

Generics are not Quantificational.

A New Path from Language Models to Semantic Theory.

Abstract for Talk at ELM 4, 2026

Generic sentences express generalizations that tolerate exceptions without *explicitly* communicating information about quantities. For example, *Ravens are black* is true despite the presence of albino ravens. The sentence doesn't explicitly communicate the number or frequency of black ravens. Whether generics semantically encode information about quantities *implicitly*, a view we call *Quantificationalism*, is controversial.

A cursory look at the literature on generics in linguistics and philosophy of language reveals that Quantificationalism is the dominant view about the underlying semantic facts (Carlson and Pelletier (1995); Cohen (1999); Nickel (2016); Sterken (2015); and many others). In psychology, the story is a little different. The dominant view rejects Quantificationalism. For example, Leslie (2007, 2008) and Leslie et al. (2011) propose that generics don't express quantifier-meanings in context; rather, they give voice to a special-purpose mode of generalizing information.

Here, we take a large-scale distributional approach to the semantic debate: we compare generics to overtly quantificational sentences by analyzing in over half a million samples how the presence or absence of a quantifier affects probability distributions over tokens following the main verb. These probabilities are extracted from modern autoregressive language models (MIXTRAL-8×22B, QWEN3-8B and others).

In what follows, we briefly present the theoretical commitments of our approach, detail the methodology used, comment on our experimental results and sketch the argument challenging a quantificational understanding of generics.

The correlation principle. Bringing language models and information theory to bear on the question of Quantificationalism requires

an assumption about the relation between linguistic meaning and the distribution of text in a corpus. This assumption, which we call *correlation principle*, posits a default connection between similarity of distribution within a large enough corpus and sameness of meaning: all else equal, similarity of distribution correlates with sameness of meaning (Harris, 1982; Boleda, 2020; Grindrod, 2023).

Methodology: comparing minimal sets of sentences. The experiments that follow use minimal sets consisting of generic and quantificational sentences. As an example, consider the following:

{*Postal workers are government employees, All postal workers are government employees, Most postal workers are government employees, Some postal workers are government employees*}.

The underlying idea is that the quantifier (or absence of it in the generic case) will induce different probability distributions at the tokens after the main verb, those associated with the property that is predicated of the kind in the generic or quantified sentence: different things are said about *all* postal workers than about just *some* of them.

Following Cilleruelo et al. (2025a), given a set of quantifiers and a sentence, we will take quantifier q as the better fit for that sentence if it induces the lowest average surprisal in the post-verb tokens. In order to compare full token probability distributions, instead of just the probability of a single token, we extend this methodology with KL-divergence. We say that two expressions are close in meaning when every single KL-divergence on the post-verb tokens is below a threshold t .

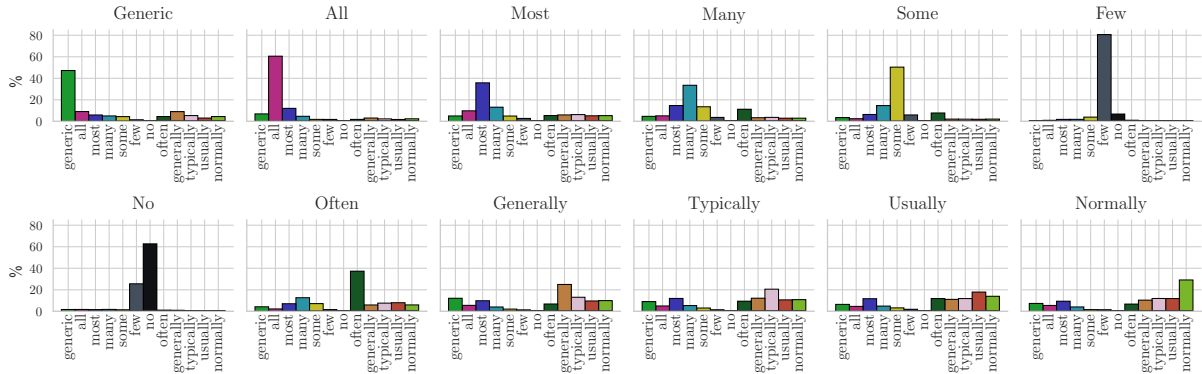


Figure 1: Percentage of terms selected as p-acceptable for each original phrase (MISTRAL-7B).

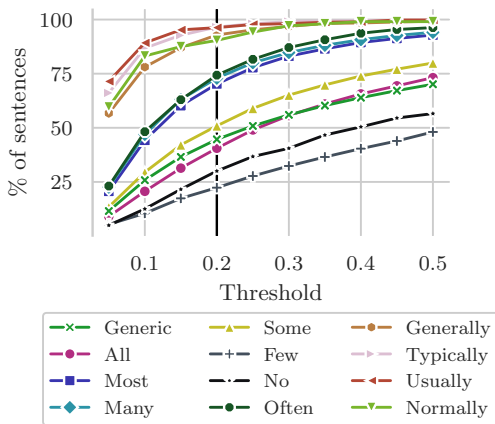


Figure 2: Percentage of sentences with *at least one* “close” quantifier per threshold (MISTRAL-8×22B).

Experiments: no (combinations of) quantifiers are synonymous to the generic. For our experiments, we sample over 600,000 quantified and generic sentences in context from the MGEN dataset (Cilleruelo et al., 2025b). We then intervene each sentence with the following quantifiers: *all*, *most*, *many*, *some*, *few*, *no*, *often*, *generally*, *typically*, *usually*, *normally* and the generic. This set of candidate quantifiers is designed to cover the spectrum of quantificational strengths and flavors the generic is posited to have in the literature (Carlson and Pelletier, 1995; Almotahari, 2022).

For the comparison with p-acceptability (surprisal), we count the percentage of sentences where each quantifier has the lowest surprisal on the property tokens given the original quantifier of the sentence (Figure 1). The plot of the generic resembles that of quantifier with more distinct meanings (*all*, *few*, *some*, *no*), while prevalence quantifiers have very flat spreads.

The comparison with KL-divergence depends on a similarity threshold t . In Figure 2, we plot the percentage of sentences for each original quantifier for which *some other quantifier* is close in meaning, for each given threshold. Consider $t = 0.2$, for which all prevalence-expressing quantifiers have over 75% sentences for which another quantifier is synonymous. For that same threshold, only 40% of generic sentences have a similar quantifier, at similar rates than *all* and *some*. But note that in the set of quantifiers considered, both *all* and *some* are clearly distinct, while the generic should not be. If Quantificationalism were true, and sameness of meaning were correlated with similarity of distribution within a representative corpus, then we should expect the curve for generic expressions in Figure 2 to cluster with the curves for quantifiers that approximate genericity.

Generics are not quantificational. In view of these results, we formulate an interesting new challenge for Quantificationalism:

- I. There’s a significant difference in the pattern of distribution within the MGEN dataset between generics and their quantificational counterparts (significant enough, at least, to make for massive differences with respect to p-acceptability).
- II. MGEN is sufficiently representative to have evidential value.
- III. Therefore, given the correlation principle, we have defeasible but non-negligible reason to think there’s an equally significant semantic difference between generics and their quantificational counterparts.

References

- Mahrad Almotahari. 2022. [Weak generics](#). *Analysis*, 82(3):405–409.
- Gemma Boleda. 2020. [Distributional semantics and linguistic theory](#). *Annual Review of Linguistics*, 6(1):213–234.
- Greg N. Carlson and Francis Jeffrey Pelletier, editors. 1995. *The Generic Book*. University of Chicago Press.
- G. Cilleruelo, Emily Allaway, Barry Haddow, and Alexandra Birch. 2025a. [Generics are puzzling. can language models find the missing piece?](#) In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 6571–6588, Abu Dhabi, UAE. Association for Computational Linguistics.
- G. Cilleruelo, Emily Allaway, Barry Haddow, and Alexandra Birch. 2025b. [Mgen: Millions of naturally occurring generics in context](#). *Proceedings of the Society for Computation in Linguistics*, 8(1):11.
- Ariel Cohen. 1999. [Generics, frequency adverbs, and probability](#). *Linguistics and Philosophy*, 22(3):221–253.
- Jumbly Grindrod. 2023. *Distributional Theories of Meaning*, pages 75–99. Springer International Publishing, Cham.
- Zellig Harris. 1982. *A Grammar of English on Mathematical Principles*.
- Sarah-Jane Leslie. 2008. [Generics: Cognition and acquisition](#). *Philosophical Review*, 117(1).
- Sarah-Jane Leslie, Sangeet Khemlani, and Sam Glucksberg. 2011. [Do all ducks lay eggs? the generic overgeneralization effect](#). *Journal of Memory and Language*, 65(1):15–31.
- Sarah-Jane Leslie. 2007. [Generics and the structure of the mind](#). *Philosophical Perspectives*, 21:375 – 403.
- Bernhard Nickel. 2016. *Between Logic and the World: An Integrated Theory of Generics*. Oxford University Press UK, Oxford, GB.
- Rachel Sterken. 2015. [Generics in context](#). *Philosophers’ Imprint*, 15:1–30.