How can genericity be expressed? A four-language experimental study using Thurstone Scaling*

Imke Driemel, Johannes Hein, Desiré Carioti, Jakob Wünsch,
Vina Tsakali, Artemis Alexiadou, Uli Sauerland & Maria Teresa Guasti

December 2024

Abstract

Generic and kind readings can be realized by a multitude of structures within and across languages. Romance languages like Italian for example make use of definite plurals but English employs bare plurals, while both options seem to be available in German. Across languages, further additional but restricted generic readings, such as the taxonomic reading and the rule-based reading, can be expressed by the use of the definite singular and the indefinite singular, respectively. As there is little cross-linguistic empirical work on this topic, we conducted an experimental study on English, German, Italian, and Greek investigating the restrictions on the distribution of the different noun types. In doing so, we used the Thurstone method for comparative judgements, a design that is rarely used in experimental linguistics. Specifically, our results report on the universality of the Blocking Principle, the cross-linguistic robustness of the well-defined kind restriction on the definite singular, and an underlying rule-based account of genericity for the indefinite singular.

Keywords: generic, kind, well defined kind restriction, bare plural, definite plural, indefinite plural, definite singular, English, German, Greek, Italian

1 Introduction

Generalizations over a group of individuals can be expressed in different ways across languages. On the meaning side, two notions that are often distinguished are D-level genericity (1) and I-level genericity (2). The former makes reference to an entire class of individuals by

*This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (LeibnizDream, grant agreement No 856421). 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. We thank Natale Stucchi for pointing us to the Thurstone Method. *Authors' contributions:* Imke Driemel (first author): Conceptualization, Methodology, Resources, Writing – Original Draft; Johannes Hein: Conceptualization, Methodology, Resources; Desiré Carioti: Resources, Methodology, Formal Analysis, Writing – Original Draft; Jakob Wünsch: Resources, Methodology; Vina Tsakali: Resources; Artemis Alexiadou: Conceptualization, Resources, Funding acquisition; Uli Sauerland: Conceptualization, Resources, Funding acquisition; and Maria Teresa Guasti: Conceptualization, Resources, Methodology, Funding acquisition.

using kind-level predicates, whereas the latter expresses a typical situation that occurs with instances of a kind (Krifka et al. 1995). On the form side, languages encode the noun phrase in different ways in sentences expressing generic and kind statements. We exemplify the patterns with D-level genericity in (1). In English, the most common way is to use a bare plural (1a), while in Italian the definite plural is used (1b). Greek behaves in many ways like Italian and other Romance languages in making use of the definite plural (1c). In German, both the bare plural and the definite plural are reported to be possible (1d).

(1) *D-level genericity across languages (kind readings)*

a. (*The) dogs are rare. *English*

b. **(I) cani sono rari.* *Italian*
the.PL dog.PL are rare
'Dogs are rare.' (Chierchia 1998: 342)

c. **(Ta) pulja dodo ehun pleon afanisti.* *Greek*
the.PL bird.PL dodo have already disappeared
'Dodo birds have already disappeared.' (Lazaridou-Chatzigoga and Alexiadou 2019: 248)

d. *(Die) Pandabären sind vom Aussterben bedroht.* *German*
the.PL panda.PL are from extinction facing
'Pandas are facing extinction.' (Krifka et al. 1995: 68)

Interestingly, while languages differ in the way they encode the noun phrase to express D-level genericity, each individual language pattern replicates with I-level genericity, as can be seen in (2). Greek and Italian require the definite determiner on the noun phrase (2b/2c), whereas English blocks it (2a). In German, both forms seem to be permitted again (2d). Hence, the cross-linguistic picture suggests a uniform analysis for D-level and I-level genericity, when it comes to the use of bare plural and definite plural.

(2) *I-level genericity across languages (generic readings)*

a. (*The) dogs love to play. *English*

b. **(I) cani amano giocare.* *Italian*
the.PL dog.PL love to.play
'Dogs love to play.' (Chierchia 1998: 341)

c. **(I) ghates ine aksiolatrefta plasmata.* *Greek*
the.PL cat.PL are adorable creatures
'Cats are adorable creatures.' (Alexiadou et al. 2007: 165)

d. *(Die) Biber bauen Dämme.* *German*
the.PL beaver.PL build dams
'Beavers build dams.' (Longobardi 1994: 653)

To express D-level genericity, an additional option is available, which is the definite singular (Vendler 1967, Carlson 1977, Krifka et al. 1995, Dayal 2004). This is possible across languages, as is shown with our small sample in (3) for English, German, Italian, and Greek, though its application is restricted to well defined kinds.

(3) *D-level genericity with definite singulars*

a. The dodo is extinct. *English*

b. *Il dodo è estinto.* *Italian*
the.SG dog.SG is extinct
'The dodo is extinct.' (Chierchia 1998: 342)

c. *To puli dodo ehi pleon afanisti.* *Greek*
the.SG bird.SG dodo has already disappeared
'The dodo bird has alread disappeared.' (Lazaridou-Chatzigoga and Alexiadou 2019: 248)

d. *Der Pandabär ist vom Aussterben bedroht.* *German*
the.SG panda.SG is from extinction facing
'The panda is facing extinction.' (Dayal 2004: 442)

Moreover, to express I-level genericity the indefinite singular can also be used (Lawler 1973, Burton-Roberts 1977, Krifka 2013), see (4). Again, such an option is more restricted, in that it is only acceptable if the sentences express essential properties of the kind, where the statements have an analytic (4a) or normative character (4b). Thus, the use of the indefinite singular is infelicitous with kind-level predicates (4c). While most of the literature is centred on English, the generic use of the indefinite singular seems to be an option cross-linguistically, see Mari et al. (2013) for discussion.

(4) *I-level genericity with indefinite singulars*

a. A tiger climbs trees. (Burton-Roberts 1977: 155)

b. A gentleman opens doors for ladies. (Cohen 2001: 196)

c. *A lion will become extinct soon. (Krifka et al. 1995: 10)

Taken together, generic meanings can be expressed by a variety of different noun types, including bare plurals, definite plurals, definite singulars, and indefinite singulars. Given that at least two of the options do not distribute as widely as the others, it is likely that there is not only one underlying semantic structure which allows for generic inferences. This many-to-many correlation has resulted in a flourishing research field where a number of solutions have been proposed to the form-meaning pairings. Two default noun types have been identified for Indo-European languages: the bare plural and the definite plural. However, not much is known about the acceptability of the definite and the indefinite singular with respect to the defaults. Since most of the generalizations are based on introspective data and there is little quantitative work on the topic, we conducted a comparative judgement study on English, German, Italian, and Greek with two goals: (i) to verify the default nature to express kind/generic readings by the use of the definite plural and the bare plural and (ii) to investigate in how far the alternative noun types definite singular and indefinite singular become an equally good option in

certain contexts. The remainder of the paper is organized as follows: section 2 will give some background on the theoretical and empirical insights in the domain of genericity. Section 3 will introduce the methodology of our experiment and our hypotheses, while the results will be presented in section 4. We will discuss our results and their implications in section 5, before section 6 concludes.

2 Previous work

This section provides an overview over the theoretical ideas that have been proposed to account for the four noun types we find across languages in the expression of generic statements. We begin with a discussion of plural kind formation, i.e., the defaults, in section 2.1. The default status comes from the fact that these noun types spell out a kind operator, i.e., it is the most direct way to trigger a kind/generic reference. In Chierchia's (1998) terms, this is either $^{\cap}$ or $^{\wedge}\iota$ depending on the language. This section is followed by section 2.2 on definite singulars and section 2.3 on indefinite singulars, noun types which realize alternative ways to express kind or generic readings. In the discussion, we mostly focus on English, German, Italian, and Greek, as these are the languages we will investigate in our experimental study. For each section, we additionally report on quantitative studies aimed at investigating the empirical picture.

2.1 Defaults: bare plurals and definite plurals

To account for the difference between English and Italian, that is (1a/2a) vs. (1b/2b), Chierchia (1998) proposes the Nominal Mapping Parameter (NMP), a parameter that determines for each language whether NPs can denote individuals (common nouns, kinds) or properties, or both. The NMP for English is set to [+arg,+pred], thus arguments in English can be mapped to kinds (5a), formed by the kind operator $^{\cap}$, a function from worlds to the sum of all instances of the kind in that world. Arguments in Italian, however, are mapped to properties since the NMP for Italian is set to [−arg,+pred]. Consequently, NPs in Italian receive a kind reading via $^{\wedge}\iota$ (5b), that is a maximality operator ι used for definite contexts with an additional intensionalizer $^{\wedge}$. Since the definite determiner realizes ι and not $^{\cap}$, nouns combining with kind-level predicates in Italian are always overtly marked with the definite determiner.[1] In contrast, kind nouns in English are bare because the $^{\cap}$ operator receives zero spell-out. For both languages, only plural nouns are licensed since a property holding of only one individual per world is not a natural kind. Greek behaves in many ways like Italian, hence the NMP in Greek is arguably also set to [−arg,+pred].

(5) *D-level genericity in Chierchia (1998)*

a. English (1a): $\textit{rare}(^{\cap}(\textit{dogs}))$

b. Italian (1b): $\textit{rare}(^{\wedge}\iota(\textit{dogs}))$

The underlying semantics for Italian in (5b) leaves room for ι-shift to take place covertly, falsely predicting bare nouns in Italian to be able to make reference to kinds. To exclude this possibility, Chierchia (1998) introduces the principle in (6) which enforces overt over covert type-shifting.

[1] An implicit assumption is that $^{\wedge}$ is never spelled out overtly.

(6) *Blocking Principle* (Chierchia 1998: 360)
For any type shifting operation τ and any X: $*\tau(X)$ if there is a determiner D such that for any set X in its domain, $D(X) = \tau(X)$.

The cross-linguistic pattern replicates for I-level genericity, recall (2). For this reason, 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 (7a), Chierchia (1998) uses GN to analyze generic readings as quantification over situations, where situations are contextually restricted. In prose, (7a) 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. The quantification is by and large universal, though it is pragmatically restricted. The additional $^{\cup}$ operator is needed to access instantiations of the kind. For Italian, Chierchia proposes logically equivalent structures without making use of $^{\cap}$, see (7b). Again, the same reasoning can be applied to Greek. The structures in (7) predict that English makes use of bare plurals, whereas Italian (and Greek) make use of definite plurals to express I-level genericity.

(7) *I-level genericity in Chierchia (1998)*

a. English (2a): GN x, s [$^{\cup\cap}$dog$(x) \wedge C(x, s)$] $[love.play(x, s)]$

b. Italian (2b): GN x, s [$x \leq \iota$dogs $\wedge C(x, s)$] $[love.play(x, s)]$

While the NMP provides us with a theory for why the bare plural is the default to express D-level and I-level genericity in English, whereas the definite plural is the default in Italian (and Greek), it does not extend to the observations made for German, for which the default nature of the bare plural is controversial.
Another way to capture the differences across languages is proposed by Dayal (2004). Dayal abandons the NMP for number marking languages and instead extends Chierchia's semantics for English (5a)/(7a) to all languages, shown in (8a) and (8b). To account for the cross-linguistic split, a universal ranking of maximality operators is proposed: $\iota > ^{\cap}$. Intuitively, the scale indicates canonical functions of the definite determiner, where ι is more canonical than $^{\cap}$. The extent to which the scale is lexicalized can vary from language to language. The definite determiner may lexicalize either both ι and $^{\cap}$, as in Italian (and Greek), or only ι, as in English. The spell-out rules are given in (8c).

(8) *D-level and I-level genericity in Dayal (2004)*

a. English/Italian kinds:
rare $(^{\cap}(dogs))$

b. English/Italian generics:
GN x, s [$^{\cup\cap}$dog$(x) \wedge C(x, s)$] $[love.play(x, s)]$

c. Assumptions about $\iota > ^{\cap}$:
Italian: $\iota, ^{\cap} \leftrightarrow$ *il / la / i / le*
English: $\iota \leftrightarrow$ *the*

One of the benefits[2] of Dayal's approach lies in accounting for the German pattern where it was reported that both the bare plural and the definite plural are possible (1d)/(2d). Dayal (2004: 442) argues for German and Italian that both ι and $\cap$ are lexicalized, but the Blocking Principle is assumed to be inactive in German in the kind/generic domain. This follows from the fact that $\cap$ is ranked lower on the scale, and thus the Blocking Principle can be switched off in some languages for $\cap$. If the Blocking Principle is inactive, $\cap$ can but does not have to be realized as the definite determiner in (1d)/(2d), leading to the optionality between bare plurals and definite plurals.

While Dayal's account finds a way to capture the optionality between bare plurals and definite plurals in German, the adjustment comes at a cost, as it severely weakens the validity of the Blocking Principle. Hence, it is worth investigating whether such a step is necessary. Although the optionality in German is widely believed to exist in the theoretical literature (Brugger 1994, Longobardi 1994, Krifka et al. 1995, Dayal 2004, Oosterhof 2004, Behrens 2005, Schaden 2012, de Swart and Zwarts 2012), results from experimental studies are so far inconclusive (Barton et al. 2015, Czypionka and Kupisch 2019). Barton et al. (2015) conducted an acceptability judgement study where German-speaking participants were asked to judge sentences (yes/no task) containing bare plural or definite plural subjects under D-level and I-level generic readings. Sentences were presented auditorily accompanied by pictures. Each sentence was preceded by *Jedes Kind weiSS* ... ('Every child knows...') to create the generic setting. Acceptance rates and experimental items are presented in (9) and (10). Barton et al. (2015) conclude that the acceptance rates confirm the the availability of the definite plural for generic statements, but note that the acceptability of definite plurals is lower than that of bare plurals and dependent on predicate type as well as the speaker's age and education level.

(9) *D-level genericity: Jedes Kind weiSS ...* (Barton et al. 2015)

a. *Blauwale sind vom Aussterben bedroht.* 100% accepted
blue.whale.PL are of.the extinction threatened
'Blue whales are facing extinction.'

b. *Die Eisbären sind vom Aussterben bedroht.* 84.9% accepted
the.PL polar.bear.PL are of.the extinction threatened
'The polar bears are facing extinction.'

(10) *I-level genericity: Jedes Kind weiSS ...* (Barton et al. 2015)

a. *Kaninchen sind Einzelgänger.* 100% accepted
rabbit.PL are loners
'Rabbits are loners.'

b. *Die Pferde sind Herdentiere.* 61.9% accepted
the.PL horse.PL are gregarious.animals
'Horses are gregarious animals.'

[2] Another advantage of deriving kinds with the $\cap$-operator across languages is that the restriction to plural kinds follows more naturally. It is not actually clear that $^\wedge\iota$ (assumed for Italian in Chierchia's account) indeed derives this restriction since singular nouns denoting singleton properties do indeed meet the presupposition of ι, as one reviewer reminds us.

Czypionka and Kupisch (2019) conducted a slightly different acceptability judgement study on the availability of generic readings for the definite plural in German. They tested definite plural subjects under I-level generic readings, and compared them with bare plural and demonstrative plural subjects. Participants were presented with a picture that displayed either a canonical property of a species or a non-canonical one. At the same time, they were presented with an auditory sentence containing a bare plural, a definite plural, or a demonstrative plural subject. The participants' task was to determine whether the sentence they heard matched the picture. The acceptance rates are presented in (11), along with the test items.[3]

(11) *Task: Is this sentence correct?* (Czypionka and Kupisch 2019: 282)

a. *Eisbären sind weiSS.* auditory stimulus
polar.bear.PL are white
'Polar bears are white.'
→ Picture with white polar bears (canonical): 98.81% accepted
→ Picture with pink polar bears (non-canonical): 83.73% accepted

b. *Die Eisbären sind weiSS.* auditory stimulus
the.PL polar.bear.PL are white
'The polar bears are white.'
→ Picture with white polar bears (canonical): 98.41% accepted
→ Picture with pink polar bears (non-canonical): 16.87% accepted

c. *Diese Eisbären sind weiSS.* auditory stimulus
these polar.bear.PL are white
'These polar bears are white.'
→ Picture with white polar bears (canonical): 98.41% accepted
→ Picture with pink polar bears (non-canonical): 3.77% accepted

As predicted, bare plurals were accepted independently of the presented picture since they are interpreted as generic, thereby excluding an interpretation that is specific to the picture (11a). In contrast, demonstrative plurals were accepted only in the canonical condition (11c) since, as expected, they can only be interpreted as pointing towards the referent in the picture. Interestingly, the definite plurals patterned with the demonstrative plurals in that they were predominantly accepted only in the canonical condition (11b). This implies that definite plurals were not interpreted generically, and thus the results run counter to the claim that definite plurals can express genericity in addition to bare plurals. The two acceptability judgement studies (Barton et al. 2015, Czypionka and Kupisch 2019) point to partially opposite conclusion with regard to the availability of the definite plural in German generic sentences. We conclude that the default status of the bare plural in German is so far not robustly attested. While Dayal's (2004) theory allows for two defaults in German, the NMP (Chierchia 1998) predicts the bare plural to be the default noun type to express genericity.

Not only is the option of the definite plural in generic contexts still under debate for German, but its availability has also been raised for English. Acton (2019) argues that the definite determiner in English can take on a social meaning if it is used in environments where the bare plural is

[3]We are reporting the results of the paper from the appendix, which excludes 14 participants who did not understand the task (Czypionka and Kupisch 2019: 263–264).

also felicitous. In such cases the use of the definite plural triggers a distancing effect of the speaker, either by "deemphasizing their membership in the group", or by "emphasizing their nonmembership" (Acton 2019: 38). We see an illustration of this effect in (12). While both (35a) and (35b) express generic statements, (35b) seems to trigger an additional inference of speaker distance.

(12) *The distance effect with definite plurals* (Acton 2019: 37, 51)

a. Americans love cars.

b. The Americans love cars. $\rightsquigarrow$ *The speaker is not an American or wishes to express distance from Americans.*

Acton (2019) bases his account on a corpus study of the House Proceedings Corpus (Djajali 2013), which comprises a complete set of transcripts of the proceedings of the US House of Representatives from February 1993 through December 2012. An analysis of the usage of *(the) Republicans* vs. *(the) Democrats* revealed a disproportionally higher usage of definite plurals over bare plurals by representatives talking about their opposing party than talking about their own. More concretely, Acton (2019: 42) found that out of all instances of definite plurals and bare plurals for *Republicans/Democrats*, Republicans were using *the Republicans* 26.1% of the time ($N = 11,042$) and *the Democrats* 53.3% of the time ($N = 13,007$), while Democrats were using *the Republicans* 54.4% of the time ($N = 18,992$) and *the Democrats* 30.4% of the time ($N = 11,352$). Since the data was annotated mostly automatically, the data set subsumes all uses of the definite plural and the bare plural of *Republicans/Democrats*, including presumably also episodic and anaphoric uses. Hence, a controlled experiment is needed to verify if the distribution is the same in generic statements only. Furthermore, this study raises the question as to whether the same effect exists for other Germanic languages and if this effect could be one of the reasons for the perceived optionality between bare plurals and definite plurals in German.

Based on the observations in this section, we conclude that the default noun type to express genericity in Italian (and Greek) is the definite plural, while in English the default noun type is the bare plural. There is an option in English to use the definite plural to express generic statements, but this option comes with a pragmatically enriched reading in that the speaker wants to express distance to the kind denoted by the noun phrase. For German, the default status of the bare plural is still under debate.

2.2 Definite singular

Nominal expressions do not have to be marked for plural to make reference to kinds. As was shown in (3), languages across the board can use kind-level predicates with definite singular noun phrases (Vendler 1967, Carlson 1977, Krifka et al. 1995, Chierchia 1998, Dayal 2004). The use of definite singulars, however, seems to be restricted to well-defined kinds (Geoffrey and Pan 1975, Carlson 1977, Krifka et al. 1995, Dayal 2004, Ionin et al. 2011). For an illustration, let us consider the minimal pair in (13). Polar bears are natural kinds in the sense that speakers agree that they form a species. Hence, they can be predicated over with kind-level predicates (13a). White bears, however, are not an agreed upon established kind. Consequently, (13b) is only marginally accepted.

(13) *The well defined kind restriction (WDK)* (Carlson 2011: 1181)

a. The polar bear is slowly disappearing.

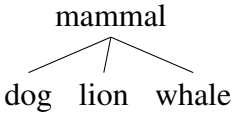
b.??The white bear is slowly disappearing.

In order to explain how singular entities can make reference to kinds, Dayal (2004) argues that definite singulars can trigger kind readings via the use of a taxonomic determiner, that is a version of the regular definite determiner (ι) which combines with a noun ranging over taxonomic entities, see (14b). The reason why in (14a) the kind-level predicate can combine with the singular noun *lion* is that the definite determiner picks out the unique taxonomic entity which is the species lion. This analysis also predicts that the definite singular can only make reference to well-defined kinds, as taxonomies have to be established.

(14) *Definite singulars as taxonomic entities* (Dayal 2004: 426)

a. The lion might become extinct.

b. $[\![the_{tax}]\!] = \lambda P.\iota X[P(X)]$, where X ranges over entities in the taxonomic domain

c. 

Interestingly, Dayal's analysis does not include the kind operator $\cap$ since it achieves a kind-like reading by letting the restrictor noun denote a taxonomic entity. As the availability of the taxonomic determiner is independent of the Nominal Mapping Parameter and/or the lexicalization of the $\cap$ operator, nothing prevents this determiner from allowing kind-level predication across Germanic, Romance, and other languages. Indeed, this is what we observed in (3). Dayal's analysis also provides a partial solution to the too-many-surface-structures puzzle. Given that the definite singular noun phrase in (14a) does not realize the $\cap$ operator, the theory does not have to be revised to accommodate kind-referring definite singulars.
One crucial prediction of the taxonomic determiner is that the WDK restriction holds cross-linguistically. For Romance languages in particular, however, the empirical status of the WDK restriction is currently under debate. Whereas non-well-defined kinds encoded as definite singulars do not seem to allow kind readings in Italian (Dayal 2004), they seem to do so in French and Spanish (Vergnaud and Zubizarreta 1992, Borik and Espinal 2015).

(15) *WDK in Romance languages*

a. *Le tigre blessé est dangereux.* $\rightsquigarrow$ kind *French*
the.SG tiger.SG wounded is dangerous
'The wounded tiger is dangerous.' (Vergnaud and Zubizarreta 1992: 644)

b. *La tigre a tre zampe è facile da cacciare.* $\not\rightsquigarrow$ kind *Italian*
the.SG tiger.SG with three legs is easy to hunt
'The tiger with three legs is easy to hunt.' (Dayal 2004: 439)

c. *El tigre herido es peligroso.* $\rightsquigarrow$ kind *Spanish*
the.SG tiger.SG wounded is dangerous
'The wounded tiger is dangerous.' (Borik and Espinal 2015: 199,fn.26)

Ionin et al. (2011) conducted an acceptability judgement study to investigate the WDK restriction in a number of languages, including English and Spanish. The study was replicated for Greek by Lazaridou-Chatzigoga and Alexiadou (2019). Participants rated 5 target sentences against a context, based on a scale from 1 (=unacceptable) to 4 (=acceptable). We report the mean ratings from both studies for the crucial conditions in (16) and (17).

(16) *Well defined kind condition*:
I really like going to the zoo. Unfortunately, there are many animals that cant be found in a zoo, or anywhere else. Its very sad. For example ...

a. The dodo bird is extinct. *English*, $M = 3.55$

b. *El pájaro dodo está en extinción* *Spanish*, $M = 3.77$
the.SG bird.SG dodo is in extinction
'The dodo bird is extinct.'

c. *To puli dodo ehi pleon afanisti* *Greek*, $M = 3.69$
the.SG bird.SG dodo has already disappeared
'The dodo bird has already disappeared.'

(17) *Non-well defined kind condition*:
My brother has been in a bad mood lately. And no wonder: his apartment is so uncomfortable, it must be very depressing to live there. And he has a very dim and unpleasant overhead light. I told him he should buy a new lamp, something pleasant. For example, I know that ...

a. The green lamp is very relaxing. *English*, $M = 2.01$

b. *La lámpara verde es muy relajante* *Spanish*, $M = 2.52$
the.SG lamp.SG green is very relaxing
'The green lamp is very relaxing.'

c. *I prasini lamba ine poli halarotiki* *Greek*, $M = 3.29$
the.SG lamp.SG green is very relaxing
'The green lamp is very relaxing.'

All in all, the results in Ionin et al. (2011) and Lazaridou-Chatzigoga and Alexiadou (2019) confirm the existence of WDK effects across languages, i.e., the definite singular is preferred with well defined kinds.[4] The mean ratings for each language were significantly worse in (17) compared to (16). For Greek, Lazaridou-Chatzigoga and Alexiadou (2019: 258) point out in their conclusion that the mean rating was unexpectedly high in (17), which warrants further investigation.

One potential weakness of the set-up in (16) and (17) is that well-defined kinds were tested with D-level genericity but non-well-defined kinds were tested with I-level genericity. However, the predictions of Dayal's theory about the taxonomic determiner only relate to the combination with kind-level predicates, i.e. D-level genericity. In fact, Dayal (2004: 431–433) notes several

[4] Ionin et al. (2011) did not find the same WDK effects for plural definites, which is expected since Dayal's theory of the taxonomic determiner does not make predictions regarding plural kind formation.

subtle differences between the bare plural and the definite singular under I-level generic readings, leading her to conclude that such readings cannot be derived by quantifying over instances of the kind. Instead, I-level generic readings of the definite singular have to be derived in an alternative way, potentially via a representative or prototypical object, though the details are left to future research as is the question whether WDK effects are predicted with I-level genericity. In any case, it is presumably more informative to test well-defined and non-well-defined kinds under D-level as well as I-level generic readings, separately.
In sum, we predict that the definite singular can be an alternative option to the default option to express kind readings across all four languages but its use is restricted to well-defined kinds. We will investigate in our study whether the WDK restrictions hold cross-linguistically, with a specific interest in Romance languages and Greek, as this has been a matter of debate. In order to do so, we tease apart D-level and I-level genericity and set up a well-defined vs. non-well defined kind contrast for each level of genericity.

2.3 Indefinite singular

Indefinite singulars have been argued to express I-level genericity, though the usage is restricted to the expression of essential properties of the kind (Lawler 1973, Burton-Roberts 1977, Greenberg 2003, Cohen 2001, Krifka 2013). The following data illustrate this restriction. While it is in general possible to form a generic statement with the indefinite singular (18a), the generic reading is blocked in (18b), instead the bare plural has to be used to trigger a generic reading (18c). The difference between (18a) and (18b) is that tree climbing is a defining property of tigers, whereas singing German arias in the shower is not a defining property of famous semanticists.

(18) *Essential properties with the indefinite singular* (Greenberg 2003: 32)

a. A tiger climbs trees. $\rightsquigarrow$ generic (repeated from (4a))

b. A famous semanticist sings German arias in the shower. $\not\rightsquigarrow$ generic

c. Famous semanticists sing German arias in the shower. $\rightsquigarrow$ generic

Some theories of such uses of indefinite singulars do not build on generically referring terms or quantification but on the assumption that the generic readings are derived from rules and regulations speakers agree on (Carlson 1995, Cohen 2001, Greenberg 2003). Cohen (2001) for example argues that the bare plural in (19a) is ambiguous between a descriptive generalization and the statement that a normative rule is in place. The indefinite singular in (19b) only has the latter interpretation.

(19) *The normative effect with indefinite singulars* (Cohen 2001: 196)

a. Gentlemen open doors for ladies.

b. A gentleman opens doors for ladies. $\rightsquigarrow$ *moral necessity*

Cohen (2001) argues for a view where the generic reading of English bare plurals in (19a) is derived by standard GN quantification, but the generic reading of the indefinite singular involves some form of a conditional describing a rule that is in effect, as shown in (20). In words, (20) denotes a rule that is in effect such that if x is a gentleman, x opens doors for ladies.

The formula in (20) neatly explains why it is the indefinite determiner that allows for such a reading, given that indefinites are treated as variables also outside of generic environments (Kamp 1981, Heim 1982).

(20) *Cohen's (2001: 197) logical form for (19b)*

in-effect(!(gentleman(x) ⇒ open-doors-for-ladies(x)))

Although the observations are so far limited to English, the argument presented here is universal and therefore cross-linguistic differences are not predicted. As with the analysis of the definite singular as a taxonomic determiner in section 2.2, the analysis of the indefinite singular does not rely on the NMP. Thus, a generic context that expresses an essential property of the kind being predicated over, possibly as a consequence of a set of normative rules speakers agree on, should allow the use of the indefinite singular as an alternative option to the default across several unrelated languages. In our study, we will set up a context enforcing a normative reading, for which we predict the indefinite singular to become an alternative to the default option.

3 A comparative judgement study

We investigate the introspective judgements reported in the literature by means of a quantative study, aiming to tease apart kind and generic readings as well as the additional effects presented in section 2. The preregistration for this study can be found here: OSF view-only link.

3.1 Alternative forced choice and Thurstone scaling

Our goal is to understand which among various options (options being sentences or sentences in a given context) is the most acceptable for speakers, but we also want to ask a more gradient question, i.e., how much one option is considered more acceptable than another. In other words, we are not only interested in ordering the options on a scale but also in quantifying the distance among options on that scale, and thus in quantifying participants' perceived acceptability.[5] This is something very hard to obtain by directly asking participants "By how much do you consider this option better than the other?" or rate a sentence on a scale, as it can be very difficult for them to furnish an explicit quantification of the acceptability concept. Accordingly, we needed a more implicit way to force participants to rate alternatives for understanding how much these alternatives differed in terms of acceptability in each language. In other words, our goals require us to use a method that does not require participants to explicitly provide an acceptability judgement, yet that enables us to derive it.

In the linguistic literature on acceptability ratings, different designs have been used: rating of a single sentence using (1-5 or 1-7) Likert scales, rating all sentences as a multiple of a given standard, as in Magnitude estimation (Stevens 1956), and a forced choice between two sentences (see Goodall 2021 for a discussion of all these methods). Among these, research

[5] As pointed out in Dillon and Wagers (2021), among others, linguistic acceptability is a cognitive attribute that can be measured and modelled on a par with other psycho-physical judgments. Linguistic acceptability is a complex attribute affected by several factors, among which is grammaticality, as pointed out by Sprouse et al. (2013), i.e., the property that a sentence can or cannot be generated by a mental grammar.

on linguistic acceptability ratings has identified the two-alternative forced choice task as a particularly sensitive, reliable method (Sprouse et al. 2013, Marty et al. 2020), which is easier, as humans are used to judge comparatively (Nunnally 1967). However, as noted in Langsford et al. (2018), two-alternative forced choice tasks are limited as they provide ordinal information (which of two sentences, say A and B, is more acceptable), but do not indicate the degree of acceptance of one over the other, i.e., do not quantify the distance between two sentences. In other words, the proportion of acceptance of A over B does not offer any clue as to how to obtain the distance in acceptability between A and B.[6]

A solution to these limits is offered by Thurstones method of multiple paired comparisons. Thurstones method of multiple paired comparisons (Thurstone 1927a,b) is a methodological procedure for psychophysical judgments. It is considered the gold standard for collecting, analyzing, and interpreting participants judgements (see Cattelan 2012, Montag 2006, Parraga 2015 for reviews). To be concrete, if we have some related but slightly different syntactic sentences and we want to understand their different degree of acceptability, with this method we can present all the alternatives in all the possible pairs and ask participants to choose for each pair the preferred one (this is possible only if we consider the order of presentation of pair of sentences not informative). Then, for each sentence, we can obtain the frequency with which an option has been preferred in each pair and generate a proportion matrix. From the proportion matrix, z-scores for each option can be calculated. In this way, we can derive a scaling of subjects' preferences, quantifying the distance between options, in terms of z-scores. The quantitative scale of z-scores thus obtained is not predetermined and has potentially infinite granularity, like magnitude estimation. Notice that, unlike with magnitude estimation, the participant is not responsible for the quantification of the acceptability in the Thurstone method (i.e., the participant does not provide an explicit rate of acceptability) since the acceptability is computed a posteriori based on the preferences in each pair. In this way, we avoid the problem of the mental representation of the scale, which is hard to explicitly state in a numeric quantification for a concept as acceptability (Sprouse 2011a), and that brings up also the issues related to the person-specific response styles and tendencies. Thus, Thurstone's method allows one to go beyond these limits. This scaling method seems to fit very well with the aims of our study, as it allows us to understand not just whether a sentence is considered more acceptable by participants, but also how much more or less acceptable the various options among them are. This would be very hard to realize by asking the participant to rate the degree of acceptability of each sentence on, for example, a Likert scale (especially for sentences that are not frankly bad). In fact, using a Likert scale would imply to ask participants to quantify explicitly their evaluation about the acceptability of an option. But it is hard to explicitly quantify how much one option is more (un)acceptable than the other, which is our aim in this study. Notice that very similar ratings for two options on a Likert scale could be regarded as a function of a precise evaluation of the participant about their preferences or a function of the participants inability to quantify the even little differences attributed to the two options. In other words, a participant could not

[6]This can be illustrated with the following example. Consider two participants, one that choses B because she considers B slightly more acceptable than A and another participant selects B because B is far more acceptable than A. One could argue that the proportion of acceptance of A over B obtained by a larger sample of participants could provide the degree of acceptability or distance between sentences. However, the problem we pointed out in the case of two participants is still present with a larger sample. Suppose that we are faced with 70% acceptance of a sentence. This proportion still does not allow us to know whether for participants the distance between the two sentences in terms of percentage is large or small. This is a critical point in linguistics where some sentences can be frankly bad, but others can be just a bit less acceptable than alternative sentences.

be aware of the fact that, although considering the options (un)acceptable, she tends anyway to prefer one of the two. Looking at the preference behavior rather than the explicit opinion could be a better strategy to reach our research goal. The Thurstone method allows us to highlight the difference between all options and does not request a direct answer of participants, but extracts from their choices or preferences participants implicit evaluation about the acceptability of each option in each context. In addition, the Likert scale also introduces a bias with regard to the participants conception of quantities since it forces the compression of the judgment in a delimited space, the pre-determined scale 1–5 or 1–7 (Sprouse 2011b, Stadthagen-González et al. 2018) and requires participants to be comfortable in using the proposed number line (Sprouse 2008). Therefore, for the reasons outlined, we chose Thurstone's method of multiple paired comparison to obtain a scale of acceptability and measure the distance between the various options.

As clearly stated in Stadthagen-González et al. (2018), Thurstone's method is particularly suitable to study acceptability judgments in linguistics, although it has not been employed very much and has largely been disregarded in the face of the most frequent use of Likert scales and magnitude estimation. However, some recent studies in linguistics seem to support a renewed interest in Thurstone's scaling method, even when compared to Likert scales. Indeed, as reported by Bellamy et al. (2018) as well as by Stadthagen-González et al. (2018, 2019), Thurstone's method would result in a more informative and preferable approach in certain studies. In the case of a code-switching study, for example, the cited researchers compared Thurstone scaling with Likert scales. They underlined how Thurstone's method of multiple paired comparisons offered more granularity than Likert scales (see Bellamy et al. 2018 for a direct comparison). Furthermore, by comparing several types of scaling and analysis designs (target pairs, random pairs, Thurstone's, Likert, and Magnitude Estimation) for levels of test-retest reliability on various types of sentences from *Linguistic Inquiry*, Langsford and colleagues (2018) found that Likert and Thurstone's are the best-performing measures for acceptability judgments. Accordingly, Thurstone's reported very good levels of within and between-participants reliability, consistent with those found for Likert scales (Langsford et al. 2018).

Considering all these pieces of evidence, Thurstone's method is a reliable and valuable method for acceptability judgments that should not be systematically preferred to Likert one, but that should be chosen when we need to account for differences in responses of participants that are difficult to be explicitly quantified by the participants themselves, as it is the case in our study.

3.2 Participants

602 adult participants aged between 18 and 58 years took part in the study. Of these, 152 were native speakers of English (age on average = 29.54, SD = 8.83, range =18–55), 155 native speakers of German (age on average = 24.53, SD = 5.1, range = 18–58), 144 native speakers of Italian[7] (age on average = 23.47, SD = 4.93, range = 18–48) and 152 native speakers of Greek (age on average = 23.92, SD = 4.38, range = 18–50). Participants were recruited online, through the Prolific platform (German, English, Greek) and the SONA System (Italian).

[7] 6 participants were removed: 2 participants with Italian as L2, 4 with previous SLI.

3.3 Materials and Method

The material consisted of 9 contexts and 4 sentences as possible completions for each of the contexts. 3 contexts were set up as controls. The control contexts were set up in order to target the relevant noun phrases in their non-generic and non-kind uses. Apart from generic/kind readings, definite noun phrases in singular (21) and plural (22) can also denote individuals who are known to be unique either by context or by world knowledge. Furthermore, bare plurals and indefinite singulars additionally trigger existential readings which arise in episodic contexts (23). For the test conditions, we constructed a generic context and a kind context with well-defined kinds (24)–(25), and with non-well-defined kinds (26)–(27). In addition, we tested a generic normative context (28) and a generic distance context (29). The target sentences presented in each context only differed in that they started with one of the four different nominals (definite singular, indefinite singular, bare plural, definite plural). As it can be noticed, we used only one item per context and this may be a limit that can be overcome in future work. However, we should point out that using more "repetition" for context is not exempt from problems (see Sprouse et al. 2013).

(21) *Unique singular context (control):*
I recently bought a telescope and last night, for the first time, I wanted to watch the cosmic bodies around the earth up close. I had to hurry setting up the telescope because...

a. Moons were already rising at 11 pm.

b. The moons were already rising at 11 pm.

c. The moon was already rising at 11 pm.

d. A moon was already rising at 11 pm.

(22) *Unique plural context (control):*
Considering the amount of plastic in them, all oceans are by now dangerously polluted.

a. Oceans today are overcrowded with plastic bottles.

b. The oceans today are overcrowded with plastic bottles.

c. The ocean today is overcrowded with plastic bottles.

d. An ocean today is overcrowded with plastic bottles.

(23) *Episodic context (control):*
I dont have any pets, so I planted lots of beautiful flowers in my garden for me to enjoy. When I went outside yesterday, they were all ruined and there were paw prints of many different sizes all over the ground. I think the following happened...

a. Cats had ruined my flowers during the night.

b. The cats had ruined my flowers during the night.

c. The cat had ruined my flowers during the night.

d. A cat had ruined my flowers during the night.

(24) *Kind, well defined context (D-level genericity):*
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.

(25) *Generic, well defined context (I-level genericity):*
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.

(26) *Kind, non-well defined context (D-level genericity):*
Life is tough for animals in the wild. Any weakness can lead to sudden death by diseases, blood loss, or predators. That is why...

a. wounded elephants are rare in the wild.

b. the wounded elephants are rare in the wild.

c. the wounded elephant is rare in the wild.

d. a wounded elephant is rare in the wild.

(27) *Generic, non-well defined context (I-level genericity):*
My brother is seeking medical advice but he has no health insurance. Unfortunately, I cannot help him. But I know of a party happening later, where some of the guests might be doctors and he could find some help there. But I also tell him that doctors always get drunk at parties. So he should be careful since...

a. drunk doctors give bad advice.

b. the drunk doctors give bad advice.

c. the drunk doctor gives bad advice.

d. a drunk doctor gives bad advice.

(28) *Generic, normative context (I-level genericity):*
A child was sent home for misbehaving in school. The parents are upset, and scold their child. They say in the end: "Remember, ...

a. children respect their teachers."

b. the children respect their teachers."

c. the child respects his/her teachers."

d. a child respects his/her teachers."

(29) *Generic, speaker distance context (I-level genericity):*
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."

Note that while we are presenting all 4 options in (23)–(29) for illustrative purposes, participants saw only two options per trial. Specifically 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. Each context is presented within a block. All items are presented to all participants in every condition resulting in a blocked within-participant within-item design.[8]

3.4 Hypotheses

Based on the discussion in section 2, we will now introduce our main hypotheses and motivate the selection of languages we targeted in our study. Previous work on genericity revealed that English speakers use bare plurals to express generic readings (Carlson 1977). Italian speakers (as well as speakers of other Romance languages and Greek) use definite plurals (Chierchia 1998, Alexiadou et al. 2007), while German speakers have been claimed to make use of both (Krifka et al. 1995, among others). A recent corpus study, however, reveals that definite plurals can be used for generic readings even in English but only if the speaker wants to either express distance to the kind denoted by the argument or present members of that kind as a monolithic bloc without internal differences (Acton 2019). It is likely that a similar reasoning is at work for the German cases. Two more cases are frequently discussed under the notion of genericity across languages: definite singulars and indefinite singulars. Dayal (2004) argues that definite

[8]The position of the sentences in a pair of sentences (top vs. bottom), the order of pairs of sentences in a block, and the order of blocks per participant are drawn from uniform distributions created with the JavaScript 'Math.random()' (https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Math/random) function by the experimental software Labvanced.

singulars give the impression of being used for generic statements because the definite determiner can also range over taxonomic entities. This predicts that definite singulars can only be used for well-defined kinds. The prediction was confirmed by an acceptability judgement study on English, Spanish, and Brazilian Portuguese (Ionin et al. 2011); the results were confirmed for Greek in a replication of this study (Lazaridou-Chatzigoga and Alexiadou 2019). Finally, the use of indefinite singulars is often felicitous in generic contexts because they can express a normative rule or a conditional statement (Cohen 2001).

Our study aims to examine core cases of genericity (bare plurals and definite plurals) and pseudo-cases of genericity (definite singulars and indefinite singulars). The core cases can vary across languages between bare plurals and definite plurals, and they can be subject to pragmatic distance effects. The pseudo-cases result from structures denoting propositions about taxonomic entities or normative rules, which can nevertheless be felicitous in generic contexts, thus creating the impression that they denote generic statements. They are expected not to vary across languages. We intend to examine how languages pattern with respect to the expression of kind/genericity, establish a scale of acceptability comprising core cases and pseudo-cases of genericity and measure the distance between options. Table 1 (BP = bare plural, DP = definite plural, DS = definite singular, IS = indefinite singular, WD = well defined, NWD = non-well defined) presents an overview of our predictions, where the default is always listed first, and the potential alternative second.

	Italian	Greek	English	German
(24) kind, WD	$DP_{Chierchia}$, DS_{Dayal}	$DP_{Chierchia}$, DS_{Dayal}	$BP_{Chierchia}$, DS_{Dayal}	$BP_{Chierchia}$ (BP/DP_{Dayal}), DS_{Dayal}
(25) gen, WD	$DP_{Chierchia}$, $DS_{?Dayal}$	$DP_{Chierchia}$, $DS_{?Dayal}$	$BP_{Chierchia}$, $DS_{?Dayal}$	$BP_{Chierchia}$ (BP/DP_{Dayal}), $DS_{?Dayal}$
(26) kind, NWD	$DP_{Chierchia}$	$DP_{Chierchia}$	$BP_{Chierchia}$	$BP_{Chierchia}$ (BP/DP_{Dayal})
(27) gen, NWD	$DP_{Chierchia}$	$DP_{Chierchia}$	$BP_{Chierchia}$	$BP_{Chierchia}$ (BP/DP_{Dayal})
(28) gen, norm	$DP_{Chierchia}$, IS_{Cohen}	$DP_{Chierchia}$, IS_{Cohen}	$BP_{Chierchia}$, IS_{Cohen}	$BP_{Chierchia}$ (BP/DP_{Dayal}), IS_{Cohen}
(29) gen, dist	$DP_{Chierchia}$	$DP_{Chierchia}$	DP_{Acton}	DP_{Acton}

Table 1: Predictions

Hypothesis I: If Chierchia (1998) and Dayal (2004) are correct in assuming that the difference between Italian and English is due to the presence/realization of an intensional maximality operator, we expect English participants to chose bare plurals over definite plurals in (24)–(28) and we expect Italian/Greek participants to chose definite plurals over bare plurals in (24)–(28). Moreover, Chierchia (1998) predicts German to pattern with English, while Dayal (2004) assumes that the Blocking Principle is inactive in German, therefore predicting both bare plural and definite plural to be equally possible options in (24)–(28).

Hypothesis II: If Cohen (2001) is correct that indefinite singulars license the expression of a normative rule, we predict indefinite singulars to be an option only in (28). This prediction holds across all four languages.

Hypothesis III: If Dayal (2004) is correct that definite singulars trigger generic readings if the extensional MAX operator ranges over taxonomic entities, we predict definite singulars to be an option only for (24) and potentially also (25). This prediction holds across all four languages.

Hypothesis IV: If Acton (2019) is correct that definite plurals can be used in languages which normally do not realize the intensional MAX operator and that the use is triggered by the speaker's intention to signal distance to the expressed kind, we predict definite plurals to be the only option for (29). This prediction holds across all four languages but is only informative for English and German.

3.5 Analysis

We used Thurstone scaling to derive a linear rating of all forms from the multiple paired comparisons. First, for each language and each condition, we obtained a proportion matrix of the frequencies with which an option has been chosen in each pair. Then, from the proportion matrix, z-scores for each option were calculated. After obtaining z-scores we subtracted from each of them the lower score in order to align to 0 the worst option selected for each context. This was done to avoid a negative value when considering the between z-score differences, increasing interpretability of results. After extracting z-scores and thus obtaining a quantitative measure of acceptability for each option in each context and language, we performed cross-language comparisons. The pattern of selected preferences (the proportion matrix) was compared across languages by means of two-by-two χ^2 comparisons corrected for multiple comparisons by the Bonferroni method. This analysis was meant to establish whether the general pattern of answers in a language was different from the other. To better explore the between-languages significant differences, we then looked at the differences between z-scores. In other words, our analysis based on z-scores are to be seen as a sort of ancillary post-hoc analysis after χ^2 comparisons. In this line, we discuss effect sizes only of languages between which a significant difference emerges. In particular, cross-language comparisons were useful to understand whether participants reported similar patterns of answers, while language-internal z-score differences were used to further explore cross-linguistic differences. As based on z-scores, the distances between two sentences were interpreted as effect sizes (differences between z-scores are considered in terms of standard deviation from the center of the (normal) distribution, i.e., the average mean). In particular, as we could consider the standardized differences between the two options obtained through Thurstone's method similar to standardized differences between two means, we evaluated this measure assuming the thresholds set by Cohen for effect sizes interpretation (Cohen 1988). Accordingly, following Cohen's thresholds (Cohen 1988: 40), we classified a difference between 0.2 and < 0.5 as small, a difference between 0.5 and < 0.8 as medium, and a difference ≥ 0.8 as a large effect. We assumed no between-option differences when differences between z-scores resulted in < 0.2. In other words, in that case, we interpreted the two alternatives as equally plausible in the opinion of participants (see also Sprouse and Almeida 2017, Marty et al. 2020). We further interpreted the relative acceptability of each option based on the previously mentioned threshold. In particular, if the first rated option shows a distance ≥ 0.8 we consider it the only acceptable option for participants. These choices are well-grounded, as by and large these thresholds are also assumed in other experimental linguistic studies (e.g., Sprouse and Almeida 2017 beyond studies in cognitive and social sciences). They are useful to have a sense of the distance between options, but they remain choices. All analyses were performed in R studio (R Core Team 2022).

4 Results

Results will be reported for each context separately by first focusing on the cross-linguistic comparisons and then exploring the acceptability rating obtained in each language. First, we report the three control contexts and then the six experimental ones. In the supplementary material, Table S2 reports the degree of preference expressed in z-scores for each option in a given context in the four languages, while Table S1 reports the between-options distances calculated

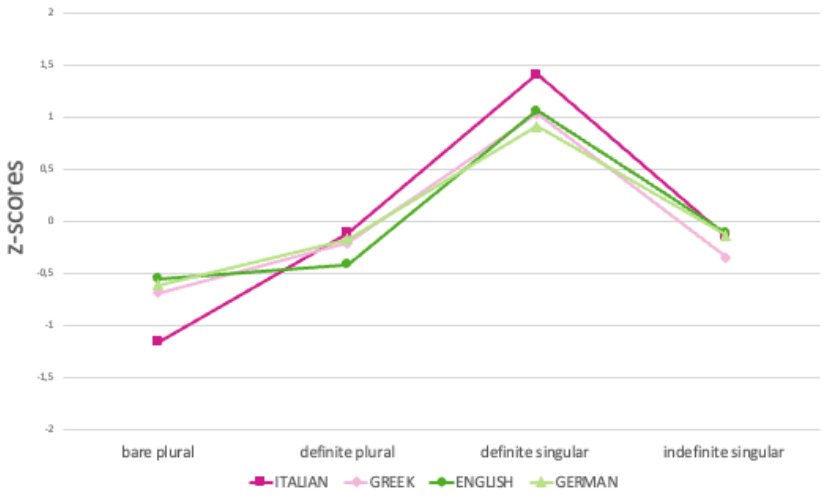
on the basis of Table S2.[9] Following the suggestion of a reviewer, we provide overview tables for each test condition which visualize the ranking between the four different candidates. Between each candidate we indicate the size distance, taken from Table S1 in the supplementary file (where $>$ *large* $>$ signals large, $>$ *medium* $>$ signals medium, $>$ *small* $>$ signals small, and $>>>$ signals equal distance). Furthermore, bar plots are provided in the supplementary material based on Table S2.

4.1 Controls

4.1.1 The unique singular context

We present the results of the definite singular control context in Figure 1. All participants in all languages showed a higher preference for the definite singular, evaluating as not acceptable all the other options (between-options distance $>$ |1| z-scores). Differences that emerge across languages concern the other options. In particular, Italian and Greek-speaking participants differed in their choices ($\chi^2_{(3)}$ = 37.18, p-value $<$ 0.001, p-value$_{\text{Bonferroni}}<$ 0.001), while English and German speaking participants reported very similar ratings ($\chi^2_{(3)}$ = 6.29, p-value = 0.09, p-value$_{\text{Bonferroni}}$= 0.54).

Figure 1: Unique singular context



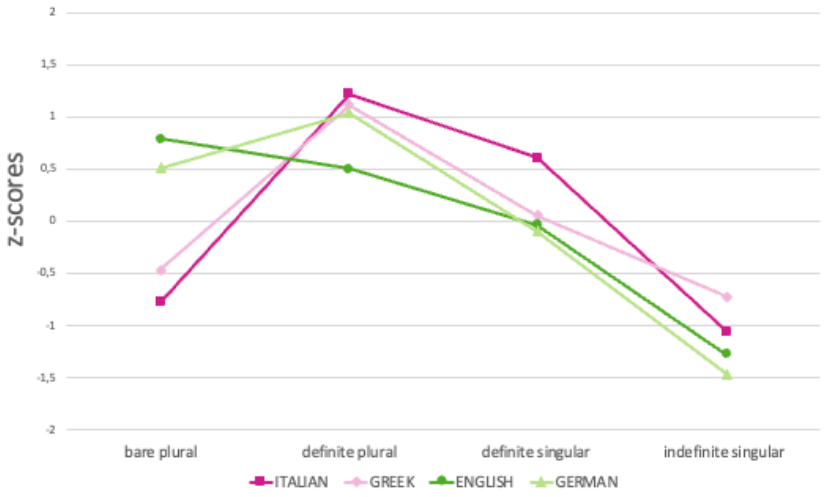
4.1.2 The unique plural context

The next control context was set up for the definite plural; results are shown in Figure 2. Significantly different patterns of preferences emerged in all the languages examined. English and German speakers differed ($\chi^2_{(3)}$ = 13.49, p-value $<$ 0.001, p-value$_{\text{Bonferroni}}<$ 0.001), as German speakers reported a higher preference for the definite plural, while English speakers preferred

[9] All data and statistical analysis scripts are available on the OSF platform: https://osf.io/e7xku/?view_only=f88bc349d60647f8b52135738ef35057.

the bare plural. In particular, German speakers considered the definite plural as the most acceptable option and the bare plural as a second option (between-options distance > |0.54| z-scores), while English speakers considered the bare plural as the most acceptable option and the definite plural as the second option (between-options distance > |0.30| z-scores). On the other hand, although both Italian and Greek speakers indicated the definite plural as the most acceptable option, they reported a generally different pattern of choices for the other options ($\chi^2_{(3)}$ = 18.39, p-value < 0.001, p-value$_{\text{Bonferroni}}$ < 0.001). Italian speakers considered the definite singular as a second option (between-options distance > |0.61| z-scores), while Greek speakers considered as acceptable only the first option, i.e., the definite plural. Since we found a large distance between the definite plural and the definite singular (between-options distance > |1.06| z-scores) and also all other options, we conclude that Greek speakers considered the definite plural as the only acceptable option.

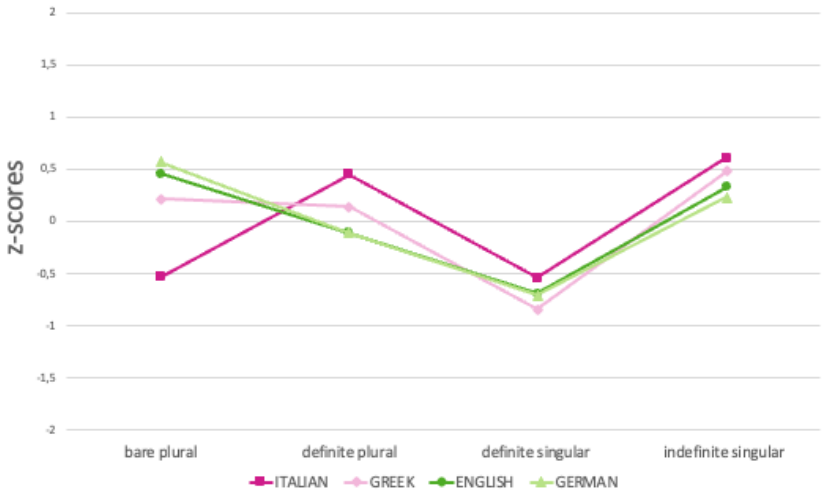
Figure 2: Unique plural context



4.1.3 The episodic context

The results in Figure 3 present our last control context, which favors the use of indefinites, that is bare plurals and indefinite singulars. English and German-speaking participants reported on average the same pattern of preferences ($\chi^2_{(3)}$ = 1.63, p-value = 0.65, p-value$_{\text{Bonferroni}}$= 1), while Italian and Greek participants reported significantly different ratings from each other ($\chi^2_{(3)}$ = 95.95, p-value < 0.001, p-value$_{\text{Bonferroni}}$ < 0.001). English and German speakers considered bare plurals and indefinite singulars very similar. Italian speakers considered the definite plural and the indefinite singular as equally best options (between-options distance = |0.16| z-scores), while Greek participants considered the indefinite singular as the best option and both definite and bare plural as the second best options (between-options distance = |0.08| z-scores).

Figure 3: Episodic context

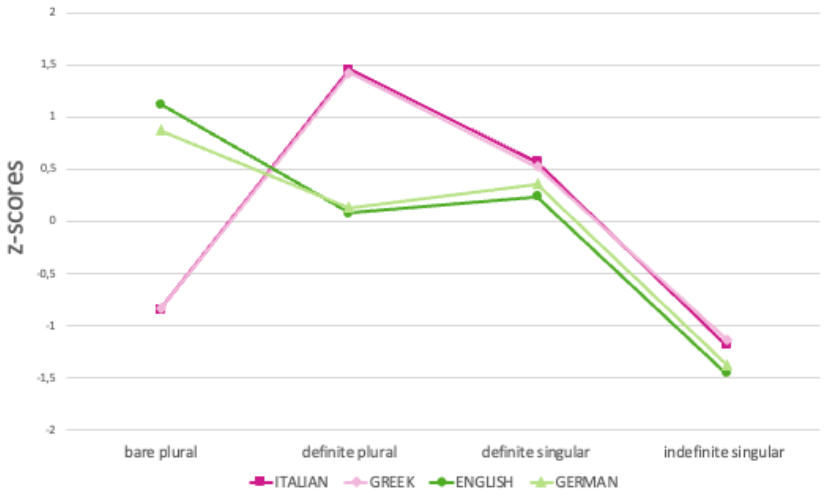


In the next sections, we turn to the experimental contexts. By and large, in the experimental contexts, languages are split into two groups: English and German cluster together and so do Greek and Italian. Some differences in the different contexts are also found among languages that we will discuss in detail below.

4.2 The kind context with well-defined kinds (D-level genericity)

Figure 4 clearly shows that German and English speakers pattern together ($\chi^2_{(3)}$ = 5.22, p-value = 0.15, p-value$_{\text{Bonferroni}}$ = 1), as do Italian and Greek speakers ($\chi^2_{(3)}$ = 3.06, p-value = 0.38, p-value$_{\text{Bonferroni}}$ = 1). English and German speakers showed the same pattern of choices, considering bare plural as the most acceptable option. Similarly, Italian and Greek speakers displayed the same pattern. Since the z-scores of the definite plural are far from those reported for the other options (between-options distance $\geq$ |0.8| z-scores), we conclude that this is the only plausible option. In all four languages, the definite singular emerged as the second best option, also for languages where it was not considered an acceptable option.

Figure 4: Kind context, well defined kinds



	Italian	Greek	English	German
1.	definite plural	definite plural	bare plural	bare plural
	v *large (.80)* v	v *large (.91)* v	v *large (.89)* v	v *medium (.51)* v
2.	definite singular	definite singular	definite singular	definite singular
	v *large (1.52)* v	v *large (1.37)* v	v v v	v *small (.23)* v
3.	bare plural	bare plural	definite plural	definite plural
	v *small (.34)* v	v *small (.30)* v	v *large (1.54)* v	v *large (1.51)* v
4.	indefinite singular	indefinite singular	indefinite singular	indefinite singular

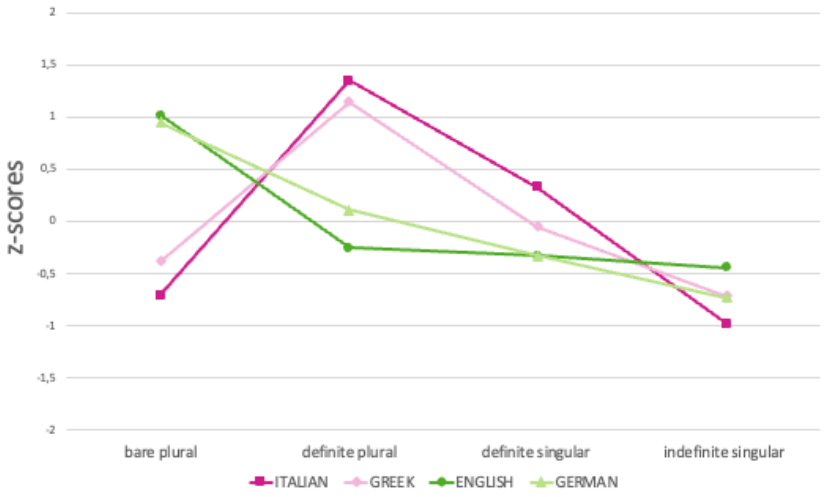
Table 2: Kind context, well defined kinds

4.3 The generic context with well defined kinds (I-level genericity)

As can be seen in Figure 5, the bare plural is the best option in German and English, while the definite plural is the best choice in Italian and Greek. Nevertheless, the patterns of preferences of English and German speakers differed ($\chi^2_{(3)} = 21.98$, p-value < 0.001, p-value$_{\text{Bonferroni}} < .001$), as did the ones of Italian and Greek speakers ($\chi^2_{(3)} = 18.2$, p-value < 0.001, p-value$_{\text{Bonferroni}} < 0.001$). Indeed, English speakers considered as the most acceptable option the bare plural. Based on z-scores, the bare plural was rated far better than all the other options (between-options distance $> |1|$), which in turn are very close to each other and, thus, can be considered

equally unacceptable (see Table 3). German speakers also strongly prefer the bare plural. Since the z-scores for the bare plural show a distance of > |0.8| to all other options, we conclude that German speakers consider it the only acceptable option. Furthermore, among the unacceptable options, they considered the definite plural as more acceptable than the definite singular (between-option distance = |0.44|). Greek and Italian speakers considered the definite plural as the most acceptable option (between-options distance > |1| z-scores). Although Greek and Italian speakers showed a very similar pattern of preference, Italian speakers had a stronger preference for definite plural compared to all other options. They also rated the definite singular differently when compared to the Greek speakers. The definite singular was a second option, very far from the others (about |1| z-score, see Table 3), while Greek speakers produced a smaller difference between the definite singular and the third option, bare plural (|0.33| z-scores).

Figure 5: Generic context, well defined kinds



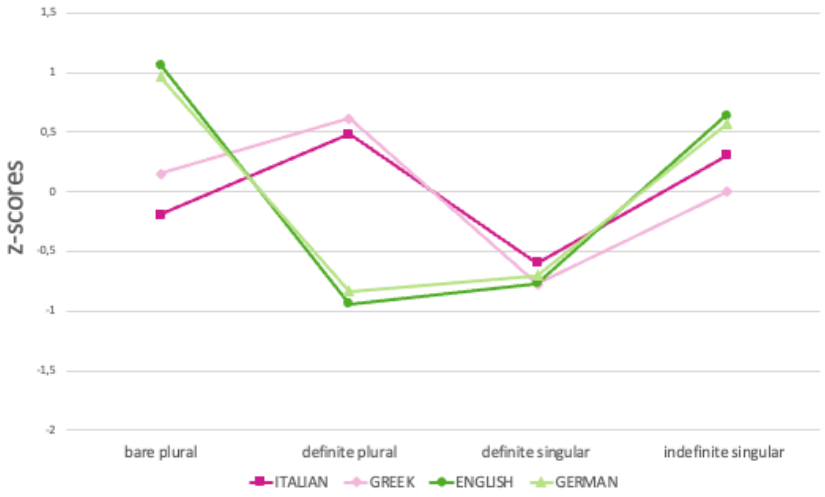
	Italian	Greek	English	German
1.	definite plural	definite plural	bare plural	bare plural
	v *large (1.00)* v	v *large (1.19)* v	v *large (1.28)* v	v *large (.85)* v
2.	definite singular	definite singular	definite plural	definite plural
	v *large (1.12)* v	v *small (.33)* v	v v v	v *small (.44)* v
3.	bare plural	bare plural	definite singular	definite singular
	v *small (.20)* v	v *small (.33)* v	v v v	v *small (.40)* v
4.	indefinite singular	indefinite singular	indefinite singular	indefinite singular

Table 3: Generic context, well defined kinds

4.4 The kind context with non-well-defined kinds (D-level genericity)

As depicted in Figure 6, English and German speakers reported the same rating patterns ($\chi^2_{(3)}$ = 1.19, p = 0.75), while Italian and Greek speakers showed significant differences ($\chi^2_{(3)}$ = 28.24, p-value < 0.001, p-value$_{\text{Bonferroni}}$ < 0.001). More concretely, English and German participants considered the bare plural as the best option, the indefinite singular as a second relevant option (between-options distance < |0.5| z-scores), and the other two as unacceptable. Based on z-scores, Italian speakers considered the definite plural and the indefinite singular as equally good options (between-options distance = |0.17|). In turn, Greek speakers showed a higher preference for definite plural, but rated bare plural and indefinite singular as equally acceptable second options (between-options distance = |0.15|).

Figure 6: Kind context, non-well defined kinds



	Italian	Greek	English	German
1.	definite plural	definite plural	bare plural	bare plural
	v v v	v *medium (.61)* v	v *small (.42)* v	v *small (.40)* v
2.	indefinite singular	bare plural	indefinite singular	indefinite singular
	v *medium (.50)* v	v v v	v *large (1.43)* v	v *large (1.28)* v
3.	bare plural	indefinite singular	definite singular	definite singular
	v *small (.41)* v	v *large (.93)* v	v v v	v v v
4.	definite singular	definite singular	definite plural	definite plural

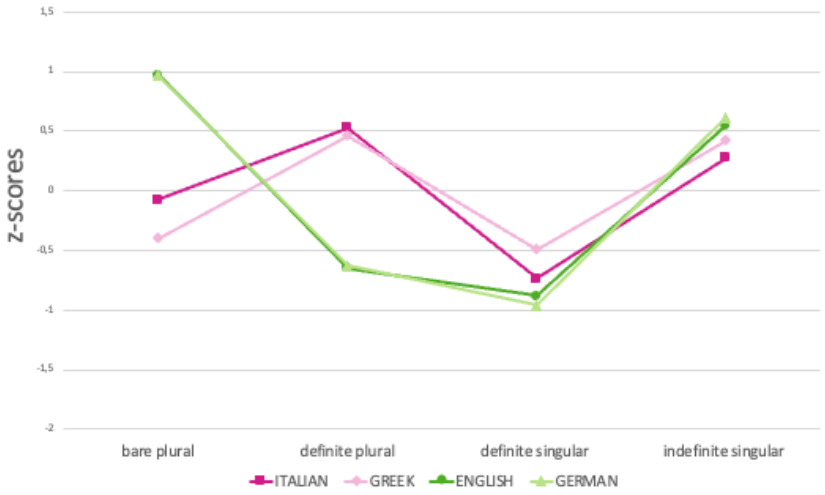
Table 4: Kind context, non-well defined kinds

4.5 The generic context with non-well-defined kinds (I-level genericity)

As with the previous context, ratings from German and English speakers turned out as very similar ($\chi^2_{(3)}$ = 0.49, p-value = 0.91, p-value$_{\text{Bonferroni}}$ = 1), while Italian and Greek speakers showed a significantly different pattern of preferences ($\chi^2_{(3)}$ = 25.02, p-value < 0.001, p-value$_{\text{Bonferroni}}$ < 0.001), see Figure 7. English and German speakers agreed in considering the bare plural as the best option, and the indefinite singular as the second option. Greek speakers considered the definite plural and the indefinite singular as equally best options (between-options distance = |0.04| z-scores), whereas Italian speakers showed a preference for the definite plural, but rated

the indefinite singular as a very close second option (between options distance = |0.25|). Greek speakers considered the bare plural and the definite singular as equally unacceptable third options (between options distance = |0.09| z-scores). In turn, as we found a medium distance between z-scores of bare plural and definite singular in Italian (between options distance = |0.66| z-scores), we conclude that the bare plural in this context is considered as more acceptable than the definite singular.

Figure 7: Generic context, non-well defined kinds



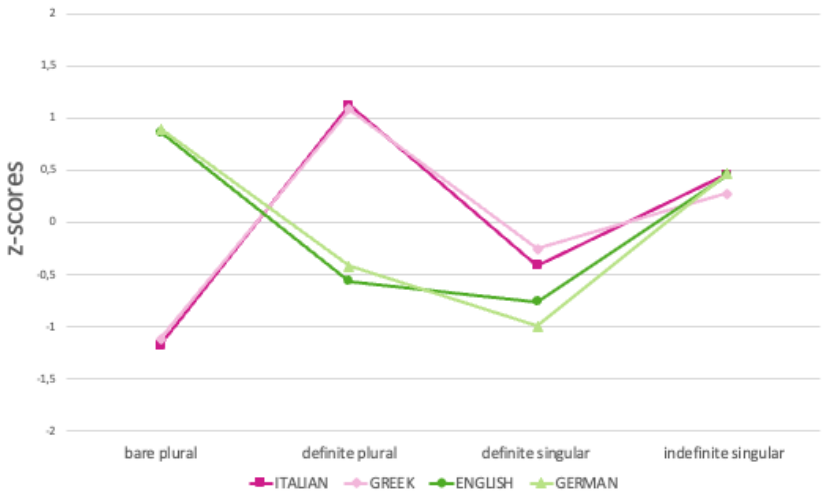
	Italian	Greek	English	German
1.	definite plural	definite plural	bare plural	bare plural
	v *small (.25)* v	v v v	v *small (.42)* v	v *small (.35)* v
2.	indefinite singular	indefinite singular	indefinite singular	indefinite singular
	v *small (.35)* v	v *large (.82)* v	v *large (1.44)* v	v *large (1.58)* v
3.	bare plural	bare plural	definite singular	definite singular
	v *medium (.66)* v	v v v	v *small (.23)* v	v *small (.33)* v
4.	definite singular	definite singular	definite plural	definite plural

Table 5: Generic context, non-well defined kinds

4.6 The generic context with normative flavor (I-level genericity)

As shown in Figure 8, answer patterns of the normative context for English and German speakers did not show significant differences ($\chi^2_{(3)}$ = 10.8, p-value = 0.012, p-value$_{\text{Bonferroni}}$ = 0.07). Both groups rated the bare plural as the most acceptable option. No differences were found between Italian and Greek ($\chi^2_{(3)}$ = 3.77, p-value = 0.28, p-value$_{\text{Bonferroni}}$ = 1). Both groups rated the definite plural as the most acceptable option. As can be seen in Table 6 and Figure 8, all four languages rated the indefinite singular as the second best option.

Figure 8: Generic context, normative



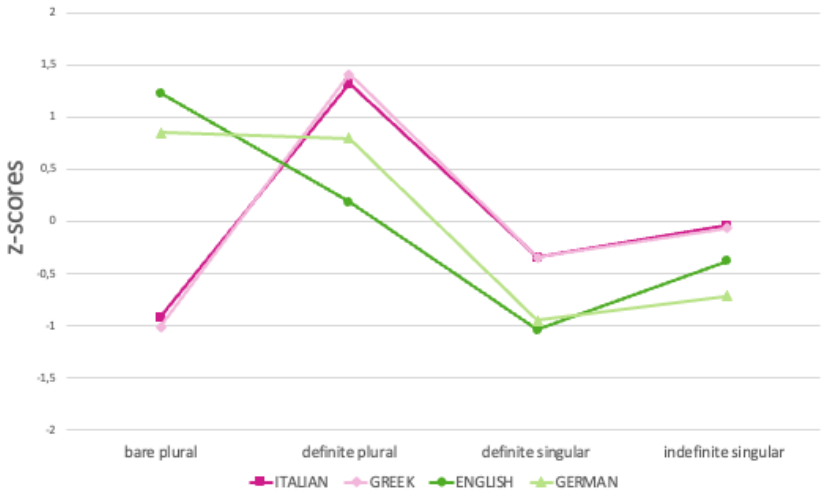
	Italian	Greek	English	German
1.	definite plural	definite plural	bare plural	bare plural
	v *medium (.66)* v	v *large (.80)* v	v *small (.42)* v	v *small (.43)* v
2.	indefinite singular	indefinite singular	indefinite singular	indefinite singular
	v *large (.87)* v	v *medium (.54)* v	v *large (1.02)* v	v *large (.88)* v
3.	definite singular	definite singular	definite plural	definite plural
	v *large (.77)* v	v *large (.87)* v	v v v	v *medium (.53)* v
4.	bare plural	bare plural	definite singular	definite singular

Table 6: Generic context, normative

4.7 The generic context with speaker distance (I-level genericity)

As can be observed in Figure 9, in the generic distance context, German and English speakers showed a significantly different pattern of preferences ($\chi^2(3) = 24.84$, p-value < 0.001, p-value$_{\text{Bonferroni}} < 0.001$), while Italian and Greek speakers provided very similar ratings ($\chi^2(3) = 0.38$, p-value $= 0.94$, p-value$_{\text{Bonferroni}} = 1$). In particular, German speakers considered the bare plural and the definite plural as equally best options (between-options distance = |0.05| z-scores), while English participants considered the bare plural as the only acceptable option (between-options distance > |1| z-scores). The definite plural emerged as the most acceptable option for Italian and Greek speakers (between-options distance > |1| z-scores).

Figure 9: Generic context, speaker distance



	Italian	Greek	English	German
1.	definite plural	definite plural	bare plural	bare plural
	v *large (1.37)* v	v *large (1.48)* v	v *large (1.05)* v	v v v
2.	indefinite singular	indefinite singular	definite plural	definite plural
	v *small (.30)* v	v *small (.28)* v	v *medium (.57)* v	v *large (1.52)* v
3.	definite singular	definite singular	indefinite singular	indefinite singular
	v *medium (.58)* v	v *medium (.66)* v	v *medium (.65)* v	v *small (.22)* v
4.	bare plural	bare plural	definite singular	definite singular

Table 7: Generic context, speaker distance

5 Discussion

Hypotheses I and II were by and large confirmed by our results, i.e., two language clusters emerged with the bare plural chosen as the default candidate in English/German and the definite plural in Italian/Greek, while we noticed an additional rise of the indefinite singular in the normative context. We could also find partial evidence for hypothesis IV, in that the definite plural was accepted as much as the bare plural in the distance context in German. Hypothesis III was not confirmed, though we notice tendencies towards a WDK restriction. After we briefly discuss the results of the controls, we will take a closer look at each hypothesis in turn, relating them to the results presented in section 3.

Controls. The results of the control context for the definite singular (Fig. 1) came out as expected since all languages chose it as the best candidate with much distance to all other options. As for the definite plural control, only the results for German and English are presumably informative since the definite plural is reported as a default term also for generic readings in Italian and Greek. For German, we find that although the definite plural is chosen as the best candidate, the distance to the bare plural is only medium (=.5 z-scores). For English, the pattern is even flipped: the bare plural is chosen as the best candidate, though the distance to the definite plural is small (=.3 z-scores). We can conclude that the definite plural control was not set up successfully. Another look at the context in (22) also reveals why German and English speakers considered both the bare plural and the definite plural as viable options. Given that the targeted meaning of the definite controls was uniqueness based, the context was set up so that the predicate in (22) applies to all referents of the noun phrase. Hence, the predicate can easily be read as generic or as applying to the maximal individual instantiated by the noun phrase. This problem did not arise with the definite singular control since a property holding of only one individual does not qualify as a natural kind.

The control for the existential reading originally targeted bare plurals, though the addition of *paw prints of different sizes* in (23) was potentially not salient enough to trigger the choice of bare plurals over indefinite singulars. As was shown in Fig. 3, both options were valid for English and German speakers, where German speakers slightly favored the bare plural. For Italian speakers, bare plurals are not an option, confirming the independent observations in the literature that unmodified bare nouns in subject position are not licensed for independent (syntactic) reasons (Longobardi 1994, Zamparelli 2002). Interestingly, bare plural subjects in Greek seem to be an option, contra e.g. the claims in Roussou and Tsimpli (1994). We speculate that this may be related to the fact that in Greek bare nouns are widely available, unlike in Italian, see the discussion in Alexopoulou and Folli (2019).

Hypothesis I: Bare plural vs. definite plural—the defaults. Given the discussion in the previous literature (Carlson 1977, Chierchia 1998, Dayal 2004, Alexiadou et al. 2007), we predict two language clusters to emerge: Italian and Greek speakers use the definite plural as the default to express kind and generic readings, whereas English and German speakers make use of the bare plural as the default. This prediction was confirmed, as our test conditions (Fig. 4 for D-level and Fig. 5 for I-level genericity) revealed the bare plural to be the best candidate in German/English, and the definite plural to be the best candidate in Italian/Greek. In (30) and (31), we demonstrate the mapping from semantic structure to surface forms for each language cluster, representatively with English and Italian, according to Dayal (2004). With the assumption that $\cap$ is mapped to $\varnothing$ in German/English, but to the definite determiner in Italian/Greek, we can derive the two clusters across D-level and I-level genericity. We also see

this prediction play out in the normative context (Fig. 8), where German and English speakers choose the bare plural as the most optimal, whereas Italian and Greek speakers pick the definite plural as the optimal candidate. For the kind and the normative condition, the languages showed no significant difference within their clusters. This was not the case in the generic condition, though the formation of the two language clusters is still visible in Figure 5.

(30) Kind, well-defined (24): *extinct*($^{\cap}$(*pandas*))

a. *Pandas are almost extinct.*

b. *I panda sono quasi estinti.*

(31) Generic, well-defined (25): GN x, s [$^{\cup\cap}$*mosquito*(x) $\wedge$ $C(x, s)$] [*give-bites*(x, s)]

a. *Mosquitos give us itchy bites in the summer.*

b. *Le zanzare ci pungono in estate.*

Interestingly, non-well defined kinds can also be accommodated with (default) plural kind formation across languages. In both the kind and the generic context with non-well-defined kinds (Fig. 6, Fig. 7), bare plurals are the best candidate and were chosen over definite plurals by German and English speakers, whereas definite plurals are the best option and were chosen over bare plurals by Greek and Italian speakers. While the preferred use of the bare plural in German/English has to be interpreted with caution as bare plurals can also denote existentials and hence allow for the possibility of episodic readings, the preferred use of the definite plural in Greek/Italian serves as robust evidence that non-well defined kinds can be accommodated with default plural kind formation across languages.

Our results speak in favor of the universality of the Blocking Principle since by and large German patterns with English. This holds in the kind condition with well defined kinds (Fig. 4) and non-well defined kinds (Fig. 6), in the generic condition with non-well defined kinds (Fig. 7), and in the normative condition (Fig. 8). The only difference between the two languages can be found in the generic condition with well defined kinds (Fig. 5), in which, among the unacceptable options, the definite plural is ranked better than the definite singular in German but not in English. If the Blocking Principle were inactive in German (Dayal 2004), we would expect the definite plural to be a much better competitor to the bare plural. Another indication that the definite plural lacks a competitor status in English, and more importantly also in German, is that with non-well defined kinds (Fig. 6, Fig. 7), the definite plural is ranked in last position in both English and German (Tables 4,5). This would be unexpected if the definite plural were able to qualifiy as a default expression for kind formation. That being said, it is noticable that the distance between bare plural and definite plural is always larger in English than in German in the conditions involving well-defined kinds, which could indicate that the definite plural does not have the same status in English and German. One aspect the study did not control for are age and dialectal differences, which are often argued to play a role for the availability of the definite plural in German kind formation (Barton et al. 2015).

Finally, our study results question the more recently developed view that kind reference is by default expressed with the definite singular noun phrase. This account is developed by Borik and Espinal (2015, 2020) for Spanish and Russian, and subsequently defended by Kwapiszewski

and Fuellenbach (2021) for Polish. We briefly discuss this account based on Spanish, as it is another Romance language which likely receives the same analysis as Italian. The basic assumption is that the definite singular makes direct reference to kinds, in the sense that common nouns denote predicates of kinds and ι applied to a common noun returns a kind individual in Spanish. Such definite kinds are numberless; their singular appearance is the result of default (singular) morphology. Besides the definite singular, Spanish also makes use of definite plurals to express kinds (like in Italian). The definite plural, however, is an indirect way to express kinds, they argue. The number morphology on definite plurals indicates the presence of a realization operator R (Carlson 1977) which shifts NP denotations from the domain of kinds to the domain of objects, making them compatible with object-level predicates. Since definite plurals, however, can also combine with kind-level predicates, such scenarios require a coercion operator, in addition to R, forming subkinds which are grouped together by some contextually triggered classification. Given this rationale, definite singulars are considered the "default way to express D-genericity" (Borik and Espinal 2015: 170), in contrast to definite plurals. Given our experimental results, this view cannot be extended to Italian, or Greek for that mattter. Clearly, the default noun type for D-level genericity is the definite plural with a large distance to the definite singular in both Italian and Greek, as is shown in Table 2.

Hypothesis II: The normative rule effect. In order to examine the distribution of the indefinite singular, we set up a context that creates a bias towards a rule-based or normative reading, recall (28). Following Cohen (2001), we predict that the indefinite singular becomes a competitor to the default noun phrase used for generic readings. Across all four languages, we find that the indefinite singular is chosen as the second best candidate in a context that filters for a normative reading (Fig. 8, Table 6). The result provides supportive evidence for the rule-based flavor of generic readings created with the indefinite singular, especially if we compare it to the status of the indefinite singular in the basic kind (Fig. 4) and generic (Fig. 5) condition, where we see that it was judged as the worst candidate. As discussed in section 2.3, Cohen (2001) suggests that the rule-based, normative reading should be modeled as a modal or a conditional statement, as shown in (32a). This does not exclude that such a reading can also be modeled with GN quantification, see (32b), though in this case we would expect a realization in the form of the default making reference to kinds, that is the bare plural in German/English and the definite plural in Italian/Greek. This is exactly what we find in Figure 8, where speakers chose the default generic term of their language as the best option, but also the indefinite singular as the second best option. Hence, the reading in (32a) is realized by the indefinite singular, specifically the modified variable $child(x)$ maps to an indefinite expression. Motivation for this assumption comes from the observation that indefinites are often modelled as variables also outside of generic contexts. In contrast, the reading in (32b) is realized by mapping $^{\cap}child(x)$ to the default generic noun type, i.e. the bare plural in English/German and the definite plural in Italian/Greek.[10]

(32) a. *A child respects her teachers*: in-effect(!($\underline{child(x)} \Rightarrow$ *respect-teacher*(x)))

b. *Children respects their teachers*: GN x, s [$^{\cup\cap}\underline{child(x)} \wedge C(x, s)$] [*respect-teacher*$(x, s)$]

Crucially, the rule-based meaning in (32a) is not prominent enough in the basic kind and generic contexts, thus the use of the indefinite singular is dispreferred. Our results also support the

[10]Note that the bare plural in Germanic is ambiguous since it can also be used as an indefinite. Consequently, it can express both (32a) and (32b).

view that the inductivist (quantification) account and the rule-based account identify different readings of generic statements, instead of two different approaches to genericity (cf. Carlson 1995).

Whereas the rise of the indefinite singular was entirely expected in the normative context, we did not expect the indefinite singular to become the second best option to express non-well-defined kinds, with small distances to the default option in English and German (Table 4 and Table 5), and an even equally good status as the default in Italian (Table 4).[11] We attribute the boost of the indefinite singular in such contexts to an avoidance strategy of the definite singular, which we will explain in more detail in the next paragraph.

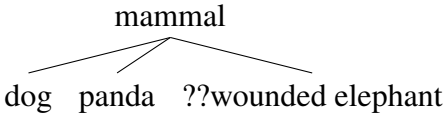
Hypothesis III: The well-defined kind restriction. The WDK restriction makes the following prediction (Dayal 2004): since the definite singular (if combined with a kind-level predicate) is encoded with a taxonomic determiner, which takes only taxonomic entities as arguments and only well-defined kinds qualify as taxonomic NPs, the definite singular should be a better option in the well-defined kind context (24) than in the non-well defined kind context (26). We illustrate this prediction in (33).

(33) a. *The panda is almost extinct.*

b.?? *The wounded elephant is rare in the wild.*

c. $[\![the_{tax}]\!] = \lambda P.\iota X[P(X)]$, where X ranges over entities in the taxonomic domain

d.



In our study, the WDK restrictions come out as tendencies rather than WDK effects across languages. While it is true that the definite singular was rated as the second best option in the well-defined kind context, the distance to the first option is never small (Fig. 4, Table 2). However, since the definite singular was not the second best option in the non-well-defined kind context and instead an indefinite raises to the second position (Fig. 6, Table 4), we can consider this contrast as a tendency towards a WDK effect. In other words, the definite singular is not an option to express non-well defined kinds due to the infelicity with the taxonomic determiner (33c). The use of the definite singular is indicated as an option in the theoretical literature, but does not emerge in our experimental study, likely because its use is restricted to an elevated register, e.g., a scientific or written register. Interestingly, based on introspective judgements we have been able to establish that in languages that are commonly used only in the spoken modality, such as two dialects spoken in Northern Italy (a variety of Milanese and of Veneto dialect), this possibility does not exist and kind readings are expressed with the definite plural. If this observation is correct, the register dimension has to be taken into account in future experimental research.

What explains the rise of indefinites with non-well defined kinds? We believe that speakers make use of the rule-based strategy, which was already at play for the normative context and is indeed also compatible with non-well defined kinds (34a). The defaults for kind expressions

[11] Greek is somewhat exceptional here as the indefinite singular is equally as available as the default option in generic contexts, but the bare plural takes over the second position in the kind context. Generally though, the results for Greek indicate that indefinites become competitors to express non-well defined kinds.

across languages are of course also available, as they can accomodate non-well defined kinds (34b), i.e., the definite plural in Italian/Greek and the bare plural in German/English.

(34) a. in-effect(!($\textit{wounded}(x) \wedge \textit{elephant}(x) \Rightarrow \textit{rare}(x)$))

b. $\textit{rare}(^{\cap}(\textit{wounded-elephants}))$

Recall that especially for Romance languages and Greek, the existence of the WDK restriction is debated (Vergnaud and Zubizarreta 1992, Ionin et al. 2011, Lazaridou-Chatzigoga and Alexiadou 2019). This observation is somewhat reflected in our results, as the WDK tendencies manifest in different ways for each language cluster. Although the distance of the definite singular to the defaults for kind reference in the non-well defined kind condition (Fig. 6) is large for each language, we can observe that the distance to the bare plural in English and German is much bigger than the distance to the definite plural in Italian and Greek. We tentatively speculate that the two routes of interpretation we sketched in (34) provide a promising way to understand such distance effects. In (34), we propose that non-well-defined kinds have to be expressed either with plural kind formation via the defaults, spelling out $^{\cap}(\textit{wounded-elephants})$ in (34b), or they have to be expressed as indefinites, spelling out $\textit{wounded}(x) \wedge \textit{elephant}(x)$ in (34a). We can now explain why we find a bigger effect in Germanic: Given that bare plurals are ambiguous between plural kinds and indefinites, the bare plural accumulates both ways a non-well-defined kind can be expressed, either via (34a) or (34b). The definite plural in Italian and Greek, however, only provides one way to express non-well-defined kinds, which is via the formation of plural kinds, i.e., (34b). We believe this is the root of the perceived weakness of WDK restrictions in Romance and Greek compared to Germanic. Since the bare plural in Germanic incorporates both the function of a plural kind and the function of an indefinite, it serves as a much stronger competitor in non-well-defined kind scenarios when speakers judge the relative acceptance of the definite singular. In Romance and Greek, however, the competitors of the definite singular include different noun types, one is the definite plural and the other the indefinite singular.

What about the availability of the definite singular in I-level generic contexts (Fig. 5, Fig. 7)? As discussed in section 2.2, Dayal (2004: 431–433) raises doubts about whether GN can access instantiations of a taxonomic entity, which would be necessary to derive I-level genericity with the definite singular. Note, however, that Ionin et al. (2011) tested WDK effects with I-level genericity, which implies that this is still an open debate. While it is true that non-well defined kinds for I-level genericity are expressed with the indefinite singular as the second best option and the definite singular is considered the worst candidate in all four languages (Fig. 7), it is, however, also clear that the definite singular is not a good candidate to express well-defined kinds for I-level genericity to begin with (Fig. 5). Whereas the definite singular is clearly the second best option in the D-level generic context (Table 2) in all four languages, this is at least not true for the I-level generic context in German and English (Table 3). This observation reflects in some way Dayal's view that the function of the definite singular in I-level generic statements is not tied to taxonomies, though we leave the details of this argument to future research.[12]

[12] Since Dayal (2004: 433) suggests that I-level genericity could potentially be licensed under the definite singular via a prototypical object associated with the species, a reviewer remarks that our target sentence in (25) might not have supported prototypicality enough to license the use of the definite singular.

Finally, our results align with the general expectation that plural kind formation is not sensitive to the well-defined kind restriction, in line with Dayal (2004). Since the WDK only arises from the assumption that singular kinds with the definite determiner map to taxonomies, no such effects are expected with plural kinds, which are derived via $^{\cap}$ or $^{\wedge}\iota$, operators that take as arguments nouns that range over individuals. This result dovetails with the findings in Ionin et al. (2011) who also did not detect WDK effects with plural kinds for English and Spanish.

Hypothesis IV: The distance effect. Following Acton (2019), we predict that a context like (29) 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. This prediction was at least partially borne out for German. As can be seen in Figure 9, German speakers chose the definite plural as often as the bare plural in the distance context. The acceptability of the definite plural contrasts with the judgements for the basic kind and generic contexts (Fig. 6 and Fig. 7), where the bare plural was clearly the best option. However, they do not indicate that the definite plural is a better alternative to the bare plural in the distance context. This result 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.

Surprisingly, however, the German pattern was not found for English. English speakers still chose the bare plural over the definite plural. We leave it open to future research why we could not replicate Acton's corpus results for *(the) Republicans/Democrats* with *(the) politicians* in our study. One difference to our study is that the addressees of the utterances in Acton's data presumably included many instances of the kind the speakers want to distance themselves from. In our study, it is likely that speakers accommodated an addressee group with members they associated with. Further reasons could include the population (due to ethical requirements, we could only recruit English speakers living in Ireland; Acton's corpus study is based on American English), or maybe also the fact that for the use of English definite plurals, it matters not only whether the kind term refers to a group that qualifies as a monolith but also whether the speaker intends to emphasize that they are not part of the group, monolith or not (only the latter is ensured by the context).

We will now turn to the question how the use of the definite determiner triggers this distance reading. Acton (2019: 58) proposes that this inference can be derived as a conversational implicature, assuming that the bare plural and the definite plural, i.e., *politicians* in (35a) and *the politicians* in (35b), are equally informative, in the sense that both forms convey an "all-or-typical interpretation". The definite plural, however, is morphologically more marked than the bare plural. Hence, its use conveys a marked meaning—presumably triggered by a manner implicature (Grice 1975, Horn 1984, Rett 2014), although this is not explicitly stated by Acton. The actual content of the marked meaning is determined by another alternative, i.e., *we politicians* in (35c), which is equally complex but more informative than *the politicians* in (35b). Acton argues further that since *we politicians* includes the speaker, its non-use in (35b) implicates speaker distance.[13] This line of reasoning is reminiscient of a scalar implicature, though this is also not explicitly stated or worked out by Acton.

(35) a. Politicians do whatever they want.

b. The politicians do whatever they want. $\rightsquigarrow$ *marked meaning: speaker not a politician*

[13] See also Alexiadou et al. (2024) who make a similar argument on the level of person feature specification.

c. We politicians do whatever we want.

Note that Acton's reasoning for the manner implicature entails that the definite plural is able to express generic readings in the first place, an assumption which is at odds with the claim that $\cap$ is not spelled out in English (Chierchia 1998, Dayal 2004). Thus, a potential way to explain the Acton results in English and our study results in German is to extend Dayal's (2004) view on German to English, such that the kind operator $\cap$ is generally spelled out in Germanic but the Blocking Principle is inactive in the generic domain. A manner implicature, however, regulates the distribution of the definite plural, making it always the more marked option unless the context specifically supports the marked meaning, as, e.g., in (29), or in statements about a political party speakers exclude themselves from. This account, of course raises the question about what happens in languages like Italian and Greek, in which bare plurals are not an option to express genericity. In such languages, either the Blocking Principle is active, or I-level and D-level genericity is not derived with $\cap$ but instead with $\wedge\iota$, as originally suggested by Chierchia (1998).

6 Conclusion

Our comparative judgement study confirms, based on quantitative data, that languages make use of different default forms for the expression of kind reference. The optimal candidate to make reference to kinds is the bare plural in German and English, and the definite plural in Greek and Italian. We could not confirm the availability of the definite plural as a second default in German, in line with Czypionka and Kupisch (2019). Moreover, we found an effect for distance marking with definite plurals in German, that aligns with corpus results for English by Acton (2019). Our investigation of the well-defined kind restriction revealed WDK tendencies across Germanic, Romance and Greek, though the effects play out differently in German/English vs. Greek/Italian. We, therefore, replicate the results by Ionin et al. (2011) and Lazaridou-Chatzigoga and Alexiadou (2019), and provide tentative support for the view of the definite singular as a taxonomic determiner (Dayal 2004). Finally, we verified that generic statements with normative flavor boost the use of the indefinite singular, supporting an analysis in which the indefinite singular encodes a rule-based strategy to express genericity (Cohen 2001, Greenberg 2003), which we argue is also at play to express non-well defined kinds. Overall, our results indicate that there has to be more than one account for generic readings, as different contexts favor different noun types, and therefore not all generic readings can be treated alike but must result from different underlying presentations, which come with their own restrictions. At the same time, we also found that in all four languages one default kind expression emerged as the best candidate across all the test conditions. We take this to support the view that these default kind expressions – bare plurals in English and German, definite plurals in Italian and Greek – directly spell out the kind operator and hence can accommodate a generic reading in context.

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.

Alexiadou, A., Hein, J., Ilić, I., and Sauerland, U. (2024). The third person is present: An argument from determiners in generic statements. Ms., to appear in Proceedings of NELS 54.

Alexopoulou, T. and Folli, R. (2019). Topic strategies and the internal structure of nominal arguments in greek and italian. *Linguistic Inquiry*, 50:439–486.

Barton, D., Kolb, N., and Kupisch, T. (2015). Definite article use with generic reference in German: an empirical study. *Zeitschrift für Sprachwissenschaft*, 34:147173.

Behrens, L. (2005). Genericity from a cross-linguistic perspective. *Linguistics*, 43:275–344.

Bellamy, K., Parafita Couto, M. C., and Stadthagen-Gonzalez, H. (2018). Investigating gender assignment strategies in mixed purepechaspanish nominal constructions. *Languages*, 3:28.

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

Borik, O. and Espinal, M. T. (2020). Numberless kinds: Evidence from Russian. *Catalan Journal of Linguistics*, 19:231–260.

Brugger, G. (1994). Generic interpretations and expletive determiners. *University of Venice Working Papers in Linguistics*, 3:1–30.

Burton-Roberts, N. (1977). Generic sentences and analyticity. *Studies in Language*, 1:155196.

Carlson, G. (1977). *Reference to kinds in English*. PhD thesis, University of Massachusetts, Amherst.

Carlson, G. (2011). Genericity. In von Heusinger, K., Maienborn, C., and Portner, P., editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 2 of *Handbooks of Linguistics and Communication Science*, page 11531185. Mouton de Gruyter, Berlin.

Carlson, G. N. (1995). Truth Conditions of Generic Sentences: Two Contrasting Views. In Pelletier, F. J. and Carlson, G. N., editors, *The Generic Book*, pages 224–237. The University of Chicago Press, Chicago/London.

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.

Cohen, A. (2001). On the Generic Use of Indefinite Singulars. *Journal of Semantics*, 18:183–209.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, volume 2nd edition. Lawrence Erlbaum Associates, New York.

Czypionka, A. and Kupisch, T. (2019). (The) polar bears are pink. How (the) Germans interpret (the) definite articles in plural subject DPs. *The Journal of Comparative Germanic Linguistics*, 22:247291.

Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy*, 27:393–450.

de Swart, H. and Zwarts, J. (2012). Optimization Principles in the Typology of Number and

Articles. In Heine, B. and Narrog, H., editors, *The Oxford Handbook of Linguistic Analysis*, pages 555–582. Oxford University Press, Oxford.

Dillon, B. and Wagers, M. W. (2021). Approaching gradience in acceptability. In Goodall, G., editor, *The Cambridge Handbook of Experimental Syntax*, pages 62–96. Cambridge University Press, Cambridge.

Djajali, A. (2013). *House Proceedings Corpus*. Stanford University, Stanford, CA. https://github.com/alexdjalali/hpc.

Geoffrey, N. and Pan, C. (1975). Inferring quantification in generic sentences. In *Papers from the 11th regional meeting, Chicago Linguistic Society*, pages 412– 428.

Goodall, G., editor (2021). *The Cambridge Handbook of Experimental Syntax*. Cambridge University Press, Cambridge.

Greenberg, Y. (2003). *Manifestations of Genericity*. Routledge, London.

Grice, P. H. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and semantics, vol. 3: Speech acts*, page 4358. Academic Press, New York.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts.

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Schiffrin, D., editor, *Meaning, form, and use in context: Linguistic applications*, page 1142. Georgetown University Press, Washington, DC.

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.

Kamp, H. (1981). A theory of truth and semantic representation. In Groenendijk, J., Janssen, T., and Stokhof, M., editors, *Formal Methods in the Study of Language*, page 277322. Mathematical Centre, Amsterdam.

Krifka, M. (2013). Definitional Generics. In Mari, A., Beyssade, C., and Prete, F. D., editors, *Genericity*, page 372389. Oxford University Press, Oxford.

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.

Kwapiszewski, A. and Fuellenbach, K. (2021). Reference to kinds and subkinds in Polish. In Blümel, A., Gajić, J., Geist, L., Junghanns, U., and Pitsch, H., editors, *Advances in formal Slavic linguistics 2018*, pages 227–259. Language Science Press, Berlin.

Langsford, S., Perfors, A., Hendrickson, A. T., Kennedy, L. A., and Navarro, D. (2018). Quantifying sentence acceptability measures: Reliability, bias, and variability. *Glossa: a journal of general linguistics*, 3:37.

Lawler, J. M. (1973). *Studies in English Generics*. PhD thesis, The University of Michigan.

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.

Mari, A., Beyssade, C., and Prete, F. D. (2013). Introduction. In Mari, A., Beyssade, C., and Prete, F. D., editors, *Genericity*, pages 1–92. Oxford University Press, Oxford.

Marty, P., Chemla, E., and Sprouse, J. (2020). The effect of three basic task features on the sensitivity of acceptability judgment tasks. *Glossa: a journal of general linguistics*, 5:1–23.

Montag, E. D. (2006). Empirical formula for creating error bars for the method of paired comparison. *Journal of Electronic Imaging*, 15:010502.

Nunnally, J. C. (1967). *Psychometyric Theory*. McGraw Hill, New York.

Oosterhof, A. (2004). Generic noun phrases in Dutch. In Karlsson, F., editor, *Proceedings of the 20th Scandinavian Conference of Linguistics*, pages 1–22. University of Helsinki, Helsinki.

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.

Rett, J. (2014). *The Semantics of Evaluativity*. Oxford University Press, Oxford.

Roussou, A. and Tsimpli, I. M. (1994). On the Interaction of Case and Definiteness in Modern Greek. In Philippaki-Warburton, I., Nicolaidis, K., and Sifianou, M., editors, *Themes in Greek linguistics*, pages 69–76. John Benjamins, Amsterdam/Philadelphia.

Schaden, G. (2012). Two Ways of Referring to Generalities in German. In Mari, A., Beyssade, C., and Prete, F. D., editors, *Genericity*, page 157175. Oxford University Press, Oxford.

Sprouse, J. (2008). The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry*, 39:686–694.

Sprouse, J. (2011a). A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language*, 87:274–288.

Sprouse, J. (2011b). A validation of amazon mechanical turk for the collection of acceptability judgments in linguistic theory. *Behavior research methods*, 43:155–167.

Sprouse, J. and Almeida, D. (2017). Design sensitivity and statistical power in acceptability judgment experiments. *Glossa: a journal of general linguistics*, 2:1–32.

Sprouse, J., Schütze, C. T., and Almeida, D. (2013). A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 20012010. *Lingua*, 134:219–248.

Stadthagen-González, H., López, L., Parafita Couto, M. C., and Párraga, C. A. (2018). Using two-alternative forced choice tasks and thurstones law of comparative judgments for code-switching research. *Linguistic Approaches to Bilingualism*, 8:67–97.

Stadthagen-González, H., Parafita Couto, M. C., Párraga, C. A., and Damian, M. F. (2019). Testing alternative theoretical accounts of code-switching: Insights from comparative judgments of adjectivenoun order. *International Journal of Bilingualism*, 23:200–220.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes: loudness. *American Journal of Psychology*, 69:1–25.

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.

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press.

Vergnaud, J.-R. and Zubizarreta, M. L. (1992). The Definite Determiner and the Inalienable Constructions in French and in English. *Linguistic Inquiry*, 23:595–652.

Zamparelli, R. (2002). Definite and Bare Kind-denoting Noun Phrases. In Beyssade, C., Bok-Bennema, R., Drijkoningen, F., and Monachesi, P., editors, *Romance Languages and Linguistic Theory 2000: Selected papers from 'Going Romance' 2000*, pages 305–343. John Benjamins, Amsterdam/Philadelphia.