

Cardinality and (In)definiteness

Yağmur Sağ

Harvard University

Introduction In every language, numeral constructions (NCs) can freely receive indefinite interpretations. This is even the case in languages that strictly disallow bare nouns in argument positions, like French (Chierchia 1998). NCs can also co-occur with the definite article in languages like English and French, but they typically require demonstratives to convey definite-like readings in articleless languages such as Mandarin and Russian (Jiang 2012). The resistance of NCs to definiteness in articleless languages is particularly puzzling because bare nouns freely allow definite interpretations in these languages (Chierchia 1998, Dayal 2004). Our study links this general indefinite characteristics of NCs to the projection of a cardinal head that may surface covertly and/or overtly, through an analysis of NCs in Turkish, an articleless optional classifier language. At the end, we also present data from Farsi, which reinforces the cross-linguistic breadth of our proposal.

Background Languages like English that distinguish between the unmarked and plural form of nouns also reflect this in their NCs. Languages like Mandarin, which lack a systematic number marking system, use the unmarked form in NCs but require a classifier, regardless whether the noun is ontologically count or mass. In English only mass nouns are incapable of directly combining with a numeral, modulo packaging coercion. This disparity has led to the view that all nouns are mass or mass-like in Mandarin-like languages, requiring some sort of fixing via classifiers to make counting possible with them (e.g., Krifka 1995, Chierchia 1998, Borer 2005).

Puzzle Turkish has a systematic number marking mechanism but NCs use the unmarked form with all numerals, also featuring an optional classifier, *tane* (1a). Both numerals and *tane* are selective for the count sense of the noun, while mass nouns obligate a mediator for counting (1b) & (1c).

- (1) a. iki (tane) kitap(*-lar) b. iki (tane) su c. iki *(damla) su
two CL book-PL two CL water two drop water
'two books' 'two waters (coerced)' 'two drops of water'

Crucial for our purposes, *tane* may seem optional syntactically, but the two forms of NCs vary in their interpretation: Both forms can be indefinite and hence yield exceptional scope readings (2). However, NCs without *tane* can also be definite freely, unlike NCs with *tane*, which instead need demonstratives to gain a definite-like interpretation (3). More precisely, while NCs with *tane* are restricted to indefiniteness, as typical for NCs cross-linguistically, NCs without *tane* draw away from the ordinary picture in also allowing definiteness.

- (2) Eğer iki (tane) proje-m seçil-ir-se, ödenek al-abil-eceğ-im.
if two CL project-1SGPOSS select-PASS-AOR-COND, funding take-ABIL-FUT-1SG
'If two of my projects (collectively) are selected, I will receive funding.' (if > two, two > if)
- (3) Polis üç (tane) hemşire-nin ölüm-ü-nü araştır-ıyor. Üç (#tane)
police three CL female.nurse-GEN death-3POSS-ACC investigate-PROG three CL
kadın-ın/ bu üç (?tane) kadın-ın ellili yaş-lar-da ol-duğ-u
woman-GEN this three CL woman-GEN fifties age-PL-LOC be-NMLZ-3POSS
tahmin ed-il-iyor.
predict-PASS-PROG
'The police are investigating the death of three nurses. It is predicted that *the three women/ these three women* were in their fifties.'

The challenges these variations introduce are twofold: (i) how the counting system works in Turkish and what role *tane* plays in this, (ii) how the presence/absence of *tane* affects interpretation, contributing to the exceptional status of the Turkish counting system.

The Cardinal Head Considering that *tane* is selective for count nouns only, we analyze it as a distinct category from obligatory classifiers. We follow Scontras (2014) in that NCs universally bear a cardinal head (CARD) that denotes a counting function. We propose that while CARD is typically only realized covertly as in English, the Turkish CARD can also be realized overtly as *tane*. (CARD can be a separate covert head above the classifier in Mandarin-like languages, or its semantics can be intertwined within the denotation of the classifier, as in Krifka (1989).) Sağ (2018) shows that Turkish unmarked nouns are ambiguous in denoting atomic properties and singular kinds like English unmarked nouns à la Dayal (2004) (also Martí 2020). Since singular kind reference is opaque to instantiation (Dayal 2004), Sağ, following Ionin & Matushansy (2006: I&M), argues that counting requires atomic properties in Turkish whether *tane* is present or not. In I&M’s view, this requirement is fulfilled by morphologically unmarked nouns in languages like Turkish, while English NCs further involve plural agreement. Adopting these views, we take CARD to uniformly presuppose a semantically singular form of the noun (drawing on I&M’s semantics of numerals):

$$(4) \quad \llbracket \text{CARD} \rrbracket = \lambda P \lambda n \lambda x : \forall y [P(y) \rightarrow AT(y)]. \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s [s \in S \rightarrow P(s)]] \\ \prod(S)(x) = 1 \text{ iff } S \text{ is a cover of } x, \text{ and } \forall z, y \in S [z = y \vee \neg \exists a [a \leq_i z \wedge a \leq_i y]]$$

A set of individuals C is a cover of an individual X iff X is the sum of all members of C .

Indefiniteness To understand the inherent indefiniteness of NCs, let us consider the interpretation of bare arguments. In the neo-Carlsonian approach, bare nouns in articleless languages can be definite via covert ι type-shifting, which is unavailable in languages with a definite article due to the Blocking Principle; it requires the use of lexical items instead of covert type-shifters (Chierchia 1998). That NCs cannot convey definiteness without demonstratives in articleless languages shows that ι type-shifting is not available for NCs, contrasting with bare nouns. In short, while indefiniteness is the default interpretation for NCs, definiteness is only available through overt marking. Jiang (2012) dedicates the source of indefiniteness to a lexical variant of numerals that has a built-in choice function variable f (Reinhart 1997). Considering the facts of Turkish NCs, we propose that this source is rather linked to CARD, as shown in (5). In Reinhart’s theory, the \exists -closure of f applies at any compositional level, deriving the exceptional scope readings of indefinites.

$$(5) \quad \llbracket \text{CARD}_f \rrbracket = \lambda P_{AT} \lambda n. f(\lambda x \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s [s \in S \rightarrow P(s)]])$$

NCs are of type e due to f built-in CARD’s denotation [see (8)]. We follow Dayal (2013) in that NCs type-shift to a predicate (via *ident*; Partee 1986) as a repair operation only in structures where a predicative meaning is required, i.e., when NCs occupy the predicate position or need to combine with an overt determiner. NCs would not undergo predicative type-shifting to shift again with ι in argument positions because they are already in the proper type to combine with whatever is taking the NC as an argument. That is, there is no motivation for the repair strategy to kick in. NCs then gain definite-like readings in articleless languages only via overt means like demonstratives.

Back to Turkish We propose that in articleless languages with overt and covert forms of CARD, the restriction of NCs to indefiniteness can be lifted by liberating one form from the choice function variable. The liberated form of CARD creates predicative NCs, which are then enriched in meaning since they can feed into covert type-shifting operators, including ι and the choice function, and allow both definite and indefinite interpretations. In Turkish, the form that is associated with f is the overt CARD, i.e., *tane*, while the covert CARD has a predicative value (6). As a result, while

NCs with *tane* are just like NCs in other languages reflecting the inherent indefiniteness, NCs with $CARD_{\emptyset}$ bring a seemingly exceptional status to the Turkish counting system [see (9)].

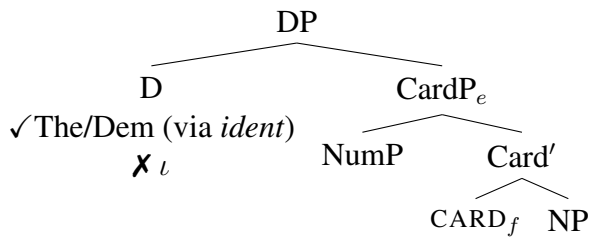
(6) $[[tane]] = (5)$: argumental vs. $[[CARD_{\emptyset}]] = (4)$: predicative

Further Support While the liberated form is the covert $CARD$ in Turkish, this is a language-specific choice. To create a comparative platform, we investigated NCs in another articleless optional classifier language, i.e., Farsi, and found that it is NCs with the classifier, i.e., *tā*, that allow definite and indefinite readings, while the classifierless form is only indefinite. Farsi NCs involve plural agreement as in English, though this is conditioned by definiteness (see also Alexiadou 2019; Smith-Stark 1975). That is, in the absence of the plural, NCs are indefinite while in the presence of the plural they are definite. While NCs with *tā* can receive a definite interpretation that also triggers plural agreement marking on the noun, NCs without *tā* are not compatible with the plural and require a demonstrative to display a definite-like behavior instead (cf. with (3)):

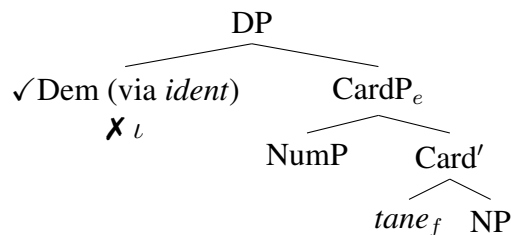
(7) Polis dar hāl-e barrasi-e marg-e se (tā) moallem-e zan-(*hā) ast. Rasāne-hā-ye mahali migooyand se *(tā) zan-hā/ #(in) se (tā) zan ke dar daheye panjah-e local say three CL woman-PL this three CL woman that in fifties-EZ zendegi-e khod budand.
 life-EZ themselves were.
 ‘The police are investigating the death of three female teachers. Local channels report that the three women/ these three women were in their fifties.’

To conclude, our study contributes to the cross-linguistic semantics of NCs by bringing novel data from Turkish and Farsi, and relates the indefiniteness of NCs with cardinality.

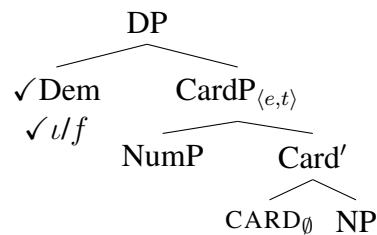
(8) Generalized Structure of NCs



(9) a. Turkish NCs with *tane*



b. Turkish NCs with $CARD_{\emptyset}$



Note: Farsi exhibits the mirror image of the structures in (9).

References. Alexiadou 2019. Morphological and semantic markedness revisited: The realization of plurality across languages. Borer 2005. Structuring Sense: Volume 1: In Name Only. Chierchia 1998. Reference to Kinds across Languages. Dayal 2004. Number Marking and Indefiniteness in Kind Terms. Dayal 2013. On the Existential Force of Bare Plurals across Languages. Ionin & Matushansky 2006. The Composition of Complex Cardinals. Jiang 2012. Nominal Arguments and Language Variation. Krifka 1989. Nominal reference, temporal constitution and quantification in event semantics. Krifka 1995. Common Nouns: A Contrastive Analysis of Chinese and English.

Martí 2020. Numerals and the Theory of Number. **Partee 1986.** Noun phrase interpretation and type-shifting principles. **Reinhart 1997.** Quantifier Scope: How labor is divided between *or* and Choice Function. **Sağ 2018.** The Semantics of Turkish Numeral Constructions. **Scontras 2022.** On the Semantics of Number Morphology. **Smith-Stark 1974.** The plurality split.