A prediction of generic *they* semantics

Dominic Schmitz

Heinrich Heine University Düsseldorf, Germany

An often-discussed topic in English linguistics are the different usages of the pronoun *they*. While in its prototypical usage it is a third person plural pronoun, there are also several attested singular usages (cf. Conrod, 2020); the most prominent being its generic use. While there is a plethora of sociolinguistic research on generic *they*, there is no account of its semantics. A first such account is the aim of the present study.

In this paper, the term "prediction" does double duty. First, from a technical perspective, the semantics of generic *they* attestations and further pronouns (*he*, *she*, plural *they*, *anyone*, *anybody*) are predicted via instance vectors (Lapesa et al., 2018). Second, from a theoretical perspective, the forms of pronouns are taken to predict their meanings, following the framework of the Discriminative Lexicon (Chuang & Baayen, 2021).

Using instance vectors in a computational implementation of the Discriminative Lexicon (Baayen et al., 2019), it was shown that plural *they* is comprehended significantly better than generic *they* ($p < 0.001$). Comparing the semantics of generic *they* to the semantics of the other pronouns, generic *they* appears to be a generic singular pronoun with remnants of plurality (cf. Table 1).

Predicting the semantics of generic *they* in a twofold fashion, the present findings provide not only a first account of its semantics, but also show that the Discriminative Lexicon is a framework fit to explore pronoun semantics.

Table 1. Mean similarities across pronoun semantics computed via cosine similarities. Higher values indicate higher similarities of vectors and, in turn, semantics.

	she	generic *they*	plural *they*	anyone	anybody
he	0.587	0.641	0.468	0.517	0.483
she		0.599	0.428	0.483	0.456
generic *they*			0.508	0.565	0.533
plural *they*				0.399	0.359
anyone					0.484

References

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. https://doi.org/10.1155/2019/4895891

Chuang, Y.-Y., & Baayen, R. H. (2021). Discriminative learning and the lexicon: NDL and LDL. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/ACREFORE/9780199384655.013.375

Conrod, K. (2020). Pronouns and gender in language. In *The Oxford Handbook of Language and Sexuality*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190212926.013.63

Lapesa, G., Kawaletz, L., Plag, I., Andreou, M., Kisselew, M., & Padó, S. (2018). Disambiguation of newly derived nominalizations in context: A Distributional Semantics approach. *Word Structure*, *11*(3), 277–312. https://doi.org/10.3366/word.2018.0131