

Semantic Variation and Pleonastic Determiners: The Case of the Plural Definite Generic

Veneeta Dayal
Rutgers University
2006

to appear in Proceedings of V Asian GLOW

Abstract:

This paper examines the status of so-called ‘pleonastic’ determiners, focusing on the plural definite generic common in Romance languages. It is noted that every language has generic or kind-denoting terms but such terms vary between bare and definite (rather than indefinite) forms across languages. Canonical and non-canonical uses of definite determiners are distinguished and language variation posited to be a choice about the cut-off points for lexicalization on a universal Scale of Specificity. Studying variation from the perspective of languages without determiners is shown to have interesting consequences for the notion of ‘pleonastic’ determiners.

Keywords: Genericity, Kind Terms, Pleonastic Determiners, Definites, Semantic Variation, Cross-linguistic Variation

1: Introduction¹

Cross-linguistic work within the tradition of Montague Grammar is a relatively recent phenomenon, beginning in the late eighties, at a time when cross-linguistic considerations were already playing central roles in phonology and syntax. The guiding principle behind the first semantic analyses of languages other than English was that variation lay in the syntactic module. The value of studying what was then referred to as “exotic” languages was to gain insight into semantic operations needed to interpret unfamiliar constructions, but the operations themselves were thought to be universally available (see, for example, Bittner (1994) and Srivastav (1991)). The first substantive proposal that raised the possibility of variation in the semantic component was made in the early nineties. Bach et al 1995 suggested that languages may choose between A(dverb)-based and D(eterminer)-based systems of quantification. This was followed soon after

¹ I am indebted for valuable feedback to Jane Grimshaw, the audiences at Asian Glow V, The Workshop on Variation at Washington University at St. Louis, SURGE at Rutgers and Chicago Linguistic Society 42. Special thanks to Rosmin Matthew for stepping in with editorial help at the last minute. I alone remain responsible for errors and omissions.

by the proposal that there was a semantic parameter involved in the mapping of noun phrase denotations (Chierchia 1998). These initiatives proved extremely inspiring, with the result that cross-linguistic studies have now become critical in the development of semantic theory. Among the domains in which this has been most evident are analyses of Polarity/Free Choice items and genericity. This paper deals with the notion of pleonastic or expletive elements in light of our increased knowledge of the cross-linguistic expression of genericity and is based largely on Dayal (2004).

Consider the following examples, all of which can or must be understood to be generic statements about the species as a whole, rather than particular statements about a specific group of individuals. (1b) is Italian and means essentially what (1c) means in English:

- 1a. The dog barks (when it is hungry).
- b. I canni abbaiano
 the_{plural} dogs bark
- c. Dogs bark.

When we see paradigms where distinct forms map onto similar meanings, whether in the same language or in different languages, it is natural to wonder whether the locus of difference is in the semantics or the syntax. In the case of genericity, one view is to consider the determiners in (1a) and (1b) to be pleonastic, taking the determinerless form in (1c) to have the relevant structure for a kind term. Under this view, apparent differences in form are erased at the level of representation that feeds into the semantics: $\{[\text{the } [\text{dog}]], [\text{I } [\text{canni}]]\} \Rightarrow \{[\text{dog}], [\text{canni}]\} = [\text{dogs}]$. Whatever assumptions we make for the interpretation of English bare plurals as kind denoting terms transfer over seamlessly to the singular definite generic in English and the plural definite generic in Romance. The issue of semantic variation becomes moot.² An alternative approach would be to maintain the difference in forms and define their meanings in such a way that their ultimate semantic contributions converge: $(\text{the}_{\text{SING-}\delta} [\text{dog } \beta]) = \gamma$, $([\text{i } \alpha [\text{canni } \beta]]) = \gamma'$; $(f_{\mu}[\text{dogs } \beta]) = \gamma''$.³ A priori, both positions are reasonable but they have different consequences. The first approach predicts complete synonymy while the latter allows for overlap rather than identity in meaning.

² Perhaps some accommodation for the singular-plural contrast may be required in the case of (1a). For more discussion see Dayal (2004).

³ γ , γ' and γ'' are intended to denote meanings that may or may not be identical, but would at least be close enough to appear so.

In this paper I will argue for an approach of the second type, one that maintains morpho-syntactic differences in semantic interpretation. This position will be argued for on the basis of linguistic evidence relating to the Romance plural definite and the English bare plural. The English singular definite generic is not discussed in detail for reasons of space (see Dayal 2004). But first, I would like to make a sociological comment on the use of the notion ‘pleonastic’ in linguistic analyses, using data from another domain.

The situation we see with pleonastic generic determiners is reminiscent of a controversy that has been associated with the scope marking/partial wh movement construction:

- 2a. [Who_i [does Carl think [Maria talked to t_i]]?
b. [Was glaubt Karl [mit wem_i Maria gesprochen hat_i]]
 what think Karl with whom Maria spoken has

Here too there is a perceived synonymy between (2a), the more familiar construction, and (2b), an “exotic” structure found in German, Hindi, Romani among many other languages. The view that (2b) has the same structure at LF as (2a), that is taking German *was* and its counterparts in other scope marking languages to be pleonastic, has the immediate advantage of being able to apply an available semantic analysis to a new construction. The alternative approach is to treat (2a) and (2b) on their own terms, taking (2a) to have a structure in which an embedded wh expression takes wide scope, and (2b), to have a structure in which two contentful wh expressions combine in unexpected ways to yield a meaning similar to (2a). It is not my intention to reargue the case of scope marking here, but simply to point out that the original impetus to treat *was* as a pleonastic came from the desire to bring an unfamiliar construction in line with a better understood construction. But a closer comparison of the two constructions soon revealed subtle differences between them that gave the edge to an approach that maintained observable structural distinctions (see Dayal 2000 for specific arguments).

The idea that there are pleonastic generic determiners, I would like to suggest, may have a similar genesis. It is interesting to speculate how we would view the English bare plural had Carlson’s (1977) highly influential work on generics started with the Romance definite generic instead of the English bare plural. Would we have been tempted to posit a null definite determiner for English bare plurals since our frame of reference would have been languages in

which the definite determiner did double duty as a standard definite and a generic?⁴ In fact, one could go further, and speculate how we would view the regular English definite determiner had Frege and Russell started with South Asian languages. In those languages bare nominals are used to pick out contextually salient entities and to refer anaphorically, functions standardly associated with a definite determiner. Coming from such a perspective, would they not have been tempted to treat English *the* as a pleonastic? I will not pursue these counterfactuals any further but leave them on the table as a somewhat light-hearted note of caution on methodological assumptions that may be at work when we consider constructions where distinct forms converge on meaning. In the rest of this paper, I discuss ways in which definites and bare nominals yield generic readings, and make a concrete proposal about the cross-linguistic variations that have been noted in this domain. My claims are limited to the phenomena under discussion here. I leave open the question of whether natural languages can have other pleonastic expressions.

2: Generics/Kind Terms, Definites and Indefinites

All languages distinguish between generic and particular statements and in all languages genericity depends upon the interaction of properties of the verbal and the nominal systems. As already indicated, the focus of this paper is on the nominal system. Since the terms definite, indefinite and generic are sometimes used to talk about the form and sometimes to talk about the meanings associated with those forms, it might be useful to review the basic assumptions and terminology. In general, I use the terms definite and indefinite to refer to the form, and the terms contextually anchored and existential readings to refer to meaning. I use bare nominal to refer to the form, reserving the terms kind denotation/reading and generic reading to refer to meaning. A kind term, in principle, can have any form, bare nominal, definite or indefinite, as long as it denotes a kind entity.

2.1. English Kind Terms

Carlson (1977) proposed some diagnostics to separate out kind terms from other nominals that also may yield generic readings (see also Krifka et al 1995). The examples in (3) and (4) use

⁴ Or if the definite singular was taken as the frame of reference, one could in principle argue for bare plurals to have a structure in which there is an empty definite determiner and the plural morphology is semantically vacuous. Thanks to Jane Grimshaw for raising this possibility.

predicates that apply meaningfully only to the species as a whole, not to individuals of the ordinary sort – objects in his terms.⁵

- 3a. Dinosaurs are extinct.
 - b. *The dinosaurs are extinct.
 - c. *Some dinosaurs are extinct.
 - d. *Fido and Roxy are extinct.
-
- 4a. The dinosaur is extinct.
 - b. *A dinosaur is extinct.
 - c. * Fido is extinct.
 - d. *Dinosaur is extinct.

The results, as we can see, vary between the plural and the singular forms. Using this diagnostic, we can say that English has two kind denoting terms, the bare plural and the singular definite. I will not go into the differences between these two in this paper, nor will I discuss at any length the status of mass terms like *water* or *rice* which behave like singular terms for purposes of syntactic agreement, but like plural terms for purposes of definiteness marking. I refer the reader to Dayal (2004) for further discussion of these issues.

When object level predicates occur in imperfective aspect, bare plurals, singular definites as well as singular indefinites yield generic readings. The singular definite, because it can function as a kind term or a regular referring term, is ambiguous between a generic and a habitual reading.⁶ English lacks a bare singular:

- 5a. Dogs bark (when they are hungry).
- b. A dog barks (when it is hungry).
- c. The dog barks (when it is hungry).

We take English to have only two genuine kind terms, the bare plural and the singular definite, based on the diagnostic of kind-level predication, rather than on generic statements involving object level predicates.⁷

⁵ 3b 3c and 4b are acceptable sentences of English under a taxonomic interpretation, where they refer to subspecies of the kind.

⁶ The generic reading translates roughly as *All dogs bark*; the habitual reading can be rendered as *There is a unique contextually salient dog and it is in the habit of barking*. A plural definite, because it is not a kind term, only has the habitual reading.

⁷ Actually, the case of indefinites is complicated but for present purposes I follow the conclusion in Krifka et al (1995) that indefinites are not kind denoting. See Dayal (2004) and Mueller-Reich (2006) for a different analysis.

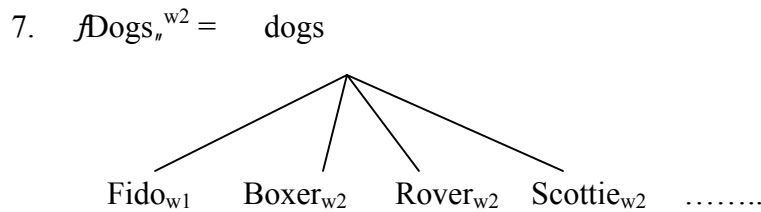
To complete the discussion we note that although the bare plural and the singular definite are both kind terms, their behavior differs in statements where aspect supports an episodic interpretation. The bare plural lends itself to an existential interpretation while the definite singular kind term does not. It is only a statement about a contextually salient dog:

- 6a. Dogs are barking. = *Some dogs are barking.*
 b. The dog is barking. ≠ *A dog is barking.*

The point that I hope to have highlighted in this discussion is that the correlation between form and meaning is not perfect. Bare plurals are kind terms, as demonstrated by their compatibility with kind-level predication, but can be used to make particular statements about members of the species. And indefinite singular terms, which are not kind-denoting, can be used in statements about the whole species. This distinction between form and meaning will be important when we try to look at the ways in which different natural languages navigate the divide.

2.2. Analyses of Kind Terms

Carlson’s well-known treatment of bare plurals takes them to be names of kinds, where kinds are individuals of a special sort. They differ from ordinary individuals in having instantiations scattered over worlds/situations, as shown schematically below:



This view of kinds is fairly well-established and there is general agreement that kind-level predication requires a kind-level argument. It is also probably not controversial that these kind individuals are built up from basic property-level meanings. That is, a common noun typically denotes a property, a function from worlds/situations to sets of individuals that have the property. The trick is to get from there to the type of meaning associated with noun phrases, an entity-type

meaning that would capture the ontological view of kinds given in (7). This has been formalized in a system such as Chierchia (1998) where the kind-formation operator NOM is defined as a function from worlds/situations to the maximal entity that instantiates the property at that world/situation:

- 8a. $\text{NOM}(\hat{\cdot}) : \lambda P_{\langle s \langle e, t \rangle \rangle} \lambda s \text{ ix } [P_s(x)]$
 b. Dinosauurs are extinct = extinct ($\hat{\cdot}$ dinosauurs)

Note that the operation $\text{NOM} : \hat{\cdot}$ has the same type as ordinary determiners, a function from property meanings ($\langle s \langle e, t \rangle \rangle$) to NP type meanings, ($\langle s, e \rangle$). While it is posited by Chierchia as a covert type shift, it could as easily be the meaning of a lexical determiner. If, for example, we had good reason to posit an empty determiner in (8b), we could define its meaning as NOM with no difference in results. Semantic commitment to this way of deriving kind terms, therefore, is independent of syntactic commitments to treating bare plurals as NPs rather than DPs. We will return to this question later in the paper.

The second aspect of Carlson's proposal for bare plurals has received less universal acceptance. According to him, even object level predication involving bare plurals makes reference to kinds. That is, bare plurals always denote kinds but they allow semantic operations access to the individual instantiations of the kind and the lexical and aspectual specification on the verb determines whether the quantification over those instantiations will be universal or existential. Again, we can use Chierchia's formalization, given in (9), to make things concrete. The basic idea is that a kind argument is of the right type (noun phrase type: $\langle s, e \rangle$) but the wrong sort, a kind rather than an object, to be used with predicates that apply to ordinary individuals. The repair operation *Derived Kind Predication* appeals to the inverse of NOM: ie $\text{PRED} : \cup$, which takes a kind term at a given index, world or situation, and returns the set of individuals that are part of the maximal entity denoted there. Default existential quantification then comes into play and, depending on the aspect, gets either generic or existential quantificational force:⁸

⁸ Note that in Chierchia's system the existential quantifier is, in effect, erased in the presence of a generic operator. The precise details are not directly relevant here.

9a. *DKP*: If P applies to objects (ie. ordinary individuals) and k denotes a kind, then

$$P(k) = \exists x [\cup k(x) \wedge P(x)]$$

b. PRED (\cup): $\lambda k_{\langle s, e \rangle} \lambda x [x \leq k_s]$

10a. Dogs bark.

b. Gen s, x [\cup dogs_s(x)] [bark_s(x)]

11a. Dogs are barking.

b. $\exists x [\cup$ dogs_s(x) & barking_s(x)]

The alternative view of bare plurals, inspired by the DRT approach to indefinites, is that they are ambiguous (Wilkinson 1991, Diesing 1992 and Gerstner and Krifka 1993, for example). They are kind denoting terms and can therefore be arguments of kind-level predicates. And they are ordinary indefinites so they can be arguments of object level predicates. Like other indefinites they introduce discourse referents, which can be caught by generic operators or existential closure, depending on the specification on the verb:

12a. Gen s, x [dogs_s(x)] [bark_s(x)]

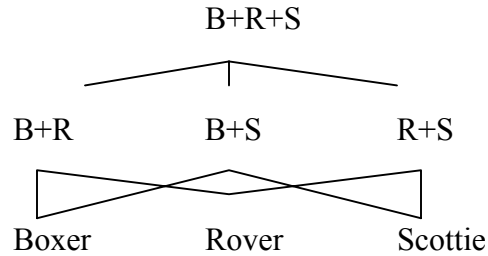
b. $\exists x$ [dogs_s(x) & barking_s(x)]

For present purposes, it is not crucial to choose between these two approaches. We will make our case based on kind-level predication, on which the two approaches agree (see Krifka et al 1995 and Dayal 2004 for further discussion).

2.3. Analyses of Definites and Indefinites

Let us turn now to definites and indefinites and consider their contribution to semantics. A very standard view of definites is that they refer to the maximal entity that meets the description. We take the domain of individuals to include atomic individuals as well as their sums and assume that singular morphology restricts the domain of quantification to the set of atomic individuals, while plural morphology brings in plural individuals into the set. We take *the* to denote a function IOTA:_t which takes a set of entities and returns the unique maximal entity in that set. A maximal entity is one that includes all others (see Link 1983 and Landman 1989)

13a.



- b. $f\text{The dogs}_n^{w2} = \iota(\{B, R, S, B+R, B+S, R+S, B+R+S\}) = B+R+S$
- c. $f\text{The dog}_n^{w2} = \iota(\{B, R, S\}) = \text{undefined}$
- d. $f\text{The dog}_n^{w1} = \iota(\{F\}) = F$

As we can see, the maximality requirement ensures that plural definites will pick out the whole group while singular definites will denote the unique individual with the relevant property or be undefined. In (13b) B+R+S is the only individual that includes all others in the set, so *the* is defined. In (13c) no such individual exists since B, R and S are all atomic and none of them include the others. The maximality requirement delivers the intuition that the only situation in which a singular definite will be defined is one in which there is only one individual in the domain.

The claim that maximality is part of the meaning of the definite determiner is based on examples like (14a)-(14b), based on a diagnostic in Loebner 1985.⁹ Indefinites differ from definites in not requiring maximality. The semantic operation associated with indefinites is EXIST:∃ which simply picks out some entity from the domain:¹⁰

- 14a. *The dogs are sleeping but the dogs are not.
- b. *The dog is sleeping but the dog is not.

- 15a. Some dogs are sleeping but some dogs are not.
- b. Some dog is sleeping but some dog is not.

16a. $f\text{Some dogs}_n^{w2} = \exists(\{B, R, S, B+R, B+S, R+S, B+R+S\}) =$
 $B+R \text{ or } B+S \text{ or } R+S \text{ or } B+R+S$

⁹ Of course, a relativized notion of maximality is needed to account for the fact that we can use definites even when they are non-unique. The point of examples such as the ones in (14) is to show that it is not possible for there to be two equally salient entities or group of entities denoted by a definite description.

¹⁰ Indefinites typically will not denote the maximal entity in the set, though the semantics allows for it. We take this to be a pragmatic effect, an implicature arising from the availability of a definite or a universal which would unambiguously refer to the full set. This implicature crucially does not arise when a new entity is introduced since definites are not compatible with them.

b. $\neg \text{Some dog}_i^{w2} = \exists(\{B, R, S\}) = B \text{ or } S \text{ or } R$

It is perhaps worth pointing out that the maximality effect is also not part of the semantics of demonstratives, which are attested universally. So the claim that a language does not have definite determiners rests on the diagnostic of maximality as a test.¹¹

In addition to maximality, definites and indefinites also differ in their relation to discourse (Kamp 1981 and Heim 1982). Definites presuppose familiarity, indefinites novelty, as illustrated in garden variety anaphoric contexts like the following:

17. I saw some dogs_i. The dogs_i /*Some dogs_i were barking.

There are many attempts in the literature to streamline the semantics of definiteness by eliminating one of these parameters, keeping only maximality or only familiarity/novelty but in this paper I will retain both. As noted in Dayal (2004), Hindi bare nominals can be considered definites along the dimension of maximality, but not along the dimension of discourse familiarity. We will see further evidence of the need to keep both aspects of the definite determiner in play when we examine cross-linguistic patterns in sections 3 and 4.

What we have so far, then, is the following table of correlations between form and meaning in English. Only bare plurals and singular definites yield kind readings, all forms except the definite plural yield generic readings, only definites yield contextually anchored or anaphoric readings and are incompatible with existential readings.¹² Of course, there are places where the lines between the various cells blur and we'll discuss some of them in the next section but this simplified picture of the distinction between form and meaning is useful to have as we proceed.

¹¹ See Robinson (2005) for the view that demonstratives may encode a non-maximality requirement.

¹² A note of clarification. It is possible for definites to be anaphorically linked to kind terms but particular examples with object denoting antecedents can be used to establish the facts.

	READINGS			→
FORMS	KIND <u>see 3,4,8</u>	GENERIC <u>see 5, 10</u>	C-ANCHORED <u>see 17 a, b</u>	EXISTENTIAL <u>see 6, 11, 12</u>
Bare Plural	√	√		√
Definite Singular	√	√	√	
▼ Definite Plural			√	
Indefinite Singular		√		√

Table 1: Form-Meaning Correlations in English

One final point is worth making. We have seen that bare plural kind terms have \exists readings in episodic contexts and are essentially synonymous with regular indefinites in those contexts. And this synonymy is what the analyses presented for them capture. Though the neo-Carlsonian approach represents them differently, the way the operations are defined, the statement with a bare plural has identical truth conditions to the statement with an indefinite. The DRT based approaches, of course, do not even posit a difference at the representational level. However, there is a difference between them that shows up in the following contexts:

18a. Dogs are barking outside. # Their names are Fido and Boxer.

b. Some dogs are barking. Their names are Fido and Boxer.

19a. #Dogs, namely Fido and Boxer, are barking outside.

b. Some dogs, namely Fido and Boxer, are barking outside.

This distinction between kind terms and regular indefinites has not featured in discussions of genericity and indefiniteness but the resistance of bare plurals to specification is significant in that it tells us that there is a residue of ‘kind’-like properties that remains in bare plurals even when they occur in episodic statements about the here and now.

3. *Cross-linguistic Variation*

Moving beyond English, even the briefest survey of other languages shows that the correlation between form and meaning outlined in table 1 does not hold universally. In this section we will

identify some generalizations about the patterns that emerge when we look at other languages and propose a framework for accounting for those patterns.

3.1. Some Generalizations

The first rather striking generalization, one that has been known for some time, is that no language has a determiner that is exclusively kind denoting. The variation is strictly between bare and definite forms. The absence of a dedicated kind determiner, namely one that is used only for referring to kinds, is surprising given that it is a semantic operation that is widely, if not universally, attested. To the best of my knowledge, no explanation for this generalization had been proposed till Dayal (2004).

There are two other stable cross-linguistic generalizations noted in Dayal (2004) and I present them here with greater explicitness. It appears from the available literature that kind terms vary between bare nominals and definites, depending on two factors. One is whether the NP involved is singular or plural, in languages that encode a number distinction in the noun phrase. The other is the form of the noun phrase used for contextually anchored readings. Not all possible combinations of form and meaning are attested.¹³ Table 2 lays out the three attested language types with the intended readings in the left-most column and the attested forms in the corresponding rows:

Table 2: Three attested language types

	<i>Determiner-less Languages</i> <i>HINDI</i>	<i>Mixed-Pattern Languages</i> <i>ENGLISH</i>	<i>Fully-Definite Languages</i> <i>ITALIAN</i>
Contextually Anchored	Bare	Definite	Definite
Singular Kind	Bare	Definite	Definite
Plural Kind	Bare	Bare	Definite

The table above is derived from examples discussed in earlier sections. I repeat the relevant examples from English:

¹³ We do not consider the case of mass terms in this paper, which seem to pattern with plural kind terms. Bare nominals in languages without number morphology, by and large, pattern with bare plurals.

- 20a. Some children came in. The children / *children sat down.
 b. The lion/*lion is an endangered animal.
 c. Dinosaurs/*the dinosaurs are extinct.

The translation of (20a) in Hindi would have a bare nominal in the relevant position, while its translation into Italian would have the determiner *il*. The same holds for the translation of (20b) where *il*, the singular version of the definite determiner, would be used in Italian. The translation of (20c) into Hindi would have a bare nominal but the translation into Italian would have the plural definite determiner.

For cross-linguistic generalizations to be interesting, however, it is essential to consider not only what is attested but, in addition, what is not attested. Table 3 includes possible language types that have not been documented in the literature so far.

Table 3. Possible Language Types

	TYPE - I	TYPE- II	Type- III	a	b	c	d	e
Anchored	B	D	D	B	B	D	D	B
Sing Kind	B	D	D	B	D	B	B	D
Plural Kind	B	B	D	D	D	B	D	B
	Hindi	English	Italian	*	*	*	*	*

One generalization that emerges from a consideration of table 3 is that a language that uses the definite determiner for plural kind formation also uses it for contextually anchored/anaphoric readings. In other words, if a language uses bare nominals for contextually anchored/anaphoric readings, then it also uses it for plural kind formation. Language types a and b, which do not adhere to this generalization, are not attested.

Another generalization we see in Table 3 is that contextually anchored/anaphoric readings and singular kind-formation agree in lexicalization. Either a language uses bare NPs for both or definites for both. Language types b, c, d and e, which do not adhere to this generalization, are not attested.

To sum up, there are three stable cross-linguistic generalizations that can be identified:

21. Generalization 1: Natural languages have no dedicated kind determiner.

Generalization 2: If a given language uses bare nominals for contextually anchored/anaphoric readings, then it also uses them for plural kind formation; if a language uses definites for plural kind formation it also uses them for contextually anchored/anaphoric readings.

Generalization 3: Contextually anchored/anaphoric readings and singular kind formation must agree in lexicalization. In a given language they will either both be bare or both definite.

These generalizations are remarkably stable and striking because there are no independent logical reasons for them.

3.2. *The Scale of Specificity and Linguistic Variation*

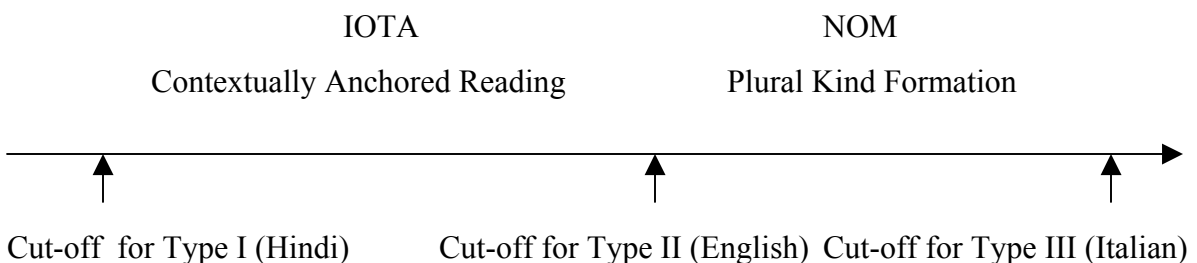
The explanation for Generalization 1 in (21) is remarkably simple. If we consider the semantic operations associated with plural kind-formation and regular definiteness, we see that the two differ only in intensionality; NOM is simply an intensional version of the maximality operator associated with the definite determiner:

22a. Plural Kind Formation : NOM: $\lambda P_{\langle s \langle e, t \rangle \rangle} \lambda s \iota x [P_s(x)]$
 b. Regular Definiteness: IOTA: $\lambda P_{\langle s \langle et \rangle \rangle} \iota x [P_s(x)]$

There are no dedicated kind determiners, then, because languages do not lexicalize extensional/intensional distinctions. The explanation, simple though it is, captures a deep connection between form and meaning as it relates to the expression of genericity.

The explanation for the second generalization turns on the claim that there is a scale of specificity that grammatical phenomena are sensitive to. The two operations we are interested in, IOTA and NOM, I claim map along this scale, with IOTA being more specific than NOM, in a sense to be elaborated upon in section 3.3.

Table 4. The Scale of Specificity



Variation arises because languages can choose distinct points on the scale for lexicalization, proceeding from left to right. Determiner-less languages are determiner-less because their cut-off point is at the extreme left. Both NOM and IOTA function covertly in such languages. Mixed languages are mixed because their cut-off point is in the middle so that IOTA is lexicalized but NOM is a covert type shift. Fully definite languages are those in which the cut-off is at the extreme right, encoding both IOTA and NOM lexically.

Language types a and b are ruled out by the proposed direction of lexicalization. In order for them to have a lexical determiner for plural kind formation, their cut-off point would have to be at the extreme right. This would mean, according to the current proposal, that IOTA could not be covert. That is, the unattested languages are those where lexicalization would not conform to the scale in Table 4.

As mentioned earlier, we will not go into an explanation of the third generalization in this paper for reasons of space. However, I will lay out some key features of the analysis here. Singular kind formation is not a by-product of plural kind formation. As stated in Dayal (1992), singular morphology in combination with the maximality requirement built into NOM would force the kind to have a unique instantiation in every world, something that clashes with the conceptual notion of a kind (see also Chierchia (1998)). The claim here is that singular kinds are atomic entities that belong in a taxonomic hierarchy and are subject to normal rules of quantification, resulting in what we recognize as taxonomic readings. The singular kind term is a particular instance of a taxonomic reading.¹⁴ We illustrate with two concrete examples:

23a. Every mammal gives live birth.

b. The dog is a mammal.

24a. Mammal-Subkinds = {DOG, LION, HORSE...}

b. $\forall x [x \in \{\text{DOG, LION, HORSE...}\} \rightarrow \text{give-live-birth}(x)]$

c. $\text{give-live-birth}(\iota x (\{\text{DOG, LION, HORSE...}\}(x)))$

¹⁴ Singular kind terms differ from plural kind terms in being atomic and cannot be a simple correlate of bare plurals created by NOM. One concrete piece of evidence is that the former is not compatible with accidental generalizations while the latter is: *The Rutgers professor seems to be born on a weekday vs. Rutgers professors seem to be born on weekdays (example due to Edwin Williams (p.c.)).

(23a) essentially says that every subkind of a mammal has the property of giving live birth and (23b) says that the unique subkind that is a DOG-subkind has that property. In other words, we recognize an ambiguity in common nouns with regard to whether they denote in the regular domain of individuals or in the taxonomic domain but not in the type of quantification involved. There are several issues that arise in connection with this approach to singular kind formation which are dealt with at greater length in Dayal (2004).

Focusing on the cross-linguistic generalization here, the explanation is obvious. Languages, as we have seen, may or may not lexicalize IOTA. Those that do, will use a lexical definite determiner for contextually anchored readings. They will also use it for singular kind terms, since the same operation is at issue. Those that do not, will use covert type shift regardless of whether the domain of quantification consists of ordinary individuals or of taxonomic entities. In each case, uniformity of lexicalization between contextually anchored and singular kind readings is predicted. That is, language types b, c, d and e in Table 3 are ruled out.¹⁵

3.3. Motivating the Scale

In the preceding section I placed IOTA at the left of the Scale of Specificity since NOM is defined in terms of IOTA. That is, the scale was motivated on the basis of theoretical considerations. However, we can also think of the scale in intuitive terms as a scale of diminishing specificity, taking into account the fact that we refer to individuals in two ways, in identity-oriented ways as well as property-oriented ways. One difference between the two is that identity-oriented reference allows for an accidental connection between the property and the predication while property-oriented reference requires an essential connection between them. A definite description is an apt vehicle for identity-oriented ways of reference. A kind term, on the other hand, clearly denotes in property-oriented ways, the identity of the individual members of the kind term is not at issue (cf. 18 and 19). English definite descriptions, then, can be considered more definite in this sense than bare plurals. Thus, working our way from intuitions too, we come to the same conclusion. IOTA, the semantic operation involved in definite

¹⁵ If there turn out to be languages in which contextually anchored readings and singular kind terms differ, it would require a modification of the analysis of singular kind formation. My hunch is that if this turns out to be the case, contextually anchored/anaphoric readings will be lexically expressed while singular kind formation will be bare.

descriptions, is at the left end of the scale of diminishing specificity while NOM, the semantic operation associated with bare plurals, falls at the right end of the scale.¹⁶

It may be worth clarifying that although definite descriptions are readily used for identity oriented modes of reference, they are not limited to such use. As was pointed out by Donnellan (1966), definite descriptions can also have attributive uses, where the identity of the individual is not particularly relevant. In such situations, interestingly, definite descriptions may be almost synonymous with bare plurals. Consider for example, (25a) uttered by a gardener whose lettuce plants are under attack. She may equally felicitously use the bare plural or the definite description to give vent to her frustration:

- 25a. The ground hogs/Groundhogs are eating up my lettuce.
b. eating-up-my-lettuce($\iota x[\text{groundhogs}_s(x)]$)
c. $\exists x [\cup \text{groundhogs}_s(x) \ \& \ \text{eating-up-my-lettuce}(x)]$

The existence of groundhogs can be easily accommodated in a context about gardening, in certain parts of America at least. The presupposition of IOTA being satisfied, it picks out the maximal entity in the set of groundhogs. Under the group reading of plural definites, it is not required that predication distribute down to all individual groundhogs. The group can be held responsible without any implication that each individual member engaged in the destructive act.¹⁷ Turning to NOM, it undergoes sort adjustment since a kind term occurs in an object level statement. This is accompanied by existential quantification over the instances of groundhogs in the situation (see section 2.2). The truth conditions of the two coincide.

Another situation where the bare plural and the definite seem to converge has been discussed by Condoravdi (1997). She notes that in a discourse like (26), the bare plural has a universal rather than the existential reading expected in non-generic contexts. Condoravdi calls this a functional reading of the bare plural. Note that the definite is a near equivalent option in these cases.

¹⁶ The reader may have noticed a certain kinship with the kinds of scales discussed by Haspelmath (1997) in connection with indefinites. Identity-oriented reference is akin to what he calls specific known and property-oriented reference bears some resemblance to what he classifies as irrealis. The in-between cases discussed below fall within his class of specific unknown. In discussing the scale I have substituted the term specificity for definiteness, used in earlier versions of the paper. As pointed out to me by Jane Grimshaw that use of the term definite was incompatible with the declaration at the beginning of section 2.

¹⁷ Sentences like “At the press conference, the reporters asked the President questions” illustrates the non-universal reading of plural definites. Very few reporters need have asked questions in order for the statement to be true. These issues are discussed in detail in Brisson (1998).

Mention of a campus allows for the easy accommodation of students, making the definite felicitous in these cases. NOM is, of course, possible since there are no specific students at issue:

26. There was a ghost on campus. The students/students were afraid.

Regardless of whether there is an independent functional reading involved in such cases, we can draw the following conclusion. We can consider IOTA more specific than NOM because it has the potential for identity-oriented reference but it is not restricted to such reference. There are situations in which an attributive plural definite and a bare plural are both acceptable. These are situations in which the existential presupposition of the definite is satisfied without there being the kind of specific reference that the bare plural is incompatible with. To sum up, IOTA and NOM are located at discrete points on the Scale of Specificity, but their semantics does not predict complementary distribution for expressions associated with them.

4. Some Further Issues

In this section we consider some further issues in connection with the view that cross-linguistic variation in genericity is predictable on the basis of lexicalization options defined on a universal scale of specificity.

4.1. Variations in Plural Kind Terms: a Refinement

The Italian plural definite generic and the English bare plural, I have claimed, encode the same semantic operation, namely NOM. This would lead one to expect that they would have identical distribution, but they do not. As expected, they align on kind and generic readings but, surprisingly, they part company on existential readings. (27c) only has the contextually anchored reading we associate with the regular definite determiner:

- 27a. *(I) cani sono diffusi
the dogs are widespread
“Dogs are widespread”
b. *(I) cani abbaiano
The dogs bark.
“Dogs bark.”

- c. I cani stanno abbaiano
 the dogs are barking
 “The dogs are barking.” not “(Some) dogs are barking”

The resistance to existential interpretations also holds true for more complex cases where the sentence is generic but the bare plural is expected to map into the nuclear scope. In (28), for example, the plural definite yields a reading whereby the same group of cats habitually runs across the garden:

- 28a. gatti corrono sul mio prato ogni giorno
 the cats run across my garden every day
 “The cats run across my garden every day” NOT “Cats run across my garden everyday.”
 b. * $\forall s [\text{day}(s)] \exists x [\text{cat}(x) \ \& \ \text{run-across-my-garden-in-}s(x)]$

Under the account sketched here we can explain this disparity between bare plurals and plural definites by appealing to the fact that the definite determiner, when it encodes IOTA, has a familiarity presupposition. I claim that a lexical definite determiner, even when it encodes NOM in addition to IOTA, retains a weak presupposition of existence and cannot be interpreted inside the nuclear scope. Presuppositions, we know, require lexical or structural triggers, so that a similar effect is not detected when these operations are covert.¹⁸

The inability of the Romance plural definites to yield existential interpretations, though well known in the literature, is not successfully accounted for in alternative accounts. For example, in accounts such as Vergnaud and Zubizarreta (1992) or Longobardi (1994, 1999) where the Romance plural definite determiner is considered a pleonastic it is not easy to rule out existential interpretations for them. The proposal made here attempts a principled account for the overlap in distribution between English bare plurals and Italian plural definites. Their uniform behavior is attributed to their having the same semantics, the difference due to the way in which presuppositions interact with the form in which they are encoded in the language.

4.3. The Plural Definite Generic and the Bare Plural in German

¹⁸ Note that the plural definite generic is acceptable in Italian translations of 25 and 26, cases where a weak existential presupposition is satisfiable by accommodation. See also Zamparelli 1998 for discussion of these issues.

The approach to variation proposed here suggests that for any given semantic operation, in particular IOTA and NOM, a language chooses either to encode them in a lexical determiner or to have them function as covert type shifts. A language like German, however, seems to present an apparent counterexample to this proposal. The goal of this subsection is to see what adjustment to the theory is needed in order to accommodate languages of this type.

As noted in Krifka et al (1995), plural and mass kind terms may be bare nominals or definites in (at least some dialects of) German. This is illustrated below:

- 29a. (Die) Pandabären sind vom Aussterben bedroht.
 “Pandas are facing extinction.”
 b. (Das) Gold steigt im Preis.
 “Gold is getting more expensive.”

Interestingly, this does not hold for contextually anchored readings or singular kind terms (Dayal 2004):

- 30a. Ich habe ein hund. Der Hund / *Hund bellt
 “I have a dog. The dog is barking.”
 b. Der Pandabär / *Pandabär ist vom Aussterben bedroht.
 “The Panda is facing extinction.”

We can maintain the account sketched so far if we constrain lexicalization along the scale of specificity as laid out in section 3, without precluding the possibility of covert type shift for a specific lexicalized operation. In order to do that, we appeal to the Blocking Principle (Chierchia 1998), for which empirical justification is given in (32):

31. *Blocking Principle*: For any type shifting operation π and any X: * $\pi(X)$ if there is a determiner D such that for any set X in its domain, $D(X) = \pi(X)$.

- 32a. kuch bacce andar aaye. bacce bahut khush the
 some children inside came children very happy were
 “Some children came in. The children were very happy.”
 b. Some children came in. *Children/The children were very happy.

A bare nominal in English does not have the option of type-shifting by IOTA since that operation is lexically encoded, but a corresponding nominal in Hindi is able to do so since there are no lexical options available.

The advantage of separating the lexicalization process from the issue of complementarity between overt and covert type shifts is that we can now make the Blocking Principle sensitive to canonical vs. non-canonical meanings. We can say that IOTA is the canonical meaning of the definite determiner in any language and the Blocking Principle prevents a language that lexicalizes IOTA from also having it as a covert type shift. This leaves open the possibility for the Blocking Principle to be suspended for its non-canonical meaning NOM. Under this perspective, German would have the same cut-off point as Romance, lexicalizing both IOTA and NOM, but would differ from Romance in enforcing the Blocking Principle only for IOTA.

It bears emphasizing that this move does not rob the proposal of predictive power. It still rules out a number of logically possible language types. For example, it rules out languages in which a lexical determiner would be needed for plural kind terms but not for contextually anchored readings, as well as languages in which contextually anchored readings could be expressed by bare or definite nominals but plural kind terms would be obligatorily definite. The approach to variation argued for remains restrictive and empirically testable.

5. Conclusion

I have argued in this paper that the Romance plural definite generic is a full-fledged determiner denoting a function from properties to individual concepts, a kind, and having a presupposition that precludes its being mapped into the nuclear scope. That is, I treat the Romance kind term as a bona fide DP. Let us now consider the morpho-syntactic status of its counterpart, the English bare plural kind term. Briefly, there are three possible syntactic structures that we might consider, a structurally reduced NP in which no Det node is projected, a full DP structure with an empty Det and a full DP structure in which N occurs in D:

- 33a. $[_{NP} N]$
 b. $[_{DP} [_{Det} e] [_{NP} N]]$
 c. $[_{DP} [_{Det} N_i] [_{NP} t_i]]$

If we take bare plurals to be NPs as in (33a), we can posit the covert type shift NOM as a repair operation that mediates when a property denoting expression occurs in an argument position. This is, in essence, the position of Carlson 1977, Partee 1987 and, modulo parameterization, also Chierchia 1998. Under this view, a language that does not lexicalize NOM (or IOTA, for that matter) simply uses it to repair type mismatches when NPs occur as arguments.

Alternatively, we can take bare plurals to be DPs with an empty determiner denoting NOM, as shown in (33b). This option is, in effect, the mirror image of the pleonastic determiner option for Romance definites. It makes the structure of the bare plural and the plural definite isomorphic, up to lexicalization. While this approach would work quite well as far as the facts considered in this paper go, there are other reasons for rejecting this possibility. Typically, empty determiners are supposed to require external licensing, resulting in subject-object asymmetries. This is the case, for example, with Italian bare plurals, which typically are restricted to governed positions.¹⁹ Absent evidence of similar restriction, positing a structure like (33b) for English bare plurals is unwarranted.

Finally, we have the possibility of a DP structure, with N moving to D as in (33c). Again, the covert type shift could be associated with the need for DP's to have argument-like meanings. Movement of N to D has been forcefully argued for by Longobardi (1994, 1999) for Italian on the basis of word order facts. Briefly, Italian allows [Det Adj N] or [N Adj] but not [Adj N] orders, a fact that can be explained if it obligatorily projects D and forces N to move to it when no determiner is present. Since these facts are not replicated in English, N to D movement would have to occur at LF. The problem is in finding independent evidence for this.

To sum up, then, it appears that semantic considerations under-determine syntactic structure. All three options, we have seen, are equally compatible with covert type shifts. (33a) appears to be the null hypothesis, the one which requires the least additional assumptions. (33b) and (33c) are plausible candidates but the required adjustments to the system would have to be justified. A consequence of adopting (33a) is that we do not require uniformity of structure for nominal arguments across or within languages. An argument with a lexical determiner may be a DP and one without such a determiner an NP, even if they mean (almost) the same thing. I have argued here, on empirical and theoretical grounds, that this may not be an undesirable consequence after all. Morpho-syntactically distinct forms, even when they appear to be synonymous, typically only overlap in meaning. Accounts that maintain such distinctions at the level that feeds into semantics, because they do not predict complete semantic identity, are better able to deal with such paradigms.

¹⁹ Chierchia (1998) analyses the empty D as encoding NOM while Longobardi 1994, 1999 and Zamparelli 1998 take it to be a narrow scope existential quantifier.

REFERENCES

- Bach, E., E. Jelinek, A. Kratzer and B. Partee (eds), 1995. *Quantification in Natural Language*, Kluwer Academic Press.
- Bittner, M.: 1994. *Case, Scope and Binding*. Kluwer Academic Publishers.
- Brisson, C.: 1998, *Distributivity, Maximalization and 'Floating Quantifiers'*, Ph.d. dissertation, Rutgers University.
- Carlson, G.: 1977, *Reference to Kinds in English*, Ph.D. Dissertation, University of Massachusetts, Amherst.
- Chierchia, G.: 1998, 'Reference to Kinds Across Languages', *Natural Language Semantics* **6**, 339-405.
- Condoravdi, C.: 1997, *Descriptions in Context*, Ph.D. Dissertation, Stanford University.
- Dayal, V.: 1992, 'The Singular-Plural Distinction in Hindi Generics', *Proceedings of SALT II*.
- Dayal, V.: 2000. "Scope Marking: Cross-linguistic Variation in Indirect Dependency", in U. Lutz, G. Mueller and A. von Stechow (eds.) *Wh-Scope Marking*, John Benjamins, pp.157-193.
- Dayal, V.: 2004, 'Number Marking and (In)definiteness in Kind Terms', *Linguistics and Philosophy* **27.4**, pp. 393-450.
- Diesing, M.: 1992, *Indefinites*, MIT Press, Cambridge.
- Donnellan, K. (1966) "Reference and Definite Descriptions", *Philosophical Review* **75**, pp. 281-304.
- Gerstner, C. and M. Krifka: 1993, 'Genericity', in J.Jacobs, A. von Stechow, W. Sternefeld and T. Venneman (eds) *Handbuch der Syntax*, de Gruyter, Berlin.
- Haspelmath, M.: (1997), *Indefinite Pronouns*, Clarendon Press, Oxford.
- Heim, I.: 1982, *The Semantics of Definite and Indefinite NP's*. Ph.D. Dissertation, University of Massachusetts, Amherst.
- Kamp, H.: 1981, 'A Theory of Truth and Discourse Representation', in J. Groenendijk, T. Janssen and M. Stokhof (eds.) *Formal Methods in the Study of Language*, Mathematical Centre, Amsterdam.
- Krifka, M., F.Pelletier, G. Carlson, A. ter Meulen, G. Chierchia and G. Link: 1995, 'Genericity: an introduction', in G.Carlson and F.J. Pelletier (eds.) *The Generic Book*, The University of Chicago Press, Chicago.
- Landman, F.: 1989, Groups I, *Linguistics and Philosophy* **12**, 559-605.
- Link, G.: 1983, 'The Logical Analysis of Plural and Mass Nouns: A Lattice Theoretic Approach', in R. Bauerle, C. Schwarze and A. von Stechow (eds.), *Meaning, Use and Interpretation of Language*, de Gruyter, Berlin.
- Loebner, S.: 1985, 'Definites', *Journal of Semantics* **4**, 279-326.
- Longobardi, G.: 1994, 'Reference and Proper Names', *Linguistic Inquiry* **25**(4), 609-665.
- Longobardi, G.: 1999, 'How Comparative is Semantics? A Unified Parametric Theory of Bare Nouns and Proper Names', Unpublished University of Trieste ms.
- Mueller-Reichau, O.: 2006, *Sorting the world -- on the relevance of the kind-level/object-level distinction to referential semantics*, Ph.D. thesis, University of Leipzig.
- Partee, B.: 1987, 'Noun Phrase Interpretation and Type-Shifting Principles', in J. Groenendijk et al. (eds.) *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Foris, Dordrecht.
- Robinson, H.: 2005, *Unexpected (In)definiteness: Romance Plural Determiners in Generic Contexts*, Ph.D. dissertation, Rutgers University.

- Srivastav, V.: 1991. *Wh Dependencies in Hindi and the Theory of Grammar*, Ph.D. dissertation, Cornell University.
- Vergnaud, J.R. & M. L. Zubizarreta: 1992, 'The definite determiner in French and in English', *Linguistic Inquiry* **23**(4).
- Wilkinson, K.: 1991, *Studies in the Semantics of Generic NP's*, Ph.D. Dissertation, University of Massachusetts, Amherst.
- Zamparelli, R. 1998, 'A Theory of Kinds, Partitives and Of/Z Possessives', in Alexiadou and Wilder (eds.) *Possessors, Predicates and Movement in the Determiner Phrase*, John Benjamins, Linguistics Today, 22.