Finally, we will show which lexemes and categories of role types contribute most to this constant rise (based on the categorisation of active, passive, prestige and population nouns in Bühlmann, 2002, p. 174).

Through our corpus study, we provide an overview of the use and distribution of gender-inclusive orthographies. By combining two search approaches, we can make differentiated observations about

both the nature and the sources of language change processes in press texts. This can contribute to a more informed discussion about gender-inclusive orthographies and their actual distribution in every-day media.

References

Bühlmann, R. (2002). Ehefrau Vreni haucht ihm ins Ohr... Untersuchung zur geschlechtergerechten Sprache und zur Darstellung von Frauen in Deutschschweizer Tageszeitungen. *Linguistik Online*, *11*(2), 163–187. https://doi.org/10.13092/lo.11.918

Friedrich, M. C. G., Drößler, V., Oberlehberg, N., & Heise, E. (2021). The Influence of the Gender Asterisk ("Gendersternchen") on Comprehensibility and Interest. *Frontiers in Psychology*, *12*, 760062. https://doi.org/10.3389/fpsyg.2021.760062

Körner, A., Abraham, B., Rummer, R., & Strack, F. (2022). Gender Representations Elicited by the Gender Star Form. *Journal of Language and Social Psychology*, 1–19. https://doi.org/10.1177/0261927X221080181

Kotthoff, H. (2017). Von Syrx, Sternchen, großem I und bedeutungsschweren Strichen. Über geschlechtergerechte Personenbezeichnungen in Texten und die Kreation eines schrägen Registers. *Sprache Und Geschlecht. Band 1: Sprachpolitiken Und Grammatik*, 91–115.

Krome, S. (2021). Gendern zwischen Sprachpolitik, orthografischer Norm, Sprach-und Schreibgebrauch. Bestandsaufnahme und orthografische Perspektiven zu einem umstrittenen Thema. *Sprachreport*, *37*, 22–29. https://doi.org/10.14618/SR-2-2021-KROM

Krome, S. (2022). Gendern in der Schule: Zwischen Sprachwandel und orthografischer Norm. *Mitteilungen des Deutschen Germanistenverbandes*, *69*(1), 86–110. https://doi.org/10.14220/mdge.2022.69.1.86

Kupietz, M., Belica, C., Keibel, H., & Witt, A. (2010). The German Reference Corpus DeReKo: A Primordial Sample for Linguistic Research. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 1848–1854.

Kupietz, M., Lüngen, H., Kamocki, P., & Witt, A. (2018). The German Reference Corpus DeReKo: New Developments – New Opportunities. *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 4353–4360.

Müller-Spitzer, C. (2022). Der Kampf ums Gendern. *Kursbuch 209*, 28–45.

Rummel, M., & Hetjens, D. (2021). Was sagt die Linguistik zum Gendern? Tendenzen in sechs Stereotypen. *Linguistische Werkstattberichte*. https://lingdrafts.hypotheses.org/2158

Sökefeld, C. (2021). Gender(un)gerechte Personenbezeichnungen: Derzeitiger Sprachgebrauch, Einflussfaktoren auf die Sprachwahl und diachrone Entwicklung. *Sprachwissenschaft*, *46*(1), 111–141.

Zacharski, L., & Ferstl, E. C. (2023). Gendered Representations of Person Referents Activated by the Nonbinary Gender Star in German: A Word-Picture Matching Task. *Discourse Processes*, 1–26. https://doi.org/10.1080/0163853X.2023.2199531
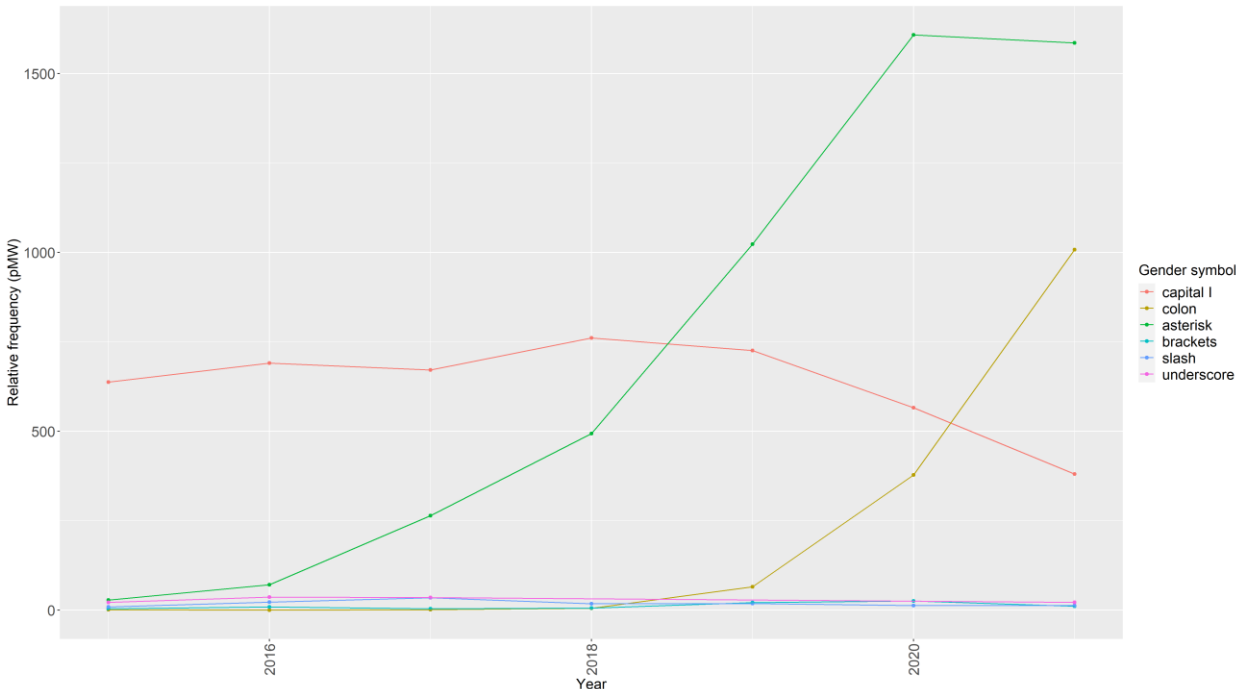
Figures

Figure 1: Use of gender symbols in German press texts from 2015-2021 (Regex search)