

An experimental study on kind and generic readings across languages: bare plural vs. definite plural*

Imke Driemel^{1†}, Johannes Hein^{1‡}, Desiré Carioti^{2§}, Jakob Wünsch^{3¶}, Vina Tsakali^{4||},
Artemis Alexiadou^{3,1,*}, Uli Sauerland^{3,††} and Maria Teresa Guasti^{2‡‡}

¹ Humboldt University, Berlin, Germany

`imke.driemel@hu-berlin.de, johannes.hein@hu-berlin.de`

² University Milano-Bicocca, Milan, Italy

`desire.carioti@unimib.it, mariateresa.guasti@unimib.it`

³ Leibniz ZAS, Berlin, Germany

`wuensch@leibniz-zas.de, sauerland@leibniz-zas.de, artemis@leibniz-zas.de`

⁴ University of Crete, Crete, Greece

`tsakali@uoc.gr`

Abstract

To express kind and generic readings, Romance languages like Italian have been reported to use definite plurals but Germanic languages like English make use of bare plurals (Krifka et al., 1995; Chierchia, 1998), where Greek patterns with Romance (Alexiadou et al., 2007). German is discussed as an exception, as both bare and definite plurals are used to express kind (Krifka et al., 1995) and generic (Longobardi, 1994) readings. We present results from an experimental study comparing English, German, Italian and Greek, focusing on two novel findings: i) German and English both express kind/generic readings with bare plurals; ii) the effect of speaker distance (Acton, 2019) makes the definite plural an additional option in German.

1 Background

Generic and kind readings can be realized by a multitude of structures within and across languages. Italian and Greek, e.g., use definite plurals but English makes use of bare plurals (Krifka et al., 1995; Chierchia, 1998; Alexiadou et al., 2007), compare (1) to (2)/(3).

(1) English (Chierchia, 1998)

a. (*The) dogs are rare.

kind

b. (*The) dogs love to play.

generic

*We thank Natale Stucchi for pointing us to the Thurstone Method. Ethical review of the experimental protocols was done by the review boards of the University of Milan-Bicocca (RM-2020-229), the German Linguistic Society (#2019-20A-220622), and the ERC Executive Agency. Consent was obtained from participants. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (LeibnizDream, grant agreement No 856421).

†Contributions: Conceptualization, Methodology, Resources, Writing – Original Draft.

‡Contributions: Conceptualization, Methodology, Resources.

§Contributions: Resources, Methodology, Formal Analysis, Writing – Original Draft.

¶Contributions: Resources, Methodology.

||Contributions: Resources.

*Contributions: Conceptualization, Resources, Funding acquisition.

††Contributions: Conceptualization, Resources, Funding acquisition.

‡‡Contributions: Conceptualization, Resources, Methodology, Funding acquisition.

- (2) Italian (Chierchia, 1998)
- a. *(I) cani sono rari. *kind*
 the dogs are rare
 ‘Dogs are rare.’
- b. *(I) cani amano giocare. *generic*
 the dogs love to play
 ‘Dogs love to play.’
- (3) Greek (Lazaridou-Chatzigoga and Alexiadou, 2019; Alexiadou et al., 2007)
- a. *(Ta) pulja dodo ehun pleon afanisti. *kind*
 the birds dodo have already disappeared
 ‘Dodo birds have already disappeared.’
- b. *(I) ghatēs ine aksiolatrefta plasmata. *generic*
 the cats are adorable creatures
 ‘Cats are adorable creatures.’

German is reported to be exceptional, as both definite plurals and bare plurals license generic/kind readings, shown in (4).

- (4) German (Krifka et al., 1995; Longobardi, 1994)
- a. (Die) Pandabären sind vom Aussterben bedroht. *kind*
 the pandas are from extinction facing
 ‘Pandas are facing extinction.’
- b. (Die) Biber bauen Dämme. *generic*
 the beavers build dams
 ‘Beavers build dams.’

To account for (1) vs. (2/3), Chierchia (1998) proposes that arguments in English are mapped to kinds, formed by the kind operator \sqcap , a function from worlds to the sum of all instances of the kind in that world. Arguments in Italian (and Greek one can argue), however, are mapped to properties, thus they receive a kind reading via $\wedge\iota$, see (5). Since the definite determiner realizes ι and not \sqcap and the Blocking Principle enforces overt over covert type-shifting, Italian/Greek kind nouns are always overtly marked for definiteness, in contrast to English. Moreover, only plural nouns are licensed since a property holding of only one individual per world is not a natural kind.

- (5) a. English (1a): $rare(\sqcap(dogs))$
 b. Italian (2a): $rare(\wedge\iota(dogs))$

Dayal (2004) adopts Chierchia’s semantics in (1a) for Germanic, and crucially also for Romance languages. To account for the cross-linguistic split, she then proposes the ranking $\iota > \sqcap$ where the definite determiner may lexicalize either both ι and \sqcap , as in Romance, or only ι , as in English. The benefit of Dayal’s approach lies in accounting for German, where it is argued that both ι and \sqcap are lexicalized but the Blocking Principle is assumed to be inactive in the kind/generic domain, leading to the observed optionality in (4).

The observations regarding kinds can also be found in the generic domain since generic readings are often argued to be built on kinds, though they crucially also involve a GN operator (Carlson, 1977; Krifka et al., 1995). As shown in (6a), Chierchia (1998) uses GN to analyze generic readings as quantification over situations, where situations are contextually restricted. In prose, (6a) expresses that for every individual x and situation s such that x is a dog and s stands in relation C to x , x loves to play in s . For Italian, Chierchia proposes logically equivalent structures without making use of \cap , see (6b).

- (6) a. English (1b): GN x, s [$\cup \cap \text{dog}(x) \wedge C(x, s)$] [*love.play*(x, s)]
 b. Italian (2b): GN x, s [$x \leq \iota \text{dogs} \wedge C(x, s)$] [*love.play*(x, s)]

The empirical facts about English, however, have recently been challenged. Acton (2019) observed that generic readings can be expressed by definite plurals in English after all, though with the effect that the speaker distances themselves from the kind expressed, shown in (7). A natural question that arises is whether the same effect exists for other Germanic languages like German and if this effect could be one of the reasons for the perceived optionality between bare plurals and definite plurals in (4).

- (7) The distance effect (Acton, 2019)
 a. Americans love cars.
 b. The Americans love cars. \rightsquigarrow *speaker distance*

We investigate the introspective judgements reported in the literature with a quantitative study, aiming to tease apart kind and generic readings as well as the additional distance effect shown in (7).

2 Experiment

Little experimental work has been done on the contrasts reported in the previous section. Barton et al. (2015) conducted an acceptability judgement study (yes-no task) for German arguing for the optionality in (4). Ionin et al. (2011) used an acceptability judgement rating task (Likert scale) to verify the Germanic–Romance split, compare (1) and (2), and the well defined kind restriction (Carlson, 1977; Dayal, 2004) for English, Spanish and Brazilian Portuguese. Lazaridou-Chatzigoga and Alexiadou (2019) replicate this study for Greek. We present results from an acceptability judgement study that makes use of pairwise forced choice comparison. Unlike previous studies, we investigate the acceptability of sentences using Thurstone’s method of paired comparison (Thurstone, 1927a). We chose this task over a yes-no or rating task, as it is often difficult to explicitly accept or reject an utterance, and to quantify the acceptability of a sentence. Thurstone proposed that the choice between the preferred options of a pair can be a useful solution to order options on an objective scale independent from the individual observer (Thurstone, 1927a). The languages investigated for the kind/generic contrasts are English, German, Italian, and Greek. As both definiteness and number seem to be relevant in expressing kinds/generics, we tested 4 different noun types: singular definite, singular indefinite, bare plural, definite plural.

2.1 Participants

581 adult participants aged between 18 and 60 years took part in the study: 152 native speakers of English (age on average = 29.54, SD = 8.83), 155 native speakers of German (age on average = 24.53, SD = 5.1), 122 native speakers of Italian (age on average = 23.41, SD = 5.17), and 152 native speakers of Greek (age on average = 24.55, SD = 8.93). They were recruited online, through the Prolific platform (German, English, Greek) and the SONA System (Italian).

2.2 Materials and Method

Each participant was presented with 9 contexts and 4 sentences as possible completions for each of the contexts. We focus on the results of 3 contexts in this paper: the generic context (8), the kind context (9), and the generic distance context (10).¹ The target sentences presented in each context only differed in that they started with one of the four different nominals (singular definite, singular indefinite, bare plural, definite plural). Note that while we are presenting all 4 options in (8)–(10) for illustrative purposes, participants saw only two options per trial.

- (8) Generic context: There are many pests in the world that make our lives difficult. They eat our supplies, disturb our sleep, or plainly get on our nerves. For example:
- a. Mosquitos give us itchy bites in the summer.
 - b. The mosquitos give us itchy bites in the summer.
 - c. The mosquito gives us itchy bites in the summer.
 - d. A mosquito gives us itchy bites in the summer.
- (9) Kind context: The constant growth of the human population on earth has taken and still is taking its toll on other life on the planet, plant or animal. For example...
- a. Pandas are almost extinct.
 - b. The pandas are almost extinct.
 - c. The panda is almost extinct.
 - d. A panda is almost extinct.
- (10) Generic distance context: There is a place in town where people meet for a drink and a chat after work. As there are federal elections coming up soon, a lot of the discussions and debates revolve around politics. Yesterday, one guest seemed very upset and continuously complained that “voting is meaningless because ...
- a. politicians do whatever they want after the election anyway.”
 - b. the politicians do whatever they want after the election anyway.”
 - c. the politician does whatever s/he wants after the election anyway.”
 - d. a politician does whatever s/he wants after the election anyway.”

¹The other contexts constitute i) controls; ii) a context filtering for a normative reading; iii) two non-well defined kind/generic contexts set up to investigate the well defined kind restriction. For more information, see the preregistration <https://osf.io/wzxqf>.

For each context, all possible paired combinations of the 4 sentences (6 pairs) appeared successively, resulting in 54 trials per participant. Participants had to indicate their preferred option for each sentence pair.

2.3 Analysis

We used Thurstone scaling to derive a linear rating of all four nominal forms from two-way comparisons. Thurstone’s method of paired comparisons (Thurstone, 1927a,b) is a methodological procedure for psychophysical judgments. It is considered the gold standard for collecting, analyzing, and interpreting subjective introspective data (see Montag 2006; Cattelan 2012; Parraga 2015 for reviews). This method allows deriving the scaling of subjects’ preferences, quantifying the distance between two options. As a consequence, we obtain not just the preference order but also, from an interpretative perspective, the degree of acceptability of each option compared to the other. Once we obtained the percentage degree in which an option has been preferred in each pair, a z-score has been derived for each option. Between-options distances were interpreted based on z-scores as effect sizes. Here we only report the between-options results. All the analyses were performed in the R software environment for statistical computing (R Core Team, 2022).

2.4 Results

In Fig. 1, we show the results for the kind and the generic context for every language sample, based on the z-scores. Overall, we can see two clusters, English and German on the one hand and Italian and Greek on the other. In both contexts, bare plural was the best alternative for both English- and German-speaking participants, while plural definite was the best choice for Italian- and Greek-speaking participants. Thus, there is a clear distinction between the two sets of languages in the way they express genericity and kind readings.

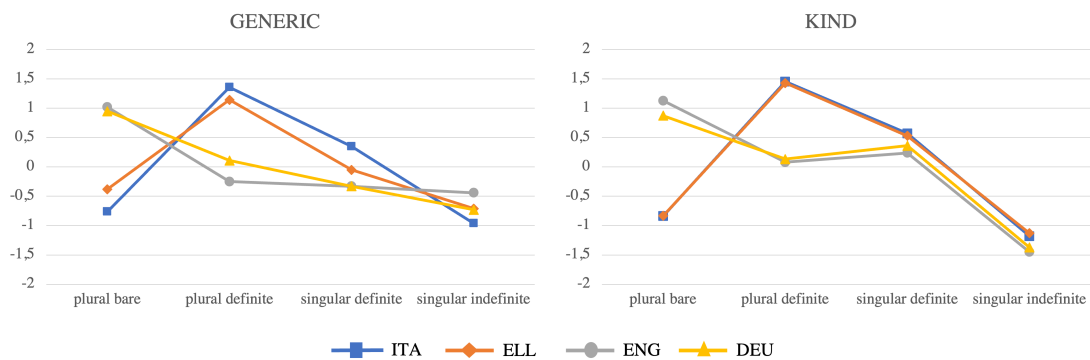
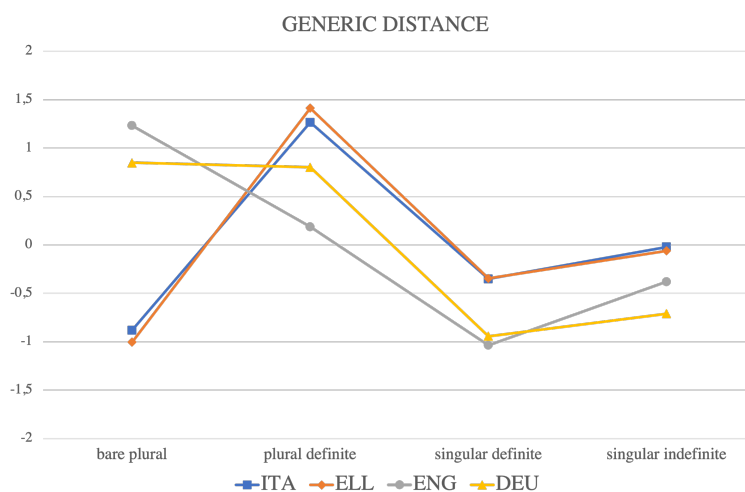


Figure 1: z-scores plotted on y -axis

In the distance context, shown in Fig. 2, German-speaking participants considered bare plural and plural definite equally acceptable. For English-speaking participants in the same context, the plural definite has a lower probability of being chosen than the bare plural, although it still remains the second best option. This means that in the distance context plural definite is a valid option for both languages, and more so in German than in English.

Figure 2: z-scores plotted on y -axis

3 Discussion

The results in Fig. 1 indicate that German is similar to English in expressing kinds/generics with bare plurals. In this sense, German is not exceptional (*pace* Longobardi 1994; Krifka et al. 1995; Dayal 2004) but patterns with English considering kind and generic contexts. This seems to be a welcome result, as the perceived optionality in German for (4) was explained by Dayal (2004) in terms of weakening the Blocking Principle in the kind/generic domain for German. Given that we find a clear distinction between Italian/Greek speakers, who use the definite plural, and English/German speakers, who use the bare plural as a default generic expression, our results support the universality of the Blocking Principle in the sense of Chierchia (1998).

The interesting finding comes from the distance context in Fig. 2 where in particular German speakers make use of definite plurals in addition to bare plurals. As expected by Acton (2019), a context which singles out a reading where speakers presumably distance themselves from the kind expressed in the target sentence boosts the use of the definite plural compared to the other kind/generic contexts shown in in Fig. 1. The default bare plural, however, is also used to an equal amount. This indicates either that only half of the speakers distance themselves from the kind expressed, or, more likely, that the expression of speaker distance is optional. The distance effect was also observed for English albeit to a much lower degree, which comes as a surprise. It could be explained by the fact that, due to ethical requirements, we could only recruit English speakers living in Ireland. Acton's corpus study, however, is based on speeches given in the US House of Representatives, thus representing American English. It is likely that register and/or cultural differences influence the use of distance marking.

References

- Acton, E. K. (2019). Pragmatics and the social life of the English definite article. *Language*, 95:37–65.
- Alexiadou, A., Haegeman, L., and Stavrou, M. (2007). *Noun Phrase in the Generative Perspective*. Mouton de Gruyter, Berlin.
- Barton, D., Kolb, N., and Kupisch, T. (2015). Definite article use with generic reference in German: an empirical study. *Zeitschrift für Sprachwissenschaft*, 34:147–173.
- Carlson, G. (1977). *Reference to kinds in English*. PhD thesis, University of Massachusetts, Amherst.
- Cattelan, M. (2012). Models for Paired Comparison Data: A Review with Emphasis on Dependent Data. *Statistical Science*, 27:412–433.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6:339–405.
- Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy*, 27:393–450.
- Ionin, T., Montrul, S., and Santos, H. (2011). An experimental investigation of the expression of genericity in English, Spanish and Brazilian Portuguese. *Lingua*, 121:963–985.
- Krifka, M., Pelletier, F. J., Carlson, G. N., ter Meulen, A., Link, G., and Chierchia, G. (1995). Genericity: An introduction. In Pelletier, F. J. and Carlson, G. N., editors, *The Generic Book*, pages 1–124. The University of Chicago Press, Chicago/London.
- Lazaridou-Chatzigoga, D. and Alexiadou, A. (2019). Genericity in Greek: an experimental investigation. In Gattnar, A., Hörnig, R., Störzer, M., and Featherston, S., editors, *Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory*, pages 245–260. The University of Tübingen, Tübingen.
- Longobardi, G. (1994). Reference and proper names: A Theory of N-movement in Syntax and Logical Form. *Linguistic Inquiry*, 25:609–665.
- Montag, E. D. (2006). Empirical formula for creating error bars for the method of paired comparison. *Journal of Electronic Imaging*, 15:010502.
- Parraga, C. A. (2015). Perceptual Psychophysics. In Cristóbal, G., Perrinet, L., and Keil, M. S., editors, *Biologically Inspired Computer Vision*, pages 81–108. Wiley.
- R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Thurstone, L. L. (1927a). A law of comparative judgment. *Psychological Review*, 34:273.
- Thurstone, L. L. (1927b). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, 21:384.