Dominic Schmitz

Pronoun comprehension from a discriminative perspective: A proof of concept

Abstract: The English pronoun system is not as straightforward as often stated by grammarians. The third-person singular pronouns *he* and *she* are not only used to refer to male and female individuals, but are also used to refer to individuals of any gender. Similarly, the pronoun *they*, commonly found as a third-person plural pronoun, is increasingly often used in its generic sense to refer to single individuals of any gender. Yet, there is little research available on the semantics and the comprehension of such specific and generic pronouns. Making use of naive discriminative learning, instance vectors, and linear discriminative learning, the present paper proposes a novel computational approach to the study of pronoun semantics and comprehension. As an exemplary case, the comprehension of generic *they* is compared to the comprehension of specific *he*, specific *she*, and plural *they*. Measures extracted from the comprehension process deliver meaningful insights into the comprehension of the given pronouns, rendering the proposed methodology a promising one for future research.

Keywords: discriminative learning, distributional semantics, pronoun comprehension, pronoun semantics

1 Introduction: Third-person pronouns in contemporary English

In contemporary English, "[g]ender classes can be differentiated only on the basis of relations with pronouns" (Huddleston and Pullum 2002: 485). Pronouns are functional elements of a language which are used to replace or serve the function of noun phrases. In their most common use, pronouns refer directly to entities in the world and are, with that, assumed to adopt the meaning of the entities they refer to (Conrod 2020).

The pronominal gender system of English traditionally differentiates three grammatical genders: masculine (*he*), feminine (*she*), and neuter (*it*). The three

Dominic Schmitz, Department of English Language and Linguistics, Heinrich Heine University Düsseldorf, e-mail: dominic.schmitz@uni-duesseldorf.de

https://doi.org/10.1515/9783111388694-006

grammatical genders are distinguished semantically, based on the two features of humanness and the biological sex of the relevant referent (Siemund 2008), making the gender system a covert one in which morphophonological information cannot be used to determine a noun's gender (Corbett 1991). Following the semantic basis, male antecedents are referred to by masculine pronouns (e.g., *the man and **his** dog*), female antecedents are referred to by feminine pronouns (e.g., *the woman and **her** dog*), and all other antecedents are referred to by neuter pronouns (e.g., *the house and **its** roof*).

However, while the gender system predominantly functions on the basis of humanness and biological sex, there is a considerable grey area (Siemund 2008; Audring 2009; McConnell-Ginet 2015; Conrod 2019). For example, within groups of homosexual men, the use of *she* instead of *he* to refer to other group members has been observed to express positive as well as negative character traits (Rudes and Healy 1979). Mathiot (1979) reports on the use of *he* instead of *she* for women, most likely used to highlight their competence, hinting at the assumed superiority of men. Both examples cast doubt on the absolute nature of biological sex as the definitive feature of gender in English. There are also cases which question the definitive nature of humanness as basis of gender assignment; several of them are mentioned by Siemund (2008). Usage of *he* and *she* instead of *it* is frequently found for animals, especially for higher domestic animals and for animals for which the outer appearance clearly identifies biological sex (e.g., lions). In some cases, the usage of *he* and *she* has become highly conventionalized, ignoring the actual biological sex of the pertinent animal: Dogs are referred to by *he* and cats by *she*, respectively. A case of reversed nature is the use of *it* to refer to babies, toddlers, and children. Finally, the pronouns *he* and *she* may even be used to refer to antecedents which are neither human nor show biological sex, be they of supernatural nature (e.g., gods, apparitions; Poutsma 1914) or inanimate. An example of the latter case are ships and countries, for which speakers regularly use *she* instead of *it*. While the exact nature of this kind of pronoun use is still unclear, one idea is that the use of *he* and *she* personifies referents, i.e., it shows their importance, while the use of *it* for humans underlyingly shows the importance of biological sex within society (Siemund 2008). In sum, humanness and biological sex apparently do neither straightforwardly nor exhaustively capture the variation of usage within the English pronominal gender system.

A further case in which the typically associated gender of a pronoun and the biological sex of referents do not coincide is the use of epicene pronouns.[1] Epicenes

1 Yet another case is the use of neopronouns. However, this case is beyond the scope of the present study.

are understood as pronouns that can refer to all individuals regardless of gender, as epicenes supposedly give no information or notional content on the referent's biological sex (e.g. Baron 1981; Newman 1997; Baranowski 2002). When used as epicenes, *he* and *she* are typically called 'generic *he*' and 'generic *she*', with the generic use of *she* being far less frequent than that of *he* (Hekanaho 2020; Conrod 2020). Examples for generic *he* and *she* are given in Examples (1) and (2).

(1) They voted for whoever looked like **he** would finally fix the plumbing. (OED 2023: s.v. <he>)

(2) When someone works for less than **she** can live on..then **she** has made a great sacrifice for you. (OED 2023: s.v. <she>)

Traditionally, grammarians have declared the masculine pronoun, i.e., *he*, to be the default unmarked gender of English for more than two centuries (Corbett 1991). However, regardless of its stipulated gender-neutral nature, the use of *he* as default epicene does not come without a long history of debate about its appropriateness (Conrod 2020). McConnell-Ginet (2015) argues that the use of generic *he* cannot be considered to be gender-neutral, as gender assumptions or associations connected to non-epicene *he* inhibit true gender-neutrality. Indeed, for the generic usage of *he* (and also for that of *she*) a number of studies have shown that utterances intended to be gender-neutral were comprehended as not gender-neutral. Instead, both pronouns biased participants' associations towards the masculine (*he*) or feminine (*she*) interpretation, respectively (Martyna 1978; MacKay and Fulkerson 1979; Miller and James 2009). Thus, the generic usage of *he* (and *she*) apparently is not perceived as gender-neutral.

Several alternatives to generic *he* have been devised in English over the last decades, however, not with much success (Baron 1981; Livia 2001). Some of the more known alternatives are phrases like *he or she*, contractions like *s/he*, and singular *they* (Adami 2009). Singular *they* used to refer to single antecedents has been part of the English language at least since the 15th century. However, grammarians commonly disapprove of its usage (Conrod 2020). Nonetheless, several different types of singular *they* developed within English (Conrod 2019). Two of them are considered to be generic, the difference between them being whether they refer to an indefinite or definite antecedent, see Examples (3) and (4).

(3) Someone ran out of the classroom, but **they** forgot **their** backpack. (Conrod 2019: 81)

(4) The ideal student completes the homework, but not if **they** have an emergency. (Conrod 2019: 81)

While research on singular *they* has increased during the last decades, most studies are concerned with singular *they* from a sociolinguistic or syntactic perspective, asking which types of singular *they* there are (e.g. Bjorkman 2017; Conrod 2019, 2020), to which degree they are accepted among speakers (e.g. Bradley 2020), and how singular *they* can be accounted for in terms of syntactic agreement (e.g. Conrod 2019, 2022).

To summarize, English pronoun gender is assumed to rely on semantics. In any of the aforementioned cases, the usage of a pertinent pronoun can be retraced to certain semantic features of its antecedent. However, the classification of potential antecedents for pertinent pronouns appears to be not as straightforward as often stated. Several pronouns, i.e., *he*, *she*, and *they*, are used for very different types of antecedents, i.e., human and non-human, animate and inanimate, singular and plural, non-generic and generic, as well as male, female, and neuter antecedents. With one pronoun referring to semantically different antecedents, the question arises whether the comprehension of pronouns themselves is different depending on their antecedents. For instance, does generic *they* bear semantics similar to those of plural *they*? That is, are generic *they* and plural *they* treated similarly in the comprehension process? There are only very few studies available that investigated related questions or pronoun semantics and their comprehension to begin with. Notably, Sanford and Filik (2007) conducted an eye-tracking study, in which the authors found that generic *they* preserved portions of the plural semantics of plural *they*.

The aim of the present study is the proposition of a novel approach to the analysis of pronoun comprehension by means of computational methods. These methods constitute a framework novel to the research on pronouns and will allow future studies to investigate similarities and dissimilarities in comprehension between different pronouns but also between specific usages of a pertinent pronoun form. As an exemplary case, the comprehension of generic *they* will be compared to the comprehension of specific *he*, specific *she*, and plural *they*.

2 Method

The methodological approach proposed consists of five main parts. First, a text corpus containing attestations of the pronouns under investigation is required. Second, based on this corpus, vector representations of the semantics of words within the corpus are created. Third, individual semantic vectors for each occurrence of the pronouns under investigation are computed. Fourth, using the semantic vectors of the respective pronouns and further words, a mental lexicon and the comprehension processes therein are modeled. Fifth, from this simulated lexicon, measures are extracted to analyze the comprehension of the respective pronouns. The corpus, data, and scripts are available at the OSF: https://osf.io/bvxu5 (accessed: 21 January 2024).

2.1 A corpus of third-person pronouns

To work on the comprehension of third-person pronouns using the approach proposed in this paper, one first requires a corpus of the usage of third-person pronouns. The corpus used for the present study is based on newspaper articles contained in the Contemporary Corpus of American English (Davies 2008). For non-generic *he*, *she*, and *they*, attestations are easily found. For generic *they*, however, finding attestations is somewhat more difficult. This is mainly due to two reasons. First, plural *they* is far more frequent than generic *they*. Second, corpora are commonly not tagged in a way that reflects whether a token of *they* is plural or generic.

For this reason, instead of searching directly for attestations of *they*, the following search queries were used: *one must*, *anybody*, *anyone*, *someone*, *somebody*, and *person*. While the indefinite pronouns frequently co-occur with the use of generic *they* (cf. Example (3) above), the noun *person* invites general observations, which in turn may call for the use of an epicene expression.

For each pronoun under investigation, i.e., *he*, *she*, generic *they*, and plural *they*, the aim was to collect at least 50 attestations. In the created corpus, there are 100 attestations of *he*, 52 attestations of *she*, 139 attestations of plural *they*, and 81 attestations of generic *they*. To enrich the corpus with linguistic material beyond pronouns, randomly selected sentences from COCA without attestations of the pronouns under investigation were added to arrive at a total number of 1000 sentences.

As a prerequisite for the next steps, the corpus was automatically tagged using the RNNTagger software (Schmid 1999). The software identifies the lemma of each word; these lemmas, instead of inflected forms, were then used in the following step.

2.2 Capturing semantics with vectors

Capturing the semantics of lexical units with vectors is an idea that originates in distributional semantics, which follows the distributional hypothesis: Differences in meaning are represented by differences in distribution (Harris 1954; Boleda 2020). If words occur in similar contexts, their semantics are expected to be similar. If words, however, frequently occur in different contexts, their semantics are expected to be different as well. Following this idea, semantics can be captured by mathematical vectors, which are computed based on the distribution of words in large amounts of text. There are several algorithms with which one may arrive at such semantic vectors, e.g., Word2Vec (Mikolov et al. 2013), GloVe (Pennington et al. 2014), fastText (Bojanowski et al. 2016), or naive discriminative learning (henceforth NDL; Baayen et al. 2011; Baayen and Ramscar 2015).

The latter, NDL, is grounded in psychological theory on cognitive mechanisms (Pearce and Bouton 2001; Rescorla 1988), which has been shown to successfully model important learning effects in humans and animals alike (Kamin 1969; Ramscar et al. 2010). Following the Rescorla-Wagner rules (Rescorla and Wagner 1972; Wagner and Rescorla 1972), learning is understood as a result of informative relations within events, which in turn consist of cues and outcomes. With each new event encountered, the associations between cues and outcomes are constantly recalibrated. The associations of a given outcome and all cues encountered thus far at a given point in time can be understood as the outcome's relation to the world around the learning individual. The recalibration of associations happens 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. Once the learning process is finished, i.e., once all available events have been encountered, the final associations represent the interrelations of a pertinent outcome with all cues encountered during the learning process. Transferring this concept onto 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; derivational suffixes like *−ee*, *−ation*, and *−ment*) found in a text corpus annotated according to the needs of the pertinent investigation. For the present study, NDL was used to compute semantic vectors for all lemmas of the corpus introduced in Section 2.1.

However, computing semantic vectors using NDL leads to one critical issue for the present investigation: For words with identical forms, their senses are conflated into one vector representation (Lapesa et al. 2018). That is, one vector for the form *they* will be computed, conflating the semantics of plural *they* and generic *they*. Facing a similar issue in a study on generic and non-generic masculine role nouns in German, Schmitz (2024) made use of instance vectors, following ideas by Lapesa et al. (2018). Instance vectors are "vector representations for individual instances of words, i.e., tokens, rather than lemmas, i.e., types" (Lapesa et al. 2018: 291f). Following Lapesa et al. (2018), for the computation of each instance vector, a window of n context words around a given target word is taken into account. The mean of the n preceding and following context words' vectors constitutes the given target word's instance vector. For the present investigation, a context word window of $n = 2$ was used, which, according to Lapesa et al. (2018), captures true semantic similarity. Following common practice in distributional semantics, function words were not counted as context words. That is, instance vectors were created on the basis of content words. The vectors of context words can be taken from any approach to distributional semantics. For the present investigation, instance vectors were computed for all target pronouns attestations based on the NDL vectors of their surroundings.

2.3 Modelling the lexicon with discriminative learning

To model comprehension – and thereby gain insight into the semantic features of the pronouns under investigation – the present study makes use of linear discriminative learning (henceforth LDL; Baayen et al. 2019). In an implementation of LDL, the mental lexicon is simulated by generating a system of form-meaning relations that discriminates between different forms and meanings. Such an implementation allows the detailed investigation of entries and their relationship to each other within the mental lexicon.

For the simulated mental lexicon, LDL implementations generally require two things: semantics and forms of lexical entries. Semantics are most commonly represented by semantic vectors, such as those introduced in Section 2.2. Forms are also represented as vectors, which most commonly contain binary-coded information on whether certain *n*-graphs or *n*-phones are part of a given word form. For each word form's individual form vector *c*, the presence of an *n*-graph/*n*-phone cue is marked with 1, while the absence of a given cue is marked with 0. The interested reader is referred to Heitmeier et al. (2021) for a detailed overview of different design choices in the implementation of LDL models.

The entirety of semantic vectors to be included constitutes the *S* matrix and the entirety of form vectors to be included constitutes the *C* matrix of the respective implementation. Thus, each entry of the lexicon is represented twice: Once by a semantic vectors, which contains its meaning, and once by a form vector, which describes its phonology or orthography.

With the *S* and the *C* matrix at hand, one can compute the comprehension and production processes of the mental lexicon via multivariate multiple regression. Comprehension is modelled by simple linear mappings from the form matrix *C* to the semantic matrix *S* and production is modelled by likewise mappings from the semantic matrix *S* to the form matrix *C*. These mappings specify how strongly input nodes are associated to output nodes. A comprehension weight matrix *F* is obtained solving

$$S = CF, \tag{1}$$

while a production weight matrix G is obtained solving

$$C = SG. \tag{2}$$

In full-sized implementations of LDL, CF and SG are approximations of the S and C matrix due to their high dimensionality and the workings of linear regression. That is,

$$\hat{S} = CF \tag{3}$$

contains the predicted semantic vectors and

$$\hat{C} = SG \tag{4}$$

contains the predicted form vectors. It is these predicted vectors in $\hat{S}$ that we are interested in, as they are the result of the simulated comprehension process. In other words, the result of mapping forms onto meanings is an approximation to the original input semantics. These approximated meanings are taken as the result of the comprehension process, i.e., the result of the interrelations of forms and meanings in the mental lexicon. From these predicted meanings, then, one may derive an array of semantic measures directly connected to the comprehension of the pertinent lexical entry.

While architecture of LDL, as it is illustrated in Figure 1, might appear simplistic to some readers, previous research has shown that such linear mappings result in overall high accuracies (e.g. Baayen et al. 2018, 2019) and that semantic as well as phonological measures derived from such an implementation can explain a variety of empirical measures, e.g., acoustic duration (e.g., Chuang et al. 2021; Schmitz et al. 2021; Stein and Plag 2021), but also real word and pseudoword semantics (e.g., Chuang et al. 2021; Schmitz et al. 2021), and the male comprehension bias in generic masculines (Schmitz et al. 2023).

2.4 Including pronouns in the discriminative lexicon

To model the comprehension of third-person pronouns in the mental lexicon, not one but multiple implementations of LDL were needed. That is, as explained in Section 2.2, the semantics of each target pronoun was not captured in one semantic vector but in as many vectors as there were corpus attestations of the respective pronoun. Including all of these vectors within the same semantic matrix S and, with that, the simulated mental lexicon would go against common assumptions of the latter, i.e., one would not assume a single pronoun to have multiple entries in the mental lexicon.

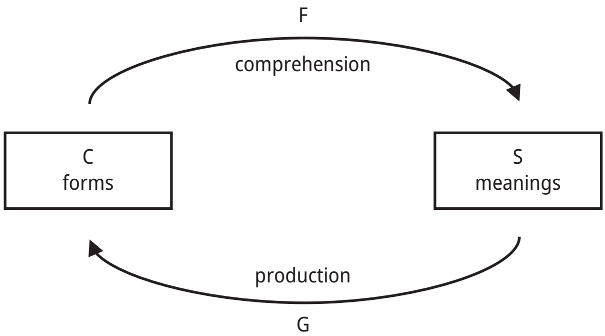

Fig. 1: Illustration of the comprehension and production mapping.

The LDL models were implemented in Julia using the JudiLing package (Luo et al. 2021) for the implementation itself and the JudiLingMeasures package (Heitmeier 2023) to compute measures based on the implementations. Besides the four pronoun vectors per LDL model, all other words of the corpus were included. In total, each implementation simulated a mental lexicon based on 3,524 entries. While the semantics of entries was represented by the NDL and instance vectors as explained above, form vectors were based on the orthography of entries and coded as trigraphs.

Measures extracted from each implementation of LDL were semantic CO-ACTIVATION, comprehension UNCERTAINTY, and semantic neighborhood DENSITY. These measures have been shown to successfully reflect the comprehended semantics of lexical entries and to model related experimental data (e.g., Schmitz et al. 2021; Stein and Plag 2021; Chuang et al. 2021; Heitmeier et al. 2023). CO-ACTIVATION is the Euclidean norm, that is the square root of the sum of the squared values, of a given predicted semantic vector. Higher values indicate higher degrees of co-activation. UNCERTAINTY represents the degree of dissimilarity between corresponding rows, with higher values indicating greater uncertainty. DENSITY is the mean cosine similarity of a target word's predicted vector with its ten most similar neighbors. Similarity is computed using cosine similarity.

3 Results

The distribution of the three comprehension measures across pronouns within the fifty LDL implementations is given in Figure 2. The measures were compared between pronouns using Bonferroni-corrected pairwise Wilcoxon tests. The results of these comparisons are given in Table 1.

Tab. 1: Results of Bonferroni-corrected pairwise Wilcoxon tests comparing the three comprehension measures across pronouns.

	CO-ACTIVATION			UNCERTAINTY			DENSITY		
	he	*she*	$they_g$	*he*	*she*	$they_g$	*he*	*she*	$they_g$
she	1.000			1.000			0.803		
$they_g$	0.234	0.329		0.314	1.000		0.001	0.038	
$they_p$	<0.001	<0.001	0.740	0.002	0.003	0.781	1.000	1.000	0.067

For CO-ACTIVATION, the values of *he* and *she* are significantly different from those of plural *they*, with plural *they* showing on average higher values, i.e., higher levels of semantic co-activation. As for generic *they*, it shows levels of co-activation in-between those of *he*/*she* and plural *they*, rendering it non-significantly different from the other pronouns.

For UNCERTAINTY, similar differences are found. The uncertainty values of plural *they* are significantly different from those of *he* and *she*, but not from those of generic *they*. Again, plural *they* comes with the highest mean value, followed by generic *they*, which in turn is followed by *he* and *she*. Hence, plural *they* shows the highest degree of uncertainty, while *he* and *she* show the lowest degree of uncertainty.

For DENSITY, differences pattern differently. Here, *he* and *she* are significantly different from generic *they*, while plural *they* is not significantly different from generic *they*. Overall, plural *they* comes with the densest semantic neighborhoods on average, followed by *he* and *she*, which in turn are followed by generic *they*. That is, *he*, *she*, and plural *they* come with the most semantically similar direct neighbors. For generic *they*, one finds a rather broad distribution of semantic neighborhood densities. Comparing interquartile ranges, this impression is confirmed: $IQR_{he} = 0.146$, $IQR_{she} = 0.201$, $IQR_{pluralthey} = 0.275$ vs. $IQR_{genericthey} = 0.319$.

4 Discussion

The present study set out to showcase an approach novel to the research on pronouns. Using computational methods – naive discriminative learning, instance vectors, and linear discriminative learning – the semantics of *he*, *she*, generic *they*, and

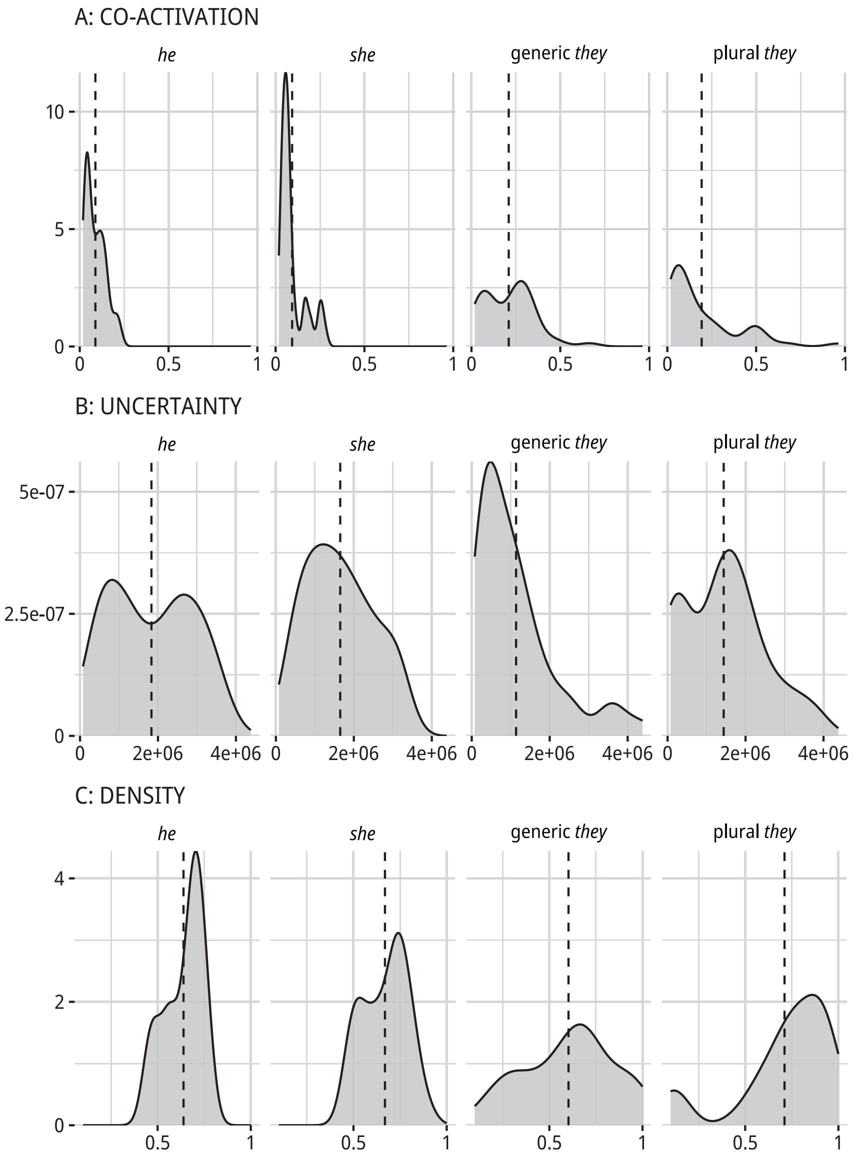

Fig. 2: Distribution of semantic CO-ACTIVATION (Panel A), comprehension UNCERTAINTY (Panel B), and semantic neighborhood DENSITY (Panel C) across pronouns within the fifty LDL implementations. Dashed lines indicate the pronoun-specific mean.

plural *they* were captured, and their comprehension was simulated. From fifty such simulations, measures on the comprehension of the pronouns under investigation were extracted and compared. But what do these measures tell us about the respective pronouns and their comprehension?

For semantic co-activation, it was found that plural *they* showed the highest degree of co-activation across the four pronouns under investigation. Hence, plural *they* co-activates the most entries in the lexicon when it is retrieved during the comprehension process. As plural *they* is the only third-person plural pronoun, all entries that have a plural form should be somewhat connected to plural *they*, as it is the pronoun used to refer to these plurals. For *he* and *she*, on the other hand, the degree of semantic co-activation was lowest. Taking into account that the set of referents for *he* and *she* is clearly more restricted than those of plural *they* (cf. Section 1), fewer co-activated lexical entries appear sensible. For generic *they*, a degree of semantic co-activation between that of plural *they* and that of *he* and *she* was found. One potential explanation for this finding is the set of potential referents: The set should at least contain all potential referents of *he* and *she* that fall into the categories of indefinite and definite non-specific antecedents, see Examples (3) and (4).

For comprehension uncertainty, plural *they* came with the highest values. Again, as the set of referents of plural *they* is clearly bigger than that of the singular pronouns, comprehending its semantics is relatively more uncertain, because there is a bigger set of potential referent semantics to chose from. Analogously, the comprehension uncertainty for *he* and *she* is lowest, while that of generic *they* is in-between the uncertainty levels of plural *they* and *he* and *she*.

For semantic neighborhood density, it was found that *he*, *she* and plural *they* come with denser semantic neighborhoods than generic *they*. That is, the ten nearest neighbors of *he*, *she* and plural *they* are significantly closer to their meaning than the ten nearest neighbors of generic *they*. As for *he*, *she*, and plural *they*, their sets of potential referents are rather clear (cf. Section 1). These potential referents, then, make up their near neighborhood. For generic *they*, however, the set of referents is much less clear or restricted. One might interpret this finding as a direct manifestation of the generic part of the generic *they*.

5 Conclusion

In sum, three main findings emerged from the present study. First, the bigger the set of potential referents, the higher the degree of semantic co-activation during the comprehension process. Second, the bigger the set of potential referents, the more uncertainty is found in the comprehension process. Third, the more restricted a set of referents, the denser the semantic neighborhood of a pronoun. These findings were obtained using computational methods novel to the research on pronouns. However, these findings should also be taken with a grain of salt due to the rather limited number of observations. Future studies will show whether the results of the present study are confirmed by studies making use of a more substantial amount of data, and how similar investigations play out for other pronouns. Additionally, the emergence of ready-to-use LLMs like GPT and BERT allows for the inclusion of context-dependent vectors for each pronoun token. Using such semantic representations may render the vectors used to simulate a mental lexicon via LDL somewhat more realistic. A step beyond replication and expansion will be the validation of the comprehension measures against behavioral data. That is, it will be shown whether such measures are able to explain, for example, differences between reaction times in studies on pronoun choice or pronoun fit. Overall, the methods proposed in the present study will allow future studies to investigate comprehension similarities and dissimilarities between different pronouns but also between specific usages of a pertinent pronoun form.

References

Adami, Elisabetta. 2009. "To each reader his, their or her pronoun". Prescribed, proscribed and disregarded uses of generic pronouns in English. In Antoinette Renouf and Andrew Kehoe (eds.), *Corpus linguistics*, vol. 69, 281–308. Amsterdam: Brill. https://doi.org/10.1163/9789042025981_016.

Audring, Jenny. 2009. *Reinventing pronoun gender*. Amsterdam: University of Amsterdam dissertation.

Baayen, R. Harald, Yu-Ying Chuang and James P. Blevins. 2018. Inflectional morphology with linear mappings. *The Mental Lexicon* 13. 230–268. https://doi.org/10.1075/ml.18010.baa.

Baayen, R. Harald, Yu-Ying Chuang, Elnaz Shafaei-Bajestan and James P. Blevins. 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. https://doi.org/10.1155/2019/4895891.

Baayen, R. Harald, Petar Milin, Dusica Filipović Đurđević, Peter Hendrix and Marco Marelli. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118. 438–481. https://doi.org/10.1037/a0023851.

Baayen, R. Harald and Michael Ramscar. 2015. Abstraction, storage and naive discriminative learning. In Ewa Dabrowska and Dagmar Divjak (eds.), *Handbook of cognitive linguistics*, 100–120. Berlin, München, Boston: De Gruyter. https://doi.org/10.1515/9783110292022-006.

Baranowski, Maciej. 2002. Current usage of the epicene pronoun in written English. *Journal of Sociolinguistics* 6. 378–397. https://doi.org/10.1111/1467-9481.00193.

Baron, Dennis E. 1981. The epicene pronoun: The word that failed. *American Speech* 56. 83. https://doi.org/10.2307/455007.

Bjorkman, Bronwyn M. 2017. Singular they and the syntactic representation of gender in English. *Glossa: a journal of general linguistics* 2. https://doi.org/10.5334/gjgl.374.
Bojanowski, Piotr, Edouard Grave, Armand Joulin and Tomas Mikolov. 2016. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics* 5. 135–146. https://doi.org/10.48550/arxiv.1607.04606.
Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics* 6. 213–234. https://doi.org/10.1146/annurev-linguistics-011619-030303.
Bradley, Evan D. 2020. The influence of linguistic and social attitudes on grammaticality judgments of singular 'they'. *Language Sciences* 78. 101272. https://doi.org/10.1016/J.LANGSCI.2020.101272.
Chuang, Yu-Ying, Marie Lenka Vollmer, Elnaz Shafaei-Bajestan, Susanne Gahl, Peter Hendrix and R. Harald Baayen. 2021. The processing of pseudoword form and meaning in production and comprehension: A computational modeling approach using linear discriminative learning. *Behavior Research Methods* 53. 945–976. https://doi.org/10.3758/s13428-020-01356-w.
Conrod, Kirby. 2019. *Pronouns raising and emerging*. Seattle: University of Washington dissertation.
Conrod, Kirby. 2020. Pronouns and gender in language. In Kira Hall and Rusty Barrett (eds.), *The Oxford handbook of language and sexuality*, Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190212926.013.63.
Conrod, Kirby. 2022. Abolishing gender on D. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 67. 216–241. https://doi.org/10.1017/cnj.2022.27.
Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139166119.
Davies, Mark. 2008. The corpus of contemporary American English (coca). https://www.english-corpora.org/coca. Accessed: 9 March 2025.
Harris, Zellig S. 1954. Distributional structure. *WORD* 10. 146–162. https://doi.org/10.1080/00437956.1954.11659520.
Heitmeier, Maria. 2023. Judilingmeasures: Calculate measures for LDL models developed with JudiLing. https://github.com/MariaHei/JudiLingMeasures.jl. Accessed: 9 March 2025.
Heitmeier, Maria, Yu Ying Chuang and R. Harald Baayen. 2021. Modeling morphology with linear discriminative learning: Considerations and design choices. *Frontiers in Psychology* 12. 4929. https://doi.org/10.3389/FPSYG.2021.720713.
Heitmeier, Maria, Yu-Ying Chuang and R. Harald Baayen. 2023. How trial-to-trial learning shapes mappings in the mental lexicon: Modelling lexical decision with linear discriminative learning. *Cognitive Psychology* 146. 101598. https://doi.org/10.1016/j.cogpsych.2023.101598.
Hekanaho, Laura. 2020. *Generic and nonbinary pronouns: Usage, acceptability and attitudes*. Helsinki: University of Helsinki dissertation.
Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316423530.
Kamin, Leon J. 1969. Predictability, surprise, attention, and conditioning. In Byron A. Campbell and Russell M. Church (eds.), *Punishment and aversive behavior*, 279–296. New York: Appleton-Century-Crofts.
Lapesa, Gabriella, Lea Kawaletz, Ingo Plag, Marios Andreou, Max Kisselew and Sebastian Padó. 2018. Disambiguation of newly derived nominalizations in context: A distributional semantics approach. *Word Structure* 11. 277–312. https://doi.org/10.3366/WORD.2018.0131.
Livia, Anna. 2001. *Pronoun envy: Literary uses of linguistic gender*. New York, Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780195138528.001.0001.

Luo, X., Yu-Ying Chuang and R. Harald Baayen. 2021. Judiling: An implementation in Julia of linear discriminative learning algorithms for language model. https://megamindhenry.github.io/JudiLing.jl/stable. Accessed: 9 March 2025.

MacKay, Donald G. and David C. Fulkerson. 1979. On the comprehension and production of pronouns. *Journal of Verbal Learning and Verbal Behavior* 18. 661–673. https://doi.org/10.1016/S0022-5371(79)90369-4.

Martyna, Wendy. 1978. What does 'he' mean? Use of the generic masculine. *Journal of Communication* 28. 131–138. https://doi.org/10.1111/J.1460-2466.1978.TB01576.X.

Mathiot, Madeleine. 1979. Sex roles as revealed through referential gender in American English. In Madeleine Mathiot (ed.), *Ethnolinguistics*, 1–48. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110804157-002.

McConnell-Ginet, Sally. 2015. Gender and its relation to sex: The myth of 'natural' gender. In Greville G. Corbett (ed.), *The expression of gender*, 3–38. Berlin, Boston: De Gruyter.

Mikolov, Tomas, Kai Chen, Gregory S. Corrado and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. In *International conference on learning representations*, https://doi.org/10.48550/arxiv.1301.3781.

Miller, Megan M. and Lori E. James. 2009. Is the generic pronoun he still comprehended as excluding women? *The American Journal of Psychology* 122. 483–496. https://doi.org/10.2307/27784423.

Newman, Michael. 1997. *Epicene pronouns: The linguistics of a prescriptive problem*. New York: Garland.

OED. 2023. *Oxford English dictionary online*. Oxford: Oxford University Press.

Pearce, John M. and Mark E. Bouton. 2001. Theories of associative learning in animals. *Annual Review of Psychology* 52. 111–139. https://doi.org/10.1146/annurev.psych.52.1.111.

Pennington, Jeffrey, Richard Socher and Christopher Manning. 2014. GloVe: Global vectors for word representation. In Alessandro Moschitti, Bo Pang and Walter Daelemans (eds.), *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532–1543. Doha, Qatar: Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1162.

Poutsma, Hendrik. 1914. *A grammar of Late Modern English*. Groningen: Noordhoff.

Ramscar, Michael, Daniel Yarlett, Melody Dye, Katie Denny and Kirsten Thorpe. 2010. The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science* 34. 909–957. https://doi.org/10.1111/j.1551-6709.2009.01092.x.

Rescorla, Robert A. 1988. Pavlovian conditioning: It's not what you think it is. *American Psychologist* 43. 151–160. https://doi.org/10.1037/0003-066X.43.3.151.

Rescorla, Robert A. and Allan R. Wagner. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Abraham H. Black and William F. Prokasy (eds.), *Classical conditioning II: Current research and theory*, 64–99. New York: Appleton-Century-Crofts.

Rudes, Blair A. and Bernard Healy. 1979. Is she for real? The concepts of femaleness and maleness in the gay world. In Madeleine Mathiot (ed.), *Ethnolinguistics*, 49–62. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110804157-003.

Sanford, Anthony J. and Ruth Filik. 2007. They as a gender-unspecified singular pronoun: Eye tracking reveals a processing cost. *Quarterly Journal of Experimental Psychology* 60. 171–178. https://doi.org/10.1080/17470210600973390.

Schmid, Helmut. 1999. Improvements in part-of-speech tagging with an application to German. In Susan Armstrong, Kenneth Church, Pierre Isabelle, Sandra Manzi, Evelyne Tzoukermann and David Yarowsky (eds.), *Natural language processing using very large corpora*, 13–25. Springer. https://doi.org/10.1007/978-94-017-2390-9_2.

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, Ingo Plag, Dinah Baer-Henney and Simon David Stein. 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. https://doi.org/10.3389/fpsyg.2021.680889.

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.

Siemund, Peter. 2008. *Pronominal gender in English: A study of English varieties from a cross-linguistic perspective*. London: Routledge.

Stein, Simon David and Ingo Plag. 2021. Morpho-phonetic effects in speech production: Modeling the acoustic duration of English derived words with linear discriminative learning. *Frontiers in Psychology* 12. https://doi.org/10.3389/fpsyg.2021.678712.

Wagner, Allan R. and Robert A. Rescorla. 1972. Inhibition in Pavlovian conditioning: Application of a theory. In Robert A. Boakes and M. S. Halliday (eds.), *Inhibition and learning*, 301–334. Academic Press Inc.