

No genericity in sight: An exploration of the semantics of masculine generics in German

Dominic Schmitz, Department of English and American Studies, English Language and Linguistics, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, dominic.schmitz@uni-duesseldorf.de

Viktoria Schneider, Department of English and American Studies, English Language and Linguistics, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, viktoria.schneider@uni-duesseldorf.de

Janina Esser, div-ling – Association for Diversity in Linguistics, Cologne, Germany, janina.esser@div-ling.org

Findings of previous behavioural studies suggest that the semantic nature of what is known as the ‘masculine generic’ in Modern Standard German is indeed not generic but biased towards a masculine reading. Such findings are the cause of debates within and outside linguistic research, as they run counter to the grammarian assumption that the masculine generic form is gender-neutral. The present paper aims to explore the semantics of masculine generics, relating them to those of masculine and feminine explicit counterparts. To achieve this aim, an approach novel to this area of linguistic research is made use of: discriminative learning. Analysing semantic vectors obtained via naive discriminative learning, semantic measures calculated via linear discriminative learning, and taking into account the stereotypicality of the words under investigation, it is found that masculine generics are semantically much more similar to masculine explicit than to feminine explicit. The results presented in this paper thus support the notion of a masculine bias in masculine generics. Further, new insights into the semantic representations of masculine generics are provided and it is shown that stereotypicality does not modulate the masculine bias.



1. Introduction

Modern Standard German has three grammatical genders: the feminine, the masculine, and the neuter. In contexts in which the sex of the referent(s) is (a) unknown, (b) not of importance, or (c) mixed (i.e. there are referents of different sexes/genders), speakers of Modern Standard German make use of the so-called *generisches Maskulinum* ‘masculine generic’. The generic nature of the masculine generic refers to the grammarian notion of it being gender-neutral, independently of its grammatical gender (cf. Doleschal, 2002).¹ Masculine generics are used in the singular and in the plural, as illustrated by (1) and (2), respectively. In both examples, referents can be of any gender.

(1) Wird heute ein Professor an eine
 be.PRS.3SG today.ADV DET.INDF.M.SG professor.M.NOM.SG to DET.INDF.F.SG
 Universität berufen, kommt dieser oft
 university.F.ACC.SG appoint.PTCP.PRS come.PRS.3SG DET.DEF.M.SG often
 mit einem ganzen Forschungsteam.
 with DET.INDF.M.SG whole.ADJ research team.N.DAT.SG
 ‘When a professor is appointed to a university nowadays, they often come with a whole research team.’

(2) Die Professor-en der regulären Schweizer Uni-s.
 DET.DEF.PL professor-M.NOM.PL DET.DEF.PL regular.ADJ Swiss.ADJ uni-F.GEN.PL
 ‘The professors of the regular Swiss unis.’

In contrast, the word forms in (3) and (4) clearly denote male referents and are read as explicit masculines due to the unambiguous contexts:

(3) Michael Rosenberger ist Professor für Moraltheologie.
 Michael Rosenberger be.PRS.3SG professor.M.ACC.SG for moral theology.F.ACC.SG
 ‘Michael Rosenberger is a professor of moral theology.’

(4) Hans-Peter und Volker Stenzl [...] als Professor-en.
 Hans-Peter and Volker Stenzl as professor-M.ACC.PL
 ‘Hans-Peter and Volker Stenzl [...] as professors.’

Considering examples (1) and (3), and (2) and (4), respectively, one finds that masculine generics and explicit masculines share their orthographic, and thus also their phonological form.

¹ In this paper, we will use the term gender-neutral as an umbrella term concerning both *sex*, i.e. the categorical biological perspective, and *gender*, i.e. the social and cultural perspective. We acknowledge that both terms – *sex* and *gender* – are not identical, nor are forms of sex and gender clearly correlated or matched up. For the present case, however, it is of negligible importance whether we specifically refer to sex or gender.

To form a counterpart explicitly denoting a single female referent, the feminine gender suffix *-in* is added to the masculine form, as illustrated in (5). For a plural counterpart, the feminine suffix is added between the masculine form and the plural suffix *-en*, with a reduplication of <n> to account for vowel quality, as is shown in (6).

- (5) Sie ist eine Professor-in von vielen.
 she.F.NOM.SG be.PRS.3SG DET.INDF.F.SG professor-F.NOM.SG among many
 ‘She is one professor among many.’
- (6) Eine Förderung von bis zu drei Professor-in-nen.
 DET.INDF.F.SG funding.F.NOM.SG for up to three professor-F.DAT.PL
 ‘Funding for up to three professors.’

In the above examples, two word forms are identical in their form: the masculine generic *Professor* and the masculine explicit *Professor*. This is true for all masculine generic and masculine explicit word pairs within number, and even across number if they are a word form derived from a base via the *-er* suffix. In such cases, the generic and explicit masculine plurals share their form with their singular counterparts. As an example, consider the German word for ‘teacher’, *Lehrer*. This word form is used as singular and plural masculine generic and explicit, as the plural is marked by a zero morpheme.

With two, or in some cases even four, semantically distinct but closely related word forms in both the singular and the plural sharing their orthographic and phonological makeup, one question naturally suggests itself: how semantically distinct are the masculine generic and the masculine explicit? This question has been investigated by previous research on the masculine generic. That is, how can a form allegedly be gender-neutral if it shares its surface representation with masculine explicit forms? Are masculine generics truly gender-neutral, or is this idea a misconception? A new approach to the exploration of these questions is the focus of the present paper.

The remainder of this paper is structured as follows. Section 2 will present previous results on the matter of masculine generics in German as well as introduce the theoretical framework made use of in the present paper. Section 3 will explain the methodology used in this paper, while Section 4 will present the analyses and results of our investigation. Section 5 will discuss our findings and conclude this paper.

2. Theoretical background

This section aims to first provide an overview of findings by previous research on the nature of the masculine generic in Modern Standard German. Then, the frameworks of naive and linear discriminative learning are introduced.

2.1 Previous research on masculine generics in German

The question whether masculine generics in Modern Standard German are truly gender-neutral has been investigated by a growing body of literature during the last decades. We aim at giving a cursory overview of this literature here, but also refer to Irmen and Linner (2005) and Gyga et al. (2009) for further concise overviews.

One of the earliest studies on the matter was conducted by Irmen and Köhncke (1996). In their study, participants were presented with sentences containing either a masculine or a feminine form. The masculine form could be understood as either generic or explicit. After sentence presentation, participants were asked to quickly indicate whether the pertinent sentence referred to a man or a woman. An analysis of overall decisions and the corresponding reaction times showed that masculine generics are less often interpreted as referring to a woman and if they are, reaction times are longer.

In a more subtle methodological approach by Braun et al. (1998), participants read a short text about either an ecotrophology (which was judged to be stereotypically female) or a geophysics conference (which was judged to be stereotypically male). Within the two texts, five phrases were either given as masculine generics, as *Beidnennung*,² or as a neutral noun.³ Participants were then to guess the percentage of female attendees. Results showed that the percentage of female attendees given for texts without masculine generics was higher than for those participants who read texts with masculine generics. Stereotypicality of the conferences' fields did show an influence; however, masculine generics in both contexts elicited significantly higher percentages of male conference attendees.

Rothermund (1998) gave brief descriptions of situations to participants. The descriptions either contained an explicitly male or female referent or a masculine generic. After reading the description, participants were prompted to decide whether items in a list of words were part of the description. Half of such items were part of the description, while the other half were not. Some of the newly added items were stereotypically female and male distractor words. It was found that it took participants longer to decide that a stereotypically male distractor was not part of the original description if the description contained a masculine generic. The opposite effect, however, was found for the same setup but with plural instead of singular referents.

² A *Beidnennung* 'mentioning of both' refers to a common phrase which is considered to be more gender-neutral than the generic masculine. An example is *Anwälte und Anwältinnen* 'lawyers (male) and lawyers (female)'.

³ Neutral words or nouns are role nouns without a counterpart of the opposing gender. An example is *Rechtsvertretung* instead of *Anwalt* 'lawyer'. Note, however, that most neutral replacements for masculine generic role nouns are not true synonyms.

Providing their participants with a cloze task, Rothmund and Scheele (2004) found that clozes are more often resolved with male referents in contexts with masculine generics as compared to contexts with other forms, such as *majuscule-I*⁴ or neutral nouns.

Heise (2000) confronted participants with beginnings of stories which contained as protagonists either masculine generics, alternative forms, such as the *majuscule-I* or the *slash-form*,⁵ or neutral nouns. Participants then had to give names to the protagonists. It was found that for stories with masculine generics and neutral nouns, participants more often used typically male names for the protagonists.

Stahlberg and Sczesny (2001) asked participants to name their favourite painter, potential candidates for the German chancellery, and celebrities. Questions were formulated with either the masculine generic or alternative forms, such as the *majuscule-I* or the *slash-form*, or neutral nouns. The authors found that questions containing masculine generics led to significantly fewer answers containing female referents. In a very similar study by Stahlberg et al. (2001), participants filled in questionnaires asking for their favourite protagonists in novels, real life, and history, and for their favourite famous athletes, singers, and politicians, among other categories. Again, results showed that when presented with masculine generics, participants replied significantly less often with female individuals.

Gygax et al. (2008) asked participants to determine whether a presented sentence is a meaningful continuation of a previously shown sentence. Participants were to judge the meaningfulness by considering the masculine generic in the first sentence and the explicitly gendered noun in the second sentence. The authors found that the proportion of positive judgements was higher for male continuations and that there was no effect of stereotypicality. Additionally, reaction times for male continuations were significantly shorter.

Irmen and Kurovskaja (2010) had participants rate sentences in terms of correctness and customariness. Sentences contained either a masculine or feminine role noun as well as either an explicitly masculine or feminine form which referred to the preceding role noun. Sentences with feminine role nouns and gender incongruent referents were rated as less correct and less customary than those with masculine forms and incongruent referents. Additionally, reaction times were slower for sentences with feminine role nouns and gender incongruent referents.

Sato et al. (2016) confronted participants with either two male faces or with a female and a male face. As language stimuli, plural forms of generic masculines were presented. Participants

⁴ *Majuscule-I* refers to a rather new affix which is considered to make word forms more gender-neutral. Commonly, the feminine inflectional suffix *-in* is added to the masculine form with an uppercase *i*. An example is *AnwältInnen* 'lawyers (of both binary sexes)' instead of the generic masculine *Anwälte* 'lawyers'.

⁵ The *slash-form* is yet another rather new alternative for the masculine generic. Here, a slash is added between the masculine form and the feminine inflectional suffix. An example is *Anwält/innen* 'lawyers (of both binary sexes)' instead of the generic masculine *Anwälte* 'lawyers'.

were asked to judge whether a given word form referred to the given faces. The authors found that responses to two male faces were given more quickly than to faces of mixed sex. They concluded that this facilitation of reaction times reflected the ease in interpreting role nouns in the masculine form to be masculine explicit rather than generic.

Misersky et al. (2019) ran an ERP study with sentences in which a role noun introduced a group of people, followed by a congruent or incongruent continuation. Role nouns were either grammatically masculine or feminine; the continuation was congruent if its noun shared the grammatical gender of the preceding role noun. For both types of incongruent continuation (masculine > feminine and feminine > masculine), a P600 was observed. That is, even though the masculine form, in theory, is assumed to be generic, its female continuation led to the same effect as an incongruent continuation of a female form, which, in theory, is not considered to be generic.

In sum, previous research overall agrees on the nature of masculine generics. They show masculine biased readings, resulting in, among other things, a higher percentage of male responses, quicker responses for male continuations, and a lower level of female representation. However, most studies do not come without issues, of which we mention the two most crucial ones here. First, a non-negligible number of studies investigating the nature of the masculine generic make use of students as participants (e.g. Gygax et al., 2008; Heise, 2000; Stahlberg & Sczesny, 2001). Students are not only particularly prone to progressive change (e.g. Bailey & Williams, 2016), but make up a rather low percentage of all language users. Thus, including only students as participants might influence results to an unknown extent. Second, with a few exceptions, studies tend to ignore the potential influence of stereotypes, which might influence the nature of pertinent generic masculine forms.

The aim of the present study is therefore twofold. First, it is investigated whether a masculine bias in masculine generics is found when one's method does not directly rely on participants and their language use specifically elicited for linguistic analysis. This not only allows for doing without the potential influence of specific social groups (e.g. students), but at the same time provides further insight into the semantic nature of masculine generics as well as of masculine and feminine explicit forms in a more general language use. Second, the potential influence of stereotypicality on the masculine bias is accounted for in our statistical analyses to ensure that said bias is not the result of stereotypes. The twofold aim is operationalised using the framework of the Discriminative Lexicon (Baayen, Chuang, Shafaei-Bajestan, et al., 2019; Chuang & Baayen, 2021) with its two computational implementations: naive and linear discriminative learning. The Discriminative Lexicon constitutes a framework which entails that linguistic knowledge and words' features are a product of speakers' experience, and, in turn, resonance processes between entries in the mental lexicon. The computational implementations are introduced in the following subsection.

2.2 Naive discriminative learning

Naive discriminative learning (henceforth NDL; Baayen et al., 2011; Baayen & Ramscar, 2015) is grounded in psychological theory on cognitive mechanisms (Pearce & Bouton, 2001; Rescorla, 1988), which has been shown to successfully model important learning effects in humans and animals (Kamin, 1969; Ramscar et al., 2010). Following the so-called Rescorla-Wagner rules (Rescorla & Wagner, 1972; Wagner & Rescorla, 1972), learning is understood as a result of informative relations within events, which in turn consist of *cues* and *outcomes*. The associations between cues and outcomes are constantly recalibrated when new events are encountered. At any stage of the learning process, the associations of a given outcome and all cues encountered thus far can be taken as the outcome's relation to the world around the learning subject. Association weights are recalibrated in such a way that weights of an association increase every time the involved cue and outcome co-occur, while association weights decrease if a pertinent cue occurs without a given outcome. At the end of this process, a stable state with final association weights is reached. These final outcomes represent the interrelations of a pertinent outcome with all cues encountered during the learning process. Adopting this reasoning to language, cues and outcomes may, for example, be content and function words as well as inflectional and/or derivational functions (e.g. singular vs. plural, specific vs. generic; derivational suffixes like *-ee*, *-ation*, and *-ment*) found in a text corpus annotated according to the needs of the respective investigation. Once the stable end state is reached, each outcome's association weights with all cues constitute the pertinent outcome's semantic vector. In comparison to other models of semantic vector computation, e.g. fastText (Bojanowski et al., 2016) or Word2Vec (Mikolov, Chen, et al., 2013), the vectors computed by NDL are linguistically transparent. For German role nouns in the present investigation, this process is straightforward. For each role noun, vectors of its semantic and formal components, e.g. its base meaning, number, gender, and genericity, will be contained within the vector space computed by NDL. The resulting vector space can be made use of in statistical analyses and also in further computational implementations, such as described in the following subsection.

2.3 Linear discriminative learning

Just like NDL, linear discriminative learning (henceforth LDL; e.g. Baayen, Chuang, Shafaei-Bajestan, et al., 2019) is part of the Discriminative Lexicon. LDL simulates a mental lexicon by generating a system of form-meaning relations by discriminating between different forms and meanings. It allows the researcher to investigate in detail the relationship between entries, i.e. their forms and meanings, in the mental lexicon. Notably, as we will outline in the following paragraphs, LDL networks are simple two-layer networks which are linguistically transparent and interpretable.

In an LDL implementation, forms are represented by numerical vectors. Such form vectors typically consist of binary-coded information on whether certain n -gram or n -phone cues are contained within a given word form. For each word form's individual form vector c , the presence of a n -gram/ n -phone cue is marked with 1, while the absence of such a cue is marked with 0. The form vectors of all word forms of a given set of words constitute the so-called form or cue matrix C , with each row corresponding to the form vector of a pertinent word form and each column representing a unique form cue. As a toy example, let us assume that we have a small corpus of only three German words: *Wind* 'wind', *Kind* 'child', and *Rind* 'bovine'. Using triphones, the corresponding C matrix then looks as follows:

$$(7) \quad C = \begin{array}{rcccccccc} & & \#wI & wIn & Ind & nd\# & \#kI & kIn & \#rI & rIn \\ \begin{array}{l} Wind \\ Kind \\ Rind \end{array} & = & \begin{array}{cccccccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{array} \end{array}$$

Meaning is also represented by numerical vectors. Meaning vectors or semantic vectors can be incorporated from any model generating such vectors, for example, NDL, fastText (Bojanowski et al., 2016), or Word2Vec (Mikolov et al., 2013). The semantic vectors of all word forms and, if applicable, inflectional and derivational functions of a given set of words constitute the so-called meaning or semantic matrix S . In S , each row corresponds to the semantic vector s of a pertinent word form or function, and each column represents a semantic dimension. Most commonly, the semantic vectors of content words are the sum of the semantic vectors of their individual parts. For example, the semantic vector of the word form *Kinder* 'children' is the sum of the semantic vectors of *Kind* 'child' and PLURAL. For our toy example, let us assume the following semantic matrix S :

$$(8) \quad S = \begin{array}{rccc} & Wind & Kind & Rind \\ \begin{array}{l} Wind \\ Kind \\ Rind \end{array} & = & \begin{array}{ccc} 1.0 & 0.4 & 0.1 \\ 0.3 & 1.0 & 0.5 \\ 0.1 & 0.2 & 1.0 \end{array} \end{array}$$

With both the form matrix C and the semantic matrix S available, one can compute comprehension and production by means of multivariate multiple regression. That is, comprehension and production are modelled by simple linear mappings from the form matrix C to the semantic matrix S , comprehension, and from the semantic matrix S to the form matrix C , production. For comprehension, thus, the following equation is solved:

$$(9) \quad S = CF$$

Let C' denote the Moore-Penrose⁶ generalised inverse of C . Then, solving for F :

$$(10) \quad F = C'S$$

F is the transformation matrix used to map C onto S . As a so-called comprehension weight matrix, F specifies how strongly nodes in the C and S matrix are associated. Similarly, for production, the following equation is solved:

$$(11) \quad C = SG$$

For G , then

$$(12) \quad G = S'C$$

Accordingly, for our toy example, we can solve $F = C'S$ with

$$(13) \quad F = \begin{matrix} & \begin{matrix} Wind & Kind & Rind \end{matrix} \\ \begin{matrix} \#wI \\ wIn \\ Ind \\ \#kI \\ kIn \\ \#rI \\ rIn \end{matrix} & \begin{bmatrix} 0.32 & 0.00 & -0.15 \\ 0.32 & 0.00 & -0.15 \\ 0.17 & 0.20 & 0.20 \\ 0.17 & 0.20 & 0.20 \\ -0.02 & 0.30 & 0.05 \\ -0.02 & 0.30 & 0.05 \\ -0.12 & -0.10 & 0.30 \\ -0.12 & -0.10 & 0.30 \end{bmatrix} \end{matrix}$$

and $G = S'C$ with

$$(14) \quad S = \begin{matrix} & \begin{matrix} \#wI & wIn & Ind & nd\# & \#kI & kIn & \#rI & rIn \end{matrix} \\ \begin{matrix} Wind \\ Kind \\ Rind \end{matrix} & \begin{bmatrix} 1.13 & 1.13 & 0.78 & 0.78 & -0.48 & -0.48 & 0.13 & 0.13 \\ -0.31 & -0.31 & 0.34 & 0.34 & 1.24 & 1.24 & -0.59 & -0.59 \\ -0.05 & -0.05 & 0.85 & 0.85 & -0.20 & -0.20 & 1.11 & 1.11 \end{bmatrix} \end{matrix}$$

For the present toy example, CF is exactly equal to S and SG is exactly equal to C . In full-sized implementations of LDL, however, CF and SG are approximations of the S and C matrix due to their high dimensionality. That is,

⁶ The inverse of a matrix needs not exist, rendering such a matrix a singular one. Most matrices used in LDL implementations are singular matrices. Thus, an approximation of the inverse must be used instead of an inverse itself. One such approximation is the Moore-Penrose generalised inverse (Moore, 1920; Penrose, 1955).

$$(15) \quad \hat{S} = CF$$

and

$$(16) \quad \hat{C} = SG$$

While LDL relies on only multivariate multiple regression, previous studies have found that such simple linear mappings result in high overall accuracies (e.g. Baayen, Chuang, Shafaei-Bajestan, et al., 2019; Baayen et al., 2018). Once comprehension is modelled, its accuracy is assessed by comparing a given word's predicted semantic vector to all observed semantic vectors, commonly by using Pearson correlation. If a given word's predicted vector is most highly correlated with its observed vector, the model's prediction is taken to be correct. For production, an additional step is required, as predicted n -grams/ n -phones need to be assembled into potential word forms. As the present study does not make use of the production part of LDL, we refer the interested reader to the detailed accounts in Baayen, Chuang, Shafaei-Bajestan, et al. (2019), Chuang et al. (2020), and Heitmeier et al. (2021).

Finally, LDL offers a variety of measures calculated via observed and predicted matrices which allow for insight into a variety of semantic features. Such measures take into account the predicted semantic vectors, as these vectors are the result of the mapping process. That is, these vectors are assumed to reflect the interrelations of the entries of the simulated mental lexicon. LDL measures have been shown to not only successfully model, e.g. acoustic duration (Chuang et al., 2021; Schmitz et al., 2021; Stein & Plag, 2021), but also real word and pseudoword semantics (Chuang et al., 2021; Schmitz et al., 2021). Because of their successful applications, measures derived from an LDL implementation will be made use of in the analysis of masculine generics as well as masculine and feminine explicit in the present paper.

2.4 Research questions

Because, as far as the authors know, there is no comparable work available on the semantics of role nouns in German in terms of distributional semantics and especially regarding discriminative learning and measures derived from an implementation of LDL, no informed predictions on how semantic vectors or LDL measures may differ between different paradigm member types (see Section 3.1) are given. Instead, the following research questions are investigated:

- RQ1: Do semantic vectors computed via NDL show semantic (dis-)similarities between paradigm member types, in line with the masculine bias found in previous research?
- RQ2: Do measures derived from an implementation of LDL reflect semantic differences between paradigm member types?
- RQ3: Do measures derived from an implementation of LDL successfully predict paradigm member types, even when stereotypicality of paradigms is accounted for?

3. Method

The following section will first introduce the text corpus used as the basis for the naive and linear discriminative learning implementations. Second, the annotation conventions used to process the corpus are illustrated. Third, the implementation of the naive discriminative learning network to train semantic vectors is presented. Finally, the implementation of linear discriminative learning is given.

3.1 Target words

For the present investigation, a set of 120 target words was adopted from Gabriel et al. (2008). In their study, the authors investigated the influence of stereotypical and grammatical information on the presentation of gender in language. For their investigation, they chose roles and their pertinent role nouns which are, from a stereotypical perspective, rather strongly associated with either males or females (e.g. *mason* and *beautician*), as well as role nouns which are neither stereotypically male nor female (e.g. *author*). Thus, Gabriel et al.'s set of items presents the perfect selection of target words for the present paper. If all role nouns, independent of their stereotypical associations, show similarities in terms of their semantics, this makes any potential findings more robust.

As the present investigation is interested in the generic masculine as well as the explicit masculine and the explicit feminine, for each target item, there are three target forms: the generic masculine form, the explicit masculine form, and the explicit feminine form. We call these constellations *target word triplets*. As we include singular and plural forms in our analyses, two sets of triplets per target word are considered. We call these groups of six words *target word paradigms*. **Table 1** illustrates some of the 120 target word triplets. A list of all triplets is part of the data available for this paper.

Table 1: Target word triplets of *Kosmetiker* ‘beautician’, *Sekretär* ‘secretary’, *Jäger* ‘hunter’, and *Professor* ‘professor’.

| English translation | masculine generic | masculine explicit | feminine explicit |
|---------------------|-------------------|--------------------|---------------------|
| ‘beautician’ | <i>Kosmetiker</i> | <i>Kosmetiker</i> | <i>Kosmetikerin</i> |
| ‘secretary’ | <i>Sekretär</i> | <i>Sekretär</i> | <i>Sekretärin</i> |
| ‘hunter’ | <i>Jäger</i> | <i>Jäger</i> | <i>Jägerin</i> |
| ‘professor’ | <i>Professor</i> | <i>Professor</i> | <i>Professorin</i> |

Accordingly, **Table 2** illustrates what a complete target word paradigm looks like for the target word *Kosmetiker* ‘beautician’. Each cell of **Table 2** is what constitutes a TYPE; we will use this

term in the remainder of this paper. Thus, there are six different TYPES per paradigm, i.e. three per triplet.

Table 2: Target word paradigm of *Kosmetiker* ‘beautician’.

| number | masculine generic | masculine explicit | feminine explicit |
|----------|-------------------|--------------------|------------------------|
| singular | <i>Kosmetiker</i> | <i>Kosmetiker</i> | <i>Kosmetikerin</i> |
| plural | <i>Kosmetiker</i> | <i>Kosmetiker</i> | <i>Kosmetikerinnen</i> |

3.2 Corpus

To investigate the research questions of the present paper, a text corpus of German was required in which explicit and generic masculine role nouns were sense disambiguated. To the authors’ knowledge, there is no such corpus available. Thus, such a text corpus of German was created.

To arrive at a feasible corpus, first, sentences were extracted from the Leipzig Corpora Collection’s *news* sub-corpus (Goldhahn et al., 2012). Sampling one million sentences for each year from 2010 to 2019, a total of ten million sentences were extracted. The *news* sub-corpus was chosen to account for general variations. That is, using only texts from news websites, it was ensured that there was no influence on the representations of any masculine or feminine forms due to variations of register or genre across the sampled sentences.

Second, a sample of 830,000 sentences was extracted from the ten million sentences sample. While working with a larger corpus is generally preferable, a huge number of sentences comes with extensive computational costs at later stages of the implementation of naive discriminative learning. To keep the required computational requirements viable, we aimed at a number of sentences close to similar implementations (cf. Baayen, Chuang, Shafaei-Bajestan, et al., 2019).

The 830,000 sentences consisted of two types of sentences. For the first type, 800,000 sentences without target words were sampled. For the second type, sentences containing target words were sampled. During this process, issues with several target items became apparent, which led to the exclusion of seven target word paradigms.⁷ To account for the different frequencies of target word paradigms within our general ten million sentences sample, the target word paradigms were binned into six groups, based on their frequencies within the ten million sentences sample corpus. For each group, then, a set number of attestations was randomly sampled for the final corpus. The numbers of attestations for all frequency groups are given in

⁷ For example, several items included in Gabriel et al. (2008) did not represent masculine generics but gender-neutral forms (e.g. *Hilfskraft* ‘aide’).

Table 3. Note that some target words showed fewer than 100 attestations. In these cases, all attestations were extracted.

Table 3: Frequency groups, number of randomly sampled attestations, and number of target word paradigms per frequency group.

| frequency | | | attestations | number of paradigms |
|-----------------|----|--------|--------------|---------------------|
| up to 200 | | | 100 | 29 |
| 201 | to | 1,000 | 200 | 38 |
| 1,001 | to | 2,000 | 300 | 12 |
| 2,001 | to | 10,000 | 400 | 17 |
| 10,001 | to | 20,000 | 500 | 14 |
| 20,001 and more | | | 600 | 3 |

These 830,000 sentences made up the initial version of the text corpus. While the 800,000 sentences without target words were kept as is, parts of the 30,000 sentences with target words underwent re-sampling during the annotation process. This is explained further in the following subsection.

3.3 Annotation

The text corpus introduced in the previous section was annotated in two ways. First, all sentences were annotated automatically using the RNNTagger software (Schmid, 1999). Using the RNNTagger, inflectional features such as case, number, and tense were annotated. As the present paper is not concerned with derivational processes, no annotation based on derivation was conducted.

Second, the 30,000 sentences containing target words were manually annotated by two authors and two assistants, since, as far as the authors know, there is no automatic annotation software available that successfully sense disambiguates between explicit and generic masculine readings. All annotators were native speakers of German with an educational level comparable to A-levels or higher. Taking into account the context of each target word, the following three features were annotated: gender (masculine vs. feminine), number (singular vs. plural), and genericity (explicit vs. generic). To ensure a high level of inter-annotator agreement, a training set of 300 randomly sampled sentences was annotated before the annotation of the corpus itself. Only in 3 cases, i.e. 1%, did annotators disagree. As a consequence, it was decided that

similar and generally opaque cases were to be documented for discussion and decisions among all annotators. If, after discussion of a given sentence among annotators, it remained unclear whether a target word was used in an explicit or generic manner, this sentence was discarded and a new sentence for the pertinent target word was sampled.

Finally, for the sentences containing target words, the automatic and the manual annotations were brought together. For target words, the manual annotation was kept, while for their sentence surroundings, the automatic annotation was adopted.

3.4 Training semantic vectors

Applying NDL as implemented by the Python package `pyndl` (Sering et al., 2022), the semantic vector space that is used in the remainder of this paper was trained based on the corpus introduced in the previous subsections. Vectors were trained for content and function words as well as for inflectional functions. Crucially, the variable of interest, `GENERICITY`, that is, its two values ‘generic’ and ‘explicit’, were treated as inflectional functions as well. Overall, the semantic vectors were trained on 830,000 sentences to a total of 49,044,960 tokens. Following Baayen, Chuang, Shafaei-Bajestan, et al. (2019), for each sentence of the corpus, each individual base, function word, and inflectional function within the sentence (outcomes) was predicted by the other bases, function words, and inflectional functions (cues) of the same sentence. As a result, semantic vectors not only for words but also for inflectional functions were obtained straightforwardly.

The resulting square matrix was of dimension $30,887 \times 30,887$. The diagonal of the matrix was then set to 0, as the present work focuses on semantic similarity (cf. Baayen, Chuang, Shafaei-Bajestan, et al., 2019). However, as a matrix of such dimensionality requires excessive computational power, the matrix was reduced before entering the next implementational step. Reduction took place by removing those columns whose variance was below the median variance of all columns. Such columns can be removed from the matrix without a loss of accuracy, as their discriminative power is negligible (cf. Baayen, Chuang, Shafaei-Bajestan, et al., 2019). The resulting matrix was of dimension $30,887 \times 15,023$.

3.5 Training a comprehension network

To train an LDL implementation using the `WpmWithLdl` package⁸ (Baayen, Chuang, & Heitmeier, 2019), the semantic vector space created by NDL was taken as a starting point. Based on the

⁸ Note that the `WpmWithLdl` package is no longer maintained. We wish to point those who are interested in implementing NDL or LDL to the `JudiLing` package (Luo et al., 2021) for Julia. The `JudiLing` package is not only steadily maintained but also offers significantly faster computation times. Find the `JudiLing` package here: <https://megamindhenny.github.io/JudiLing.jl/stable/>.

semantic vectors, a semantic matrix S was created. Following the reasoning introduced in 2.3, semantics of content words consisted of the sum of their pertinent parts, e.g. for *Kinder* ‘children’: $Kinder = Kind + PLURAL$. For the focus of the present paper, this meant that masculine generics such as *Anwalt* ‘lawyer’ are represented by their base meaning $Anwalt_{base}$ as well as the function vectors for number (i.e. *SINGULAR* or *PLURAL*), gender (*MASCULINE*), and genericity (*GENERIC*). Masculine explicit, in contrast, contain for genericity the semantic vector of *EXPLICIT*, and feminine explicit consist of their base meaning, a number vector, *FEMININE* for gender, and *EXPLICIT* for genericity.

The semantic matrix was created for 10,222 word forms. We arrived at this number by combining our target paradigm members and their respective semantic vectors, and the semantic vectors for which entries in CELEX (Baayen et al., 1995) were found. The reason entries in CELEX were required is found in the construction of the form matrix. The present study makes use of triphones, as previous studies have found strings of three elements (i.e. phones or grams) to capture the variability of neighbouring phonological information well for a number of languages (e.g. Baayen, Chuang, Shafaei-Bajestan, et al., 2019; Chuang et al., 2021; Milin et al., 2017; Schmitz et al., 2021). Triphones were chosen over trigrams with follow-up research in mind, which requires phone representations (see the discussion for more on future research directions). Using triphones, phonological transcriptions of all word forms were required. These transcriptions were adopted from CELEX.

While the resulting cue matrix of dimension $10,222 \times 9,320$ on its own is still within reasonable size, e.g. concerning resulting computation times, the size of the semantic matrix of dimension $10,222 \times 15,023$ would have led to issues, especially concerning not only computation times but also the hardware resources required. Following Baayen, Chuang, Shafaei-Bajestan, et al. (2019), who found that working with at least 4,000 columns suffices, we reduced the dimensionality of the semantic matrix. We proceeded similarly to the reduction of the matrix given in 3.4. That is, checking all columns for their variance, we disregarded those below the median variance, as they overall contribute least to the accuracy and discriminatory features of vectors. The final semantic matrix was of dimension $10,222 \times 7,511$.

Using the semantic and form matrices, the comprehension part of the LDL network was modelled as introduced in Section 2.3. Then the following measures were extracted:

COMPREHENSION QUALITY. Computing the correlation of a given word’s observed and predicted semantic vector, a measure of **COMPREHENSION QUALITY** is obtained. Higher values indicate a higher **COMPREHENSION QUALITY**.

SEMANTIC NEIGHBOURHOOD DENSITY. To compute a word's SEMANTIC NEIGHBOURHOOD DENSITY, its predicted vector and the semantic vectors of its eight nearest neighbours⁹ were checked for their correlation coefficients. The mean of these coefficients is taken as the value of this measure. Higher values indicate denser semantic neighbourhoods.

SEMANTIC ACTIVATION DIVERSITY. This measure consists of the square root of the sum of the squared values of a given word's predicted vector, i.e. the Euclidean norm of that vector. Higher values imply stronger links to many other entries, indicating a word's SEMANTIC ACTIVATION DIVERSITY.

4. Analyses and results

In the following subsections, first the semantics of generic and explicit forms are compared to investigate RQ1. Then, a closer look at the aforementioned LDL measures is given to explore RQ2. Lastly, LDL measures and stereotypicality are brought together in predicting paradigm member TYPES in line with RQ3, providing further insight into paradigm members' semantic nature.

All analyses were conducted in R (R Core Team, 2021). Data and scripts are available at: <https://osf.io/z6t85/>.

4.1 Semantic similarity across paradigm member TYPES

To analyse the semantic similarity across paradigm member TYPES, we made use of cosine similarity. Cosine similarity expresses the (dis-)similarity of two given vectors with values typically in a range of $[-1,1]$. Higher values indicate a high similarity, while lower values indicate a higher dissimilarity. As semantic vectors reflect words' semantics, cosine similarity values allow for judgements concerning semantic similarities. As such, cosine similarity has been regularly used in the context of distributional semantics (e.g. Huyghe & Wauquier, 2020; Sitikhu et al., 2019). Cosine similarity was computed using the *gds* package (Schmitz & Schneider, 2022).

In the present case, cosine similarities were used to investigate how similar the semantics of two given paradigm member TYPES are. For example, it was used to check how similar the set of vectors of singular masculine generics is to the set of singular masculine explicit. Such a comparison was conducted for all three possible combinations of TYPES across all paradigm triplets within number. The vectors used in this analysis are those which were used as vectors in the observed *S* matrix in the LDL implementation in 3.5. The results of these comparisons are given in **Table 4**.

⁹ Note that the number of neighbours taken into consideration is a parameter. Other studies may use different numbers, e.g. 10 or 20. We chose 8, as this is the default setting of the *WpmWithLdl* package which was used to compute the LDL measures.

Table 4: Mean cosine similarity values and standard deviations for TYPES computed across all paradigm triplets within number. SG = singular; PL = plural.

| | number | masculine explicit | feminine explicit |
|--------------------|--------|--------------------|-------------------|
| masculine generic | SG | 0.996 (9.95e-7) | 0.934 (6.20e-4) |
| | PL | 0.991 (3.56e-6) | 0.822 (7.03e-5) |
| masculine explicit | SG | | 0.939 (6.17e-4) |
| | PL | | 0.835 (6.60e-5) |

Vectors of masculine explicit are most similar to vectors of masculine generics, while vectors of feminine explicit are most similar to vectors of masculine explicit. While this overall pattern is true for both the singular and plural, differences are more pronounced in the latter. Using Wilcoxon-Mann-Whitney-Tests, it was found that the differences between TYPES are highly significant, with $p < 0.001$ in all cases, indicating that masculine generics are most similar to masculine explicit.

While these results already provide a first insight into the semantic nature of the masculine generic, they consist of rather simple statistical analyses of bare semantic vectors. Hence, in the following subsections, the measures extracted from the LDL implementation as well as stereotypicality judgements are used to further analyse the semantic representations of masculine generics. Stereotypicality judgements are incorporated, as previous studies have shown that semantic vectors may be subject to biases as well (e.g. Bolukbasi et al., 2016; Caliskan et al., 2017). The stereotypicality judgements used in our analysis were taken from Gabriel et al. (2008), the same study the target word paradigms for the present investigation were adopted from. Stereotypicality here refers to the assumed extent to which given groups are made up of women or men, while groups were presented to participants by both masculine and feminine German role nouns.

4.2 LDL measures

Let us now have a look at the LDL measures extracted for the individual paradigm member TYPES. The measures, COMPREHENSION QUALITY, SEMANTIC NEIGHBOURHOOD DENSITY, and SEMANTIC ACTIVATION DIVERSITY, are illustrated by **Figure 1**.

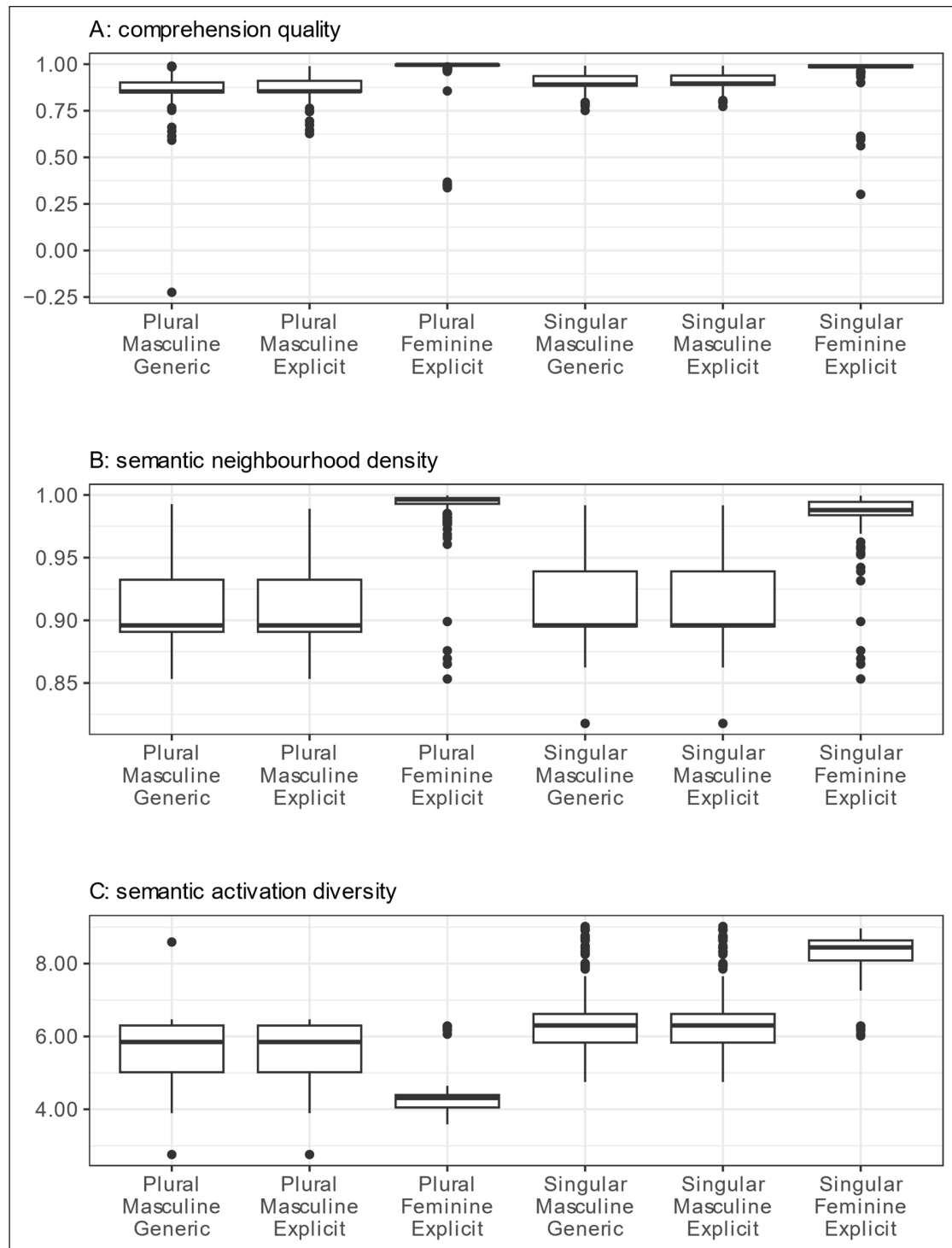


Figure 1: LDL measures per paradigm member TYPE as computed by the LDL implementation. Panel A: COMPREHENSION QUALITY; Panel B: SEMANTIC NEIGHBOURHOOD DENSITY; Panel C: SEMANTIC ACTIVATION DIVERSITY.

Apparently, certain paradigm members show very similar values for certain measures. Applying Wilcoxon-Mann-Whitney-Tests, we took a closer look at the individual values. For all three measures, masculine generics and explicit are highly similar within number, with all p -values above 0.17. Feminine explicit, on the other hand, are significantly different to either TYPE of masculine in either number, with all p -values below 0.001. Feminine explicit are significantly different across number for all measures, with all p -values below 0.001. In sum, the semantic measures apparently reflect semantic differences between paradigm member TYPES.

4.3 Predicting paradigm member TYPES

To obtain greater insight of how the three LDL measures and stereotypicality might influence the masculine bias found in 4.1, we tested whether the measures and ratings were able to successfully predict the six paradigm member TYPES.

First, the independent variables under discussion were checked for their correlation. It was found that the LDL measures COMPREHENSION QUALITY and SEMANTIC NEIGHBOURHOOD DENSITY showed a high correlation coefficient ($\rho = 0.67$). Using highly correlated variables within the same regression model may lead to issues of collinearity, rendering the estimates of the model unreliable (cf. Tomaschek et al., 2018). We thus decided to perform a principal component analysis (PCA; see e.g. Schmitz et al., 2021; Tomaschek et al., 2018). In such an analysis, the dimensionality of the data is reduced by transforming the included variables into principal components. This transformation results in linear combinations of the predictors that are orthogonal to each other, thus not correlated. Both highly correlated LDL measures entered the PCA. Once the PCA is computed, one may decide which principal components to retain for further analysis. For a meaningful decision, we followed three rules of thumb (cf. Baayen, 2008; O'Rourke et al., 2005). First, the eigenvalue of a component should be higher than 1, as an eigenvalue of 1 or more indicates that a component explains more variance than it introduces. Second, the cumulative variance explained by candidate components should at least be higher than 80 % to explain a sufficient amount of variance overall. Third, components are only useful for further analyses if their makeup is interpretable. Following these rules, we retained the first principal component. As for its interpretation, the component is highly positively loaded with both COMPREHENSION QUALITY and SEMANTIC NEIGHBOURHOOD DENSITY. Thus, the component can be considered a measure of quality of comprehension and neighbourhood density, indicating that a higher degree of comprehension quality comes with denser semantic neighbourhoods. We will henceforth call this component SEMANTIC QUALITY & DENSITY.

Then, TYPE entered a multinomial regression analysis as dependent variable using the `nnet` package (Venables & Ripley, 2002), while STEREOTYPICALITY ratings, SEMANTIC ACTIVATION DIVERSITY, as well as SEMANTIC QUALITY & DENSITY were included as predictor variables.

STEREOTYPICALITY did not reach significance, while SEMANTIC ACTIVATION DIVERSITY and SEMANTIC QUALITY & DENSITY reached significance. The coefficients, standard errors, z-values, and p-values of the fitted model are given in **Table 5**. Effect sizes as well as confidence intervals are given in the OSF supplementary material.

Table 5: Summary of the fitted multinomial regression model.

| | Estimate | Std. Error | z-value | p-value |
|-------------------------------|----------|------------|---------|---------|
| Intercept | 2.261 | 0.514 | 4.397 | < 0.001 |
| STEREOTYPICALITY | 0.201 | 0.183 | 1.094 | 0.274 |
| SEMANTIC ACTIVATION DIVERSITY | 2.011 | 0.516 | 3.900 | < 0.001 |
| SEMANTIC QUALITY & DENSITY | -1.408 | 0.311 | -4.532 | < 0.001 |

The model fit is high, with an R^2 value of 0.32. This R^2 value is the so-called McFadden's pseudo R^2 value, i.e. a value which indicates the goodness of fit of a multinomial regression model (McFadden, 1974). According to McFadden (1979), values between 0.2 and 0.4 represent excellent model fit.

Let us now take a closer look at the individual effects of SEMANTIC ACTIVATION DIVERSITY and SEMANTIC QUALITY & DENSITY. The effects of both predictor variables are illustrated in **Figure 2**. For SEMANTIC ACTIVATION DIVERSITY, one finds that plural feminine explicit forms show the highest probability for lowest values (approx. 2.7), while singular feminine explicit forms behave contrarily, with the highest probability for highest values of SEMANTIC ACTIVATION DIVERSITY (approx. 9). Plural masculine forms, that is, both generics and explicit forms, show a very similar pattern in terms of their probabilities. Their highest probability is reached at a semantic activation diversity level of approx. 4.8. Analogously, both singular masculine forms show a very similar pattern as well. Their highest probability is reached at a semantic activation diversity level of approx. 6.9.

For SEMANTIC QUALITY & DENSITY, masculine forms within number also display very similar patterns. Plural masculine forms show their highest probability at the lower end of SEMANTIC QUALITY & DENSITY, at approx. -7.3. Singular masculine forms show their highest probability at the higher end of SEMANTIC QUALITY & DENSITY, at approx. 1.3. The feminine forms here also show a similar pattern. However, their probabilities are below those of the masculine forms for all values of SEMANTIC QUALITY & DENSITY. Their highest probability, which is only approx. 0.07, is reached at a SEMANTIC QUALITY & DENSITY value of approx. 1.3.

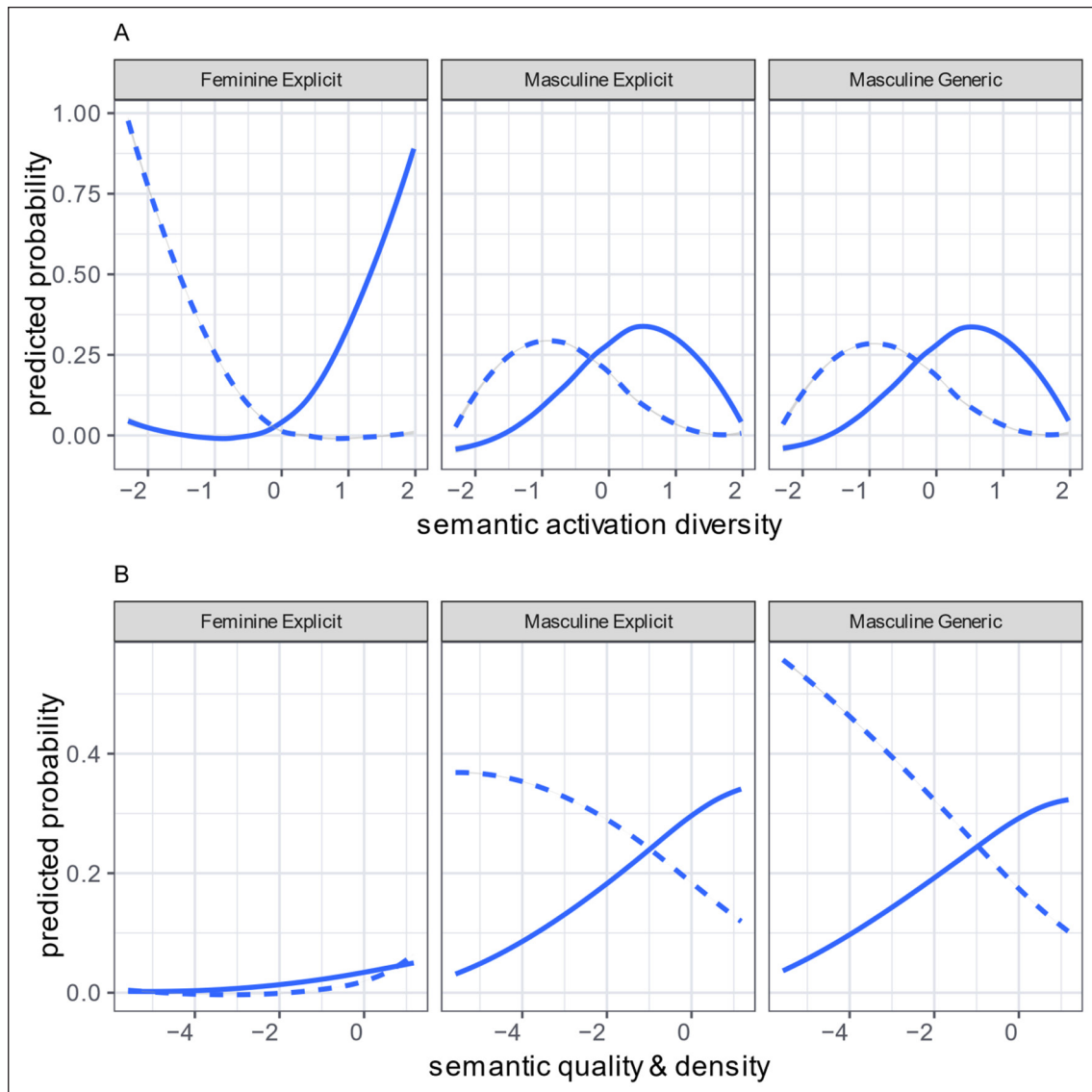


Figure 2: Predicted probabilities of the six paradigm member TYPES as modelled by SEMANTIC ACTIVATION DIVERSITY (Panel A) and SEMANTIC QUALITY & DENSITY (Panel B). Dashed lines = singular; solid lines = plural.

Note that the multinomial regression analysis of TYPE disregarded the interrelations of the six paradigm member TYPES. Three TYPES are singular, three TYPES are plural. Four TYPES are masculine, two TYPES are feminine. Two TYPES are generic, four TYPES are explicit. Hence, the levels of TYPE are not independent of each other but related, due to shared features. To check whether these relations show an influence on the results presented thus far and, if so, to see what the nature of these influences is, we fitted three separate logistic regression models. Each model was fitted to predict one of the three features: NUMBER, GENDER, and GENERICITY. Analogously to the multinomial regression model, STEREOTYPICALITY, SEMANTIC ACTIVATION DIVERSITY, and

SEMANTIC QUALITY & DENSITY were included as predictor variables alongside the two features which were not the dependent variable of the pertinent logistic regression model. For instance, NUMBER was predicted by the three variables of interest and by GENDER and GENERICITY.

For the prediction of NUMBER, it was found that the LDL measures, SEMANTIC ACTIVATION DIVERSITY and SEMANTIC QUALITY & DENSITY, show significant effects. The higher the values of these measures, the more likely a form is to be singular. For the prediction of GENDER, it was found that SEMANTIC QUALITY & DENSITY shows a significant effect. The higher the value of this predictor is, the more likely a form is masculine. For the prediction of GENERICITY, no significant effects were found. While this may seem surprising at first, a closer look at the paradigm member straightforwardly explains this finding. First, there are masculine generics but no feminine generics. Second, taking into account the observed values of the LDL measures, the two levels of GENERICITY, explicit and generic, appear to be almost identical. Hence, these variables cannot help disambiguate explicit and generic forms. In sum, the three individual analyses of NUMBER, GENDER, and GENERICITY are in line with the results of the multinomial regression analysis of TYPE. The coefficients, standard errors, z-values, and p-values of the three fitted generalised linear models are given in **Table 6**; the effect sizes are part of the OSF supplementary material.

Table 6: Summary of the three logistic regression models.

| dependent variable: NUMBER | Estimate | Std. Error | z-value | p-value |
|-----------------------------------|-----------------|-------------------|----------------|----------------|
| Intercept (baseline: plural) | -0.150 | 0.352 | -0.426 | 0.670 |
| STEREOTYPICALITY | -0.002 | 0.102 | -0.016 | 0.988 |
| SEMANTIC ACTIVATION DIVERSITY | 2.578 | 0.226 | 11.403 | 0.000 |
| SEMANTIC QUALITY & DENSITY | 0.438 | 0.152 | 2.871 | 0.004 |
| GENERICITY.G | -0.001 | 0.221 | -0.005 | 0.996 |
| GENDER.M | 0.676 | 0.398 | 1.700 | 0.089 |
| dependent variable: GENDER | Estimate | Std. Error | z-value | p-value |
| Intercept (baseline: feminine) | 0.341 | 0.258 | 1.321 | 0.187 |
| STEREOTYPICALITY | 0.133 | 0.124 | 1.068 | 0.286 |
| SEMANTIC ACTIVATION DIVERSITY | -0.252 | 0.187 | -1.349 | 0.177 |
| SEMANTIC QUALITY & DENSITY | -1.863 | 0.163 | -11.401 | 0.000 |
| GENERICITY.G | 18.907 | 616.001 | 0.031 | 0.976 |
| NUMBER.S | 0.579 | 0.450 | 1.286 | 0.198 |

(Contd.)

| dependent variable: GENERICITY | Estimate | Std. Error | z-value | p-value |
|---------------------------------------|-----------------|-------------------|----------------|----------------|
| Intercept (baseline: explicit) | -19.511 | 714.964 | -0.027 | 0.978 |
| STEREOTYPICALITY | 0.008 | 0.095 | 0.085 | 0.932 |
| SEMANTIC ACTIVATION DIVERSITY | 0.028 | 0.146 | 0.190 | 0.849 |
| SEMANTIC QUALITY & DENSITY | -0.073 | 0.126 | -0.579 | 0.563 |
| GENDER.M | 19.482 | 714.964 | 0.027 | 0.978 |
| NUMBER.S | 0.002 | 0.211 | 0.008 | 0.994 |

In sum, the present analyses found that STEREOTYPICALITY does not have a significant influence on any TYPE of paradigm member. The LDL measures, however, show significant effects. Concerning the quality of comprehension and semantic neighbourhood density, as entailed in SEMANTIC QUALITY & DENSITY, singular masculine forms are found to have the highest values. Masculine plurals, however, show an opposite picture. Concerning SEMANTIC ACTIVATION DIVERSITY, plural feminines are found to have the highest values. Feminine singulars, however, show an opposite effect. Thus, SEMANTIC QUALITY & DENSITY, and hence its underlying measures, appear to modulate masculine forms more extensively than feminine forms, while SEMANTIC ACTIVATION DIVERSITY clearly distinguishes feminine forms but masculine forms to a far lesser extent.

5. Discussion and conclusion

The goal of the present paper was to explore the semantic nature of masculine generics and their semantic relations to masculine and feminine explicit forms by means of naive and linear discriminative learning. Importantly, the present investigation made use of corpus data instead of language data elicited specifically for linguistic analyses, as well as of stereotypicality judgements to account for a potential influence of stereotypes. In total, three research questions were investigated.

Regarding RQ1, it was found that comparing semantic vectors computed with NDL via cosine similarities, semantic (dis-)similarities between paradigm member types can be observed. Masculine generic and explicit forms were highly similar in the singular and the plural. Explicit feminine forms were significantly different when compared to either masculine form, but explicit feminines were more similar to explicit masculines than to generic masculines. This is interesting for two reasons. First, from a general perspective, this seems counterintuitive if one assumes the masculine generic to be a gender-neutral form. If it was gender-neutral, it should be as similar to masculine explicit forms as to feminine explicit forms. Second, from a computational perspective, this suggests that the explicit vector, which is part of both the masculine and the feminine explicit forms, shifts these forms towards the same direction within the vector space, leading to the feminine explicit forms' greater similarity to masculine explicit forms than to masculine generics.

In light of RQ2, it was found that semantic measures derived from our implementation of LDL reflect the semantic (dis-)similarities found for RQ1. For all three measures under investigation, COMPREHENSION QUALITY, SEMANTIC NEIGHBOURHOOD DENSITY, and SEMANTIC ACTIVATION DIVERSITY, it was found that masculine generics and explicit forms are highly similar within number. Feminine explicit forms, however, show not only significantly different values for all three measures when compared with either masculine form, but also when compared with feminine explicit forms of different number.

In light of RQ3, we made use of the aforementioned LDL measures (with SEMANTIC NEIGHBOURHOOD DENSITY and SEMANTIC ACTIVATION DIVERSITY as parts of one principal component) and stereotypicality judgements to predict the six different paradigm member types in a multinomial regression analysis. For stereotypicality judgements, no significant effect was found. Hence, the effects of the LDL measures, as presented in 4.3 and as discussed in the following, are not confounded by stereotypicality biases.

For SEMANTIC ACTIVATION DIVERSITY, it was found that masculine forms, both explicit forms and generics, show similar effects, with singular forms being related to somewhat higher values than plural forms. Plural feminine explicit forms show the highest probabilities with the lowest values of SEMANTIC ACTIVATION DIVERSITY, while singular feminine explicit forms show the highest probabilities with the highest values of SEMANTIC ACTIVATION DIVERSITY. These differences in SEMANTIC ACTIVATION DIVERSITY appear to reflect some sort of form competition. With most feminine singular role nouns ending in *-in*, they share their final triphone or trigram with entries in the lexicon which are not role nouns, e.g. *Kinn* ‘chin’. Across all word forms used in the present implementation, 235 end in *-in*. Out of these 235 word forms, 47% are feminine singular target words. Hence, this form component is subject to competition, which in turn leads to higher levels of uncertainty and thus a higher degree of SEMANTIC ACTIVATION DIVERSITY. Contrarily, the form of feminine plural role nouns is a good cue for their feminine feature, as the *-innen* part is not found in many other words. In our data, 96% of words ending in *-innen* are feminine plural target words. Hence, there is less uncertainty and therefore a lower degree of SEMANTIC ACTIVATION DIVERSITY. For masculine role nouns, the situation is somewhat different. They show different endings, and these endings are also found in words which are not role nouns (e.g. *Bohrer* ‘drill’). In our data, 33% of words ending in *-eur*, 28% of words ending in *-er*, 20% of words ending in *-or*, 12% of words ending in *-nt*, 9% of words ending in *-ist*, and 7% of words ending in *-ar* are masculine target words. On average, masculine role noun targets make up approx. 18% of the words with similar endings. While this value is clearly lower than that of feminine singular role nouns, we assume that the variable nature of the pertinent endings (e.g. *-er* is used not only for agent nouns but also for, among other things, instance nouns, plurals, comparatives and other adjective inflection) lowers the degree of form competition. Thus, the degree of SEMANTIC ACTIVATION DIVERSITY in masculine role nouns is somewhat at a medium level.

In terms of COMPREHENSION QUALITY and SEMANTIC NEIGHBOURHOOD DENSITY, as combined as SEMANTIC QUALITY & DENSITY, it was found that plural masculine forms show highest probabilities for lower values, while singular masculine forms show highest probabilities

for higher values. For feminine forms, SEMANTIC QUALITY & DENSITY shows overall little predictive value. The differences found between masculine and feminine role nouns indicate that feminine role nouns ‘live’ in their own semantic space. To see whether this is indeed the case, we made use of t-Distributed Stochastic Neighbour Embedding (t-SNE) (Maaten & Hinton, 2008), a dimension reduction technique that can map high-dimensional data on just two dimensions. Visualising the resulting two dimensions has proven highly successful in cluster-detection (Arora et al., 2018). For the analysis of the word vectors, we follow Shafaei-Bajestan et al. (2022) in adapting their t-SNE settings¹⁰ for use with the Rtsne package (Krijthe, 2015) and the gdsrn package (Schmitz & Schneider, 2022). The resulting two dimensions for the entire predicted semantic matrix of the current LDL implementation are shown in **Figure 3**.

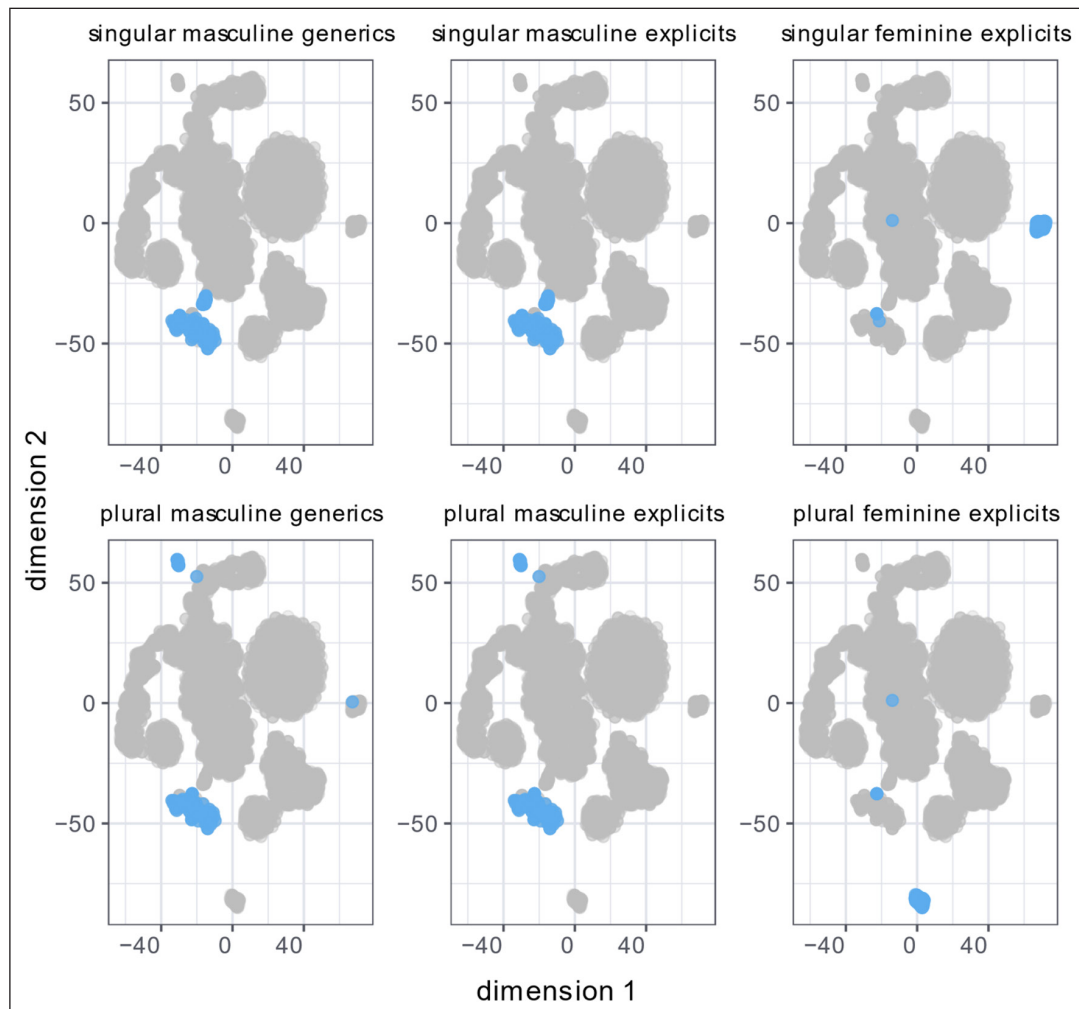


Figure 3: Display of the two dimensions computed with the t-SNE technique. Each panel highlights the locations of target words belonging to one of the six paradigm member types in blue; the highlighted type is given on top of the panels. Grey dots represent all other entries of the lexicon.

¹⁰ The adapted t-SNE settings are as follows: perplexity = 35, number of iterations = 4000, exaggeration factor = 12, learning rate = 200, and initialisation = random.

Unsurprisingly, masculine forms cluster together across genericity and number, with only a few exceptions between number. Singular feminine explicit and plural feminine explicit cluster as well but not across number, again with only a few exceptions. Notably, the masculine clusters are closer to other entries of the lexicon than both feminine clusters are. These findings support the initial idea of feminine forms living in their own remote area of the semantic space. Thus, not only their form (as discussed above) but also the meaning of their relevant form component, i.e. the feminine gender suffix, leads to significant differences between feminine and masculine forms. In other words, feminine role nouns show an interpretable exponent of their grammatical gender, which in turn is connected to a shift in semantic space, as is illustrated by the semantic measures derived via LDL.

In sum, the present paper demonstrated that masculine vectors, even when trained on a corpus with differentiated semantics for masculine explicit and generic, show very similar vectors and thus have very similar semantics. Feminine explicit, on the other hand, are less semantically similar. This finding is true for both the singular and the plural. Using measures derived from an implementation of LDL, the semantic differences between masculine and feminine role nouns were further explored. Due to their makeup, feminine forms ‘live’ in their own area of the semantic space with significantly different degrees of competition.

As grammarians have postulated that the masculine is the ‘generic gender’ in German, one would hope that the difference between an explicit and a generic masculine would be learnable. Hence, genericity with its two levels, explicit and generic, should be representable in the form of a function, as was done in the present paper. However, genericity is not formally marked – explicit and generic masculines share their surface representation.¹¹ Thus, their activation diversities and their neighbourhood densities are similar, if not identical. Authors of previous research on the nature of masculine generics already hinted at the reasons for the masculine generic’s bias. Stahlberg et al. (2001) assume that masculine generics have a semantic component of ‘maleness’ due to their similar form and masculine grammatical gender, while Irmen and Linner (2005) speak of a *Resonanzprozess* ‘process of resonance’, in which masculine generics are influenced by the resonance of masculine explicit in and with the lexicon. Sato et al. (2016), in line with Gygas et al. (2012) and (2021), argue that even though the masculine generic form should function purely as grammatically masculine, it is nonetheless semantically linked to the masculine explicit. Indeed, a post-hoc analysis shows that the predicted vectors of masculine explicit are just as similar to the observed vectors of masculine explicit as to the observed vectors of masculine generic (Wilcoxon-Mann-Whitney-Tests, $p = 0.32$). Hence, it appears that masculine explicit and generic are semantically not distinguishable.

¹¹ The orthographic and phonological representations of masculine explicit and generic are identical. However, there is no account of whether their phonetic forms are identical as well.

The present findings, as well as the discussion in Sato et al. (2016), give rise to a question that lends itself to be investigated with the present methodology. That is, how do other forms that are supposedly gender-neutral perform? As an example, take the rather new ‘gender star’ form, which inserts an asterisk before the feminine gender suffix, e.g. *Professor*in* ‘professor (of any sex or gender)’. The asterisk is realised as glottal stop, a phone not uncommon in German, but indeed rather unusual as the onset consonant of a suffix. Using triphones for form representations in the LDL setup of the present paper, an integration of such alternative forms is straightforward and should thus be a goal of future research.

To summarise, masculine generics and explicit forms show highly similar semantic features, while feminine forms live in their own parts of the semantic space. Thus, when a generically intended masculine form is encountered, its explicit masculine counterpart is co-activated to a high degree – its feminine counterpart is not. This, in turn, is an explanation for the masculine bias in masculine generics observed in previous studies and the present one. Overall, this paper brought forward a robust case for the masculine bias of masculine generics in German, controlling for both the influence of specifically elicited language data by participants from single social groups and the influence of societal stereotypes.

Abbreviations

Glossing abbreviations follow the Leipzig Glossing Rules.

| | |
|------|-----------------------|
| 3SG | third person singular |
| ACC | accusative |
| ADJ | adjective |
| ADV | adverb |
| DAT | dative |
| DEF | definite |
| DET | determiner |
| F | female |
| GEN | genitive |
| INDF | indefinite |
| M | masculine |
| NOM | nominative |
| PL | plural |
| PRS | present |
| PTCP | participle |
| SG | singular |

Data accessibility statement

The data and the script to analyse the data with are available at: <https://osf.io/z6t85/>.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for the partial funding of this research (Grant PL 151/11-1 ‘Semantics of derivational morphology’ to Ingo Plag). The authors are grateful to Laureen Schmitz and Dennis Dahlhausen for assisting in the manual annotation of sentences, and to the Department of English Language and Linguistics at Heinrich Heine Universität Düsseldorf as well as to the audiences of the 28th Germanic Linguistics Annual Conference, the Second International Conference on Error-Driven Learning in Language, InSemantiC 2022, the 3rd International Twitter Conference on Linguistics, and the 45. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft for valuable feedback.

Competing interests

The authors have no competing interests to declare.

Author contributions

DS and JE were responsible for the conceptualisation of the study. DS and VS implemented the discriminative networks and extracted the required data. DS and JE were responsible for data curation. DS conducted the statistical analysis with feedback by VS and JE. DS wrote the manuscript; VS and JE reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

References

- Arora, S., Hu, W., & Kothari, P. K. (2018). *An analysis of the t-SNE algorithm for data visualization*. ArXiv. DOI: <https://doi.org/10.48550/arXiv.1803.01768>
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511801686>
- Baayen, R. H., Chuang, Y.-Y., & Blevins, J. P. (2018). Inflectional morphology with linear mappings. *The Mental Lexicon*, 13(2), 230–268. DOI: <https://doi.org/10.1075/ml.18010.baa>
- Baayen, R. H., Chuang, Y.-Y., & Heitmeier, M. (2019). *WpmWithLdl: Implementation of word and paradigm morphology with linear discriminative learning* (1.3.17.1) [R package].
- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, 2019, 4895891. DOI: <https://doi.org/10.1155/2019/4895891>
- Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118(3), 438–481. DOI: <https://doi.org/10.1037/a0023851>
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. University of Philadelphia.
- Baayen, R. H., & Ramscar, M. (2015). Abstraction, storage and naive discriminative learning. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 100–120). De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110292022-006>
- Bailey, M., & Williams, L. R. (2016). Are college students really liberal? An exploration of student political ideology and attitudes toward policies impacting minorities. *The Social Science Journal*, 53(3), 309–317. DOI: <https://doi.org/10.1016/j.soscij.2016.04.002>
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. DOI: https://doi.org/10.1162/tacl_a_00051
- Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Proceedings of the 30th International Conference on Neural Information Processing Systems*, 4356–4364.

- Braun, F., Gottburgsen, A., Sczesny, S., & Stahlberg, D. (1998). Können Geophysiker Frauen sein? Generische Personenbezeichnungen im Deutschen. *Zeitschrift für Germanistische Linguistik*, 26(3), 265–283. DOI: <https://doi.org/10.1515/zfgl.1998.26.3.265>
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186. DOI: <https://doi.org/10.1126/science.aal4230>
- Chuang, Y.-Y., & Baayen, R. H. (2021). Discriminative learning and the lexicon: NDL and LDL. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. DOI: <https://doi.org/10.1093/acrefore/9780199384655.013.375>
- Chuang, Y.-Y., Lõo, K., Blevins, J. P., & Baayen, R. H. (2020). Estonian case inflection made simple: A case study in Word and Paradigm Morphology with Linear Discriminative Learning. In L. Körtvélyessy & P. Štekauer (Eds.), *Complex words* (pp. 119–141). Cambridge University Press. DOI: <https://doi.org/10.1017/9781108780643.008>
- Chuang, Y.-Y., Vollmer, M. L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., & Baayen, R. H. (2021). The processing of pseudoword form and meaning in production and comprehension: A computational modeling approach using linear discriminative learning. *Behavior Research Methods*, 53(3), 945–976. DOI: <https://doi.org/10.3758/s13428-020-01356-w>
- Doleschal, U. (2002). Das generische Maskulinum im Deutschen. Ein historischer Spaziergang durch die deutsche Grammatikschreibung von der Renaissance bis zur Postmoderne. *Linguistik Online*, 11(2). DOI: <https://doi.org/10.13092/lo.11.915>
- Gabriel, U., Gygax, P., Sarrasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, 40(1), 206–212. DOI: <https://doi.org/10.3758/BRM.40.1.206>
- Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building large monolingual dictionaries at the Leipzig Corpora Collection: From 100 to 200 languages. *Proceedings of the 8th International Language Resources and Evaluation (LREC'12)*.
- Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23(3), 464–485. DOI: <https://doi.org/10.1080/01690960701702035>
- Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2009). Some grammatical rules are more difficult than others: The case of the generic interpretation of the masculine. *European Journal of Psychology of Education*, 24(2), 235–246. DOI: <https://doi.org/10.1007/BF03173014>
- Gygax, P., Gabriel, U., Lévy, A., Pool, E., Grivel, M., & Pedrazzini, E. (2012). The masculine form and its competing interpretations in French: When linking grammatically masculine role names to female referents is difficult. *Journal of Cognitive Psychology*, 24(4), 395–408. DOI: <https://doi.org/10.1080/20445911.2011.642858>
- Gygax, P., Sato, S., Öttl, A., & Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: A challenge for our cognitive system. *Language Sciences*, 83, 101328. DOI: <https://doi.org/10.1016/j.langsci.2020.101328>

- Heise, E. (2000). Sind Frauen mitgemeint? Eine empirische Untersuchung zum Verständnis des generischen Maskulinums und seiner Alternativen. *Sprache & Kognition*, 19(1/2), 3–13. DOI: <https://doi.org/10.1024//0253-4533.19.12.3>
- Heitmeier, M., Chuang, Y.-Y., & Baayen, R. H. (2021). Modeling morphology with linear discriminative learning: Considerations and design choices. *Frontiers in Psychology*, 12, 4929. DOI: <https://doi.org/10.3389/fpsyg.2021.720713>
- Huyghe, R., & Wauquier, M. (2020). What's in an agent? A distributional semantics approach to agent nouns in French. *Morphology*, 30(3), 185–218. DOI: <https://doi.org/10.1007/s11525-020-09366-2>
- Irmen, L., & Köhncke, A. (1996). Zur Psychologie des 'generischen' Maskulinums. *Sprache & Kognition*, 15, 152–166.
- Irmen, L., & Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology*, 57(5), 367–375. DOI: <https://doi.org/10.1027/1618-3169/a000044>
- Irmen, L., & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen. *Zeitschrift für Psychologie/Journal of Psychology*, 213(3), 167–175. DOI: <https://doi.org/10.1026/0044-3409.213.3.167>
- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). Appleton-Century-Crofts.
- Krijthe, J. H. (2015). *Rtsne: t-distributed Stochastic Neighbor Embedding using a Barnes-Hut implementation* (0.16) [R package]. <https://github.com/jkrijthe/Rtsne>
- Luo, X., Chuang, Y.-Y., & R. H. Baayen. (2021). *JudiLing: An implementation in Julia of linear discriminative learning algorithms for language model* (0.7.0) [Julia package].
- Maaten, L. van der, & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(86), 2579–2605.
- McFadden, D. L. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press.
- McFadden, D. L. (1979). Quantitative methods for analysing travel behaviour of individuals: Some recent developments. In D. Hensher & P. Stopher (Eds.), *Behavioural travel modelling* (pp. 279–318). Routledge. DOI: <https://doi.org/10.4324/9781003156055-18>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations, ICLR 2013 – Workshop Track Proceedings*. DOI: <https://doi.org/10.48550/arxiv.1301.3781>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS'13) – Volume 2*. <http://arxiv.org/abs/1310.4546>
- Milin, P., Divjak, D., & Baayen, R. H. (2017). A learning perspective on individual differences in skilled reading: Exploring and exploiting orthographic and semantic discrimination cues. *Journal*

- of *Experimental Psychology: Learning, Memory, and Cognition*, 43(11), 1730–1751. DOI: <https://doi.org/10.1037/xlm0000410>
- Misersky, J., Majid, A., & Snijders, T. M. (2019). Grammatical gender in German influences how role-nouns are interpreted: Evidence from ERPs. *Discourse Processes*, 56(8), 643–654. DOI: <https://doi.org/10.1080/0163853X.2018.1541382>
- Moore, H. E. (1920). On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society*, 26, 394–395. DOI: <https://doi.org/10.1090/S0002-9904-1920-03332-X>
- O'Rourke, N., Hatcher, L., & Stepanski, E. J. (2005). *Using SAS for univariate & multivariate statistics*. SAS Institute Inc.
- Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, 52(1), 111–139. DOI: <https://doi.org/10.1146/annurev.psych.52.1.111>
- Penrose, R. (1955). A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3), 406–413. DOI: <https://doi.org/10.1017/S0305004100030401>
- R Core Team. (2021). *R: A language and environment for statistical computing* (4.0.4). R Foundation for Statistical Computing. <https://www.r-project.org/>
- Ramsar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, 34(6), 909–957. DOI: <https://doi.org/10.1111/j.1551-6709.2009.01092.x>
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43(3), 151–160. DOI: <https://doi.org/10.1037/0003-066X.43.3.151>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Rothermund, K. (1998). Automatische geschlechtsspezifische Assoziationen beim Lesen von Texten mit geschlechtseindeutigen und generisch maskulinen Text-Subjekten. *Sprache & Kognition*, 17, 183–198.
- Rothmund, J., & Scheele, B. (2004). Personenbezeichnungsmodelle auf dem Prüfstand. *Zeitschrift für Psychologie/Journal of Psychology*, 212(1), 40–54. DOI: <https://doi.org/10.1026/0044-3409.212.1.40>
- Sato, S., Gygas, P. M., & Gabriel, U. (2016). Gauging the impact of gender grammaticization in different languages: Application of a linguistic-visual paradigm. *Frontiers in Psychology*, 7. DOI: <https://doi.org/10.3389/fpsyg.2016.00140>
- Schmid, H. (1999). Improvements in part-of-speech tagging with an application to German. In S. Armstrong, K. Church, P. Isabelle, S. Manzi, E. Tzoukermann, & D. Yarowsky (Eds.), *Natural language processing using very large corpora* (pp. 13–25). Springer. DOI: https://doi.org/10.1007/978-94-017-2390-9_2
- Schmitz, D., Plag, I., Baer-Henney, D., & Stein, S. D. (2021). Durational differences of word-final /s/ emerge from the lexicon: Modelling morpho-phonetic effects in pseudowords with linear discriminative learning. *Frontiers in Psychology*, 12. DOI: <https://doi.org/10.3389/fpsyg.2021.680889>

- Schmitz, D., & Schneider, V. (2022). *gdsm: General functions for Distributional SeMantics* (0.1) [R Package]. <https://github.com/dosc91/gdsm>
- Sering, K., Weitz, M., Shafaei-Bajestan, E., & Künstle, D.-E. (2022). pyndl: Naïve Discriminative Learning in Python. *Journal of Open Source Software*, 7(80), 4515. DOI: <https://doi.org/10.21105/joss.04515>
- Shafaei-Bajestan, E., Moradipour-Tari, M., Uhrig, P., & Baayen, R. H. (2022). *Semantic properties of English nominal pluralization: Insights from word embeddings*. ArXiv. <https://arxiv.org/abs/2203.15424v1>
- Sitikhu, P., Pahi, K., Thapa, P., & Shakya, S. (2019). A comparison of semantic similarity methods for maximum human interpretability. *International Conference on Artificial Intelligence for Transforming Business and Society*. DOI: <https://doi.org/10.1109/AITB48515.2019.8947433>
- Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen. *Psychologische Rundschau*, 52(3), 131–140. DOI: <https://doi.org/10.1026//0033-3042.52.3.131>
- Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name your favorite musician. *Journal of Language and Social Psychology*, 20(4), 464–469. DOI: <https://doi.org/10.1177/0261927X01020004004>
- Stein, S. D., & Plag, I. (2021). Morpho-phonetic effects in speech production: Modeling the acoustic duration of English derived words with linear discriminative learning. *Frontiers in Psychology*, 12. DOI: <https://doi.org/10.3389/fpsyg.2021.678712>
- Tomaschek, F., Hendrix, P., & Baayen, R. H. (2018). Strategies for addressing collinearity in multivariate linguistic data. *Journal of Phonetics*, 71, 249–267. DOI: <https://doi.org/10.1016/j.wocn.2018.09.004>
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer. DOI: <https://doi.org/10.1007/978-0-387-21706-2>
- Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–334). Academic Press Inc.

