Samira Ochs and Jan Oliver Rüdiger

# Of stars and colons: A corpus-based analysis of gender-inclusive orthographies in German press texts

**Abstract:** New gender-inclusive orthographies such as the asterisk or the colon are widely debated in various German public spheres, including areas of academia. However, little is known about the frequency and distribution of these neographies. This study presents a microdiachronic analysis of binary (i.e., including only men and women) and non-binary (i.e., inclusive of all genders) variants of gender-inclusive orthographies from 2015 to 2023. Using the German Reference Corpus (DeReKo), we assembled a subcorpus of press texts from 15 sources, comprising approximately 1.2 billion tokens. Our corpus-driven approach reveals that the non binary variants asterisk (*Schüler*innen* 'pupils') and colon (*Schüler:innen*) have become dominant after the 2019 change of the German civil status law to encompass a third positive gender option, *divers*. Binary forms, especially the capital *I* (as in *SchülerInnen*), have been in decline since. The non-binary forms are unevenly distributed across sources and lexemes, a pattern further explored through a lexicon-based search. We examined 131 pre-selected personal nouns with systematic gender differentiation to compare binary and non-binary variants to regular masculine and feminine forms. The results show that masculine forms remain dominant, followed by feminine forms, with both maintaining stable frequencies over time. In contrast, binary and non-binary forms remain marginal. These quantitative baselines enhance the linguistic understanding of how sources, lexemes, and extralinguistic events shape the use of gender-inclusive language. Our study thus offers a foundation for more objective discussions on the subject.

**Keywords:** corpus analysis, gender-inclusive language, gender symbols, German Reference Corpus (DeReKo), language change, neographies, non-binary language, press texts

**Samira Ochs,** Leibniz Institute for the German Language (IDS), e-mail: ochs@ids-mannheim.de
**Jan Oliver Rüdiger,** Leibniz Institute for the German Language (IDS), e-mail: ruediger@ids-mannheim.de


https://doi.org/10.1515/9783111388694-003

# 1 Introduction

The efforts to develop more (gender-)inclusive linguistic forms in German date back to the emergence of feminist linguistics in the 1970s. During this time, scholars like Pusch (1979) and Trömel-Plötz (1978) began to critically examine the pervasive dominance of masculine forms and called for greater linguistic visibility of women. Among the resulting proposals – many of which were rooted in a binary understanding of gender – was the practice of explicitly including both masculine and feminine forms within a noun phrase, such as *Bürgerinnen und Bürger* ('female and male citizens') instead of using only the generically intended masculine *Bürger* ('[male] citizens').

However, concepts of social gender – its performativity and fluid nature – have evolved significantly since then (cf., for example, Butler 1988). Today, gender is commonly understood as a continuum that includes not only male and female but also non-binary, genderfluid, or agender identities, among others (Jourian 2015; Lev 2004; Thorne et al. 2019). This shift is also reflected on the legal level: In 2019, the German civil status law was amended to include the gender category *divers* for intersex, non-binary or genderfluid individuals[1]. While binary forms of person reference sufficed for feminist linguistic efforts, emerging queer realities call for linguistic forms that have the potential to encompass a wider spectrum – or even the entirety – of gender identities.

Binary as well as gender-neutral linguistic forms (e.g., *die Person* 'the person', *der Mensch* 'the human being') are widely accepted by the German population as strategies for gender-inclusive language, as evidenced by a recent poll conducted by the German broadcaster WDR (2023). Because these forms are part of the standard language system, they are often not even explicitly perceived as gender-inclusive language (Zacharski 2024: 230). In contrast, innovative linguistic forms that are intended to include all genders frequently interfere with standard German orthography – for instance, through the insertion of typographic characters within words –- and are therefore highly debated, both in the public sphere and within academia (for a comprehensive overview of the content of these debates, cf. Acke 2023: 47–51).

To date, gender-inclusive language has primarily been investigated from psycholinguistic perspectives, focusing on aspects such as comprehensibility (Braun et al. 2007; Friedrich et al. 2021; Pabst and Kollmayer 2023; Tross 2023) and men-

**1** https://www.antidiskriminierungsstelle.de/EN/about-discrimination/grounds-for-discrimination/gender-and-gender-identity/third-option/third-option-node.html, accessed: 25 March 2024. In Austria, a similar law allowing a third gender entry was introduced in 2019. Switzerland followed in 2022.

tal representations (Keith et al. 2022; Körner et al. 2022; Kurz and De Mulder 2023; Zacharski and Ferstl 2023). These efforts are complemented by sociolinguistic studies examining attitudes and opinions towards gender-inclusive language (Adler and Plewnia 2019; Jäckle 2022; Kotthoff 2023; Löhr 2021, 2022). Quantitative, corpus-based analyses of gender-inclusive German in written texts have recently become another focus of research (Link 2024; Müller-Spitzer et al. 2024c,b; Müller-Spitzer and Ochs 2023; Sökefeld 2021, forthcoming; Waldendorf 2023). Our study contributes to this emerging field by investigating the actual use and distribution of gender-inclusive orthographies in German press texts. We analyze data from 15 sources available in the German Reference Corpus (DeReKo), comprising a total of 1.2 billion tokens and covering the period from 2015 to 2023.

Our analysis is structured along two perspectives: First, we examine the frequency of various gender-inclusive orthographies relative to one another. We highlight changes from binary to non-binary usages, especially following the amendment of the German civil status law in 2019. Second, we conduct a detailed analysis of a preselected set of personal nouns, comparing gender-inclusive variants with regular masculine and feminine forms. By providing these quantitative baselines, we aim to enhance the linguistic understanding of how sources, lexemes, and extralinguistic events shape the use of gender-inclusive language. Additionally, we seek to contribute valuable insights to ongoing discussions about gender-inclusive language.

In Section 2, we briefly outline the grammatical gender system of German, the so-called *masculine generic*, and various strategies for gender-inclusive language. Section 3 introduces the study design, including a description of the corpus and the data. The results of our analyses are presented and discussed in Section 4, followed by concluding remarks in Section 5.[2]

# 2 Gender-inclusive German

## 2.1 Grammatical gender and the masculine generic

German is a grammatical gender language with three genders: masculine, feminine, and neuter. Gender assignment is largely arbitrary but can, in some cases – particularly for monosyllabic nouns – be predicted by morphophonological criteria (Hellinger and Bußmann 2003: 143; cf. Kupisch et al. 2022 for examples of further

**2** Datasets, Python scripts and Supplementary Material are available in the OSF at https://osf.io/9mzab, accessed: 20 February 2025.

predictable structures). Personal nouns, which refer to human beings, represent a unique nominal domain where lexical-semantic factors often influence grammatical gender assignment (Hellinger and Bußmann 2003: 146). In simple terms, this means that feminine forms are typically used to refer to women, while masculine forms refer to men – for instance, the word for 'mother' (*die Mutter*) is feminine, whereas the word for 'father' (*der Vater*) is masculine. Exceptions to this pattern are sometimes systematic, such as the derogatory use of feminine nouns to describe men who deviate from traditional gender norms (Kotthoff and Nübling 2018: 85).

In this context, the concept of the masculine generic is central. In many natural and grammatical gender languages (Hellinger and Bußmann 2001), the masculine form is used not only to refer to specific men but also to groups of people whose gender is either unknown or irrelevant, as well as to mixed-gender groups (Diewald 2018: 286). In German, only personal nouns with systematic gender differentiation are affected by the ambiguity of the masculine form – specifically, masculine personal nouns from which feminine forms can be derived. Feminine derivations are typically created by adding the suffix *–in* to the masculine base (Doleschal 1992), as in *Wissenschaftlerin* 'female scientist' derived from *Wissenschaftler* 'male scientist'. This suffix is highly productive and attaches to nearly all masculine derivation bases (Fleischer and Barz 2012: 236–237). The suffixation with *–in* transforms "grammatically masculine, semantically male human nouns into grammatically feminine, semantically female ones" (Stefanowitsch and Middeke 2023: 293). Thus, referential gender is specified through morphological means (Hellinger and Bußmann 2003: 152–153). Within such pairs, the masculine form serves a dual purpose: It can be used either gender-specifically or generically. For example, consider the noun phrase *ein Wissenschaftler und zwei Wissenschaftlerinnen* ('one male and two female scientists'): Here, the masculine form is used gender-specifically, contrasting with the feminine form and creating a semantic minimal pair. In contrast, in the phrase *Sieben Wissenschaftler diskutierten* ('seven scientists.MASC.PL were discussing'), the masculine form is ambiguous. It may refer exclusively to men or to a mixed-gender group, with the exact reference determined only by context (Müller-Spitzer et al. 2024c: 3). Other types of personal nouns, such as epicene or lexical gender nouns[3], are not subject to the concept of masculine generics.

The generic use of the masculine is often linked to androcentrism and the notion of 'male as norm' (Bailey et al. 2019), which is why it is frequently criticized in (queer-)feminist contexts as exclusive of non-male gender identities (Kotthoff and

**3** Epicene nouns do not specify referential gender via grammatical gender, i.e., they can be used to refer to people of any gender (e.g., *die Person*.FEM 'the person', *der Mensch*.MASC 'the human being'). Lexical gender nouns have referential gender encoded in their lexical meaning; grammatical gender usually aligns with this (e.g., *die Mutter*.FEM 'the mother' vs. *der Vater*.MASC 'the father').

Nübling 2024: 135–144). Therefore, the adequacy of the masculine generic as a form to denote all genders is a topic of ongoing controversy in both society and academia (for a selection of perspectives, see Müller-Spitzer 2022; Pusch 1984; Simon 2022; Trutkowski and Weiß 2023). Advocates of non-discriminatory language typically reject it as a truly gender-neutral way of person reference (e.g., Acke 2019; Hellinger and Bußmann 2003). In contrast, opponents of new gender-inclusive forms argue that the masculine generic is inherently gender-neutral (e.g., Eisenberg 2020; Meineke 2023; Trutkowski and Weiß 2023). However, a growing body of psycholinguistic studies demonstrates that the masculine generic consistently carries a male bias (e.g., Glim et al. 2023; Gygax et al. 2008; Körner et al. 2022; Zacharski and Ferstl 2023). As Glim et al. put it, it "does not represent men and women equally well" (2023: 2) but instead favors the (mental) representation of men. This is supported by evidence from computational linguistics and discriminative learning, which shows that masculine generics are semantically very close to gender-specific masculines in actual language use (Schmitz et al. 2023; Schmitz 2024). It is therefore not only questionable whether women are adequately represented by the masculine generic, but also whether it can account for identities beyond the gender binary. Such limitations are at the heart of current debates on gender-inclusive language.

## 2.2 Strategies of gender-inclusive language in German

As the adequacy of the masculine generic as a gender-neutral form is increasingly questioned, alternative, more inclusive ways of person reference are being explored. In German, two broad categories of gender-inclusive language can be distinguished: implicit and explicit strategies.

Implicit strategies, sometimes referred to as gender-neutral strategies, aim to make gender linguistically 'invisible'. This is achieved through neutralizing forms such as epicenes, collective nouns, and the nominal use of plural participles and adjectives (Eisenberg 2020). Paraphrasing with pronouns, relative clauses, or passive constructions also falls into this category. These strategies are generally unobtrusive – indeed, a recent poll by the German broadcaster WDR (2023) found that most respondents held positive attitudes toward gender-neutral formulations.

Explicit strategies, by contrast, make gender visible in language, either in a binary or non-binary way. Binary forms are particularly common in political and journalistic context. Fully-fledged pair forms such as *Lehrerinnen und Lehrer* ('female and male teachers') or *Bürgerinnen und Bürger* ('female and male citizens') are widely used there (Bast et al. 2024; Müller-Spitzer et al. 2022, 2024a; Rosar 2022; Truan 2019). Shortened versions of the pair form include those using a forward slash (*Lehrer/Lehrerinnen*, *Lehrer/innen*, *Lehrer/-innen*), parentheses (*Lehrer(innen)*), or

the capitalised *I* (called *Binnen-I* 'capital *I*', as in *LeherInnen*). Notably, fully-fledged binary pair forms like *Lehrerinnen und Lehrer* have the highest acceptance rates in the WDR poll and are not perceived as intrusive ways to represent gender in language (Müller-Spitzer et al. 2022; Zacharski 2024). This suggests that feminist efforts to increase the visibility of women in language and normalize such visibility have been relatively successful.

However, current research indicates that binary forms are seen as reinforcing the gender binary and are neither representative of nor accepted by the queer community (Motschenbacher 2013; Siegenthaler 2024). In fact, when asked for acceptability judgments, queer respondents rate masculine generics higher than pair forms or the capital *I* (Löhr 2021, 2022). This underscores the limitations of binary strategies in addressing the linguistic needs of non-binary and gender-diverse individuals. Gender is increasingly understood as a spectrum rather than a purely binary concept. This change is also evident in the legal sphere, as Germany's civil status law was revised to introduce the gender category *divers* in 2019, providing recognition for gender-nonconforming individuals. As a result of these societal changes, gender-queer and other marginalized communities are seeking linguistic forms that represent this expanded understanding of gender, leading to the emergence of new gender-inclusive forms.

Similar to shortened binary forms like the capital *I*, forward slashes, or parentheses, these new forms operate on the word-internal level: A typographic character is inserted between the masculine base and the feminine suffix, thereby creating a new suffix (Völkening 2022, and in this volume). The characters currently in use in German include the asterisk (*Lehrer*innen*)[4], the colon (*Lehrer:innen*), and the underscore (*Lehrer_innen*). It is possible that additional symbols, such as the mediopoint (*Lehrer·innen*) already used in French, could be adopted in the future (Diewald and Steinhauer 2020: 127). However, our analysis focuses on the characters already in use today.[5]

**4** The term *asterisk* is used in this chapter, while Völkening and Schmitz et al. in their chapters refer to the same concept as *gender star*.

**5** We also do not consider proposals falling under so-called *exit gender* strategies, which advocate for more radical linguistic innovations. These include introducing entirely new suffixes, such as *–x* or *–ecs*, to create gender-neutral personal nouns like *Lehrx* instead of the gender-marked forms *Lehrer*.MASC and *Lehrerin*.FEM (AG Feministisch Sprachhandeln 2014: 22). Another example is the systematic use of the neuter for gender-neutral designations, combined with the elimination of the feminine suffix *–in*. This would result in paradigms such as *der Lehrer*.MASC.SG – *die Lehrer*.FEM.SG – *das Lehrer*.NEUT.SG (Pusch 1984). However, the feasibility of such innovations is questionable, as they would require profound changes to grammatical structures (Kotthoff and Nübling 2018: 221). Moreover, the lack of corresponding corpus annotations makes it impossible to identify and an-

The symbolic meanings of these characters vary, with the asterisk and underscore carrying the most interpretive weight (Kotthoff 2017: 11). The asterisk, for instance, might function as a placeholder (akin to its use in programming languages), or symbolize gender diversity through its radiating form. The underscore, by contrast, may signify a 'gap' to be filled by new gender identities or the fluidity between male and female. The colon, while lacking this level of symbolic association, shares with the others a common goal: to represent gender identities beyond the binary (Diewald and Steinhauer 2020; Friedrich et al. 2021; Genderleicht 2024; Körner et al. 2022). Both psycholinguistic (Zacharski and Ferstl 2023) and sociolinguistic research (Löhr 2021, 2022) suggests that forms with gender symbols are more inclusive – not only in terms of mental representations but also in their acceptability among members of queer communities.

To illustrate the differences between all strategies, consider the sentence 'All teachers were in school' and its various realizations in Examples (1-a) to (1-f). It is important to note, however, that not all masculine generics can be replaced by every option shown here. In particular, the availability of epicenes is not always guaranteed:

(1) a. Masculine generic:
***Alle Lehrer*** *waren an der Schule.*
b. Implicit strategy (epicene):
***Alle Lehrkräfte*** *waren an der Schule.*
c. Implicit strategy (paraphrase):
***Alle, die unterrichten,*** *waren an der Schule.*
d. Explicit strategy (binary):
***Alle Lehrerinnen und Lehrer*** *waren an der Schule.*
e. Explicit strategy (shortened binary):
***Alle Lehrer/-innen*** *waren an der Schule.*
f. Explicit strategy (non-binary):
***Alle Lehrer*innen*** *waren an der Schule.*

Gender symbols such as in example (1-f) (also called *neographies*) represent an overt strategy of gender-inclusive language that intentionally challenges conventional orthography. This has made them a subject of debate within the *Rat für Deutsche Rechtschreibung* ('Council for German Orthography'), the primary international body regulating Standard German orthography. In its most recent resolution, the Council classified gender symbols as special characters, akin to the

alyze such forms systematically in corpus-linguistic studies, particularly in standard press texts. Consequently, these strategies fall outside the methodological scope of the present study.

paragraph (§) or percent (%) signs. This means they are not considered part of core orthography. Opponents of gender symbols often use this classification to argue that they constitute "incorrect" or "poor" German (Eisenberg 2022; Zifonun 2021). This reasoning has spurred a rise in petitions to prohibit gender symbols in certain regions, and restrictive regulations have been implemented in several German federal states (see Müller-Spitzer et al. 2024a for an overview of states and petitions). A common argument supporting such bans is that gender symbols render texts cumbersome, unreadable, and overly lengthy (Kurfer 2024; Meuleneers 2024; Pfalzgraf 2024). However, empirical research has repeatedly refuted these claims (Blake and Klimmt 2010; Friedrich et al. 2021, 2024; Pabst and Kollmayer 2023). Moreover, such criticisms would only be valid if the use of gender-inclusive forms caused substantial changes to text structure or content, which a recent study by Müller-Spitzer et al. (2024c) suggests not to be the case.

A key aspect often overlooked in the debate is the actual frequency of gender symbols: How often do we encounter them in everyday written language? Are they sufficiently frequent and widely distributed to warrant claims of being a significant intrusion into the language? Where are they most commonly used, and how has their usage evolved over time? Although these questions are important, the research landscape is still limited. Some studies have examined specific lexical items and their various realizations in corpora (Adler and Hansen 2020; Bast et al. 2024; Krome 2020, 2021). Other studies focus on highly specialized sources, such as university documents (Acke 2019; for French: Burnett and Pozniak 2021) or city websites (Müller-Spitzer et al. 2024b; Müller-Spitzer and Ochs 2023).

To date, only three larger-scale corpus studies on gender-inclusive language in German exist: Sökefeld (2021) compares press texts and blog posts, annotating all possible variants of personal nouns in her data. Her findings show that while the asterisk is increasing in frequency over time, masculine generics and gender-neutral forms still dominate. Waldendorf (2023) extracts different variants of gender-inclusive language from five German newspapers using computer-linguistic methods. She observes a rise in binary pair forms as well as non-binary variants such as the asterisk and colon, linking these trends to the political orientation of the newspapers. Most recently, Link's study (2024) examines the use of gender-inclusive language from a contrastive perspective. Analyzing six conservative and left-liberal newspapers from Germany, Austria, and Switzerland, she finds that Austria employs gender-inclusive forms significantly more frequently than the other two countries. The study also identifies a marked increase in the use across all three countries between 2017 and 2021, with trends diverging thereafter. Contrary to previous findings, Link states that the political orientation of the newspapers had no significant influence on the use of gender-inclusive language.

Our study builds on and extends these foundational efforts by analyzing gender-inclusive language across a considerably broader textual basis. In contrast to these previous studies, which often focus on broader patterns of gender-inclusive language use, we exclusively examine gender symbols. This means that pair forms like *Lehrerinnen und Lehrer* and gender-neutral nouns like *Lehrkräfte* are deliberately excluded from our analysis. Instead, the first part of our study systematically compares the relative frequencies of binary forms (capital *I*, parentheses, forward slash) and non-binary forms (asterisk, colon, underscore), providing a nuanced perspective on their use and distribution. We also focus on how the use of gender symbols changed after the 2019 amendment to the German civil status law, adding an extralinguistic dimension to our analysis that has not been considered in previous studies.

In the second part, we introduce an innovative approach by analyzing gender-inclusive orthographies at the lexical level. Using a preselected set of personal nouns, we compare the frequency of their gender-inclusive realizations to that of regular inflectional forms (i.e., masculine and feminine). Unlike previous studies, which often rely on anecdotal evidence or focus on single lexemes, our approach enables a comprehensive, data-driven analysis of the relationship between regular forms and gender symbols for a broader set of terms. Our diachronic perspective (2015–2023) allows us to trace remarkable changes in usage over time and offers insights into lexical dynamics that have so far been little discussed in the field.

## 3 Corpus and data

Our analysis employs two complementary corpus approaches: First, a corpus-driven investigation extracts all gender-inclusive orthographies (Section 4.1), and second, a lexicon-based search is conducted with a pre-selected set of lexemes (Section 4.1.1). Combining these two methods enables us to a) track the development of gender-inclusive orthographies across all personal nouns in the corpus, and b) analyze the full inflectional paradigm for a set of lexemes, allowing us to compare gender-inclusive to regular inflectional forms.

The corpus is a specifically compiled subcorpus of the German Reference Corpus (DeReKo; Kupietz et al. 2018, 2010). It comprises 15 different press sources, including newspapers, magazines, and the German Press Agency (*dpa*), from Germany, Switzerland, and Austria. These sources are all published on a national level and represent a variety of publishing houses, target different audiences, and reflect diverse political orientations, thus ensuring a broad spectrum of German-language

media.[6] For our analysis, we focused on texts published between 2015 and 2023, resulting in a total of approximately 2.3 million documents with 1.2 billion tokens.[7] We decided on this time span as it covers the years just before the amendment of the German civil status law up until the most recent year available in DeReKo.

We used the raw data from DeReKo (IDS-I5/XCES ,Lüngen and Sperberg-McQueen 2012) and analyzed it using the CorpusExplorer software (Rüdiger 2023). Frequency lists and N-gram tables[8] were generated: Words with gender symbols were broken down into bi- or trigrams. For example, the word *Lehrer*innen* would be split into 1) *Lehrer* 2) *, and 3) *innen*, allowing these components to be extracted directly from the N-Gram table. Forms with a capital *I* were retrieved as bigrams. For example, *LehrerIn* would be represented as 1) *Lehrer*, and 2) *In*. The data was further filtered using a separate Python script.[9] Following this, the corpus-driven data was manually reviewed and cleaned, especially to eliminate false positives. 10,187 types and 24,438 tokens were removed, which accounts for 31.51% of the originally retrieved types and 8.86% of the tokens. This included proper names like *LinkedIn* (false positive for the capital *I*) or trigrams in which punctuation marks were followed by the preposition *in* (e.g., *Berlin : In*). After data cleansing, a total of 22,600 types and 250,730 tokens remain in the corpus-driven dataset. Manual cleansing was not necessary for the lexicon-based search, which involved a pre-selected set of search terms, ensuring no false positives were retrieved.

While it is also possible to use regular expressions (Regex) for data extraction (Sökefeld forthcoming; Waldendorf 2023), we opted against this approach after testing it in a small preliminary study. We found that the tokenization of gender symbols in DeReKo is inconsistent. While forms such as the capital *I* are rather unproblematic, punctuation characters like the slash and colon are inconsistently to-

**6** Table S1 in the Supplementary Material gives an overview of all sources and their category, main topics, target audience and political orientation. This information was retrieved either from the website of the source, specific media information platforms like AdAlliance, from the Wikipedia, or from the *Institut für Medien- und Kommunikationspolitik* (IfM).

**7** Tables S2 and S3 in the Supplementary Material show the amounts of documents and tokens per source and year. Unfortunately, the zwi16 corpus (*ZEIT Wissen* from 2016) is not available in DeReKo.

**8** N-grams are contiguous sequences of n items (typically words or characters) from a given text or speech corpus. For example, a bigram consists of two consecutive words, a trigram consists of three, and so on. N-gram tables provide a way to count and analyze the frequency of these sequences within a corpus, which helps in identifying patterns and structures in the data.

**9** The scripts (annotated Jupyter notebooks) are provided in the OSF. We publish a simple version of the code which works well with smaller data. We also publish the version we used, which employs *Ray*. *Ray* is a Python package that enables a strong parallelization of tasks and performs well in big data scenarios. The datasets for both the corpus-driven and the lexicon-based approach can also be downloaded from the repository.

kenized. While it is possible to account for this with Regex, it increases complexity and leads to a high rate of false positives, particularly for non-binary forms. In contrast, N-grams offer a clearer, simpler solution that avoids these issues, while providing a computationally more efficient and easily parallelizable solution.

# 4 Results and discussion

## 4.1 Corpus-driven approach

To capture all personal nouns with gender-inclusive orthographies in our corpus, we conducted a corpus-driven search that covered both binary (capital *I*, parentheses, slash) and non-binary (asterisk, colon, underscore) strategies. This approach does not allow for comparisons with their corresponding regular inflectional forms (i.e., masculine and feminine forms), as the basic population of nouns detected by our search remains unknown. This limitation applies to personal nouns in general. In practical terms, this means we cannot determine whether gender-inclusive forms are becoming more frequent in relation to the unknown baseline population, or if the entire category of personal nouns is simply growing. Nevertheless, if we assume that the proportion of personal nouns in standard press texts does not change significantly over time, we can reasonably infer that gender-inclusive forms are indeed increasing in frequency. So far, only two studies have addressed the question of a quantitative baseline for personal nouns, both estimating their proportion at approximately 3% in their respective corpora (Müller-Spitzer et al. 2024c; Sökefeld et al. 2023). Large-scale diachronic studies on the frequency of personal nouns have yet to be conducted. Keeping this limitation in mind, our corpus-driven approach offers an overview of the development and source-specific distribution of various gender inclusive strategies.

### 4.1.1 Development over time

Figure 1 shows the development of all six gender-inclusive orthographies from 2015 to 2023 across all sources. The vertical line marks the year 2019, when the German civil status law was amended to include a third positive gender option, *divers*. We regard this legal change as an extralinguistic event that could impact the use of gender-inclusive language, particularly with respect to non-binary forms. From 2015 to 2018, the binary capital *I* was the most common form, but has been steadily decreasing since. Parentheses remain infrequent throughout the years, and slashes

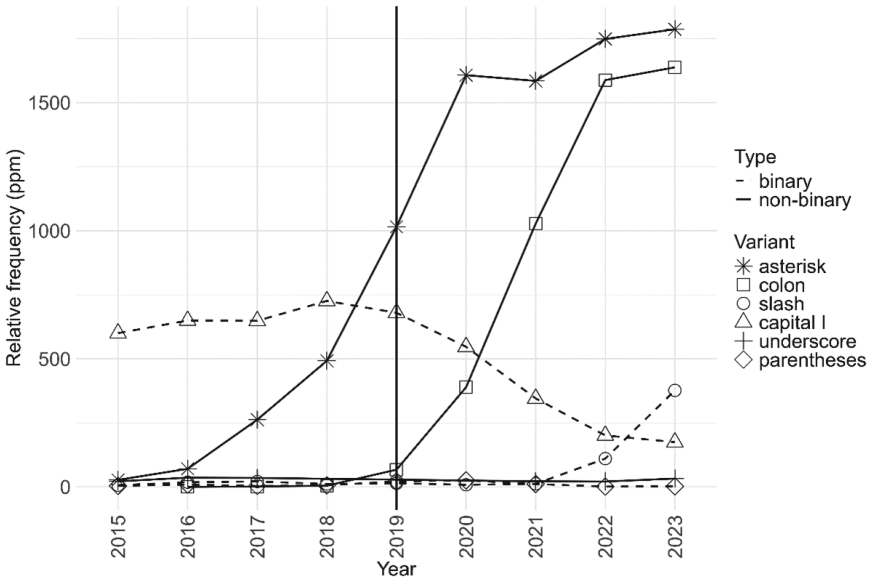


**Fig. 1:** Relative frequencies of gender-inclusive orthographies per year.

were rare before experiencing an increase after 2022, which we explain in the following discussion. Currently, the non-binary asterisk is dominant and continues to rise, although a flattening of the curve is observable in 2023. In 2022 and 2023, the colon nearly matches the asterisk, but still remains less frequent. This finding contrasts with Link's (2024: 10–11) corpus data, in which the colon is already more frequent than the asterisk, highlighting how the choice of textual basis can lead to different interpretations of the same phenomenon. The third non-binary option, the underscore, plays almost no role in our corpus.

To investigate whether the use of gender-inclusive orthographies differs significantly before and after the change in the civil status law, the data was split into two groups: one before 2019 and one from 2019 onward. A linear regression model was fitted to predict the relative frequency based on the interaction between the variables VARIANT and TIME GROUP. Since the relative frequency is not normally distributed, it was logarithmically transformed for the model. The asterisk was chosen as the reference level. The model shows a strong fit, with an R-squared value of 0.8503, meaning it explains approximately 85% of the variance in the data. The overall model is highly significant, as indicated by the F-statistic (19.63) and its p-value ($p < 0.001$).

The main effects of the different variants, which represent their impact before 2019 relative to the reference level (asterisk), are significant for all forms, except for

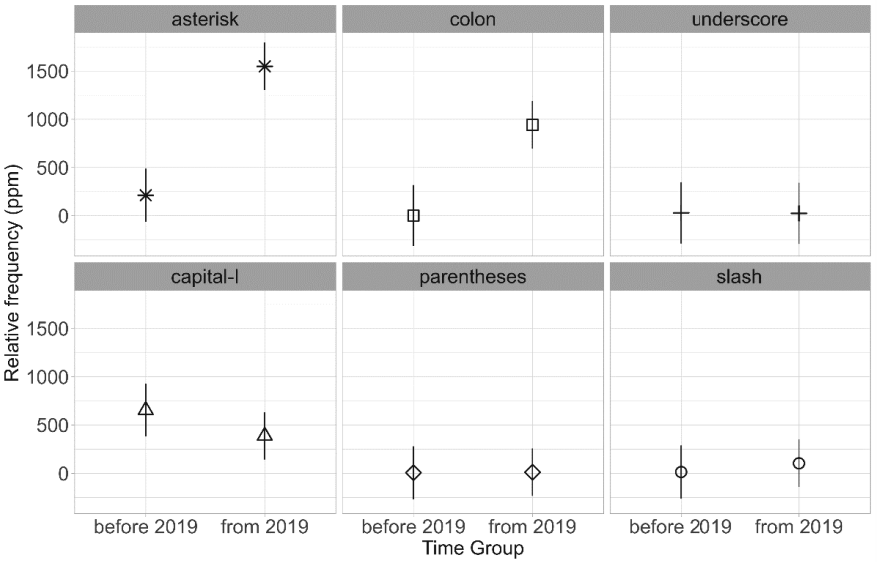


**Fig. 2:** Interaction plot for the linear regression model of relative frequency (per million words) predicted by the interaction of time group and variant. The top row shows the non-binary variants, while the bottom row displays the binary variants.

the underscore. The negative coefficients for colon (−4.8652), parentheses (−3.1387), and slash (−2.2517) indicate that they were significantly less frequent ($p < 0.01$) compared to the asterisk before 2019, while the capital *I* (1.6494) was significantly more common ($p < 0.05$). The main effect for TIME GROUP is also highly significant ($p < 0.001$), meaning that the reference value asterisk is approximately 12 times more frequent after 2019 than before.

The interaction terms illustrate how the influence of the variants shifted after 2019. The coefficient for the colon (3.9154) reflects a significant increase ($p < 0.001$) in frequency after 2019 relative to the asterisk. In contrast, the capital *I* experienced a substantial decrease in frequency ($p < 0.01$) after 2019, with its relative frequency dropping by approximately 96%, as indicated by the negative coefficient of −3.1474. The underscore and parentheses also exhibited a significant decline in frequency after 2019 ($p < 0.05$), whereas the slash remained largely unchanged ($p = 0.13$). Figure 2 visualizes these effects.

This analysis confirms that the non-binary variants asterisk and colon increased significantly after the change in the civil status law, underscoring the role of linguistic practices in reflecting societal change. In contrast, the underscore has remained largely irrelevant for representing non-binarity in our corpus, both be-

fore and after 2019. Similarly, parentheses and slash, which could theoretically function as binary markers, play no significant role in either time group. Before 2019, the binary capital *I* was the most common form, but its usage declined sharply afterward.

This shift suggests that non-binary neographies fulfil a linguistic need that binary forms cannot address. Binary forms lack the same socio-indexical significance (Kotthoff 2023: 209) and, at least theoretically, can always be replaced by pair forms. Waldendorf's data (2023: 8, Figure 2) confirm that fully-fledged binary pair forms (e.g., *Autorinnen und Autoren*) were the most frequent strategy in 2021. This indicates that binary strategies have not disappeared, but have predominantly shifted to pair forms rather than relying on word-internal strategies. Non-binary representations, by contrast, are currently limited to word-internal neographies. This distinction may signal an emerging division between binary strategies realized at the noun-phrase level and non-binary strategies confined to word-internal constructions.

### 4.1.2 Source-wise distribution

Next, we examine the use of gender symbols at the source level. Figure 3 shows the distribution of binary and non-binary variants across the 15 sources, sorted by their maximum frequency on the y-axis. The sources are arranged from top left to bottom right, starting with *taz* (y.max = 2,041 ppm) in the top left corner and ending with *Börsen-Zeitung* (y.max = 2.99 ppm) in the bottom right corner. From this, it is clear that the sources arrange on a spectrum, with *taz* being the most prominent driver of non-binary neographies. *Taz* is an alternative-left daily newspaper critical of societal structures (taz 2025). It contributes 94.38% of all tokens with non-binary neographies, aligning with research suggesting that the use of gender-inclusive language is often linked to political positioning (Bast et al. 2024; Jäckle 2022; Kotthoff 2017). Notably, *taz* was also the first newspaper to adopt the colon in 2016, with other sources following only in 2020. It can thus be considered a 'pioneer' in the use of non-binary gender-inclusive forms and serves as a reference point for what is maximally possible in terms of the frequency of gender symbols in journalistic texts.

However, the inclusion of *taz* raises important methodological considerations. Its exceptional contribution to non-binary neographies demonstrates its unique position, but excluding such a source could lead to a biased corpus selection. For instance, excluding sources based on the frequency of specific elements – whether high, as in *taz*, or low, as in *NZZ* or *Gala* – would raise significant questions about where to set thresholds for inclusion or exclusion. Would sources like *Brigitte* or

*Couch*, which also contribute a large number of non-binary tokens, then also need to be excluded? Such decisions risk introducing arbitrary biases into the corpus selection and make it challenging to define a generalizable middle.

The spectrum-like arrangement of sources underscores the need for separate analysis of individual sources to account for their unique contributions and contextual differences. For example, a comparison of the regression model presented in 4.1.1 with and without *taz* shows that while *taz* modulates certain effects – such as the prominence of non-binary neographies and the significance of interaction terms – significant changes remain observable across the remaining 14 sources for all variants in interaction with time. This indicates that *taz* is not the sole driver of linguistic variation within the corpus but rather a major contributor whose influence should be contextualized rather than overstated or dismissed. Link (2024: 4–13), for instance, deliberately excluded *taz* from her study comparing gender-inclusive language use in Germany, Austria, and Switzerland. Her rationale was that *taz* lacks equivalents in the other two countries and could distort her analysis of Germany. While this approach avoids over-representation, it also limits the conclusions she draws, such as her claim that political orientation has no significant influence on the use of gender-inclusive language (Link 2024: 1) – an assertion that might be questioned given that the most politically positioned source, *taz*, was not included. In summary, while *taz* undeniably shapes the upper bound of non-binary neographies, its inclusion highlights both its unique role as a driver of linguistic innovation and the methodological challenges in balancing representativeness and variability in corpus studies on gender-inclusive language.

The analysis also reveals that *taz* experiences a considerable decrease in binary forms over the years, indicating a substitution process for non-binary variants beginning after the change of the German civil status law. When only binary options were available, *taz* made extensive use of them. However, as soon as non-binary options became available and politically relevant, these began to be employed instead. As a result, *taz* stands out as the source with the strongest affinity for using word-internal strategies, regardless of whether they are binary or non-binary. This contrasts with *Brigitte*, Germany's most widely read women's magazine.[10] There, non-binary forms have increased significantly since 2019, while binary forms were not employed in meaningful numbers before non-binary options became available. *Brigitte* may therefore represent another key driver of change in the adoption of non-binary variants. Similarly, from 2020 onwards, the fashion and lifestyle magazine *Couch* also saw a noticeable rise in non-binary forms. This highlights how genre

**10** https://www.ad-alliance.de/cms/portfolio/print/portfolio.html?p=/print/portfolio/brigitte, accessed: 27 January 2025.

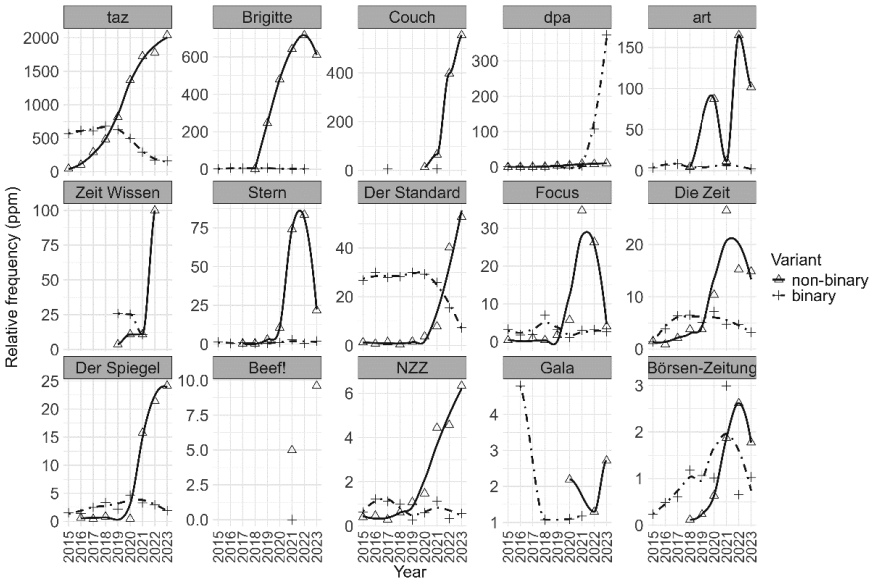


**Fig. 3:** Relative frequencies of binary and non-binary forms by source and year.

and audience – in this case, "young and modern women"[11] – might influence the use of gender-inclusive strategies. However, the status of *Couch* as a driver of change is questionable due to its significantly smaller audience and reach, especially when compared to *Brigitte* (0.19 million readers vs. 1.25 million readers).

Beyond that, only a few other sources show a clear upward trend in the use of non-binary forms over time, though at much lower frequencies (as indicated by the y-axis labels). These sources include the left-leaning Austrian daily *Der Standard*, the news magazine *Der Spiegel*, and the Swiss daily *Neue Zürcher Zeitung* (NZZ).[12] While NZZ has traditionally been viewed as liberal, it is increasingly criticized for its conservative or right-wing leanings (Eigenmann and Loser 2017). It frequently publishes critical articles on gender-inclusive language, making the increase in non-binary forms surprising. This shift may be explained by meta-linguistic usage, where gender-inclusive forms are used ironically or critically in articles discussing

**11** https://www.ad-alliance.de/cms/portfolio/print/portfolio.html?p=/print/portfolio/couch, accessed: 27 January 2025.

**12** Until 2022, the popular science magazine *ZEIT Wissen* also exhibited an increase in non-binary forms. However, in 2023, neither binary nor non-binary forms are attested in our subcorpus. This might be due to new editorial guidelines that we do not have access to, or mistakes in the DeReKo compilation.

gender issues: Waldendorf (2023: 16) highlights similar critical or mocking uses in the conservative German newspaper *Die Welt*.

For binary variants, most sources exhibit a generally decreasing or flattening trend. The only exception is *dpa*, the German Press Agency,[13] which shows a strong increase in the use of binary forms and remains the only source where binary variants outnumber non-binary ones. This increase can be attributed to the use of slashes, particularly in the forms *Autor/in* and *Autor/innen* ('author(s)'), which appeared 25,680 times in 2022 and 2023. In 2021, *dpa* committed to avoiding masculine generics and adopting gender-inclusive language.[14] Since authors' names are typically included in every press release, it is likely that masculine generics like *Autor* and *Autoren* were replaced with slash constructions following this policy change. Thus, *dpa* serves as an example of regulated change, where editorial decisions notably shape linguistic practices. Interestingly, despite the official acceptance of the slash as the only gender-inclusive form by the Council for German Orthography, its use remains rare outside of *dpa* (see Figures 1 and 2). This highlights the limited influence of normative regulations when it comes to conveying social meaning through language.

The distribution of gender symbols across text types has also been shown to vary in other corpora. For example, in the case of the city website of Hamburg, the majority of gender symbols are limited to specific subpages (Müller-Spitzer et al. 2024a). Therefore, claims about the frequency of gender symbols must always account for the textual basis and the inconsistent distribution across different text types. Ideally, the use of gender symbols should always be understood in the context of all personal nouns in a given text to accurately capture the relationships between them (Müller-Spitzer et al. 2024c). The following section addresses this by comparing the frequency of gender-inclusive orthographies with that of regular inflectional forms for a set of preselected nouns.

## 4.2 Lexicon-based approach

As outlined in the previous section, one of the challenges in researching personal nouns using corpus linguistic methods lies in the fact that the overall population of these forms remains unknown. Consequently, it is difficult to contextualize differ-

**13** The German Press Agency (*dpa*) is a news agency that provides current and neutral information to other media outlets, rather than publishing directly for end readers like newspapers or magazines do. It serves as a central source of news that editorial teams can process and publish individually.

**14** https://www.presseportal.de/pm/8218/4947122, accessed: 25 March 2024.

ent realizations of the same personal noun or to comprehensively relate all person references within a given corpus (Sökefeld et al. 2023: 34). Another issue pertains to the dual interpretation of the masculine base form as either gender-specific or generic. This distinction depends on contextual cues, which can be analyzed manually in small-scale corpus data (Müller-Spitzer et al. 2024c; Waldendorf 2023) but pose significant challenges in large-scale corpora such as the one employed in our study.

The analysis that follows compares binary and non-binary gender-inclusive orthographic forms of a preselected set of 131 personal nouns with their masculine base forms – without differentiating between their gender-specific and generic usages – and with the feminine derivations formed by *–in*. Thus, the entire inflectional paradigm of each item is accounted for and can be extracted from the corpus, enabling more meaningful frequency comparisons. For instance, the paradigm for the noun *Autor* ('author') includes:

– masculine (generic and gender-specific): *Autor*, *Autors*, *Autoren*
– feminine: *Autorin*, *Autorinnen*
– binary: *Autor/-in*, *Autor/-innen* (including variants with capital *I* and parantheses)
– non-binary: *Autor*in*, *Autor*innen* (including variants with colon and underscore)

The search items were derived from a list of personal nouns provided by the *Duden* editorial team. This original list contains 10,000 personal nouns with systematic gender differentiation. From this, we selected the 145 nouns with the highest frequency scores in DeReKo. However, post-corpus analysis revealed the need to exclude 14 items. This exclusion applied to population nouns, which are predominantly used as attributive adjectives (e.g., *Stuttgarter Flughafen* 'Stuttgart Airport', *Frankfurter Bahnhof* 'Frankfurt Central Station') and only superficially resemble masculine personal nouns (e.g., *(der) Stuttgarter* '(the) man from Stuttgart'). Additionally, two lexemes (*Zähler* 'meter/counter' and *Wetter* 'weather') were excluded as they are primarily used with their other, non-human meaning, with rare exceptions where they refer to individuals (in the sense of 'someone who counts' and 'someone who bets'). The remaining 131 nouns, along with their inflectional forms and frequency data from DeReKo, are detailed in Table S4 of the Supplementary Material.

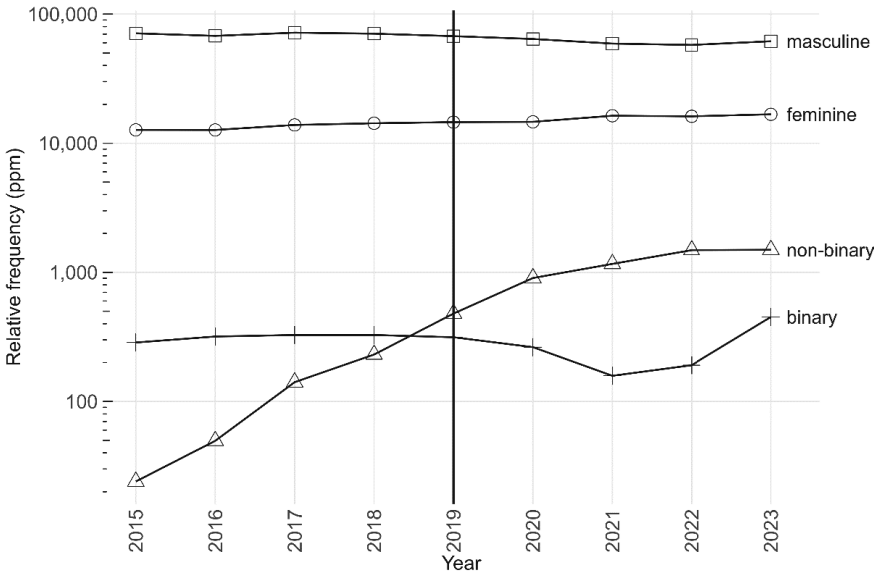


**Fig. 4:** Log-transformed relative frequencies of masculine, feminine, non-binary, and binary realizations for all 131 lexemes combined.

### 4.2.1 Development over time

Figure 4 illustrates the distribution of masculine, feminine, binary, and non-binary forms. The y-axis is log-transformed to prevent the curves for gender-inclusive variants from being obscured along the x-axis when raw frequencies are plotted.[15] Despite this transformation, the axis labels indicate raw frequencies, underscoring that masculine forms represent by far the most frequent and robust realization among the 131 lexemes. When aggregating all years, 84.85% of all tokens ($N = 9,787,168$) are masculine, 13.83% are feminine, 0.69% are non-binary, and 0.63% are binary.

As previously noted, the masculine forms may either refer to specific males or to mixed-gender groups – a distinction that can only be made through manual annotations and which is not represented in Figure 4. Feminine forms follow in sec-

**15** Figure S1 in the Supplementary Material presents raw frequency distributions, highlighting the relative rarity of gender-inclusive variants compared to feminine forms and, in particular, masculine forms, but obscuring trends for binary and non-binary forms due to their low frequency. A detailed view of individual gender-inclusive forms is provided in Figure S2, also using a log-transformed scale, for cases where granular trend analysis is required.

ond place, with an average of one tenth of the frequency of masculine forms. Non-binary forms exhibit a comparable exponential, though flattening, growth pattern, as observed in the corpus-driven analysis, surpassing binary forms in frequency between 2018 and 2019. From 2019 onwards, the frequency of the non-binary forms is constantly higher than that of the binary forms.

Similar to the analysis in Section 4.1.1, we applied a linear regression model to examine changes in forms before and after 2019, focusing on their relationship to the masculine base form after the amendment of Germany's civil status law. To address the non-normal distribution of relative frequencies, the data were log-transformed prior to analysis. The model demonstrates a strong fit, with an R-squared value of 0.9786, indicating that approximately 98% of the variance in the data is explained. Furthermore, the overall model is highly significant, as evidenced by the F-statistic (183) and its p-value ($p < 0.001$).

The main effects for the feminine, non-binary, and binary variants are all highly significant ($p < 0.001$), with negative coefficients confirming that these forms were used less frequently than the masculine base form prior to 2019. The main effect of the time group (before vs. after 2019) is not significant, indicating that the frequency of the masculine base form remained stable over time.

The interaction terms reveal how the use of gender-inclusive forms shifted after 2019. The feminine and binary variants show no significant change compared to the masculine base form, suggesting their relative stability over time. In contrast, the non-binary forms exhibit a significant increase ($p < 0.001$) in relation to the masculine base form after 2019. While non-binary variants remain a rare phenomenon overall, this substantial relative growth highlights their emerging relevance in the linguistic landscape.

Interestingly, when excluding the newspaper *taz* from the model, the binary variant also shows a significant increase after 2019 ($p < 0.05$), whereas it is not significant in the model that includes *taz*. The increase of non-binary variants remains highly significant, also without *taz*. This again suggests that while *taz* modulates certain patterns, linguistic variation persists across the remaining sources. Therefore, the observed changes in gender-inclusive forms are not solely driven by *taz* but are distributed more broadly within the corpus.

These findings underscore the importance of comparing different realizations of personal nouns and accounting for cross-source variation. This allows us to distinguish between absolute surface frequencies and relative developments in relation to base forms as well as source-specific effects, providing deeper insights into patterns of gender-inclusive language use.

### 4.2.2 Lexical variance

Next, we use the data from the lexicon-based approach for detailed analyses on the lexical level. This enables an examination of the distribution of masculine, feminine, binary, and non-binary realizations for each of the 131 items. In the following discussion, items are presented as lemmas in small capitals.

Masculine realizations are overwhelmingly dominant for the majority of lexemes: 126 out of 131 items (96.18%) appear as masculine in at least 50% of cases. Figure 5 illustrates proportions for the 20 most frequent items in our corpus, while Table 1 gives an overview of the top ten items in each category. Only four items exhibit a majority (i.e., more than 50%) of feminine forms. Two of these are political terms: *Kanzler* ('chancellor', 63.27% feminine) and *Bundeskanzler* ('federal chancellor', 51.21% feminine) – a pattern explained by Angela Merkel's 16-year tenure as German chancellor (2005–2021). This period also influenced other linguistic phenomena, such as compound formation with feminine first elements (Ochs 2024). Additionally, two job titles traditionally associated with women, according to stereotype research (Misersky et al. 2014), show feminine forms in more than 50% of instances: *Kassierer* ('cashier', 61.64%) and *Sekretär* ('secretary', 60.19%).

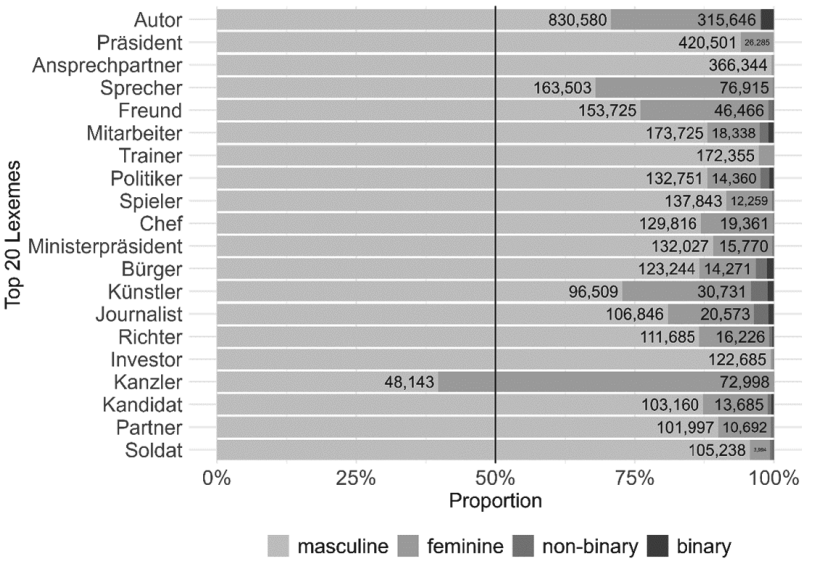


**Fig. 5:** Proportions of masculine, feminine, non-binary and binary realizations for the 20 most frequent items. Raw frequencies are indicated within the columns.

An exception to this trend is the lexeme *Erzieher* ('preschool teacher'), where neither feminine (41.01%) nor masculine (48.02%) realizations exceed 50%. This is not coincidental: *Erzieher* has the highest proportion of gender-inclusive realizations among all items, with 10.96% when binary and non-binary forms are summarized. The final column of Table 1 lists the remaining most frequent items based on the proportion of gender-inclusive forms. The data snippet highlights that gender-inclusive realizations remain a marginal phenomenon overall, with proportions almost universally below 10%. Notably, however, at least one gender-inclusive token (binary or non-binary) is attested for each of the 131 lexemes.

These distributions indicate that, in addition to considering the corpus base and sources when interpreting the frequency and patterns of gender-inclusive orthographies, it is essential to account for lexical differences as well (Müller-Spitzer et al. 2024c: 11). Analyzing only one single lexeme as a case study – for example, as done by Adler and Hansen (2020) or Krome (2021) – especially if, by chance, it represents an extreme case such as *Erzieher*, would not provide sufficient evidence to draw general conclusions about the broader use of gender-inclusive orthographies. It is therefore necessary to approach gender-inclusive orthographies from both aggregated perspectives, to understand overarching trends, and with detailed analyses, to capture the nuances of individual lexemes.

In the final step, we focus on the micro-diachronic development of a set of items, as so far we have only reported the overall distributions across all years. We selected the ten lexemes with the highest overall proportions of gender-inclusive forms to track their development over time (see Table 1, final column). Figure 6 displays the results of this analysis, starting in the upper left corner with *Erzieher* (10.96% gender-inclusive forms) and ending with *Teilnehmer* in the bottom right corner (3.85% gender-inclusive forms).

For all items, the proportions in 2023 are as follows: 1) masculine, 2) feminine, 3) non-binary, 4) binary (with the exception of *Tänzer* 'dancer', which has a slightly higher proportion of feminine forms). This pattern mirrors the overall distribution shown in Figure 4. The proportion of masculine forms is generally decreasing, while non-binary realizations are slowly gaining ground across this set of items. The most notable increase is seen with *Erzieher*, which reached 13.79% non-binary forms in 2023.

Interestingly, for some lexemes, the proportions of masculine and feminine forms are beginning to converge, most prominently for *Schüler*, *Leser* and *Künstler*. For *Erzieher* and *Tänzer*, the masculine and feminine forms are nearly equally frequent in 2023. We interpret this as an increase in the use of pair forms, which are captured as two separate tokens (one masculine, one feminine) in our search.

**Tab. 1:** Overview of the top ten items with the highest overall proportions of masculine, feminine, binary, and non-binary forms. For clarity, items are presented in their lemma form.

| rank | masculine | feminine | binary | non-binary | binary + non-binary |
|---|---|---|---|---|---|
| | *Hersteller* | *Kanzler* | *Erzieher* | *Mieter* | *Erzieher* |
| | 99.75% | 63.26% | 6.74% | 5.93% | 10.96% |
| **1** | 'producer' | 'chancellor' | 'preschool teacher' | 'tenant' | 'preschool teacher' |
| | *Coach* | *Kassierer* | *Anwohner* | *Anwohner* | *Mieter* |
| | 99.66% | 61.63% | 3.55% | 4.83% | 9.25% |
| **2** | 'coach' | 'cashier' | 'resident' | 'resident' | 'tenant' |
| | *Ansprechpartner* | *Sekretär* | *Mieter* | *Erzieher* | *Anwohner* |
| | 99.59% | 60.19% | 3.32% | 4.23% | 8.38% |
| **3** | 'contact (person)' | 'secretary' | 'tenant' | 'preschool teacher' | 'resident' |
| | *Anbieter* | *Bundeskanzler* | *Bewohner* | *Bewohner* | *Bewohner* |
| | 99.49% | 51.21% | 2.38% | 3.82% | 6.19% |
| **4** | 'provider' | 'federal chancellor' | 'inhabitant' | 'inhabitant' | 'inhabitant' |
| | *Sponsor* | *Sänger* | *Schüler* | *Schüler* | *Schüler* |
| | 99.42% | 47.27% | 2.37% | 3.61% | 5.98% |
| **5** | 'sponsor' | 'singer' | 'pupil' | 'pupil' | 'pupil' |
| | *Investor* | *Sozialminister* | *Autor* | *Leser* | *Leser* |
| | 99.37% | 42.48% | 2.26% | 3.46% | 5.28% |
| **6** | 'investor' | 'social minister' | 'author' | 'reader' | 'reader' |
| | *Anleger* | *Schauspieler* | *Radfahrer* | *Tänzer* | *Tänzer* |
| | 99.22% | 41.42% | 1.97% | 3.22% | 4.95% |
| **7** | 'investor' | 'actor' | 'cyclist' | 'dancer' | 'dancer' |
| | *Händler* | *Tänzer* | *Wähler* | *Künstler* | *Radfahrer* |
| | 98.57% | 41.05% | 1.93% | 2.88% | 4.34% |
| **8** | 'retailer' | 'dancer' | 'voter' | 'artist' | 'cyclist' |
| | *Aufsteiger* | *Erzieher* | *Leser* | *Wissenschaftler* | *Künstler* |
| | 98.57% | 41.01% | 1.82% | 2.78% | 3.93% |
| **9** | 'riser' | 'preschool teacher' | 'reader' | 'scientist' | 'artist' |
| | *Arbeitgeber* | *Schriftsteller* | *Zuhörer* | *Radfahrer* | *Teilnehmer* |
| | 97.96% | 35.94% | 1.73% | 2.37% | 3.85% |
| **10** | 'employer' | 'writer' | 'listener' | 'cyclist' | 'participant' |

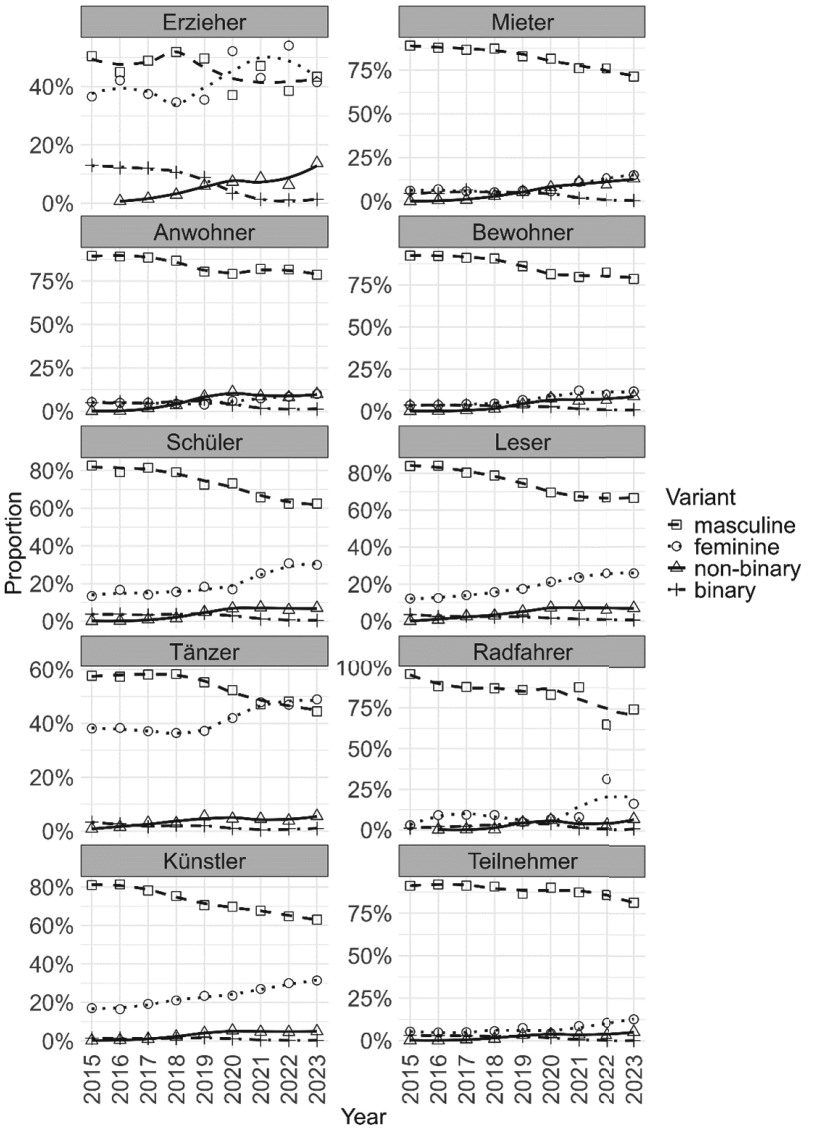


**Fig. 6:** Proportions of masculine, feminine, non-binary, and binary realizations for the ten lexemes with the highest overall share of gender-inclusive forms, with the trend reading from top-left to bottom-right.

It is also noteworthy that these ten items are primarily general or unspecific role nouns: *Mieter*, *Anwohner*, *Bewohner*, *Schüler*, *Leser*, *Radfahrer* and *Teilnehmer*. A possible interpretation is that general personal nouns, which often refer to unspecific groups, are predominantly used in masculine generic forms. These can now be replaced by gender-inclusive variants or pair forms. In contrast, lexemes that are mostly used to refer to specific individuals – such as authors in news articles or political figures like *Kanzler* or *Präsident* – are less likely to adopt gender-inclusive (or rather, gender-unspecific) variants. For *Kanzler*, only 0.06% of all tokens are gender-inclusive (binary or non-binary), for *Bundeskanzler*, it is even less (0.04%); and for *Präsident*, it is only 0.01%. Data like this help us understand the lexical domains in which gender-inclusive forms are being used. To date, there is little empirical data on these lexical distributions. Further investigations into the semantic classes of personal nouns could help clarify this trend.

The lexemes with the highest proportion of masculine forms (Table 1, second column) are also predominantly general role nouns.[16] However, they differ from those discussed before in that, while they can refer to humans, they are more likely to denote non-human agents. Examples include *Hersteller*, *Ansprechpartner*, *Anbieter*, and *Sponsor*, which are more frequently applied to entities such as companies or sports clubs rather than individual humans. However, our data does not allow us to disentangle the different uses of masculine forms, as we cannot distinguish between them solely based on their surface form. This issue affects not only the distinction between human and non-human references, but also between gender-specific and generic uses of masculine forms. Only by differentiating between these two uses can we determine whether generic masculine forms are being replaced by gender-inclusive alternatives, while specific masculines may remain stable over time.

Currently, automatic detection of masculine generics is not possible (Sökefeld et al. 2023: 38). To our knowledge, only one study has addressed this issue through manual annotations of German press texts, differentiating human and non-human as well as gender-specific and generic uses of the masculine (Müller-Spitzer et al. 2024c). Given the extensive nature of our dataset, manual annotations are not feasible. However, focused investigations on smaller text samples or specific lexemes could provide valuable insights into this phenomenon. For instance, Waldendorf (2023: 7) conducted a manual analysis of a small sample of her corpus, finding that masculine generics still dominate, although her case study does not take a

**16** *Coach* is an exception in this context, as it holds a special status as an English loanword and is rarely, if ever, used in its feminine derivation, *Coachin* (appearing in only in 0.33% of cases according to Kopf 2022: 89).

diachronic approach. Moreover, her analysis does not address masculine specifics, which would be necessary to contextualize the use of masculine generic forms. Sökefeld (forthcoming) analyzed a small sample of lexemes, showing that masculine generics are declining to varying degrees, with an accompanying rise in gender-inclusive forms. This suggests a similar trend may be occurring in our data. However, based on our dataset, we can only conclude that masculine and feminine forms are generally stable over time, while non-binary forms remain a rising but still marginal phenomenon – at least in the written press texts analyzed here. Certain lexical items show a stronger tendency towards gender-inclusive usage, while others, particularly those ambiguous between human and non-human references, remain largely unaffected by these new variants.

## 5 Conclusion and outlook

Generating linguistic visibility and representing gender diversity is a key goal in both feminist and queer linguistic endeavors. While feminist approaches have established binary linguistic forms, the recognition of gender identities beyond the binary calls for new modes of linguistic representation. Our corpus-driven diachronic analysis of German press texts demonstrates that, at present, this is predominantly achieved through the use of word-internal gender symbols. The asterisk (*Schüler*innen*) and the colon (*Schüler:innen*) are by far dominant, while the underscore (*Schüler_innen*) plays a negligible role in our corpus. Binary word-internal forms, such as the capital *I* (*SchülerInnen*), are generally on the decline, especially since the amendment of the German civil status law in 2019. In contrast, other studies have found fully-fledged binary pair forms (*Schülerinnen und Schüler*) to be rising at a rate comparable to that of the non-binary asterisk and colon (Waldendorf 2023: 8). This suggests a domain-specific differentiation: Binary gender representations appear at the noun phrase level, while non-binary identities are conveyed word-internally.

Our lexicon-based approach further supplements these findings by providing insights into the frequency of gender-inclusive forms relative to conventional inflectional forms. We observe that masculine forms remain the dominant realization across lexemes and years, followed by feminine forms, with both binary and non-binary variants being relatively rare. Statistical analyses, however, showed that non-binary forms experience a significant increase in relation to the masculine base form after 2019, highlighting the importance of distinguishing between absolute surface frequencies and changes relative to base forms. Besides that, lexical choices are crucial when examining gender-inclusive orthographies. Our data indicate that

the realization of gender-inclusive forms is highly dependent on individual lexemes and their semantic domains.

The influence of specific sources on the development of gender-inclusive language also warrants further investigation. For non-binary forms, the alternative-left newspaper *taz* is the primary contributor, followed by the women's magazine *Brigitte* and the lifestyle magazine *Couch*. Besides that, only a few sources show a clear upward trend in the use of non-binary forms. However, fitting the linear regression models without *taz* revealed that all changes in the use of gender-inclusive orthographies remained significant for the remaining 14 sources after 2019. Therefore, while *taz* modulates certain effects and may be a leading innovator, neographies spread more widely across the corpus. It remains to be seen whether this spread continues, whether trends begin to decrease again, or whether neographies become niche phenomena in certain media.

With this study, we provide quantitative baselines to the growing body of research on gender inclusive German. Understanding the development, frequency, and distribution of new orthographical forms is essential for fostering more productive discussions on the topic, both in public discourse and academic fields. Furthermore, our findings can serve as an empirical foundation for developing usage-oriented guidelines on gender-inclusive language, such as those used in universities and state institutions. Further analyses of additional press sources (e.g., queer magazines) and other text types (particularly social media) will help illuminate the contexts and domains in which gender-inclusive forms are most commonly used.

## References

Acke, Hanna. 2019. Sprachwandel durch feministische Sprachkritik. *Zeitschrift für Literaturwissenschaft und Linguistik* 49(2). 303–320. https://doi.org/10.1007/s41244-019-00135-1.

Acke, Hanna. 2023. Excessiveness in a German social media debate on gender-fair language. *Neuphilologische Mitteilungen* 124(1). 46–77. https://doi.org/10.51814/nm.122738.

Adler, Astrid and Karolina Hansen. 2020. Studenten, Studentinnen, Studierende? Aktuelle Verwendungspräferenzen bei Personenbezeichnungen. *Muttersprache* 130(1). 47–63.

Adler, Astrid and Albrecht Plewnia. 2019. Die Macht der großen Zahlen. Aktuelle Spracheinstellungen in Deutschland. In Ludwig M. Eichinger and Albrecht Plewnia (eds.), *Neues vom heutigen Deutsch. Empirisch—Methodisch—Theoretisch*, 141–162. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110622591-008.

AG Feministisch Sprachhandeln. 2014. Anregungen zum antidiskriminierenden Sprachhandeln. https://feministisch-sprachhandeln.org/wp-content/uploads/2015/04/sprachleitfaden_zweite_auflage.pdf. Accessed: 2 February 2025.

Bailey, April H., Marianne LaFrance and John F. Dovidio. 2019. Is man the measure of all things? A social cognitive account of androcentrism. *Personality and Social Psychology Review* 23(4). 307–331. https://doi.org/10.1177/1088868318782848.

Bast, Jennifer, Jürgen Maier, Georg Albert and Jan Georg Schneider. 2024. Gendered debates? The use of gender-sensitive language in German televised debates, 1997–2022. *European Journal of Politics and Gender* 1–25. https://doi.org/10.1332/25151088Y2024D000000023.

Blake, Christopher and Christoph Klimmt. 2010. Geschlechtergerechte Formulierungen in Nachrichtentexten. *Publizistik* 55. 289–304. https://doi.org/10.1007/s11616-010-0093-2.

Braun, Friederike, Susanne Oelkers, Karin Rogalski, Janine Bosak and Sabine Sczesny. 2007. "Aus Gründen der Verständlichkeit …": Der Einfluss generisch maskuliner und alternativer Personenbezeichnungen auf die kognitive Verarbeitung von Texten. *Psychologische Rundschau* 58(3). 183–189. https://doi.org/10.1026/0033-3042.58.3.183.

Burnett, Heather and Céline Pozniak. 2021. Political dimensions of gender inclusive writing in Parisian universities. *Journal of Sociolinguistics* 25(5). 808–831. https://doi.org/10.1111/josl.12489.

Butler, Judith. 1988. Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal* 40(4). 519–531. https://doi.org/10.2307/3207893.

Diewald, Gabriele. 2018. Zur Diskussion: Geschlechtergerechte Sprache als Thema der germanistischen Linguistik – exemplarisch exerziert am Streit um das sogenannte generische Maskulinum. *Zeitschrift für germanistische Linguistik* 46. 283–299. https://doi.org/10.1515/zgl-2018-0016.

Diewald, Gabriele and Anja Steinhauer. 2020. *Handbuch geschlechtergerechte Sprache. Wie Sie angemessen und verständlich gendern*. Dudenverlag.

Doleschal, Ursula. 1992. *Movierung im Deutschen: Eine Darstellung der Bildung und Verwendung weiblicher Personenbezeichnungen*. München: Lincom Europa.

Eigenmann, David and Philipp Loser. 2017. Nach Deutschland, nach rechts. *Tages-Anzeiger* https://www.tagesanzeiger.ch/nach-deutschland-nach-rechts-957458637553. Accessed: 2 February 2025.

Eisenberg, Peter. 2020. Zur Vermeidung sprachlicher Diskriminierung im Deutschen. *Muttersprache* 130(1). 3–16.

Eisenberg, Peter. 2022. Weder geschlechtergerecht noch gendersensibel. *APuZ Aus Politik und Zeitgeschichte* 5–7. https://www.bpb.de/shop/zeitschriften/apuz/geschlechtergerechte-sprache-2022/346091/weder-geschlechtergerecht-noch-gendersensibel. Accessed: 2 February 2025.

Fleischer, Wolfgang and Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110256659.

Friedrich, Martin C. G., Vanessa Drößler, Nicole Oberlehberg and Elena Heise. 2021. The influence of the gender asterisk ("Gendersternchen") on comprehensibility and interest. *Frontiers in Psychology* 12. 1–11. https://doi.org/10.3389/fpsyg.2021.760062.

Friedrich, Martin C. G., Sebastian Gajewski, Katharina Hagenberg, Claudia Wenz and Elena Heise. 2024. Does the gender asterisk ("Gendersternchen") as a special form of gender-fair language impair comprehensibility? *Discourse Processes* 61(9). 439–461. https://doi.org/10.1080/0163853X.2024.2362027.

Genderleicht. 2024. Genderstern, Doppelpunkt und Binnen-I kurz und knapp erklärt. https://www.genderleicht.de/genderzeichen. Accessed: 2 February 2025.

Glim, Stefanie, Anna Körner, Holden Härtl and Ralf Rummer. 2023. Early ERP indices of gender-biased processing elicited by generic masculine role nouns and the feminine–masculine pair form. *Brain and Language* 242. 1–7. https://doi.org/10.1016/j.bandl.2023.105290.

Gygax, Pascal, Ute Gabriel, Oriane Sarrasin, Jane Oakhill and Alan Garnham. 2008. Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes* 23(3). 464–485. https://doi.org/10.1080/01690960701702035.

Hellinger, Marlis and Hadumod Bußmann (eds.). 2001. *Gender across languages: The linguistic representation of women and men, volume 1*. Amsterdam, Philadelphia: John Benjamins.

Hellinger, Marlis and Hadumod Bußmann. 2003. Engendering female visibility in German. In Marlis Hellinger and Hadumod Bußmann (eds.), *Gender across languages: The linguistic representation of women and men. Volume 3*, 141–174. Amsterdam, Philadelphia: John Benjamins Publishing Co.

Jourian, T. J. 2015. Queering constructs: Proposing a dynamic gender and sexuality model. *The Educational Forum* 79(4). 459–474. https://doi.org/10.1080/00131725.2015.1068900.

Jäckle, Sebastian. 2022. Per aspera ad astra – Eine politikwissenschaftliche Analyse der Akzeptanz des Gendersterns in der deutschen Bevölkerung auf Basis einer Online-Umfrage. *Politische Vierteljahresschrift* 63(3). 469–497. https://doi.org/10.1007/s11615-022-00380-z.

Keith, Nina, Kristine Hartwig and Tobias Richter. 2022. Ladies first or ladies last: Do masculine generics evoke a reduced and later retrieval of female exemplars? *Collabra: Psychology* 8. https://doi.org/10.1525/collabra.32964.

Kopf, Kristin. 2022. Ist Sharon Manager? Anglizismen und das generische Maskulinum. In Gabriele Diewald and Damaris Nübling (eds.), *Genus – Sexus – Gender*, 65–104. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110746396-003.

Kotthoff, Helga. 2017. Von Syrx, Sternchen, großem I und bedeutungsschweren Strichen. Über geschlechtergerechte Personenbezeichnungen in Texten und die Kreation eines schrägen Registers. In *Sprache Und Geschlecht. Band 1: Sprachpolitiken Und Grammatik*, 91–115. Duisburg: Universitätsverlag Rhein-Ruhr.

Kotthoff, Helga. 2023. Gendern unter soziolinguistischer Perspektive. *Deutsche Sprache* 3. 181–220. https://doi.org/10.37307/j.1868-775X.2023.03.02.

Kotthoff, Helga and Damaris Nübling. 2018. *Genderlinguistik: Eine einführung in Sprache, Gespräch und Geschlecht*. 1st edn. Tübingen: Narr.

Kotthoff, Helga and Damaris Nübling. 2024. *Genderlinguistik: Eine Einführung in Sprache, Gespräch und Geschlecht*. 2nd edn. Tübingen: Narr.

Krome, Sabine. 2020. Zwischen gesellschaftlichem Diskurs und Rechtschreibnormierung: Geschlechtergerechte Sprache als Herausforderung für gelungene Textrealisation. *Muttersprache* 130(1). 64–78.

Krome, Sabine. 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.

Kupietz, Marc, Cyril Belica, Holger Keibel and Andreas Witt. 2010. The German reference corpus DeReKo: A primordial sample for linguistic research. In *Proceedings of the seventh international conference on language resources and evaluation (LREC 2010)*, 1848–1854.

Kupietz, Marc, Harald Lüngen, Pawel Kamocki and Andreas Witt. 2018. The German reference corpus DeReKo: New developments – new opportunities. In *Proceedings of the 11th international conference on language resources and evaluation (LREC 2018)*, 4353–4360.

Kupisch, Tanja, Miriam Geiß, Natalia Mitrofanova and Marit Westergaard. 2022. Structural and phonological cues for gender assignment in monolingual and bilingual children acquiring German. Experiments with real and nonce words. *Glossa: A Journal of General Linguistics* 7(1). 1–37. https://doi.org/10.16995/glossa.5696.

Kurfer, Tobias. 2024. Schlecht informiert: Gendersprache in den Medien. In Ewa Trutkowski and André Meinunger (eds.), *Gendern – auf Teufel*in komm raus?*, 235–264. Kulturverlag Kadmos.

Kurz, Pia and Hannah De Mulder. 2023. A star is born? The German gender star and its effects on mental representation. *Psychology of Language and Communication* 27(1). 384–404. https://doi.org/10.58734/PLC-2023-0018.

Körner, Anita, Bleen Abraham, Ralf Rummer and Fritz Strack. 2022. Gender representations elicited by the gender star form. *Journal of Language and Social Psychology* 41. 553–571. https://doi.org/10.1177/0261927X221080181.

Lev, Arlene Istar. 2004. *Transgender emergence: Therapeutic guidelines for working with gender-variant people and their families*. Binghamton: Haworth Clinical Practice Press.

Link, Sabrina. 2024. The use of gender-fair language in Austria, Germany, and Switzerland: A contrastive, corpus-based study. *Lingua* 308. 103787. https://doi.org/10.1016/j.lingua.2024.103787.

Löhr, Ronja A. 2021. Gendergerechte Personenbezeichnungen 2.0. Wie nichtbinäre Personen den Genderstern und andere Bezeichnungsvarianten beurteilen. *Muttersprache* 131(2). 172–182. https://doi.org/10.53371/60206.

Löhr, Ronja A. 2022. "Ich denke, es ist sehr wichtig, dass sich so viele Menschen wie möglich repräsentiert fühlen". In Gabriele Diewald and Damaris Nübling (eds.), *Genus – Sexus – Gender*, 349–380. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110746396-012.

Lüngen, Harald and C. Michael Sperberg-McQueen. 2012. A TEI P5 document grammar for the IDS text model. *Journal of the Text Encoding Initiative* 3. 1–18. https://doi.org/10.4000/jtei.508.

Meineke, Eckhard. 2023. *Studien zum genderneutralen Maskulinum*. Heidelberg: Universitätsverlag Winter.

Meuleneers, Paul. 2024. On the "invention" of the Gendersprache in German media discourse. In Falco Pfalzgraf (ed.), *Public attitudes towards gender-inclusive language: A multilingual perspective*, 159–182. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783111202280-007.

Misersky, Julia, Pascal M. Gygax, Paolo Canal, Ute Gabriel, Alan Garnham, Friederike Braun, Tania Chiarini, Kjellrun Englund, Adriana Hanulikova, Anton Öttl, Jana Valdrova, Lisa Von Stockhausen and Sabine Sczesny. 2014. Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods* 46(3). 841–871. https://doi.org/10.3758/s13428-013-0409-z.

Motschenbacher, Heiko. 2013. Gentlemen before ladies? A corpus-based study of conjunct order in personal binomials. *Journal of English Linguistics* 41(3). 212–242. https://doi.org/10.1177/0075424213489993.

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

Müller-Spitzer, Carolin and Samira Ochs. 2023. Geschlechtergerechte Sprache auf den Webseiten deutscher, österreichischer, schweizerischer und Südtiroler Städte. *Sprachreport* 39(2). 1–5. https://doi.org/10.14618/sr-2-2023_mue.

Müller-Spitzer, Carolin, Samira Ochs and Jan Oliver Rüdiger. 2024a. Auf der Suche nach Genderzeichen: Die Hamburger Volksinitiative "Schluss mit Gendersprache in Verwaltung und Bildung" und die Verwendung von Genderzeichen auf der Webseite der Stadt Hamburg. https://lingdrafts.hypotheses.org/2671. Accessed: 26 July 2024.

Müller-Spitzer, Carolin, Samira Ochs, Jan Oliver Rüdiger and Sascha Wolfer. 2024b. Die Herausbildung neuer Routinen der Personenreferenz am Beispiel der deutschen Weihnachts- und Neujahrsansprachen. In Paul Meuleneers, Lisa Zacharski, Evelyn Ferstl and Damaris Nübling (eds.), *Genderbezogene Personenreferenzen: Routinen und Innovationen*, 213–236. Helmut Buske Verlag.

Müller-Spitzer, Carolin, Samira Ochs, Jan Oliver Rüdiger and Sascha Wolfer. 2024c. Less than one percent of words would be affected by gender-inclusive language in German press texts. *Humanities and Social Sciences Communications* 11(1). 1–13. https://doi.org/10.1057/s41599-024-03769-w.

Müller-Spitzer, Carolin, Jan Oliver Rüdiger and Sascha Wolfer. 2022. Olaf Scholz gendert. Eine Analyse von Personenbezeichnungen in Weihnachts- und Neujahrsansprachen. Billet. https://lingdrafts.hypotheses.org/2370.

Ochs, Samira. 2024. Komposita mit den relationalen Zweitgliedern Gatte und Gattin – eine korpusbasierte Studie aus genderlinguistischer Perspektive. *Zeitschrift für Wortbildung / Journal of Word Formation* 8(1). 1–40. https://doi.org/10.21248/zwjw.2024.1.104.

Pabst, Laura Mathilde and Marlene Kollmayer. 2023. How to make a difference: The impact of gender-fair language on text comprehensibility amongst adults with and without an academic background. *Frontiers in Psychology, Section: Gender, Sex and Sexualities* 14. 1–9. https://doi.org/10.3389/fpsyg.2023.1234860.

Pfalzgraf, Falco. 2024. Attitudes of the purist association Verein Deutsche Sprache (VDS) towards gender-inclusive use of German. A conceptual expansion of the term linguistic purism. In Falco Pfalzgraf (ed.), *Public attitudes towards gender-inclusive language*, 183–208. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783111202280-008.

Pusch, Luise F. 1979. Der Mensch ist ein Gewohnheitstier, doch weiter kommt man ohne ihr—Eine Antwort auf Kalverkämpers Kritik an Trömel-Plötz' Artikel über 'Linguistik und Frauensprache'. *Linguistische Berichte* 63(79). 84–102.

Pusch, Luise F. 1984. *Das Deutsche als Männersprache: Aufsätze und Glossen zur feministischen Linguistik*. Berlin: Suhrkamp.

Rosar, Anne. 2022. Mann und Frau, Damen und Herren, Mütter und Väter – Zur (Ir-)Reversibilität der Geschlechterordnung in Binomialen. In Gabriele Diewald and Damaris Nübling (eds.), *Genus – Sexus – Gender*, 267–292. De Gruyter. https://doi.org/10.1515/9783110746396-009.

Rüdiger, Jan Oliver. 2023. Corpusexplorer. Computer software. http://corpusexplorer.de.

Schmitz, Dominic. 2024. Instances of bias: The gendered semantics of generic masculines in German revealed by instance vectors. *Zeitschrift für Sprachwissenschaft* 43(2). 295–325. https://doi.org/10.1515/zfs-2024-2010.

Schmitz, Dominic, Viktoria Schneider and Janina Esser. 2023. No generiaddress in sight: An exploration of the semantics of masculine generics in German. *Glossa Psycholinguistics* 2. https://doi.org/10.5070/G6011192.

Siegenthaler, Aline. 2024. *Deutsch- und französischsprachige Leitfäden für gendersensible Sprache: Normierung und Variation in Hochschulen und Behörden*. Geneva: Université de Genève dissertation.

Simon, Horst J. 2022. Sprache Macht Emotion. *APuZ Aus Politik und Zeitgeschichte* 5–7. https://www.bpb.de/shop/zeitschriften/apuz/geschlechtergerechte-sprache-2022/346087/sprache-macht-emotionen. Accessed: 2 February 2025.

Stefanowitsch, Anatol and Kirsten Middeke. 2023. Gender-marking -ess: The suffix that failed. *Zeitschrift für Anglistik Und Amerikanistik* 71(3). 293–319. https://doi.org/10.1515/zaa-2023-2029.

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

Sökefeld, Carla. forthcoming. Student*innen, Student/innen oder Student:innen? Kurzformen der graphisch markierten Movierung quantitativ und diachron. In Alexander Werth (ed.), *Die Movierung. Formen, Funktionen, Bewertungen.*, Berlin, Boston: De Gruyter.

Sökefeld, Carla, Melanie Andresen, Johanna Binnewitt and Heike Zinsmeister. 2023. Personal noun detection for German. In Harry Bunt (ed.), *Proceedings of the 19th joint acl - iso workshop on interoperable semantic annotation (isa-19)*, Nancy. https://sigsem.uvt.nl/isa19/ISA-19-proceedings.pdf.

taz. 2025. Zahlen, Fakten, Sensationen. Alles über uns, unsere Geschichte und was es sonst über das "linke Tageszeitungsprojekt" zu wissen gibt. https://taz.de/Zahlen-Fakten-Sensationen/!v=b33ef2ce-205d-4e54-abfd-a8a63bb236da. Accessed: 2 February 2025.

Thorne, Nat, Andrew Kam-Tuck Yip, Walter Pierre Bouman, Ellen Marshall and Jon Arcelus. 2019. The terminology of identities between, outside and beyond the gender binary – a systematic review. *International Journal of Transgenderism* 20(2–3). 138–154. https://doi.org/10.1080/15532739.2019.1640654.

Tross, J. 2023. *Divers versus verständlich?: Gendergerechte Formulierungen in Texten der Leichten Sprache*. Berlin: Frank & Timme GmbH. https://doi.org/10.57088/978-3-7329-8945-4.

Truan, N. 2019. The discursive construction of the people in European political discourse: Semantics and pragmatics of a contested concept in German, French, and British parliamentary debates. In Jan Zienkowski and Ruth Breeze (eds.), *Imagining the Peoples of Europe: Populist discourses across the political spectrum*, vol. 83, 201–228. John Benjamins Publishing Company. https://doi.org/10.1075/dapsac.83.09tru.

Trutkowski, Ewa and Helmut Weiß. 2023. Zeugen gesucht! Zur Geschichte des generischen Maskulinums im Deutschen. *Linguistische Berichte* 273. https://doi.org/10.46771/9783967692792_2.

Trömel-Plötz, Senta. 1978. Linguistik und Frauensprache. *Linguistische Berichte* 57(78). 49–68.

Völkening, Lena. 2022. Ist Gendern mit Glottisverschlusslaut ungrammatisch? Ein Analysevorschlag für das Suffix [ʔɪn] als phonologisches Wort. *Zeitschrift für Wortbildung / Journal of Word Formation* 6(1). 58–80. https://doi.org/10.3726/zwjw.2022.01.02.

Waldendorf, Anica. 2023. Words of change: The increase of gender-inclusive language in German media. *European Sociological Review* 40(2). 1–18. https://doi.org/10.1093/esr/jcad044.

WDR. 2023. WDR-Studie: So gendern die Deutschen. https://www1.wdr.de/nachrichten/gender-umfrage-infratest-dimap-100.html. Accessed: 2 February 2025.

Zacharski, Lisa. 2024. Using pair forms, criticizing the gender star – attitudes towards binary and non-binary gender-inclusive language in German. In F. Pfalzgraf (ed.), *Public attitudes towards gender-inclusive language: A multilingual perspective*, 209–242. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783111202280-009.

Zacharski, Lisa and Evelyn Ferstl. 2023. Gendered representations of person referents activated by the nonbinary gender star in German: A word-picture matching task. *Discourse Processes* 60(4–5). 294–319. https://doi.org/10.1080/0163853X.2023.2199531.

Zifonun, Gisela. 2021. Eine Linguistin denkt nach über den Genderstern. *Sprachreport* 37(2). 46–51. https://doi.org/10.14618/sr-2-2021-zifo.