

Mass Definite Generics

Aviv SCHOENFELD

Introduction.

Languages with definiteness and the count-mass distinction differ in whether the definite generic article is forbidden, optional or obligatory in (simplex) mass NPs, (1) (English, German, Spanish).

- (1)
- a. (#The) gold is getting more expensive.
(the gold can refer to a kind of gold)
 - b. (Das) Gold steigt im Preis.
(Dayal 2004:ex.86b)
 - c. #(El) agua se encuentra por todas partes.
(Borik & Espinal 2015:ex.31b)
the water_{refl} found for all parts
'Water is found everywhere.'

Although generic *the* is forbidden in the simplex mass NPs in (1a) and (2), it is optional in the complex ones in (3), with five sorts of modification. *Pesto* and *pesto sauce* are synonyms, so it is modification rather than meaning which licenses *the* in (3a) and (3b–e) by extension.

- (2) (#The) {pesto, hating, tuberculosis, tape, wine} is widespread.
unmodified

- (3)
- a. (The) pesto sauce is widespread.
1st noun in N-N compound
 - b. (The) electrician's tape is widespread.
modification genitive
 - c. (The) hating of minorities is widespread.
argumental genitive

d. (The) pulmonary tuberculosis is widespread.

relational adjective

e. (The) French wine is widespread.

classificative ethnic adjective

We take the licensing by modification in (2–3) to be insightful about modification in general, and we attribute the optionality of *the* in (3) to modifiers having kind-level denotations.

As for (1), the analyses of Dayal (2004) and Borik & Espinal (2015) have different implications for how languages vary, and we give a new argument for the former from diachrony.

Background.

Under Dayal (2004:§3.2), properties can shift to kinds via the type-shifts in (4).

(4)

a. $\lambda P.\cap P$ ($\cap P$ defined only if every extension of P has a maximal element)

‘The function from property P to its kind-correlate.’

b. $\lambda P.\iota(\lambda k.P_{\text{taxonomic}}(k))$

‘The function from property P to the maximal element in the set of (proper and improper) kinds of P .’

(4a) is covert in English while (4b) is vocalized as *the*, and (4b) is applicable only if (4a) is not. To illustrate, (4a) is applicable to GOLD; in every situation with gold, the sum of all gold is gold. This applicability blocks (4b) (vocalized as *the*) from applying to GOLD, explaining *the* being forbidden in (5a). By contrast, (4a) is inapplicable to LION; in situations with multiple lions, the sum of all lions is not a (singular) lion. This allows (4b) vocalized as *the* to apply to LION, explaining *the* being obligatory in (5b).

- (5)
- a. **WIDESPREAD** (\cap GOLD) (4b) blocked
(#The) gold is widespread.
 - b. \cap LION undefined **WIDESPREAD** ($(\lambda \mathbf{k} . \text{LION}_{\text{taxonomic}}(\mathbf{k}))$)
(#The) lion is widespread.

Modification. In Polish, the position of an adjective corresponds to a kind- or instance-level use.

- (6)
- a. czarny dzięcioł ‘woodpecker who is black’
(Wągiel 2014:ex.10)
 black woodpecker
instance-level use
 - b. dzięcioł czarny ‘specimen of the species *Dryocopus martius*’
 woodpecker black
kind-level use

We posit that the sorts of modifiers in (3) have the dual-use in (6), but without affecting word order in English. In support, *pesto* has a kind-level use in *Pesto Genovese is a (widespread) pesto*. Also, there is theoretical intuition that modificational genitives like *electrician’s* in (3c) involve reference to kinds (Munn 1995). Lastly, McNally & Boleda (2004) analyze *pulmonary* in (7a) as having the kind-level denotation in (8a), which we extend to the modifiers in (7b–c) via (8b–c).

- (7)
- a. Tuberculosis can be pulmonary.
(McNally & Boleda 2004:ex.33)
 - b. Hating can be of minorities.
genitive argument
 - c. This kind of wine is French.
classificative ethnic adjective
 (Arsenijević et al. 2014)

(8)

a. $\llbracket \text{pulmonary} \rrbracket = \lambda \mathbf{k}.\text{PULMONARY}(\mathbf{k})$

(McNally & Boleda 2004:ex.35b)

'The set of kinds which verify the kind-level predicate PULMONARY.'

b. $\llbracket \text{of minorities} \rrbracket = \lambda \mathbf{k}.\forall w \forall e[\text{R}_w(e, \mathbf{k}) \rightarrow * \text{MINORITY}_w(\text{TH}(e))]$

'The set of kinds s.t. every possible event (e) which they realize (R) has a plurality of minorities (*MINORITY) as its theme (TH).'

c. $\llbracket \text{French} \rrbracket = \lambda \mathbf{k}.\text{ORIGIN}(\mathbf{k}, \text{France})$

(Arsenijević et al. 2014:ex.17)

'The set of kinds which come into existence within the spatial domain of France.'

(9) implements the dual-use assumption on *pesto* as a modifier (as in *pesto sauce*).

a. $\llbracket \text{inst pesto} \rrbracket = \lambda s \lambda x.\text{PESTO}(x) \langle s, \langle e, t \rangle \rangle$

instance-level property

'The function from situations s to the set of sums of pesto in s.'

b. $\llbracket \text{subkind pesto} \rrbracket = \lambda \mathbf{J} \lambda \mathbf{k}.\mathbf{J}(\mathbf{k}) \wedge \text{PESTO}(\mathbf{k}) \langle \langle e^k, t \rangle, \langle e^k, t \rangle \rangle$

kind-level modifier

'The function from sets of kinds to their intersection with the set of kinds of pesto.'

We propose that the (non-)occurrence of *the* in (3) corresponds to two derivations of equivalent propositions. The bare version of (3a) utilizes (9a), whose property-intersection (\cap_p) with SAUCE undergoes η ; it is covert in English, hence (10a) has bare *pesto sauce*. By contrast, the definite version of (3a) utilizes (9b), which prompts SAUCE to shift to SAUCE (the set of kinds of sauce), and the maximal element of the resulting set is picked out by ι denoted by *the*, (10b). In both cases, the argument of WIDESPREAD is pesto sauce as a kind.

(10)

- a. **WIDESPREAD** ($\cap \lambda w \lambda x. \text{PESTO}_w(x) \quad \cap_p \lambda w \lambda x. \text{SAUCE}_w(x)$)
WIDESPREAD ($\cap \lambda w \lambda x. \text{PESTO}_w(x) \wedge \text{SAUCE}_w(x)$)
Pesto sauce is widespread.
- b. **WIDESPREAD** ($\iota \lambda J \lambda k. J(k) \wedge \text{PESTO}(k) \text{ SAUCE}$)
WIDESPREAD ($\iota \lambda k. \text{SAUCE}(k) \wedge \text{PESTO}(k)$)
The pesto sauce is widespread.

We extend analysis (10) to Polish, which lacks a definite article but the use of adjectives manifests in word order, (11) (Wągiel p.c.).

(11)

- a. Pszenica zwyczajna jest rozpowszechniona w Europie.
wheat common is widespread in Europe.LOC
'Common wheat (*Triticum aestivum*) is widespread in Europe.'
- b. Zwyczajna pszenica jest rozpowszechniona w Europie.
common wheat is widespread in Europe.LOC
'Wheat which is ordinary for wheat is widespread in Europe.'

In our analysis, post-nominal *zwyczajna* 'common' in (11a) denotes a kind-level modifier which applies to **WHEAT** and returns a set whose maximal element is *Triticum aestivum*. This kind is picked out by ι as in (10b), except ι is covert in Polish. By contrast, pre-nominal *zwyczajna* in (11b) denotes an instance-level property which combined with **WHEAT** returns the property of instances of wheat which are ordinary for wheat. This property shifts to a kind via \cap , as in (10a).

Why does modification license mass definite generics in English? In our analysis, it is due to changing the locality facts. When **PESTO** is most local to the kind-level predicate, the mismatch is repairable via the highly-ranked \cap , which blocks lower-ranked type-shifts like (4b). By contrast, when **PESTO** is most local to the denotation of the modifier, the mismatch is unrepairable with \cap , which allows lower-ranked shifts like \cap_p in (10a) or the shift from **SAUCE** to **SAUCE** in (10b). The latter needs ι denoted by *the* to achieve reference to kinds, hence modification licenses the definite generic article in mass NPs in English.

Cross-linguistic.

(12) is Borik & Espinal's (2015:ex.63) account of the definite article being forbidden in (1a) and obligatory in (1c).

(12)

- | | | | | | |
|----|-----------------------------------|---------|-------------------------|---------|----------------------|
| a. | Basic intension of noun | English | instance-level property | Spanish | kind-level predicate |
| b. | Is $\bar{\iota}$ in the language? | | yes | | no |

Following (12), the basic intension of *gold* is a property to which $\bar{\iota}$ is applicable, hence *the* is forbidden in (1a). By contrast, the basic intension of *agua* 'water' is a set of kinds whose maximal element is water as a kind. The only way to refer to this kind is via $\bar{\iota}$ denoted by *el*, hence it is obligatory in (1c). (12) could extend to account for the optionality in (1b) by positing that German has covert $\bar{\iota}$ and *Gold* is ambiguous between a property and a kind-level predicate, i.e. the bare version of (1b) results from covert $\bar{\iota}$ applying to the property-denotation of *Gold*, and the definite version results from $\bar{\iota}$ denoted by *das* applying to the kind-predicate denotation. However, it is unclear why languages should vary in the two ways in (12).

(13) is Dayal's (2004) account of (1), which assumes that $\bar{\iota}$ is a canonical function of the definite article while $\bar{\iota}$ is non-canonical. (13a) has the denotations of the definite articles, and it follows from (13b) that although the German definite article can denote $\bar{\iota}$, this does not block covert $\bar{\iota}$. Thus, bare and definite (1b) result respectively from covert $\bar{\iota}$ and $\bar{\iota}$ denoted by *das*.

(13)

- | | | | | |
|----|--|----------------|--|-----------------------------|
| | | <i>English</i> | | <i>German</i> |
| | <i>Spanish</i> | | | |
| a. | The definite article lexicalizes | $\bar{\iota}$ | | $\bar{\iota}$ $\bar{\iota}$ |
| b. | A covert type-shift is blocked if it is equivalent to any a canonical any function of an overt determiner. | | | |

Unlike (12), diachrony offers an answer to why languages vary as in (13). English, German and Spanish represent consecutive stages in a diachronic progression where the definite article expands in use (Mainz 2020), which we formalize as acquiring \bar{n} in addition to \bar{i} . We further assume that there is a delay between acquiring \bar{n} and the universal Blocking Principle in (14) taking effect to block covert \bar{n} . Thus, the optionality in (1b) is due to German residing in the delay, while the obligatoriness in (1c) is due to (14) “catching up” after the Spanish definite article has acquired \bar{n} . Thus, we argue for (13) over (12) as an account of (1).

References:

- Arsenijević, B. et al. 2014. Ethnic adjectives are proper adjectives. *Chicago linguistic society* 46. 17–30.
- Borik, O. & M. T. Espinal. 2015. Reference to kinds and to other generic expressions in Spanish: Definiteness and number. *The linguistic review* 32(2). 167–225.
- Dayal, V. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and philosophy* 27(4). 393–450.
- Munn, Alan. 1995. *WECOL* 24, 181–195.
- Schmuck, Mirjam. 2020. The grammaticalization of definite articles in German, Dutch and English: A micro-typological approach. In *German and Dutch in contrast*, 145–178. Berlin/Boston: de Gruyter.
- Wągiel, M. 2014. From kinds to objects: Prenominal and postnominal adjectives in Polish. *Olomouc Linguistics Colloquium 2014*, 457– 476.