A is for *a*-generics: Predicate Collectivity in Generic Constructions

Carlotta Marianna Cascino
Département d'Études Cognitives, École Normale Supérieure, Paris, France
Department of Psychology, Princeton University, Princeton, NJ, USA
carlotta.marianna.cascino@ens.psl.eu

Abstract

Generic statements like *A dog has four legs* are central to encode general knowledge. Yet their form–meaning mapping remains elusive. Some predicates sound natural with indefinite singulars (*a*-generics), while others require the definite article (*the*-generics) or the bare plural (bare-plural generics). For instance, why do we say *The computer revolutionized education* but not *A computer revolutionized education*? We propose a construction-based account explaining why not all generic statements are created equal. Prior accounts invoke semantic notions like kind-reference, stage-levelness, or accidental generalization, but offer no unified explanation. This paper introduces a new explanatory dimension: predicate collectivity level, i.e. whether the predicate applies to each member of a group or to the whole group as a unit (without necessarily applying to each of its members individually). Using two preregistered acceptability experiments we show that *a*-generics, unlike *the*-generics and bare-plural generics, are dispreferred with collective predicates. The findings offer a functionally motivated, empirically supported account of morphosyntactic variation in genericity, providing a new entry point for Construction Grammar.

1 Introduction

We interact meaningfully in the world on the basis of our knowledge of categories. A key reason humans (and other animals) categorize entities is to predict how to interact with new instances. Instances of a category tend to share properties with other members of the same category and not share properties with members of competing categories (Rosch & Mervis, 1975).

One way humans explicitly inform others about properties of categories is by using certain linguistic constructions, regularly referred to as generic statements. An aspect of generic statements that has garnered a great deal of attention is that people are willing to endorse generic statements even when a property only holds of a minority of instances. For instance, most people agree with the statement in (1) (Pelletier & Asher, 1997; Leslie et al., 2011), even though only adult female ducks lay eggs.

(1) Ducks lay eggs.

Much interest in generic statements concerns this fact, which distinguishes generics from universally quantified statements (*All duck lay eggs*). As in example (1), generic categories are often expressed using a bare plural form and much work on generics focuses on this type of generic (e.g., Carlson, 1977; Cohen & Erteschik-Shir, 1997, 2002; Kiss, 1998; Nyugen, 2020).

However, generic expressions in English can be expressed in alternative ways as well. In particular, generic meaning can be expressed with the indefinite singular article (*a*) as in (2), which we refer to here as *a*-generics. A third way of expressing generic meaning involves the definite article (*the*) with a singular noun as in (3), which we refer to here as *the*-generics.

(2) *A*-generic: A duck lays eggs.

(3) *The*-generic: The duck lays eggs.

Languages rarely offer speakers a choice between constructions without the choice being meaningful. The choice of one construction over another may signal a different interpretation, context, register, or dialect (Humboldt, 1999; Clark, 1987; Goldberg, 1995). And in fact, when distinct generic constructions have been considered, researchers have posited some functional distinction or other between them.

In a comparison between *a*-generics and bare plurals, Cohen (2001) argued that *a*-generics must express a rule or a regulation. Bare plurals, instead, may either express the same type of mean-

Proceedings of the Second International Workshop on Construction Grammars and NLP, pages 109–119

ing, or simply describe the way things happen to be. Others have likewise evoked the idea that *a*-generics convey law-like, nonaccidental generalizations (Greenberg, 2003), expressing necessary ("analytic") properties (Lawler, 1973; Burton-Roberts, 1977).

Furthermore, it has been suggested that bare plural generics but not *a*-generics are compatible with conjunctions of predicates that refer to equally good, but mutually incompatible characteristic properties, none of which are satisfied by the majority of the kind, as in example (4) (Nickel, 2008; Kirkpatrick, 2022):

(4a) Computers were invented in the 20th century and perfected in the 21st century.

(4b) #A computer was invented in the 20th century and perfected in the 21st century.

It is not the case, indeed, that most individual computers were both invented in the 20th century and perfected in the 21st century.

A-generics have been claimed to be further restricted by disallowing "stage-level" predicates, which take stages of individuals as arguments (Condoravdi, 1994), as in (5):

(5a) Penguins are endangered.

(5b) #A penguin is endangered.

Guerrini (2025) has recently argued that some restrictions on *a*-generics—like not allowing for accidental generalizations (see example 6c)—stem from the claim that the singular indefinite form cannot denote a "kind," because kinds are inherently plural entities.

A second research direction concerns the distinction between *a*-generics and *the*-generics. For instance, Platteau (1980:121-122), suggesting that the basic principles of definite and indefinite reference are also applicable to generic NPs, claimed that indefinite generics "refer to a random element of a certain species", such that "the selected sample has the same default properties as all the other members of the species". On the other hand—they claim—definite generics refer to one definite entity, which is "the abstract representative of the species".

Later work by Krifka (1987) and Krifka et al. (1995) distinguished the functions of indefinite generics ("I-generics") and definite generics ("D-generics") as follows.[1] Definite generics can involve "kind" predicates (6a), i.e. predicates whose subject is a kind; dynamic predicates (6b), i.e. non-stative predicates (see also Heyer, 1985); or accidental properties (6c) (see also Lawler, 1973; Burton-Roberts, 1977; Cohen, 2001; Greenberg, 2002, 2003). On the other hand, none of these types of predicates is possible with indefinite generics (7a-c):

(6a) The lion is extinct.

(6b) The rat reached Australia in 1770.

(6c) The madrigal is popular.

(7a) ?A lion is extinct.

(7b) ?A rat reached Australia in 1770.

(7c) ?A madrigal is popular.

Krifka further argues that only indefinite generics can be applied to "kinds that are not well-established," providing the contrast in (8) (see also Carlson, 1977):

(8a) ?The lion with three legs is ferocious.

(8b) A lion with three legs is ferocious.

As for the forms, according to Krifka (1987), singular definites, plural definites and taxonomic[2] generics belong to the class of "D-generics", while singular indefinites belong to the class of "I-generics". Bare plurals and bare singular generics, instead, have none of the restrictions just mentioned (see also Krifka, 2003), occurring in both classes, as shown in the following examples:

(10a) Lions are extinct.

(10b) Bronze was invented before 2000 B.C.

(11a) Rats reached Australia in 1770.

(11b) Rice was introduced in East Africa some centuries ago.

(12a) Madrigals are popular.

(12b) Music is popular.

(12a) Lions with three legs are ferocious.

(12b) Gold which is hammered flat is precious.

Overall, within this system, D-Genericity has been analyzed as "reference to kinds, which is NP-oriented", i.e. dependent on the type of noun phrase (13a); and I-genericity has been analyzed as "default quantification" which has scope over the

[1] As will become clearer later, Krifka's categories of "D-generics" and "I-generics" are not to be equated with specific grammatical forms such as the definite singular or the indefinite singular. Rather, definite generic NPs and indefinite generic NPs merely serve as their most typical realizations (Krifka, 1987: 4).

[2] Taxonomic generics have been claimed to have themselves different forms, and to refer to subspecies of a kind (Galmiche, 1985), as in the following examples:

(9a) *One lion*, namely the Asian lion, is nearly extinct.

(9b) *This lion* (the Asian lion) is nearly extinct.

(9c) *The rice they grow in East Africa* needs little water.

A detailed treatment of their possible forms lies beyond the scope of this paper.

VP as well (Krifka, 1987), occurring in "characterizing sentences" (13b), i.e. generalizations over groups of particular episodes of facts (Krifka et al., 1995). Krifka et al.'s proposal also allows for kind-referring NPs to occur in characterizing sentences (13c), recognizing potential overlaps in both form and meaning between I-generics and D-generics:

(13a) The potato was first cultivated in South America.

(13b) A potato contains vitamin C, amino acids, protein and thiamine.

(13c) The potato is highly digestible.

In our work, instead, we distinguish three types of generic constructions, which we refer to simply as *a*-generics, *the*-generics and bare-plural generics. This differs from Krifka's use of the labels "I-generics" and "D-generics", because we presume that the morphology provides an invitation to identify functional categories, and our goal is to determine what those categories are.

As for experimental work, Driemel et al. (2025) presented cross-linguistic evidence based on an acceptability judgements study testing singular definite, singular indefinite, bare plural, and definite plural generic forms. Their results show that bare plurals are preferred in English and German for kind- and characterizing-level readings, while definite plurals dominate in Romance and Greek. Although from their graph it is possible to note that definite singulars are preferred over indefinite singulars for kind reference, the authors do not explicitly mention it.

We are unaware of other prior experimental work testing distinctions among generic morphosyntactic forms, with the exception of Fuellenbach et al. (2019), who hypothesized that *a*-generics prefer normative or essential predicates ("principled": e.g. *A fep has red wings*) rather than incidental predicates ("statistical": e.g. *A fep throws glow sticks*). In a two-alternative forced choice task, child and adult participants were first exposed to an image of a target novel animal (e.g., a *kevta*) followed by a statement of one of four types:

Kevtas / A kevta / The kevta / This kevta wears scarves.

Participants were then asked: *Which one of these is also a kevta?* Only one of the images contained the same novel animal with the predicated property (e.g., a kevta wearing a scarf). Of interest was whether participants would interpret the statement as generic, in which case other instances of the same category should also share the same property (e.g., wear a scarf). *The*-generics were instead predicted to lead to lower generalisability with statistically connected property, but not necessarily to higher generalisability with principled properties[3]. They also predicted that bare plural subjects would support both principled and statistical properties equally well. Results showed that participants were more likely to treat the statement as generic when the predicate was normative or essential (e.g., "has red wings") than when the property was incidental "throws glow sticks") regardless of the morphosyntactic form of the statement.

2 Hypotheses

Much prior work on genericity has focused on the semantic compatibility between generics and certain types of predicates. Building on this literature, we hypothesize that a key factor influencing the acceptability of different generic constructions lies in whether the predicate is construed as applying to an individual or to a group. This distinction amounts to the well-known contrast between distributive and collective predicates. A *distributive* predicate applies individually to each member or subset of a group (or parts of an entity), while a *collective* predicate applies to a group or entity as a unit, without necessarily applying to each of its members individually (Link, 1983; Landman, 1989; Champollion, 2020).

For instance, the quantifiers *each* and *every* require a distributed interpretation, while *all* allows for a collective interpretation. That is, the statements in (14a) describe some very strong children, while (14b) allows an interpretation in which the children acted as a group to raise the turkey.

(14a) *Distributive*: Each / Every child lifted a 100 pound turkey.

(14b) *Collective*: All the children lifted a 100 pound turkey.

We hypothesized that the critical distinction between *a*-generics and *the*-generics is similar. Since the indefinite singular determiner, *a*, evokes a single indefinite individual, *a*-generics require predicates that can be construed as applying to (most)

[3]The author claimed that this pattern is similar to the one predicted for *a*-generics, but they expected *the*-generics to be rated lower overall, due to their overall more restricted use.

any[4] individual of the category. The meaning of *a*-generics, is motivated, on this perspective, by the fact that the predicate applies to any randomly selected individual of a category.

The definite singular determiner, *the*, on the other hand, generally combines with identifiable, specific nouns rather than any randomly selected member of a group. Therefore *the*-generic interpretations cannot be motivated in the same way as *a*-generics. Instead, we hypothesized that *the*-generics predicate a property of a clearly identifiable group or kind.

Construction	*a*-generics	*the*-generics
Morphosyntactic Form	[*A* N'] VP	[*The* N'] VP
Functional Constraints	VP predicate is construed to apply to a randomly selected instance of the category N'	VP predicate is construed to apply to the category as a collective or group

Table 1: Two hypothesized generic constructions in English

To understand the claims in Table 1, consider the pairs of sentences in (15) and (16). As confirmed by norming, described in the next section, (15) involves a collective predicate and (16) a distributive predicate. We predict that collective predicates will be rated more acceptable with *the*-generics (e.g., 15a) than with *a*-generics (15b), while distributive predicates will be rated more acceptable with *a*-generics (e.g., 16a) than with *the*-generics (16b):

Predicate construed to apply to a collective:

(15a) The bee pollinates crops across the globe.>

(15b) A bee pollinates crops across the globe.

Predicate construed to apply to (randomly chosen) individuals:

(16a) A bee dies after it stings. >

(16b) The bee dies after it stings.

In this paper we test in two preregistered studies[5] the prediction that the distinction between collective and distributive readings impacts English speakers' preference for *the*- vs. *a*-generics. The derivations of this prediction can be schematized as follows:

i) ***a*-generics favor predicates that apply to randomly chosen individual instances of a category**: *a*-generics will be judged less acceptable when combined with properties that apply to a category construed as a collective.

ii) ***the*-generics favor predicates that apply to a specific, clearly identifiable category**: *the*-generics will be judged more acceptable when combined with properties that apply to a category construed as a collective.

iii) **bare plurals display neither restriction.**

Our proposal draws inspiration from Platteau's (1980) distinction between reference to a random element of a certain species and reference to the abstract representative of the species. We also take up Krifka's (1987) call for a functional distinction between types of genericity, but reinterpret it within a constructional perspective, associating functional distinctions with morphosyntactic distinctions. Note also that predicate collectivity is orthogonal to kind-reference in the sense described by Krifka (1987): while kindhood concerns the referential status of the NP, collectivity captures how the VP applies across instances.

In sum, previous accounts have identified a wide range of semantic constraints on *a*-generics. Rather than replacing earlier insights, our findings identify a new explanatory dimension, i.e. predicate collectivity, proposing a novel empirically-grounded and morphosyntactically-oriented perspective on genericity.

3 Experiment 1

To test these predictions empirically, we conducted two experiments. Experiment 1 focused on how predicate collectivity influences the acceptability of *a*- vs. *the*-generics across a range of naturalistic sentences.

3.1 Methods

Participants. 79 native English speakers were recruited via Prolific (47F, 32M; M = 38.2 yrs) to provide acceptability ratings. As planned, participants who failed to accurately rate acceptable fillers higher than unacceptable fillers were excluded from analyses (mean rating unacceptable fillers $\geq$ mean rating acceptable fillers). This proved a stringent criterion and 23 participants were excluded, resulting in 56 participants included in the analysis. Because of the high number of exclusions, we also

[4] We include "most any" here because of the well-known fact that generic statements need not hold of every single instance of a group to be judged felicitous (e.g. *A duck lays eggs*).

[5] Link to the preregistrations: https://aspredicted.org/hs5y-b6p7.pdf

ran the analyses on all participants. Results on the predicted interaction did not change and can be found in Appendix A.

Materials. We constructed 12 English predicates (verb phrases) and count nouns in subject position (e.g., [computer] has transformed education) (see Appendix B for items). Twenty-four stimuli were then created by instantiating each noun phrase in two versions: with an *a*-generic (e.g., *A computer has transformed education*), and with a *the*-generic (e.g., *The computer has transformed education*). We additionally included 6 filler items. Each filler was a generic sentence with a bare plural subject; 3 fillers were intended to be fully acceptable, while three others deliberately contained grammatical errors. The latter served as attention checks and exclusion criterion. To quantify the degree of collectivity vs distributivity, we normed each of the 12 predicates combined with bare plural subjects. For this, a separate group of 22 native English speakers was recruited via Prolific and paid for their time to perform a forced-choice task asking whether each sentence was about individuals or groups. An example is provided in Figure 1:

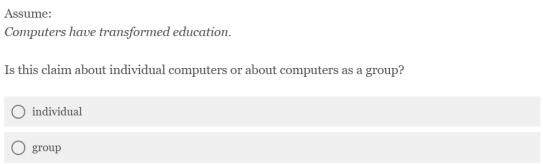

Figure 1: An example of the task in the norming experiment.

Participants who always responded with the same answer (n = 5) were excluded. The mean proportion of "group" responses from the remaining participants for each item was then used as predictor in the statistical analysis. This is a collectivity score, and ranged between 0 and 1. Predicates and their corresponding collectivity scores are reported in Appendix B.

Procedure. Each participant rated the acceptability of one version of each sentence (either an *a*-generic or *the*-generic), on a 7-point Likert scale, with generic type counterbalanced across participants. We further subdivided items so that each participant judged 6 target sentences: 3 *a*-generics and 3 *the*-generics, along with the 6 filler sentences. Items were presented in a randomized order for each participant. Instructions are provided in Figure 2.

An alternative design would have been to present participants with both versions of each sentence (with an *a*-generic and with a *the*-generic) and ask them to rank the two sentences of the pair for acceptability. A within-subjects setup of this kind typically affords greater statistical power by reducing variability across participants. However, exposure to one version would likely influence judgments of the other, turning the task into a relative rather than independent assessment. To avoid this, we adopted a design in which each participant was exposed to only one version of a given sentence, rating it independently of its alternative.

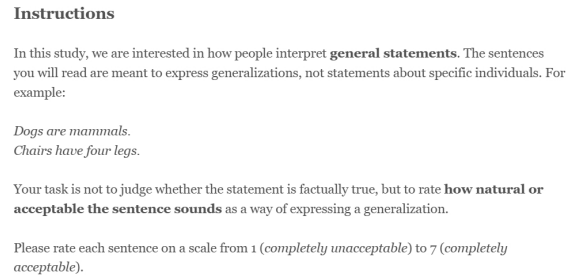

Figure 2: Instructions for Experiment 1.

3.2 Results

Results confirmed the hypotheses that *a*-generics and *the*-generics display different distributional patterns, and that *a*-generics were judged more acceptable when combined with a predicate that was more likely to be interpreted as applying to individuals (i.e., lower collectivity), rather than groups. This is shown in Figure 3.

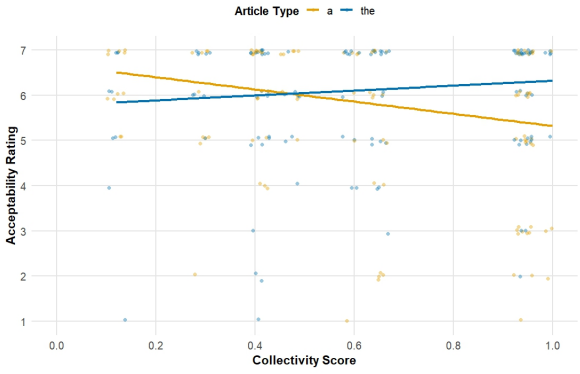

Figure 3: Acceptability ratings (1–7) as a function of collectivity scores (0–1) for *a*-generics and *the*-generics. Each point represents an individual response. Solid lines depict linear fits for each article type. Collectivity scores, normed separately, capture the degree to which a predicate was judged to refer to a collective vs. individual property. The figure illustrates that article acceptability interacts with collectivity, with *a*-generics associated with lower acceptability as collectivity increases.

This was confirmed with the planned cumulative link mixed-effects model (using the ordinal package in R [Christensen, 2019]. The output was raw acceptability ratings on the Likert scale. Article (*the* vs. *a*) and degree of collectivity were included as fixed interacting factors. Random intercepts and slopes for collectivity level were included by participant, while by-item random effects included intercepts and random slopes for article.

The model was fit to 342 observations (log-likelihood = –434.60; AIC = 901.21). The predicted interaction between article and predicate type was significant ($\beta = 2.45$, SE = 0.92, z = 2.66, p = .0077). Specifically, when the article was *a*, increasing collectivity scores significantly decreased acceptability ratings ($\beta = -1.68$, SE = 0.66, z = -2.56, p = .0105). A likelihood ratio test comparing the full model with article–collectivity interaction to a reduced additive model confirmed that including the interaction significantly improved model fit ($\chi^2(1) = 5.8367$, p = .0157), justifying its inclusion.

To examine this interaction more closely, we then fit separate models for *a*-generics and *the*-generics sentences[6]. For *a*-generics sentences, acceptability ratings decreased as collectivity increased, with a marginally significant negative effect of collectivity ($\beta = -1.81$, SE = 0.95, z = -1.9, p = .057). In contrast, for *the*-generics, collectivity showed a positive but non-significant effect on ratings ($\beta = 0.48$, SE = 0.79, z = 0.6, p = .546).

Taken together, these results indicate that the significant interaction observed in the full model is primarily driven by the sensitivity of *a*-generics to collectivity, whereas *the*-generics appear robust to this variation.

3.3 Discussion

Our results provide evidence that the morphosyntactic form of generic statements motivates their constraints. Specifically, as the predicate's collectivity score increased, *a*-generics significantly decreased in acceptability, as predicted by the constraint hypothesized in Table 1. As for *the*-generics, although the effect of collectivity did not reach statistical significance, the positive trend in the data warrants further investigation.

The contrast between collective and distributive predicates recalls Krifka et al.'s suggestion (1995; see also Krifka, 1987; Guerrini, 2025) that *a*-generics do not allow subjects that refer to kinds. The current proposal goes beyond this observation in several ways. First, we demonstrate that *a*-generics disprefer not only kind-level predicates (as suggested by Krifka et al., 1995), but collective predicates more broadly. This includes cases that do not involve reference to kinds *per se*. For instance, *a*-generics are rated significantly less acceptable with predicates such as *pollinates crops across the globe* or *hunts in packs* (e.g., *A bee pollinates crops across the globe*, *A wolf hunts in packs*), both of which attribute properties to the collective behaviour of a cateogry. Secondly, we show that not only do *a*-generics disprefer predicates that apply to groups, we positively characterize the type of interpretation *a*-generics prefer: *a*-generics prefer predicates that apply to a randomly selected instance of a category. Furthermore, we motivate why the *a*-generic construction patterns the way it does: the conventional referential profile associated with indefinite NPs in English helps explain its functional constraints in generic interpretation. As a result, we predict that languages with analogous morphosyntactic distinctions (e.g., indefinite singular vs. definite singular forms) will exhibit similar distributional tendencies, and that reversed patterns would be typologically rare or marked.

4 Experiment 2

While Experiment 1 confirmed our core prediction, it left open the behavior of bare-plural generics. Experiment 2 introduces this additional form to evaluate whether it patterns more like *the*-generics or *a*-generics in its sensitivity to predicate collectivity.

4.1 Methods

Participants. We recruited 116 native English speakers via Prolific (68F, 45M, 2NB; M=37yrs). As planned, participants whose mean rating of three ungrammatical catches was equal to or higher than the mean rating of grammatical bare plurals were excluded from analyses (n = 25, excluded). Reported analyses were therefore run on 91 participants.

Materials. The same *a*-generic and *the*-generic sentences used in Experiment 1 were included, now along with a bare plural generic as well (e.g., *Computers have transformed education*). Since bare-plural generics were a new condition, we re-

[6] In doing so, we had to drop the random slope for article by item due to convergence issues.

duced the 6 filler sentences in Experiment 1 to 3 ungrammatical catch trials (all in bare plural form). As in Experiment 1, these 3 sentences served as attention checks and exclusion criterion. The stimuli can be found in Appendix C.

Procedure. The procedure for this experiment was the same as in Experiment 1 (participants were asked to rate each sentence's acceptability on a 7-point Likert scale). Each participant saw one version of each sentence (either with *a*-generic, *the*-generic or bare plural generic), with article type counterbalanced across participants. We further subdivided items, so that each participant judged 6 target sentences—2 *the*-generics, 2 *a*-generics and 2 bare-plural generics, from one of six lists, assigned randomly, along with the 3 catch trials. Items were presented in a randomized order for each participant.

4.2 Results

Results again confirmed that the perceived naturalness of generic noun phrases is modulated by the collective properties of the predicate. As shown in Figure 4, *a*-generics received the highest acceptability ratings when combined with predicates that were less collective, while receiving the lowest acceptability with predicates that were more collective. Bare-plural generics and *the*-generics trended toward greater acceptability with more collective predicates, though this effect did not reach significance in isolation.

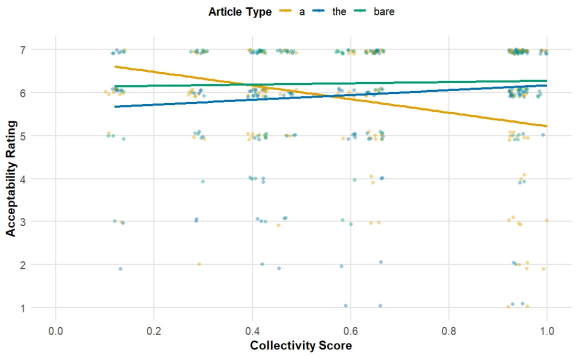

Figure 4: Acceptability ratings (1–7) as a function of collectivity scores (0–1) for *a*-generics, *the*-generics and bare generics. Each point represents an individual response. Solid lines depict linear fits for each article type. Collectivity scores, normed separately, capture the degree to which a predicate was judged to refer to a collective vs. individual property. The figure illustrates that article acceptability interacts with collectivity, with *a*-generics associated with lower acceptability as collectivity increases.

This was confirmed with the planned cumulative link mixed-effects model (fitted using the ordinal package in R; Christensen, 2019). The output was raw acceptability ratings on the Likert scale. Article (definite singular vs. indefinite singular vs. bare plural) and degree of collectivity were included as fixed interacting factors. Random intercepts and slopes for collectivity were included by participant, while by-item random effects included intercepts and random slopes for article. The model was fit to 546 observations (log-likelihood = –682.52; AIC = 1407.04). The predicted interaction between article type and predicate collectivity was statistically significant. While *a*-generics showed a negative effect of collectivity on acceptability ($\beta = -2.03$, SE = 0.9482, z = -2.137, p = .033), both bare and *the*-generics showed significantly more positive slopes compared to *a*-generics ($\beta = +3.02$, SE = 1.09, z = 2.76, p = .006, and $\beta = +3.06$, SE = 1.13, z = 2.71, p = .007, respectively), reversing the trend. Separate models for each article type[7] replicated what we saw in Experiment 1: *a*-generics showed a moderately significant decrease in acceptability as collectivity increased ($\beta = -1.62$, SE = 0.93, z = -1.735, p = .083). *The*-generics and bare plurals numerically trended in the opposite direction, but there was no significant effect of collectivity on acceptability for either *the*-generics ($\beta = +1.27$, SE = 0.80, z = 1.58, p = .114) or bare-generics ($\beta = +0.41$, SE = 0.69, z = 0.60, p = .551).

4.3 Discussion

This second experiment builds on Experiment 1 by introducing bare-plural generics, thereby allowing us to assess how they pattern with respect to predicate collectivity level. The results replicate the core finding from Experiment 1: *a*-generics decrease in acceptability as predicate collectivity increases, aligning with previous proposals that they are anchored in random instance interpretation. On the other hand, *the*-generics exhibit the opposite trend, albeit non-significantly when tested in isolation. Bare-plural generics show a positive trend similar to *the*-generics, suggesting they may prefer collective predicates more than *a*-generics. Crucially, although the positive effects of collectivity on *the*- and bare-plural generics did not reach significance in isolation, the interaction structure

[7] In fitting separate models for each article type, we had to drop the random slope for article by item and the random slope for collectivity level by participant, due to convergence issues.

of the model shows that their behavior is reliably distinct from that of *a*-generics.

This supports the broader claim that the morphosyntactic form of generic constructions modulates how it interacts with the properties of the predicate. This distinction helps motivate why languages differentiate morphosyntactic strategies for expressing genericity: each form carries its own functional constraints—whether tight or loose.

5 Conclusion

The findings presented in this paper provide the first experimental evidence that genericity is sensitive to constructional variation. By systematical comparisons, we demonstrate that different morphosyntactic forms are not interchangeable. Instead, each encodes distinct functional constraints motivated by more typical uses of the same forms. In particular, unlike *the*-generics and bare-plural generics, *a*-generics tend to combine with VPs predicating a property of any individual member of a category, but not of a collectivity. These findings support the Construction Grammar insight that it is *constructions* (i.e. form-meaning pairings), and not merely lexical items or semantic content, that shape interpretive possibilities.

Other factors may plausibly influence the choice between *a*-generics, *the*-generics and bare-plural generics. For instance, noun type (e.g. mass/count), register, or information structure might modulate acceptability judgments. Future research could explore how these factors interact with predicate collectivity level, possibly by extending the stimuli dataset. Future work may also investigate how morphosyntactic distinctions correlate with collectivity in other languages, and whether such patterns can be captured or induced in Large Language Models.

While preliminary, these findings lay the foundation for a broader empirical research agenda focused on genericity as a construction-sensitive phenomenon.

6 References

Bäck, A. (1996). Review of *Generische Kennzeichnungen*, by Gerhard Heyer (1987). *Noûs*, *30*(2), 276–281.

Bencini, G. M. L., & Goldberg, A. E. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language*, *43*(4), 640–651.

Borik, O., & Espinal, M. T. (2015). Reference to kinds and to other generic expressions in Spanish: Definiteness and number. *The Linguistic Review*, *32*, 167–225.

Burton-Roberts, N. (1977). Generic sentences and analyticity. *Studies in Language*, *1*, 155–196.

Carlson, G. N. (1977). *Reference to kinds in English* (Doctoral dissertation). University of Massachusetts, Amherst.

Carlson, G. N. (1989). On the semantic composition of English generic sentences. In G. Chierchia, B. H. Partee, & R. Turner (Eds.), *Properties, Types and Meaning* (pp. 167–192). Dordrecht: Springer.

Carlson, G. N. (1995). Truth conditions of generic sentences: Two contrasting views. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 224–237). Chicago: University of Chicago Press.

Carlson, G. N. (2011). Genericity. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An International Handbook* (Vol. II, pp. 1153–1185). Berlin: de Gruyter.

Champollion, L. (2020). Distributivity, collectivity and cumulativity. In L. Matthewson, C. Meier, H. Rullmann, & T. E. Zimmermann (Eds.), *Wiley's Companion to Semantics*.

Chierchia, G. (1995). Individual level predicates as inherent generics. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 176–224). Chicago: University of Chicago Press.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, *6*, 339–405.

Chierchia, G. (2022). 'People are fed up; don't mess with them.' Non-quantificational arguments and polarity reversals. *Journal of Semantics*, *39*, 475–521.

Clark, H. H. (1987). *Arenas of language use*. Chicago: University of Chicago Press.

Cohen, A., & Erteschik-Shir, N. (2002). Topic, focus, and the interpretation of bare plurals.

Natural Language Semantics, *10*(2), 125–165.

Cohen, A. (2001). On the generic use of indefinite singulars. *Journal of Semantics*, *18*(3), 183–209.

Condoravdi, C. (1994). *Descriptions in contexts* (Doctoral dissertation). Yale University.

Christensen, R. H. B. (2019). *ordinal — Regression Models for Ordinal Data* (R package version 2019.12-10). Retrieved from https://CRAN.R-project.org/package=ordinal

Dahl, O. (1985). *Tense and aspect systems*. Oxford: Blackwell.

Dayal, V. (2013). On the existential force of bare plurals across languages. In I. Caponigro & C. Cecchetto (Eds.), *From Grammar to Meaning* (pp. 49–80). Cambridge: Cambridge University Press.

Diesing, M. (1992). *Indefinites*. Cambridge, MA: MIT Press.

Dotlačil, J. (2010). *Anaphora and distributivity: A study of same, different, reciprocals and others* (Doctoral dissertation). Utrecht University.

Driemel, I., Hein, J., Carioti, D., Wünsch, J., Tsakali, V., Alexiadou, A., Sauerland, U., & Guasti, M. T. (2025). An experimental study on kind and generic readings across languages: Bare plural vs. definite plural. *Proceedings of the Amsterdam Colloquium*, 360–366.

Farkas, D. F., & de Swart, H. (2005). The generic article. In *Proceedings of the ESSLLI Workshop on Cross-linguistic Semantics*.

Fuellenbach, K. (2020). *A generic subject: The interplay of morphosyntax and the human conceptual system* (Doctoral dissertation). University of Oxford.

Fuellenbach, K., Gelman, S. A., Husband, E. M., & Cuneo, N. (2019). Generalising properties based on the morphosyntax of the subject: Generic and non-generic interpretations. *AMLaP 2019*, Moscow, Russia.

Galmiche, M. (1985). *Phrases, syntagmes et articles génériques*. Langage 79.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224.

Greenberg, Y. (2002). Two types of quantificational modalized genericity, and the interpretation of bare plural and indefinite singular NPs. In *Proceedings of SALT 12* (pp. 104–123).

Greenberg, Y. (2003). *Manifestations of genericity*. New York: Routledge.

Guerrini, J. (2024). Revisiting kind predication. [Manuscript under revision].

Guerrini, J. (2025). English bare plurals and distributivity. In N. Webster et al. (Eds.), *Proceedings of the 41st West Coast Conference on Formal Linguistics* (pp. 179–183).

Heyer, G. (1985). Generic descriptions, default reasoning and typicality. *Theoretical Linguistics*.

Heyer, G. (1987). *Generische Kennzeichnungen: Zur Logik und Ontologie Generischer Bedeutung*. München: Philosophia Verlag.

Humboldt, W. von. (1999). *On language: On the diversity of human language construction and its influence on the mental development of the human species* (P. Heath, Trans.; M. Losonsky, Ed.). Cambridge: Cambridge University Press. (Original work published 1836)

Kirkpatrick, J. (2022). Generic conjunctivitis. *Linguistics and Philosophy*, *45*.

Kiss, K. E. (1998). On generic and existential bare plurals and the classification of predicates. In S. Rothstein (Ed.), *Events and Grammar* (pp. 145–162). Dordrecht: Kluwer.

Kratzer, A. (1995). Stage-level and individual-level predicates. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 125–176). Chicago: University of Chicago Press.

Krifka, M. (1987). An outline of genericity. *SNS-Bericht 87–23*, University of Tübingen.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and Contextual*

Expressions (pp. 75–115). Dordrecht: Foris.

Krifka, M., Pelletier, F. J., Carlson, G. N., Meulen, A., Link, G., & Chierchia, G. (1995). Genericity: An introduction. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 1–124). Chicago: University of Chicago Press.

Krifka, M. (2003). Bare NPs: Kind-referring, indefinites, both, or neither? In *Semantics and linguistic theory XIII*, 180-203.

Landman, F. (1989). Groups, I. *Linguistics and Philosophy*, *12*(5), 559–605.

Langacker, R. W. (1997). Generics and habituals. *Current Issues in Linguistic Theory*, *143*, 191–222.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Lawler, J. (1973). Studies in English generics. *University of Michigan Papers in Linguistics*.

Leslie, S. J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, *65*(1), 15–31.

Leslie, S. J. (2015). 'Hillary Clinton is the only man in the Obama Administration': Dual character concepts, generics, and gender. *Analytic Philosophy*, *56*(2), 111–141.

Leslie, S. J., & Lerner, A. (2022). Generic generalizations. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition). Retrieved from https://plato.stanford.edu/archives/fall2022/entries/generics/

Nickel, B. (2008). Generics and the ways of normality. *Linguistics and Philosophy*, 31, 629–648.

Nunberg, G., & Pan, C. (1975). Inferring quantification in generic sentences. In *Papers from the Eleventh Regional Meeting of the Chicago Linguistic Society (CLS 11)* (pp. 412–422).

Nyugen, A. (2020). The radical account of bare plural generics. *Philosophical Studies*, *177*, 1303–1331.

Ojeda, A. (1991). Definite descriptions and definite generics. *Linguistics and Philosophy*, *14*, 367–397.

Pelletier, F. J., & Asher, N. (1997). Generics and defaults. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of Logic and Language* (pp. 1125–1179). Cambridge, MA: MIT Press.

Platteau, F. (1980). Definite and indefinite generics. In J. van der Auwera (Ed.), *The Semantics of Determiners* (pp. 112–123). London: Croom Helm.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605.

Scha, R. (1981). Distributive, collective and cumulative quantification. In J. Groenendijk, T. Janssen, & M. Stokhof (Eds.), *Formal Methods in the Study of Language*. Amsterdam: Mathematical Centre Tracts.

Wilkinson, K. (1995). The common noun kind. In G. N. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 383–397). Chicago: University of Chicago Press.

Appendices

A Analysis without exclusions (Exp. 1)

For this analysis, we fitted the same model as the one with exclusions, with the only exception that we had to drop the random intercept for item due to convergence issues. The model was fit to 486 observations (log-likelihood = –633.68; AIC = 1297.36). The predicted interaction between article and predicate type was significant ($\beta = 2.04$, SE = 0.84, z = 2.42, p = .0153).

We then fitted separate models for *a*-generic and *the*-generic sentences. Due to convergence issues, we simplified the random-effects structure by dropping the random slope for article by item. *A*-generics sentences showed negative but non-significant effects of collectivity on acceptability ratings ($\beta = -0.75$, SE = 0.81, z = -0.92, p = .358). In contrast, for *the*-generics, collectivity showed a positive but non-significant effect on ratings ($\beta = 0.76$, SE = 0.65, z = 1.16, p = .243).

B Sentence Stimuli for Experiment 1

Item	TARGET SENTENCES	Collectivity Score
1	The wolf hunts in packs. A wolf hunts in packs.	0.94
2	The wolf sharpens its teeth on bones. A wolf sharpens its teeth on bones.	0.41
3	The bee pollinates crops across the globe. A bee pollinates crops across the globe.	0.65
4	The bee dies after it stings. A bee dies after it stings.	0.12
5	The airplane revolutionized global travel. An airplane revolutionized global travel.	1.00
6	The airplane lowers its gear before landing. An airplane lowers its gear before landing.	0.59
7	The computer has transformed education. A computer has transformed education.	0.94
8	The computer boots up in seconds. A computer boots up in seconds.	0.29
9	The elephant is the largest land animal. An elephant is the largest land animal.	0.65
10	The elephant flaps its ears to cool down. An elephant flaps its ears to cool down.	0.41
11	The microwave modernized home cooking. A microwave modernized home cooking.	0.94
12	The microwave heats food in minutes. A microwave heats food in minutes.	0.47
FILLER SENTENCES		
	Cats purr them when they are content.	
	Birds build nests in spring.	
	Students often study late before exams.	
	Phones distract to people during meetings.	
	Doctors help to patients managing pain.	
	Plants grow faster in sunlight.	

C Sentence Stimuli for Experiment 2

Item	TARGET SENTENCES	Collectivity Score
1	The wolf hunts in packs. A wolf hunts in packs. Wolves hunt in packs.	0.94
2	The wolf sharpens its teeth on bones. A wolf sharpens its teeth on bones. Wolves sharpen their teeth on bones.	0.41
3	The bee pollinates crops across the globe. A bee pollinates crops across the globe. Bees pollinate crops across the globe.	0.65
4	The bee dies after it stings. A bee dies after it stings. Bees die after they sting.	0.12
5	The airplane revolutionized global travel. An airplane revolutionized global travel. Airplanes revolutionized global travel.	1.00
6	The airplane lowers its gear before landing. An airplane lowers its gear before landing. Airplanes lower their gear before landing.	0.59
7	The computer has transformed education. A computer has transformed education. Computers have transformed education.	0.94
8	The computer boots up in seconds. A computer boots up in seconds. Computers boot up in seconds.	0.29
9	The elephant is the largest land animal. An elephant is the largest land animal. Elephants are the largest land animals.	0.65
10	The elephant flaps its ears to cool down. An elephant flaps its ears to cool down. Elephants flap their ears to cool down.	0.41
11	The microwave modernized home cooking. A microwave modernized home cooking. Microwaves modernized home cooking.	0.94
12	The microwave heats food in minutes. A microwave heats food in minutes. Microwaves heat food in minutes.	0.47
FILLER SENTENCES		
	Cats drink quickly milk.	
	Students study often late before exams.	
	Doctors help to patients managing pain.	