

# On the (non-)optionality of the Turkish classifier *tane*

Yağmur Sağ

Received: date / Accepted: date

**Abstract** Languages like English that distinguish between the morphologically unmarked and plural forms of nouns also reflect this in their numeral constructions. Languages like Mandarin, which lack a systematic number marking mechanism, use the unmarked form for all numerals but require classifiers in counting. While considerable research exists on these systems, little is known about languages where counting constructions use classifiers optionally. This paper aims to fill the gap by analyzing the counting system in Turkish, a language with systematic number marking that employs unmarked nouns for all numerals and features an optional classifier, *tane*. The investigation of this little-understood ingredient of counting provides novel insights into the interpretation of numeral constructions cross-linguistically.

Numeral constructions display a robust pattern of strong indefiniteness, while typically requiring overt markers like a definite article or demonstrative for definiteness. The prevalent perspective treats numeral constructions as predicative expressions, defaulting to an existential type-shift in argument positions in the absence of an overt determiner. In contrast, this study argues that numeral constructions inherently function as arguments, with indefiniteness stemming from a cardinal head residing within their structure. The rationale behind this claim is rooted in Turkish numeral constructions, which freely allow definite and indefinite interpretations without *tane*, while numeral constructions with *tane* remain exclusively indefinite. Drawing from an agreement-based model of number marking variations in counting, I propose that Turkish features both a covert and an overt cardinal head, with *tane* spelling out overtly the default form with a built-in indefinite semantics. Notably, the covert form sets Turkish apart from the cross-linguistic pattern as it lacks this indefinite force and thus yields inherently predicative numeral constructions, which allow for definiteness via covert *iota* type-shifting, owing to Turkish being an articleless language. The analysis extends to two more optional classifier languages, Western Armenian and Farsi, providing evidence for the agreement-based approach and reinforcing the connection between cardinality and indefiniteness.

**Keywords** numeral constructions · (optional) classifiers · cardinality · (in)definiteness · number agreement

## 1 Introduction

It is a well-known fact that in every language, counting/numeral constructions (NCs, henceforth) can freely occur in argument positions, conveying indefinite interpretations. This is even the case in languages that strictly disallow bare nominal arguments, like French (Chierchia 1998). In French and other languages with articles such as English, NCs can also co-occur with the definite article and function as a definite description.

---

F. Author  
first address  
Tel.: +123-45-678910  
Fax: +123-45-678910  
E-mail: fauthor@example.com

S. Author  
second address

In articleless languages such as Mandarin and Russian, definiteness is not freely available for NCs, which typically require demonstratives instead to display a definite-like behavior (Jiang 2012). The inability of NCs to receive definite readings in these languages is particularly puzzling because bare nominals freely allow definite interpretations in articleless languages, which is standardly assumed to be achieved through covert *iota* type-shifting in the neo-Carlsonian approach.

A generally accepted view due to Link (1983) treats NCs as inherently predicative expressions, which allows them to be arguments of determiners and also occupy the predicate position. When they occupy an argument position without an accompanying determiner, NCs are assumed to undergo existential type-shifting by default, and hence the indefiniteness of NCs (e.g., Partee 1987, Link 1987, Verkuyl 1993, Carpenter 1998, Landman 2003, Ionin and Matushansky 2006, cf. Montague 1974, Bennett 1974, Barwise and Cooper 1981, Scha 1981, van der Does 1992, Dayal 2013, among others). However, the resistance of NCs to definiteness in articleless languages remains to be mysterious if NCs are inherently  $\langle e, t \rangle$  type expressions. In other words, it is not obvious why *iota* type-shifting is not equally available for NCs in these languages as it is available for bare nouns, despite that in languages like English and French, NCs are compatible with definiteness through the use of an overt definite article.

The key contribution of this study is to show that NCs are inherently argumental expressions. I argue that the source of indefiniteness in NCs cross-linguistically lies in the projection of a cardinal head that may surface covertly and/or overtly (cf. Jiang 2012), centering around an analysis of NCs in Turkish, an articleless optional classifier language. I propose that the cardinal head comes with a built-in choice function variable, adopting a theory of indefiniteness in the sense of Reinhart (1997). The predicative interpretation of NCs is not the default but derived as a repair operation only in structures requiring a predicative meaning (Dayal 2013). However, I further illustrate that inherently predicative NCs may coexist with default argumental NCs in a language featuring both overt and covert forms of the cardinal head, making definiteness possible via *iota* type-shifting if the language is articleless. Turkish as well as Farsi serve as exemplars of this phenomenon.

### 1.1 The Puzzle of Turkish Numeral Constructions

In languages like Mandarin, Cantonese, Japanese, etc., NCs obligatorily involve a classifier between a numeral and a noun regardless of whether the noun is (ontologically) categorized as count or mass. This is exemplified for Mandarin in (1) (Cheng and Sybesma 1999, pg. 514; see also Jiang 2012 and Kim 2009 for Mandarin, Japanese, and Korean, among others).

- (1) a. san \*(zhi) bi  
       three CL pen  
       ‘three pens’  
       b. san \*(ba) mi  
       three handful rice  
       ‘three handfuls of rice’

In many other languages like English, numerals directly combine with a count noun. However, counting with a mass noun still requires the existence of an intervening element, as in Mandarin-like languages:

- (2) a. three pens  
       b. three drops of water

There seems to be some sort of complementarity between the presence of systematic number marking in a language and the existence of an obligatory classifier in its NCs that involve a count noun (Sanches and Slobin 1973, Greenberg 1990).

English-like languages systematically distinguish between the singular and plural forms of the noun. Morphologically unmarked nouns like *book*, convey a singular interpretation, while plural-marked nouns like *book-s* convey a plural/number neutral interpretation (Sauerland et al 2005, Spector 2007, and Zweig 2009, among others.) This distinction is also reflected in NCs: The singular is used with the numeral ‘one’ while the plural is used with all other numerals. Mandarin-like languages, however, do not have a fine distinction between the singular and plural forms of the noun. Unmarked nouns consistently yield number neutral

readings, while plural marking is possible only under certain conditions. Accordingly, these languages use the unmarked form of the noun with all numerals but require the mediation of a classifier, as stated above.

And yet, a third group of languages, including Turkish, presents an interesting puzzle to the cross-linguistic semantics of counting because their NCs share features with these two types of languages. Although Turkish distinguishes between the unmarked and plural forms of nouns as English (e.g., *kitap* ‘book’ and *kitap-lar* ‘books’), the noun in Turkish NCs always appears in the unmarked form even with numerals other than ‘one’. Turkish NCs also feature an intervening word between the numeral and the noun, though only optionally. This word, i.e., *tane*, is known in the literature as a numeral classifier (Underhill 1976, Schroeder 1992, Lewis 2000, Göksel and Kerslake 2005, Öztürk 2005):<sup>1</sup>

- (3) a. bir (tane) kitap  
       one CL   book  
       ‘one book’  
       b. iki (tane) kitap(\*-lar)  
       two CL   book-PL  
       ‘two books’

Just as in the other groups of languages, Turkish mass nouns obligate a mediating element for counting, as shown in (4a). Numerals can directly combine with mass nouns only if some sort of covert universal packaging/sorting is at play, as in ‘two waters’ in English. Crucially, the co-occurrence of *tane* with mass nouns is also only possible if the mass noun is used in a count sense. This is illustrated in (4b), which refers to two units of water that comes in contextually determined containers.

- (4) a. iki \*(damla) su  
       two drop   water  
       ‘two drops of water’  
       b. iki (tane) su  
       two CL   water  
       ‘two waters’

The picture gets more complicated when the interpretation of Turkish NCs is considered. A difference between the forms with and without *tane* reveals that *tane* is not always a freely available option for NCs.

As stated above, NCs are freely indefinite in argument positions, though definiteness is only possible through overt marking; via the definite article in languages with articles, and via other means such as demonstratives in articleless languages. Turkish does not have an overt definite article, but only NCs with *tane* display the general restriction to indefiniteness. NCs without *tane* can be interpreted as both definite and indefinite, presenting an exceptional behavior within this cross-linguistic picture (Schroeder 1992, Öztürk 2005). I illustrate this in (5), delaying the detailed discussion of the empirical facts for now. While both forms of NCs can introduce new discourse referents as evidenced by their occurrence in the first sentence, only the form without *tane* can refer to a unique/maximal entity introduced in the preceding context, and hence the contrast in the second sentence. Let us call this the (in)definiteness puzzle of Turkish NCs.<sup>2</sup>

- (5) İçeri iki (tane) öğretmen, bir (tane) doktor ve üç (tane) mühendis gir-di.   **İki (#tane)**  
       inside two CL   teacher,   one CL   doctor and three CL   engineer enter-PAST two CL  
       **öğretmen** benim-le konuş-mak iste-di.  
       teacher   me-with speak-INF want-PAST  
       ‘Two teachers, one doctor, and three engineers entered inside. The two teachers wanted to talk to me.’

To recapitulate, there are four properties of Turkish NCs that make up our core puzzle:

<sup>1</sup> *tane* is compatible with all kinds of count nouns. There is another classifier in Turkish, i.e., *adet*, and it is only compatible with non-human count nouns. In this paper, I only discuss *tane* because the distribution of the two classifiers is the same, and *tane* is more commonly used.

<sup>2</sup> The most natural choice for referring to the two teachers in the second sentence is the plural *öğretmen-ler* ‘the teachers.’ However, the NC without *tane* is still a grammatical use, in contrast to the form with *tane*, which forces reference to two different teachers in the second sentence. See the appendix for additional information about such examples.

- i. The noun is unmarked for all numerals despite the existence of a systematic number marking system in the language.
- ii. An intervening word, i.e., *tane*, optionally surfaces between a numeral and a count noun.
- iii. NCs are generally indefinite when not accompanied by an overt determiner. In the absence of *tane*, however, both definite and indefinite interpretations are available.
- iv. Building on the previous point, *tane* may seem optional syntactically but clearly has a non-optional aspect with consequences regarding the interpretation of NCs.

The immediate challenge these properties introduce is internal to Turkish and the group of languages it belongs to, in general. We need to understand how the counting system works in Turkish compared to English-like and Mandarin-like languages and what role *tane* plays in this. A question of particular interest is how the current accounts of relatively better understood obligatory classifiers shape our understanding of the optional-looking classifier-like words such as *tane*, and vice versa. A further issue arising from this inquiry concerns the consequences for a more general cross-linguistic picture of NCs. What insights could we potentially gain from our investigation regarding number marking variations in NCs? Finally, central to the purposes of this paper, we need to understand how the presence/absence of *tane* affects the interpretation of NCs and how this contributes to the exceptional status of the Turkish counting system.

## 1.2 Overview of the Paper

Turkish NCs have received attention in several studies (e.g., Schroeder 1992, Öztürk 2005, Ionin and Matushansky 2006, 2019, Bale et al 2010, Bayırlı 2017, Sağ 2018, Alexiadou 2019, Martí 2020, Scontras 2014, 2022, Turgay 2022), but the four properties listed above have never been studied all together. This paper proposes a uniform analysis of these properties, taking *tane* as the focus of the investigation. The analysis is divided into two parts: The first part, targeting the properties in (i) and (ii), examines the optional occurrence of *tane* and variations in number marking patterns. The second part, concerning the properties in (iii) and (iv), takes up the (in)definiteness puzzle of Turkish NCs and its cross-linguistic implications.

I start the first part of the analysis by illustrating that *tane* is distinct from obligatory classifiers of Mandarin-like languages that have been argued to be some sort of repair mechanism for counting with nouns that are mass or mass-like in nature (Chierchia 1998 and Krifka 1989, 1995, 2003, among others). Building on Scontras (2014, 2022), which takes counting constructions to universally involve the projection of a cardinal head, I then propose that *tane* denotes a cardinality function by virtue of being an overtly realized form of this head in Turkish. While the English cardinal head is always covert, the Turkish cardinal head has both an overt and a covert realization. The proposal is linked to an account of the cross-linguistic variation in number marking in NCs. Building on Ionin and Matushansky's (2006, 2019) view of numerals, I argue that the cardinal head uniformly presupposes a semantically singular form of the noun. This requirement is fulfilled by morphologically unmarked nouns in languages like Turkish, while English NCs further reflect number agreement on the lexical NP. I also provide a semantic account of this agreement mechanism, drawing on the analyses proposed in Sauerland (2003) and Scontras (2014, 2022).

The second part of the analysis offers an account that will enrich our understanding of the indefinite characteristics of NCs cross-linguistically. In essence, I associate the universal restriction to indefiniteness with the cardinal head, which, I propose, involves a built-in choice function variable in its denotation (cf. Jiang 2012). I further propose that in languages with both an overt and a covert cardinal head, one form may be liberated from the choice function variable, resulting in predicative NCs, making both definite and indefinite interpretations possible. Turkish, being one such language, features NCs with the overt cardinal head as the indefinite form, whereas NCs with the covert variant realize the predicative form.

While the discussion revolves around Turkish NCs substantially, it also extends to two more optional classifier languages, Western Armenian and Farsi, for which I provide an analysis analogous to Turkish. These languages, where plural marking in NCs is constrained by factors relating to specificity and definiteness, form the central motivation behind an agreement-based approach to variation in number marking in NCs (Ionin and Matushansky 2019 and Alexiadou 2019, cf. Sigler 1996, Borer 2005, Bale and Khanjian 2008, 2014, Khanjian 2013, Martí 2020, Kalomoiros 2021 for WA and Ghomeshi 2003, Gebhardt 2009, and Mache 2012 for Farsi). The analysis of Farsi is further illuminating because Farsi NCs display the mirror image of Turkish NCs in their interpretation. That is, it is the covert cardinal head, i.e., NCs without the classifier, that reflects

exclusively indefinite characteristics in Farsi, while the overt form freely allows both definite and indefinite readings. This finding reinforces the connection between the indefiniteness of NCs and the cardinal head, affirming the validity of our analysis regarding the interpretation of Turkish NCs. Furthermore, the analysis predicts that the effects of featuring both argumental and predicative NCs only become evident in articleless languages. NCs in Western Armenian, a language with a definite article, validate this prediction.

At this point, a note on the choice of terminology is imperative: The term ‘classifier’ does not have a consistent use in the literature. Some take it to refer to obligatory classifiers in Mandarin and languages alike, some use it as a general term for all ‘quantizing’ words of NCs and measurement. Here, I adopt a descriptive use of the term ‘classifier’, i.e., an intervening element between a numeral and an (ontological) count noun in NCs. Following the convention in the literature then, I will continue calling *tane* a classifier, although it will be analyzed differently from the classifiers of Mandarin-like languages.

The outline of this paper is as follows: Section 2 reviews an existing account of Turkish number marking semantics, which will lay the foundation of the account developed here. Section 3 compares *tane* with obligatory classifiers. Section 4 is dedicated to the analysis, where I first present my account of *tane* and adopt an initial semantics of number marking in NCs and then take up the (in)definiteness puzzle and provide the core of the analysis. Section 5 discusses Western Armenian and Farsi data. Section 6 revises the agreement analysis adopted earlier. Section 7 concludes.<sup>3</sup>

## 2 Turkish Number Marking Semantics

The first step of solving our puzzles is to understand the semantics of Turkish nominals, as it will inform us about the denotation of the noun that a numeral and *tane* combine with. In this section, which serves as theoretical backdrop for the analysis, I summarize Sağ’s (2019, 2022) account of Turkish number marking system. We will see that morphologically unmarked nouns are strictly singular and plural-marked nouns are number neutral, a distinction that also has consequences for kind reference (see also Sağ 2018, Renans et al 2017, 2020).

English is one of many other languages where there is an asymmetry between morphological and semantic (un)markedness. The standard view is that while morphologically unmarked nouns are semantically marked as strictly singular, plurals have an unmarked denotation (Krifka 2003, Sauerland et al 2005, Spector 2007, and Zweig 2009). Sağ (2019, 2022) argues that Turkish is not different from English in its nominal semantics. Unmarked nouns like *kitap* ‘book’ denote an atomic set while plural-marked nouns like *kitap-lar* ‘book-s’ denote a number neutral set, inclusive of atomic and plural entities:

- (6) a.  $\llbracket \textit{kitap} \rrbracket = \{a, b, c\}$   
 b.  $\llbracket \textit{kitap} + PL \rrbracket = \{a, b, c, a \oplus b, a \oplus c, b \oplus c, a \oplus b \oplus c\}$

The evidence that Turkish plurals are number neutral comes from their behavior in downward entailing contexts and questions. As in English, Turkish plurals have a ‘one or more’ reading in these contexts even though they have a multiplicity interpretation in positive contexts. For example, for the condition in (7) to hold, it is enough if one is cheated by one man. Following the analyses of English plurals in Sauerland et al (2005), Spector (2007), and Zweig (2009), Sağ argues that Turkish plurals are number neutral and the multiplicity reading is a conversational implicature. Renans et al (2017, 2020) provide experimental evidence for this view.

- (7) Eđer **erkek-ler** tarafından aldatıldıysan, sen de biz-e katıl-abil-ir-sin.  
 if man-PL by you.be.cheated you also we-DAT join-ABIL-AOR-2SG  
 ‘If you have been cheated by men, you can join us.’ (one or more men)

<sup>3</sup> The Turkish data reflect the judgments of fifteen native speakers, including myself. The Western Armenian data discussed in this paper represents the variety spoken in Beirut, corroborated by Hossep Dolatian. For the Farsi data, ten native speakers have been consulted via informal conversations, including Amir Anvari and Masoud Jasbi. All the data was collected following the techniques reported in Mahowald et al (2016). The examples of all the other languages are sourced from the literature.

The picture is more complicated on the side of morphologically unmarked nouns. English unmarked nouns are identified as singular terms since they yield a singular interpretation consistently.<sup>4</sup> As shown in (8a) and (8b), Turkish unmarked nouns convey number neutrality in certain cases, but as shown in (8c), they are interpreted as strictly singular and definite in case-marked argument positions.

- (8) a. Ali **kitap** oku-du.  
Ali book read-PAST  
'Ali read one or more books.'
- b. Ali ve Merve **çocuk**.  
Ali and Merve child  
'Ali and Merve are children.'
- c. Ali **kitab-ı** oku-du.  
Ali book-ACC read-PAST  
'Ali read the book.'

Although this dual nature of unmarked nouns makes it hard for them to be identified as number neutral or singular terms, there is evidence for the strict singular view.<sup>5</sup> Let us start with (8a), where the unmarked noun occupies the non-case-marked argument position. This is an instance of a well-known phenomenon, i.e., pseudo-incorporation (due to Massam 2001), in Turkish (Öztürk 2005). Sağ shows that only adjectives that denote classificatory properties are compatible with the number neutral interpretation of incorporated nouns while modification does not yield a contrast in case-marked argument positions. As shown below, the incorporated noun *book* can be modified with *religious*, yielding a 'one or more books' reading. However, its modification with *old* meaning *worn-out*, as opposed to *ancient/historical*, is ungrammatical, which instead requires the indefinite or plural form of the noun.

- (9) a. Ali, ev-e geldikten sonra, *dini*/ \**eski* **kitap** oku-du.  
Ali home-DAT having.come after religious old book read-PAST  
'After he came home, Ali read one or more religious books.'  
Not: 'After he came home, Ali read one or more old (worn-out) books.'

A similar case arises in the predicate position, except that modification introduces a contrast in number interpretation. When the unmarked noun *doktor* is modified by the adjective *practitioner*, it is compatible with both a singular and plural subject. If the adjective is *handsome*, though, it is only compatible with a singular subject (requiring the plural marker, i.e., *doktor-lar*, with a plural subject):

- (10) a. Ali (ve Mehmet) *pratisyen* **doktor**.  
Ali and Mehmet practitioner doctor  
'Ali is a practitioner doctor.' 'Ali and Mehmet are practitioner doctors.'
- b. Ali (\*ve Mehmet) *yakışıklı* **doktor**.  
Ali and Mehmet handsome doctor  
'Ali is a handsome doctor.' Not: 'Ali and Mehmet are handsome doctors.'

The adjectives that are compatible with the number neutral reading of unmarked nouns define a type of the noun that they modify; religious books are types of books, and practitioner doctors are types of doctors. In contrast, the other set of adjectives does not have such a function out of the blue; both *worn-out* and *handsome* define some physical properties of books and doctors, respectively.

Sağ (2022) explains the puzzling behavior of unmarked nouns following Dayal's (2004) analysis of English definite singular kind terms like *the dinosaur* in 'The dinosaur is extinct.' English and Turkish unmarked nouns are ambiguous in denoting atomic properties of ordinary individuals and atomic properties of taxo-

<sup>4</sup> There are cases where English unmarked nouns do not receive a strictly singular reading, such as in compounds, e.g., *book-shopping* and with weak definites, *John reads the newspaper*. However, such cases have a wider distribution in Turkish compared to English.

<sup>5</sup> Bliss (2004), Bale et al (2010), and Görgülü (2012) argue that Turkish unmarked nouns denote a number neutral set, while plural nouns denote pluralities only, exclusive of atoms. While the data reviewed in this section provides evidence against this view, I refer the reader to Sağ (2022) for a more in-depth discussion. Furthermore, unmarked nouns also yield number neutrality in existential statements (Görgülü 2012). The reader can also find the discussion of this case in Sağ (2022), where the analysis is similar to those occupying the non-case-marked object position to be explained below.

nomic individuals, i.e., (sub-)kinds. English unmarked nouns can either be definite singulars at the ordinary object level or definite singular kind terms by their combination with *the*. A widely accepted view for languages without overt definite determiners such as Turkish is that their nouns can undergo covert *iota* type-shifting for definiteness.<sup>6</sup> In object-level contexts, as in (8c), then the unmarked noun *kitap* ‘book’ denotes an atomic set of ordinary book individuals, and can refer to a contextually salient unique book via *iota* type-shifting. This explains its strictly singular and definite interpretation.

- (11) a.  $\llbracket \textit{kitap} \rrbracket = \lambda x. \textit{book}(x) = \{a\}$   
 b.  $\iota: \lambda P. \iota x. P(x)$   
 c.  $\iota(\llbracket \textit{kitap} \rrbracket) = \iota x. \textit{book}(x) = a$   
 d.  $\llbracket (8c) \rrbracket = \textit{read}(\textit{Ali}, \iota x. \textit{book}(x))$

Now let us consider the kind-level characteristics of Turkish nouns to understand the cases in (8a) and (8b). In Turkish, both unmarked and plural nouns can be used in kind-level statements, as shown in (12a). However, only plurals can combine with distributive predicates applying to individual members of the species, such as *come from different regions*, as illustrated in (12b). The same contrast also holds for English as represented in the translations. This indicates that the kind reference achieved by unmarked nouns differs from the one achieved by plurals although kinds, in general, are inherently plural entities in that they are associated with atomic and plural object-level entities (Carlson 1977).

- (12) a. **Dinozor(-lar)** 250 milyon yıl önce evrimleşmiş-tir.  
 dinosaur-PL 250 million year ago evolve-PERF-GEN  
 ‘The dinosaur/Dinosaurs evolved 250 million years ago.’  
 b. **Ayı\*(-lar)** bu hayvanat bahçesin-e farklı bölge-ler-den gel-di.  
 bear-PL this zoo-DAT different region-PL-ABL come-PAST  
 ‘Bears/\*The bear came to this zoo from different regions.’

Sağ analyzes Turkish plurals as kind terms via the *nom* operator ( $\cap$ ), as claimed for English plurals by Chierchia (1998). *Nom* is a function from properties to functions from situations  $s$  to the maximal entity satisfying that property in that situation (Chierchia 1998, pg. 351). Based on this view, the plural kind term *dinozorlar* ‘dinosaurs’ in (12a) is interpreted as below:

- (13) a. For any property  $P$  and world/situation  $s$ , where  $P_s$  is the extension of  $P$  in  $s$   

$$\cap P = \begin{cases} \lambda s. \iota x. P_s(x), & \text{if } \lambda s. \iota x. P_s(x) \text{ is in } K, \text{ the set of kinds} \\ \text{undefined,} & \text{otherwise} \end{cases}$$
  
 b.  $\llbracket (12a) \text{ with plural} \rrbracket = \textit{evolved}(\lambda s. \iota x. \textit{dinosaur}_s(x))$

Plural kind terms can be type-shifted to sets of object-level entities that instantiate the kind via *pred* ( $\cup$ ). This operator applies to the extension of the kind (i.e., extension in whatever world/situation it is interpreted relative to) and returns the set of singular and plural instantiations of the kind (in that world/situation) (Chierchia 1998, pg. 350):

- (14) Let  $d$  be a kind. Then for any world/situation  $s$ , where  $d_s$  is the plural individual that comprises all of the atomic members of the kind  

$$\cup d = \begin{cases} \lambda x. x \leq d_s, & \text{if } d_s \text{ is defined} \\ \lambda x. \textit{FALSE}, & \text{otherwise} \end{cases}$$

In episodic contexts as in (12b), Derived Kind Predication (DKP) applies, which provides sort-adjustment introducing  $\exists$ -quantification over the instantiations of the kind provided by *pred* in a given situation. This

<sup>6</sup> Covert *iota* type-shifting is blocked by *the* in English due to the Blocking Principle which ensures the use of lexical items before covert type-shifting operations are resorted to (Chierchia 1998). Furthermore, type-shifting operators are assumed to be regulated by Revised Meaning Preservation, which bans  $\exists$ , and hence strong indefinite interpretation of bare nouns, in both English and Turkish (Dayal 2004). We revisit these principles in Section 4.3.1.

results in an existential reading.<sup>7</sup> The ability of plural kind terms to be type-shifted to sets of object-level entities makes them compatible with distributive predicates which require access to different parts of these entities. The plural version of (12b) means that some bear individuals that instantiate the bear kind in the relevant situation came to this zoo and the regions that these individuals came from are different, as demonstrated below:

$$(15) \quad \llbracket (12b) \text{ with plural} \rrbracket = \exists x [\cup^{\cap} \text{bear}(x) \wedge \forall y, z [[y < x \wedge z < x \wedge y \neq z] \rightarrow \iota r_1 [\text{region}(r_1) \wedge \text{came.to.zoo.from}(r_1)(y)] \neq \iota r_2 [\text{region}(r_2) \wedge \text{came.to.zoo.from}(r_2)(z)]]]]$$

Unmarked nouns in (12) denote a singleton set whose member is a taxonomic individual, i.e., a (sub-)kind. Undergoing covert *iota* type-shifting, they denote a singular kind term and refer to a unique kind. Based on this, the interpretation of (12a) with the unmarked noun is as shown below: (Taxonomic entities and properties of taxonomic entities are represented by the superscript K.)

$$(16) \quad \begin{array}{l} \text{a. } \llbracket \text{dinozor} \rrbracket = \lambda x^K. \text{dinosaur}^K(x^K) \\ \text{b. } \iota(\llbracket \text{dinozor} \rrbracket) = \iota x^K. \text{dinosaur}^K(x^K) \\ \text{c. } \llbracket (12a) \text{ with unmarked} \rrbracket = \text{evolved}(\iota x^K. \text{dinosaur}^K(x^K)) \end{array}$$

Dayal (2004) claims that singular kind terms are grammatically atomic entities, similar to group terms like *team*. They do not allow type-shifting to sets of object-level entities we intuitively associate with kinds. Hence, the derivation fails when they combine with a distributive predicate, as in (12b). Furthermore, singular kind terms do not yield an existential reading, unlike plural kind terms, as evidenced in (8c), where the unmarked noun only has a singular and definite reading. This also follows from that type-shifting to object-level entities is not available for singular kind terms. The kind-driven existential reading depends on this shift, which is ensured for plural kind terms by *pred* when DKP applies. Dayal argues that singular kinds still hold a relation to object-level entities at the conceptual level despite their atomic nature at the grammatical level, analogous to the membership relation that exists between an atomic group term and its members. Sağ calls this relation *belong-to*, as defined below:

$$(17) \quad \begin{array}{l} \text{Belong-to relation} \\ \text{belong-to}(y, x^K) \text{ is true iff } y \text{ is a member of the kind } x^K, \text{ where } x^K \text{ is a singular kind and } y \text{ is an} \\ \text{object-level entity.} \end{array} \quad (\text{Sağ 2022, pg. 764})$$

As stated above, non-case-marked unmarked nouns have been analyzed as pseudo-incorporated arguments in the Turkish literature. Sağ argues that incorporation takes place with an incorporating thematic function that establishes a *belong-to* relation between singular kinds and their object-level members, resulting in number neutrality.<sup>8</sup> The restriction in modification follows from the fact that taxonomic kinds can only be modified by adjectives that denote a classificatory/sub-kind forming property, like *religious*, as in (9a).<sup>9</sup> Based on this, the denotation of *Ali kitap okudu*. ‘Ali did book-reading.’ in (8a) is shown in (18). It means that Ali is involved in a reading event with a theme argument that belongs to the book kind.

$$(18) \quad \exists e \exists y [\text{belong-to}(y, \iota x^K. \text{book}^K(x^K)) \wedge \text{read}(e) \wedge \text{Th}(e) = y \wedge \text{Ag}(e) = \text{Ali}]$$

The modificational contrast arising in the predicate position with respect to the number interpretation of unmarked nouns also follows from their ambiguous nature. In (10a), the noun *doktor* denotes an atomic property at the ordinary object level. Therefore, it can only be modified by adjectives that describe an object-level property like *handsome* and be predicated of only singular subjects. On the other hand, *doktor* in (10b) denotes the doctor kind, and hence it is only compatible with taxonomic adjectives like *practitioner*.

<sup>7</sup> Plurals take obligatory narrow scope in their existential reading. This is ensured by DKP because the sort-adjusting  $\exists$ -quantification is introduced locally at the level of predication, and therefore takes the narrowest scope. Turkish plurals can also have a definite reading in episodic contexts, which is possible through *iota*.

<sup>8</sup> See Martí (2020) and Sağ (2018) for other accounts that dedicate the number neutrality of unmarked nouns to the semantics of incorporation.

<sup>9</sup> In Sağ’s view, what counts as a taxonomic property depends on the context and the verb. While *eski* ‘worn-out’ is an object-level property in a reading context, it may be considered a classificatory property in a selling/buying context, e.g., in the context of a second-handbook store. This is known to be the *name-worthiness* requirement of pseudo-incorporation.



Sağ proposes that singular kind reference in the predicate position is made possible through a null copula (which is overt when followed by overt tense/aspect markers (Kornfilt 1996, Kelepir 2003)). The copula establishes a *belong-to* relation between the referent of a singular kind term and the referent of a singular or plural subject term, similar to what happens in pseudo-incorporation (see also de Swart et al 2007). This phenomenon, called *kind specification*, will be crucial when we compare *tane* with obligatory classifiers in Section 3.2. Thus, for an example, the denotation of *Ali (ve Merve) çocuk*. ‘Ali is a child./Ali and Merve are children.’ in (8b) is provided in (19). (19b) and (19c) can be paraphrased as ‘Ali is a member of the child kind’ and ‘Ali and Merve are members of the child kind’, respectively.

- (19) a.  $\llbracket \text{COP} \rrbracket = \lambda x^K \lambda y. \text{belong-to}(y, x^K)$   
 b.  $\llbracket \text{Ali child} \rrbracket = \text{belong-to}(\text{Ali}, \iota x^K. \text{child}^K(x^K))$   
 c.  $\llbracket \text{Ali and Merve child} \rrbracket = \text{belong-to}(\text{Ali} \oplus \text{Merve}, \iota x^K. \text{child}^K(x^K))$

To wrap up, Turkish and English nominals are similar at the ordinary object and kind-level domains. The fact that singular kind reference extends to pseudo-incorporation and the predicate position in Turkish creates the illusion that Turkish and English nominal semantics are fundamentally different when, in fact, they only vary in distribution.

Having established our take on the semantics of Turkish nominals, we are now ready for a theoretical inspection of *tane*, compared to obligatory classifiers.

### 3 Comparing *tane* with Obligatory Classifiers

A notable distinction between English and obligatory classifier languages is the fact that while only mass nouns are incapable of directly combining with a numeral in the former (excluding ‘packaging/sorting’ coercions), all nouns, including those that we intuitively categorize as count, require the mediation of an intervening item in the latter. This disparity has led scholars to hypothesize that all nouns are mass or mass-like kind terms in Mandarin-like languages, at least as far as a grammatical level of distinction is concerned. As a consequence, classifiers have been claimed to exist as a repair mechanism in NCs (e.g., Krifka 1989, 1995, 2003 and Chierchia 1998, cf. Cheng and Sybesma 1999). Central to these views is the obligatory emergence of classifiers with counting. That is, obligatoriness points to a level of noun denotation that requires some sort of fixing to make counting possible with it.

The Turkish classifier is similar to classifiers in Mandarin-like languages in that it appears between a numeral and an (ontological) count noun. But it fundamentally differs from these classifiers in being optional, which, at first sight, suggests that it is not essential and therefore, should belong to a separate category. However, being optional is not necessarily a distinguishing factor. As sketched out above, Turkish nominals are rich in interpretation, having ordinary object and kind-level interpretations, each displaying variations based on number marking. As will be discussed below, Turkish also displays countability distinctions, and *tane* is sensitive to them. Hence, it is necessary to assess to what extent this diversity plays a role in the optionality of the classifier. The question is whether the noun can be in need of a repair mechanism when it combines with *tane*, unlike what happens when it surfaces in the absence of *tane*.

In this section, maintaining the spirit of the analyses offered for obligatory classifiers, I will provide a negative answer to this question and illustrate that the noun takes part as an atomic property of ordinary individuals in Turkish NCs irrespective of the classifier. In other words, I will confirm the intuition that *tane* is an instance of a different category.

#### 3.1 Obligatory Classifiers

Mandarin, Japanese, and Korean are articleless languages and hence their bare nouns can freely be arguments, as in Turkish (e.g., Krifka 1995, Chierchia 1998, Kurafuji 1999, Li 1999, Yang 2001, Jiang 2012, and Kim 2009). However, unlike in Turkish, their unmarked nouns consistently yield number neutral readings while the plural marker surfaces in a limited way. The Mandarin plural marker *-men* (see Yang 2001) and the Japanese plural marker *-tachi* (see Kurafuji 1999) include definiteness in their denotation, whereas the Korean plural marker *-tul* denotes specificity (Kim 2009). Consider the following contrast in Mandarin (Li 1999, pg. 78):

- (20) wo qu zhao haizi(-men).  
 I go find child-PL  
*w/o plural*: ‘I will go find a child/children/the child/the children.’  
*w/plural*: ‘I will go find the children.’

Unmarked nouns can receive kind-level, existential, and definite readings in these languages, as exemplified for Mandarin in (21a) and (21b) (Yang 2001, pg. 20, 32). They are also compatible with distributive predicates like *come from different regions*, as illustrated in (21c) (p.c. with Yi-Hsun Chen).

- (21) a. Gou juezhong le  
 dog extinct ASP  
 ‘Dogs are extinct.’  
 b. Waimian gou zai-jiao  
 outside dog be-barking  
 ‘Outside, (dogs)/(the) dog(s) are/is barking.’  
 c. Xiong cong butong de quyu laidao-le zhe-zuo-dongwuyuan  
 bear from different MOD region come-ASP this-CL-zoo  
 ‘Bears came to this zoo from different regions.’

These facts show that Mandarin unmarked nouns are similar to Turkish and English bare plurals. They are number neutral and their kind-level interpretations pattern with plural kind reference (Yang 2001). The generally accepted view due to Chierchia (1998) is that unmarked nouns of Mandarin-like languages are uniformly kind terms of type  $\langle s, e \rangle$ , which can be type-shifted to sets of object-level entities via *pred*. The availability of type-shifting to sets of object-level entities also renders them compatible with distributive predicates.

NCs in Mandarin-like languages require a mediator between the noun and the numeral because kinds are not the right type for counting purposes under the view that counting operates on the predicative meanings of nouns. In Chierchia’s account, kinds are mass-like in that they are inherently plural, and their atomic instances are not accessible for counting. Classifiers play a role of an *atomizer* function in returning the atomic instances of the kind to fix the type mismatch. Based on this, the combination of the classifier and the noun in *san zhi bi* ‘three CL pen’ in (1) can be represented as below.

- (22) a.  $\llbracket zhi \rrbracket = \lambda k \lambda x [\cup k(x) \wedge AT(x)]$   
 b.  $\llbracket bi \rrbracket = pen^k$   
 c.  $\llbracket zhi bi \rrbracket = \lambda x [\cup pen^k(x) \wedge AT(x)]$

Chierchia’s view is similar to the account provided in Krifka (1989, 1995, 2003) in requiring classifiers for access to object-level instances of kinds to make counting possible with them. However, in Krifka’s account, classifiers have a dual role of instantiation and counting; they take a kind and *measure* the number of specimens instantiating that kind. English differs from Mandarin in having a count-mass distinction, but its mass nouns are similar to Mandarin nouns and hence require a classifier. (Krifka uses the term classifier to include all quantizing nouns.) Count nouns, however, are different and directly combine with a numeral.

To capture the count-mass distinction in English, Krifka (2003) proposes that count nouns have a built-in classifier, an idea first discussed in Krifka (1989, 1995).<sup>10</sup> A count noun entering the derivation as a kind term is first shifted to an object-level denotation by a null operator. In virtue of this operator, a count noun has a number argument and denotes an extensive measure function. Based on this, the derivation of ‘three dogs’ is as shown below. *R* represents the instantiation relation holding between a kind and its object-level instantiations, *OU* corresponds to the measurement of the number of atoms or ‘an object unit’ of an entity:<sup>11</sup>

<sup>10</sup> An alternative account that is discussed and adopted in Krifka (1995) is that numerals have a built-in classifier. Due to the similarity between the two proposals, I only discuss the view adopted in Krifka (2003).

<sup>11</sup> In Krifka’s account, singular kinds participate in English NCs when the noun is count, not plural kinds. The plural marking in NCs is an instance of syntactic agreement with the number word, which does not exist in languages like Turkish and Hungarian. In the absence of a number word, the plural marker is semantically contentful and creates a property leaving the number argument unspecified. Note also that Krifka analyzes singular kinds as names of kinds, where the definite article, when present, takes the role of the identity function. That is, both *dog* and *the dog* refer to the dog kind in English. Furthermore, *R* does not distinguish between singular and plural kinds. As discussed in Section 2, we follow Dayal’s (2004) view, instead. We

- (23) a.  $\llbracket dog_{kind} \rrbracket = dog^k$   
 b. count operator:  $\lambda y^k \lambda n \lambda w \lambda x [R_w(x, y^k) \wedge OU_w(y)(x) = n]$   
 c.  $\llbracket dog_{count} \rrbracket = \lambda n \lambda w \lambda x [R_w(x, dog^k) \wedge OU_w(dog^k)(x) = n]$   
 d.  $\llbracket three\ dogs \rrbracket = \lambda w \lambda x [R_w(x, dog^k) \wedge OU_w(dog^k)(x) = 3]$

Kind terms denoted by mass nouns do not involve this inherent shifting mechanism. Hence, mass nouns lack a number argument and rely on measure terms for quantization, as *gallon* in *three gallons of milk*, illustrated in (24a). Mandarin-like languages do not have count nouns and the classifier does this job for all nouns, as exemplified for *ben* in *san ben shu* ‘three CL book’ in (24b).

- (24) a.  $\llbracket gallon \rrbracket = \lambda y^k \lambda n \lambda w \lambda x [R_w(x, y^k) \wedge GALLON_w(y^k)(x) = n]$   
 b.  $\llbracket ben \rrbracket = \lambda y^k \lambda n \lambda w \lambda x [R_w(x, y^k) \wedge OU_w(y^k)(x) = n]$

To sum up, in languages like Mandarin, classifiers are essential elements of counting because all nouns, being kinds inherently, have a mass or mass-like nature, which makes them ineligible for direct combination with numerals.<sup>12</sup> I will now show that *tane* does not fit with the proposed analyses of obligatory classifiers. We will see that *tane* does not have a repairing role in counting with kind terms or mass nouns.

### 3.2 *tane* and Kind Terms

As discussed in Section 2, Sağ (2022) claims that Turkish aligns with English in its nominal semantics, and both plurals and unmarked nouns have kind reference. Following Chierchia (1998), plural kind terms are derived by *nom* and can be type-shifted to sets of object-level instances via *pred*. Following Dayal (2004), unmarked nouns are ambiguous between atomic properties of object-level and taxonomic individuals, and their taxonomic denotation yields singular kind reference via covert *iota* type-shifting.

Given that Turkish NCs require the unmarked form of the noun, we expect two possibilities for the noun denotation: it can either be a singular kind or an atomic property of object-level individuals. If the former, we need a transformation from the kind-level denotation into an object-level one to proceed with counting. This cannot be achieved through a freely available covert type-shifting operator (like *pred*). As we have seen, singular kind terms are grammatically atomic and hence disallow type-shifting to sets of object-level instances, making certain interpretations unavailable for them unlike the case with plural kind terms (e.g., see (12b)). For this reason, we expect the noun in Turkish NCs to denote atomic properties of object-level individuals, at least in NCs without *tane*.

Could it still be possible for the noun to denote a singular kind in NCs with *tane*? An analysis to consider could be in lines with Krifka (1989, 1995, 2003) where count nouns have a built-in classifier, one function of which involves the establishment of a relation between a kind and its object-level instances. To see whether this role could potentially be attributed to *tane*, let us now return to the phenomenon of *kind specification* we have seen in Section 2.

In kind specification, a *belong-to* relation is established by the (null) copula between a singular kind and its object-level members in the predicate position. We have determined the denotation of unmarked nouns in the predicate position based on whether they undergo taxonomic or object-level modification. When subjected to object-level modification, they function as atomic predicates; when subjected to taxonomic modification, they serve as singular kind terms, as repeated below:

---

have seen that singular and plural kind reference have different distributions and hence differ in the nature of the instantiation operator/relation applicable to them.

<sup>12</sup> There are also other studies that do not take the role of kind reference as a key factor for the semantics of classifiers in NCs. But in these accounts, classifiers are also given some function that would ‘fix’ the denotation of the noun otherwise unsuitable for counting. Borer (2005), as an example, argues that “...all nouns, in all languages, are mass, and are in need of being portioned out, in some sense, before they can interact with the ‘count’ system.” (pg. 93). In her view, mass nouns represent the absence of a dividing structure, and in English-like languages, plurals, and in Mandarin-like languages, classifiers are ‘stuff dividers.’ The evidence for this view is provided based on Armenian, reported to be a language where a classifier and the plural marker can optionally occur in NCs, but never surface in the same structure. Borer takes this as the indication of identity between plural markers and classifiers. My stance against this view draws on data from Western Armenian spoken in Beirut and Farsi. As discussed in Section 4.1, these languages allow the classifier and the plural marker to co-occur. See also Doetjes (2019) for a discussion of other such languages.

- (25) a. Ali (\*ve Mehmet) *yakışıklı* doktor.  
 ‘Ali is a handsome doctor./\*Ali and Mehmet are handsome doctors.’  
 b. Ali (ve Mehmet) *pratisyen* doktor.  
 ‘Ali is a practitioner doctor./Ali and Mehmet are practitioner doctors.’

If kind specification is also possible with NCs, *tane* could be establishing a *belong-to* relation between a singular kind and object-level entities associated with it, as illustrated in (26), similar to what the copula does. If so, we expect the same modificational restrictions to surface in NCs with *tane*, but this is not attested. Unmarked nouns in NCs can receive object-level modification regardless of the classifier, as shown in (27). So, we can also conclude that NCs do not involve kind specification with singular kind terms.<sup>13</sup>

- (26)  $[[tane]] = \lambda x^K \lambda y. \textit{belong-to}(y, x^K)$   
 (27) Sevgi iki (tane) *yakışıklı* doktor-a mesaj at-tı.  
 Sevgi two CL handsome doctor-DAT text send-PAST  
 ‘Sevgi texted two handsome doctors.’

Considering the problem of covert type-shifting with singular kind terms and the lack of kind specification in NCs, *tane* could not be given an atomizer semantics in the sense of Chierchia (1998). This is because an atomizer demands access to object-level entities to extract atoms out of them. A semantics of this sort could be possible if *tane* combined with plural kind terms, because plural kinds allow type-shifting to sets of object-level instantiations (via *pred*). However, neither *tane* nor *numerals* take the plural form of the noun.

For these reasons, we do not expect singular kinds to be an option for the noun denotation in Turkish NCs in the presence or absence of *tane*.

### 3.3 *tane* and Mass Nouns

We have seen that *tane* does not have a type-fixing function on kind terms. We can then conclude that unmarked nouns participate in NCs in their ordinary object-level denotation. The next step is to see whether *tane* might have anything to do with the counting of mass nouns.

So far, we have only focused on the interpretation of ontological count nouns in Turkish. However, it is crucial to see where Turkish stands regarding the count-mass distinction compared to English and Mandarin. Whereas English grammatically distinguishes between the count and mass sense of the noun, Mandarin-like languages are known to reflect this distinction only with the choice of the classifier, not at the level of the noun (Cheng and Sybesma 1999).

I have shown in Section 1.1 that Turkish numerals cannot directly combine with a mass noun like *water* unless the mass noun is coerced into a count interpretation through covert universal packaging/sorting (see (3) and (4)). This is similar to ‘two waters’ in English. We have seen that coercion is also necessary for *tane* to co-occur with mass nouns.<sup>14</sup> The contrast between count and mass nouns in their ability to directly combine

<sup>13</sup> The order where the adjective precedes the NC, i.e., *yakışıklı iki tane doktor*, is also possible. As pointed out by a reviewer, this alternative order could be the underlying structure in which case *tane* would combine with the noun first. However, there are reasons to take the order in (27) to be the default if one is to be derived from the other. First, the adjective + NC order is marked, where the adjective is arguably topicalized within the nominal projection (Guisti 1996). The second reason comes from the numeral *bir* ‘one’, which can take on either a numerical reading or function as an indefinite article depending on its position relative to an adjective (Tat 2011). As demonstrated below, *bir* must precede the adjective to convey numerical information, suggesting that this order serves as the underlying structure for the combination of adjectives with NCs.

- (i) a. *yakışıklı bir doktor* ‘a handsome doctor’  
 b. *bir yakışıklı doktor* ‘one handsome doctor’

<sup>14</sup> Mass nouns are generally known to be capable of directly combining with a numeral if a sub-type interpretation is at issue. The famous example discussed for English is *two bloods* to mean two types of blood. The Turkish counterpart requires the intervention of *tür* ‘type/kind’ for this interpretation, i.e., *iki #(tür) kan* ‘two type blood’. However, as shown in (i), numerals and the numeral+*tane* combination are compatible with a sub-kind interpretation in general.

- (i) Bu muazzam tür-ler arasında, malesef iki (tane) kuş yok olma tehlikesi-yle karşı karşıya.  
 this astonishing kind-PL among unfortunately two CL bird extinct become danger-with facing  
 ‘Among these astonishing kinds, unfortunately, two birds are facing extinction.’

with numerals and the numeral+*tane* combination already points to a more English-like characteristic of Turkish nominals. It also shows that *tane* is selective for the count sense of the noun just like numerals. However, before reaching this conclusion, we need more evidence to confirm the status of the count-mass distinction in Turkish. Furthermore, there is an analysis we need to take into consideration.

Rothstein (2017) and Schvarcz and Rothstein (2017) claim that ontological count nouns in Brazilian Portuguese and Hungarian are flexible in their grammatical representation, analogous to *stone/stones* pair in English (see also Pires de Oliveira and Rothstein 2011 for Brazilian Portuguese). The mass denotation of ontological count nouns is identified as *furniture*-type mass nouns, which denote sets of naturally individuable units. This claim partly relies on the fact that in Brazilian Portuguese and Hungarian, most nouns are compatible with a comparative evaluation in terms of both cardinality and a non-cardinal measure dimension such as volume, just like *furniture*-type nouns in these languages. Schvarcz and Rothstein further argue that Hungarian, being an optional classifier language, reflects this ambiguity in its NCs. In its count use, a singular noun can directly combine with a numeral, but it needs an intervening classifier in its mass use.

If a similar case exists in Turkish, we can point to a difference in the type of the noun in NCs with and without *tane*. This difference can be captured by Krifka’s analysis if we assume that a count noun inherently involves a number argument, rather than obtaining it through a null operator that transforms a kind term into an object-level one. The hypothesis is then as follows: In NCs without *tane*, the noun is count, with a denotation as in (28a), and therefore, it can directly combine with a numeral. In NCs with *tane*, the noun is mass, with a denotation as in (29a), and that is why a classifier is needed for counting.<sup>15</sup>

- (28) *üç köpek* ‘three dog’:
- a.  $\llbracket dog_{count} \rrbracket = \lambda n \lambda w \lambda x [OU_w(dog)(x) = n]$
  - b.  $\llbracket three\ dog \rrbracket = \lambda w \lambda x [OU_w(dog)(x) = 3]$
- (29) *üç tane köpek* ‘three CL dog’:
- a.  $\llbracket dog_{mass} \rrbracket = \lambda w \lambda x [dog_w(x)]$
  - b.  $\llbracket tane \rrbracket = \lambda n \lambda P \lambda w \lambda x [OU_w(P)(x) = n]$
  - c.  $\llbracket three\ tane\ dog \rrbracket = \lambda w \lambda x [OU_w(dog)(x) = 3]$

I will now demonstrate that Turkish displays a three-way classification of noun denotations —count, mass, and an intermediate class of *furniture*-type nouns—but there is no sign of flexibility in the grammatical representation of ontological count nouns.

### 3.3.1 On the Count-Mass Distinction

Departing from the traditional view (ter Meulen 1981, Link 1983, Bunt 1985, Pelletier and Schubert 1989) as being solely within a two-way count-mass distinction, more recent research identifies three distinct categories of noun denotations (Doetjes 1997, Chierchia 2010, Barner and Snedeker 2005, Rothstein 2010, Landman 2011, Grimm 2012, and Deal 2017). These are roughly represented by *cat*, which we readily conceptualize as count, *blood*, which we readily conceptualize as mass, and *furniture*, which share characteristics with both.

Without getting into the theoretical stance taken, I will refer to the diagnostics used in Deal (2017) to show that this three-way distinction is also evident in Turkish, though with a slightly varying behavior compared to English. The difference between ontological count and mass nouns surfaces through five diagnostics: ability

<sup>15</sup> An analysis along these lines has been proposed for Russian optional classifiers *štuka* ‘item,’ *čelovek* ‘person,’ and *golova* ‘head’ in Khrizman (2016) (see also Aikhenvald 2000 and Khrizman et al 2015). NCs with these classifiers can combine with naturally individuable mass nouns, although *štuka* has been reported to be significantly degraded when paired with mass nouns such as *furniture* and *footwear* in Khrizman (2016). Moreover, NCs with these classifiers contrast with classifierless NCs in exhibiting properties of measuring constructions (e.g., *five liters of water*) rather than functioning as counting expressions. This is particularly evident in agreement facts and their inability to be antecedents of reciprocals. Below, I discuss the reasons why Turkish NCs should not be analyzed analogously. Nevertheless, it is essential to highlight that NCs with or without *tane* exhibit a fundamental difference by being compatible with reciprocals:

- (i) İki (tane) öğrenci birbirin-e aşık ol-muş.  
 two CL student each.other-DAT love be-EVID  
 ‘Apparently, two students fell in love with each other.’

to directly combine with numerals, choice of quantificational elements, plural marking, combination with so-called ‘count adjectives’, and comparison with respect to numerosity vs. mass/volume.

***keci ‘cat’ vs. kan ‘blood’*** We know that the ability to directly combine with numerals and *tane* is a distinguishing factor for count and mass nouns in Turkish. These nouns also differ in their choice of quantifiers. For example, *cat*-like nouns are compatible with *bir kaç* ‘a few’, *blood*-like nouns are compatible with *biraz* ‘a little’, as shown in (30) (see Göksel and Kerslake 2005 and Görgülü 2010).<sup>16</sup> Note further that *tane* remains to be an option with the count quantifier, but it cannot co-occur with the mass quantifier, as evidenced in these examples.

- (30) a. birkaç (tane)/ \*biraz (tane) kedi  
 a.few CL a.little CL cat  
 ‘a few cats’  
 b. \*birkaç (tane)/ biraz (\*tane) kan  
 a.few CL a.little CL blood  
 ‘a little blood’

Turkish also distinguishes between count and mass nouns through plural marking, though in a different way from English. Turkish mass nouns can be pluralized, but this is possible with an additional *abundance* or *unorderly scattered* interpretation, as in (31a) and (31b), a phenomenon that also exists in Greek (see Tsoulas 2009 and Kouneli 2019 for Greek). In this way, the pluralization of mass nouns contrasts with the pluralization of count nouns in Turkish.

- (31) a. Adam-ın burn-un-dan kan(-lar) akı-yor-du.  
 man-GEN nose-3SGPOSS-ABL blood-PL flow-IMPERF-PAST  
 without PL: ‘There was (some) blood flowing from the man’s nose.’  
 with PL: ‘There was a lot of blood flowing from the man’s nose.’  
 b. Yerdeki kan(-lar)-ı temizle-yeceğ-im.  
 on.the.ground blood-PL-ACC clean-FUT-1SG  
 without PL: I will clean the blood on the ground.  
 with PL: ‘I will clean the scattered amounts of blood on the ground.’

One other distinguishing factor is that the combination of nouns with the so-called ‘count adjectives’ (the term due to Quine 1960) such as *small* and *round* is possible with count nouns, but not with mass nouns:<sup>17</sup>

- (32) küçük kedi/ \*küçük kan  
 small cat small blood  
 ‘small cat’/ ‘\*small blood’

Finally, comparatives are assessed based on numerosity with count nouns and based on volume with mass nouns (Barner and Snedeker 2005, Bale and Barner 2009). (33a) is true if Sevgi has a greater number of books compared to İrem, irrespective of their volume or length. But (33b) is assessed based on who has a greater volume of water, regardless of the number of portions of the water.

<sup>16</sup> This contrast also surfaces with question words *kaç* ‘how many’ and *ne kadar* ‘how much’, though there are also other quantifiers that are compatible with both count and mass nouns like *çok* ‘a lot of’ and *bütün* ‘all’. Furthermore, count nouns can co-occur with *biraz* if pseudo-incorporated, as shown below. Görgülü (2010) notes that in such cases, quantification is over the event denoted by the verbal complex, as reflected in the translation.

- (i) Biraz kitap oku-du-m.  
 a.little book read-PAST-2SG  
 ‘I did a little book-reading.’

<sup>17</sup> Mass nouns also differ from singular count nouns in allowing predication with singular and plural individuals irrespective of taxonomic vs. object-level modification in the predicate position (see the discussion in Section 2):

- (i) a. Buradaki ve şuradaki AB türü kan. ‘This one and that one are AB type blood.’  
 b. Buradaki ve şuradaki kurumuş kan. ‘This one and one that are dry blood.’

- (33) a. Sevgi-nin İrem-den daha çok kitab-ı var.  
 Sevgi-GEN İrem-ABL more very book-3SGPOSS have  
 ‘Sevgi has more books than İrem.’  
 b. Sevgi-nin İrem-den daha çok su-yu var.  
 Sevgi-GEN İrem-ABL more very water-3SGPOSS have  
 ‘Sevgi has more water than İrem.’

These diagnostics show that Turkish count and mass nouns have distributional and semantic distinctions.

*mobilya* ‘furniture’ Nouns like *mobilya* ‘furniture’ and *mücevher* ‘jewelery’ share characteristics with both count and mass nouns, and hence constitute a third variety in Turkish, as in English, albeit with some less clear-cut differences. These nouns are degraded, if not fully ungrammatical, when they combine with a numeral or a count quantifier directly. The intervention of *parça* ‘piece’ makes it perfect, but the degradedness persists with *tane* instead, as seen in (34a). Note that *parça* would be incompatible with count nouns unless the noun is coerced into a mass use via universal grinding. That is, (34b) could not mean ‘three chickens’, as is the case with the English counterpart.

- (34) a. üç/ bir kaç ??(parça)/ (??tane) mobilya  
 three a few piece CL furniture  
 ‘three/a few pieces of furniture’  
 b. üç/ bir kaç parça tavuk  
 ‘three/a few pieces of chicken’

The co-occurrence of *mobilya* with the mass quantifier is not preferred. However, if the context supports a focus on the volume of the furniture, it could be possible. Suppose that we are going to move and need to rent a truck. The following would be felicitous in a conversation with the renting agent to assess how large a truck we would need.

- (35) Biraz mobilya-mız var, çok değil.  
 a.little furniture-1SGPOSS have much not  
 We have a little furniture, not much.’

These tests point to more mass-like characteristics of *furniture*-type nouns, but they also pattern with count nouns in that they can be pluralized without inducing special *abundance* or *unorderly scattered* interpretations. They are also compatible with count adjectives:

- (36) a. Mobilya(-lar) bugün gel-iyor.  
 furniture-PL today come-IMPERF  
 ‘The furniture will be delivered today.’  
 b. yuvarlak /küçük mobilya  
 round small furniture

Finally, *furniture*-type nouns allow comparison based on numerosity and volume, sharing characteristics with both count and mass nouns. While (37) compares quantities of furniture in terms of numbers of pieces in the most natural sense, if it is accompanied by a follow-up such as ‘She will need a larger moving truck.’ it can also compare their volume.

- (37) Sevgi-nin İrem-den daha çok mobilya-sı var.  
 Sevgi-GEN İrem-ABL more very furniture-3SGPOSS have  
 ‘Sevgi has more furniture than İrem.’

The behavior of the three varieties of nouns is summarized in Table 1 (cf. Deal’s table for English, pg. 9).

To conclude, Turkish nouns have clear distributional and semantic distinctions in countability, making it more similar to English-like languages despite some puzzling variations such as the compatibility of mass nouns with the plural marker in Turkish. As much as these variations await independent research, our goal has been to establish that Turkish departs fundamentally from obligatory classifier languages in having a three-way classification of count-mass distinction.

	keci	mobilya	kan
direct combination with numerals and the count quantifier	✓	??	*
combination with <i>tane</i>	✓	??	*
combination with mass quantifier	*	✓	✓
regular pluralization	✓	✓	*
pluralization only with special interpretations	*	*	✓
combination with count adjectives	✓	✓	*
comparison based on number	✓	✓	*
comparison based on volume	*	✓	✓

Table 1: The three-way count-mass distinction in Turkish

### 3.3.2 Interim Discussion

What mainly concerns us in this comparison is that *tane* is only compatible with count nouns like *keci* ‘cat’. While it is not obvious to what extent *furniture*-type nouns are count and to what extent they are mass, it seems that they do not (fully grammatically) allow counting without the mediation of the intervening item *parça* ‘piece’, similar to the case with mass nouns. This mediation is not (readily) provided by *tane* either.

The facts also illustrate that analyzing Turkish unmarked nouns as flexible nouns is problematic. As we have seen above, there are cases where *furniture*-type nouns have count behavior, but ontological count nouns do not display any mass behavior (modulo coercion).<sup>18</sup> In contrast to the *furniture*-type, *cat*-type nouns are not compatible with mass quantification (30a), and they do not allow comparison based on volume (33a). Crucial for our purposes, count and *furniture*-type nouns differ in their ability to combine with *tane*. These differences are not expected if unmarked count nouns have a mass denotation that groups with the *furniture*-type nouns and if the intervention of *tane* is a way of making counting possible with them.<sup>19</sup>

Eliminating the possibility that Turkish count nouns have a flexible denotation, we have ample evidence against analyzing *tane* as some sort of repair mechanism with kind terms or mass nouns on a par with obligatory classifiers in Mandarin-like languages. Based on this, I argue that unmarked count nouns in Turkish NCs uniformly denote atomic properties of object-level individuals and hence they do not need a mediator to make counting possible with them.

## 4 The Analysis

In this section, I present my analysis of Turkish NCs and *tane*. I begin by addressing the optionality of *tane* and the cross-linguistic variation in number marking patterns. Following Scontras (2014, 2022), I take NCs to universally bear a cardinal head that denotes a counting function. In Turkish, *tane* serves as the overt form of this head, distinguishing it from English where the cardinal head is only covertly realized. So, the primary role of *tane*, I propose, is to count.

Then we will shift our focus to the non-optional aspect of *tane*, i.e., the (in)definiteness puzzle, and elaborate on the cross-linguistic interpretation of counting constructions. Central to the discussion in this paper, NCs are restricted to indefinite interpretations in the absence of an accompanying overt determiner in

<sup>18</sup> It is not obvious where fruit and vegetable nouns belong in this classification. They can directly combine with numerals(+*tane*) and the count quantifier, but there are also cases where they co-occur with the mass quantifier.

(i) Derin-in bir kaç/ biraz elma-sı var.  
 Derin-GEN a few a.little apple-3SGPOSS exist  
 ‘Derin has a few/some apples.’

While it is unclear to what extent universal grinding is at play here, the mass-like behavior of such nouns cannot be a factor in the semantics of *tane*. This is because *tane* is compatible with all types of count nouns, including those that strictly display a count behavior and do not readily give rise to universal grinding coercions (e.g., human and inedible animates.)

<sup>19</sup> We have noted that *furniture*-type nouns are not fully ungrammatical in their direct combination with numerals and *tane*. The degradedness should not be taken as support for the possibility that *tane* operates on furniture-type mass nouns. Otherwise, we would expect *tane* to be perfect with these nouns, as is the case with count nouns. It instead raises the question to what extent *furniture*-type nouns are count in nature, so they are compatible with *tane* to some degree, being awkward at best.



argument positions. I account for this generalization by associating the cardinal head with a choice function variable in the sense of Reinhart (1997) (cf. Jiang 2012). In short, I analyze NCs to be inherently argumental expressions with a built-in indefinite force. I further propose that one form of the cardinal head in Turkish, i.e., the covert form, is exempt from this force and thus is enriched in meaning whereas *tane* embodies the typical inherently indefinite cardinal head in Turkish.

#### 4.1 Counting with Atoms and Number Agreement in NCs

We have concluded that NCs require the noun to denote an atomic property in Turkish. Let us take this as the first step on our way to the analysis of *tane*.

Ionin and Matushansky (2006, 2019) claim that the lexical NP appearing with a cardinal numeral is semantically singular cross-linguistically, despite the appearance of the noun in languages like English. They argue that only individuals of the same cardinality can be counted. That is, numerals, being modifiers of type  $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$ , require atomic properties as an argument because the members of a plural property do not necessarily evaluate to the same cardinality.<sup>20</sup> Based on this analysis, the semantics of the numeral *two* is illustrated in (38). The constraint ensuring the atomicity requirement of numerals is given in (39) (Ionin and Matushansky 2006, pg. 321 & 329).<sup>21</sup>

- (38)  $\llbracket two \rrbracket = \lambda P \lambda x. \exists S [\prod(S)(x) \wedge |S| = 2 \wedge \forall s \in S P(s)]$   
 a.  $\prod(S)(x) = 1$  iff  $S$  is a cover of  $x$ , and  $\forall z, y \in S [z = y \vee \neg \exists a [a \leq_i z \wedge a \leq_i y]]$   
 b. A set of individuals  $C$  is a cover of an individual  $X$  iff  $X$  is the sum of all members of  $C$ :  
 $\sqcup C = X$

- (39)  $\llbracket two \rrbracket(P)(x)$  is defined iff  $\exists n \forall z [P(z) \rightarrow |z| = n]$

Then,  $\llbracket two books \rrbracket$  can be described informally as follows:

- (40)  $\lambda x \in D_e. x$  is a plural individual divisible into 2 non-overlapping individuals  $p_i$  such that their sum is  $x$  and each  $p_i$  is a book.

Languages like Turkish, Finnish, Hungarian, and Welsh transparently reflect the atomicity requirement of numerals with the singular form of the noun in their NCs. Languages like English, though, seem to challenge this view at first glance, given that the noun occurs in the plural form instead. Ionin and Matushansky argue that the plural marking on the lexical NP is number agreement reflecting the semantic plurality of NCs. That is, *books* in *two books* is lexically singular, denoting a set of atomic individuals, but the entire NC is semantically plural denoting a set of plural individuals divisible into two nonintersecting parts. The plural marker is the indicator of this plurality.<sup>22</sup> This view then dedicates the difference in the form of the noun in English and Turkish NCs to the presence/absence of number agreement reflected on the noun.

Delaying the discussion of the evidence for this approach until Section 5.1, I will now combine insights from Ionin and Matushansky's view of numerals and Scontras's analysis of English NCs to provide an initial semantic account of number agreement in NCs. We will slightly revise this account in Section 6.

Scontras claims that number marking does not contribute to the semantics of the nominal but instead is dependent on it, similar to Ionin and Matushansky's view. That is, the plural in *three books* does not spell out the \* operator (Link 1983), which closes an atomic property under sum. Following Sauerland (2003) in that syntactic number features are located on a designated functional head, Scontras argues that there

<sup>20</sup> This generalization holds for simplex numerals. Ionin and Matushansky argue that complex numerals as in *two hundred books* are derived compositionally, i.e., *hundred books* being of type  $\langle e, t \rangle$  can be a sister to *two* that is of type  $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$ . Given the constraint in (39), it is ensured that the set denoted by *hundred books* can be an argument to the numeral *two* since the set of *hundred books* contains plural individuals of the same cardinality.

<sup>21</sup>  $\prod$  represents a partition relation between a set  $S$  and an individual  $x$ , where the members of  $S$  are *individual parts* of  $x$ . For example, the partition set of  $a \oplus b$  is the set of its non-overlapping parts, i.e.,  $\{a, b\}$ , based on the way it is defined in (38a).  $\leq_i$  establishes a relation between two individuals where one is the *improper individual part* of the other. E.g.,  $a \leq_i a \oplus b$ .

<sup>22</sup> A syntactic implementation of this approach can be found in Alexiadou (2019) (building on Landau 2016 and Wechsler and Zlatić 2003) and Bayırh (2012). See also Krifka (1989, 1995, 2003), Ortman (2019), Farkas and de Swart (2010), Sağ (2018) for studies adopting an agreement-based approach to number marking in NCs.

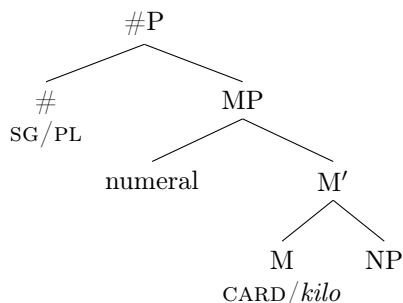
is a number head, i.e., #, that projects above NCs and hosts singular (SG) and plural (PL) features.<sup>23</sup> SG carries with it a numerical presupposition for *one-ness* of the property with which it composes and PL surfaces when the presupposition of SG fails to be satisfied. The choice between SG and PL is determined by Maximize Presupposition (Heim 1991), which requires the use of SG whenever its presupposition is met.<sup>24</sup>

In *one book*, every member of the set evaluates to 1 with respect to cardinality, and hence the lexical NP surfaces in the singular form. In *two books*, every member of the set evaluates to a number other than 1 yielding a presupposition failure for SG, and therefore probing plural morphology on the noun.

The *one-ness* presupposition of SG in English is not only determined by cardinality but also a measure that is specified by the head of a measure phrase, such as *kilo*. As evidenced in *three kilos of apples*, the specific measure supplied by the measure term determines number morphology. In *one kilo of apples*, every member of the set necessarily evaluates to 1 with respect to the kilo measure, and hence we witness the singular form *kilo*. With a different number, such as *three*, though, the plural morphology surfaces, i.e., *kilos*, due to the presupposition failure.<sup>25</sup>

For a straightforward account of number marking in constructions with cardinals and measure terms, Scontras analyzes numerals as individual-denoting expressions referring to natural numbers of type  $n$  that serve as an argument of a functional counting element, i.e., a cardinal head (CARD), or a measure term (see also Krifka 1989, 1995, 2003, Zabbal 2005, Landman 2004, and Wągiel 2018, among others). Offering a generalized Measure Phrase structure, Scontras proposes that CARD is introduced as a measure head just like other measure terms.<sup>26</sup>

(41) Generalized MP



An MP denotes a quantity-uniform (QU) property with respect to the measure expressed by the M head. That is, every member of the set denoted by the MP maps to the same number with respect to this measure. The one-ness presupposition of SG involves this information. In the case of CARD, number marking is sensitive to the quantity-uniform measure  $\mu_{CARD}$ , while in the case of measure terms, number marking is sensitive to the quantity-uniform measure internal to the measure term, e.g.,  $\mu_{kg}$ . Based on this, the semantics of English # heads is represented as below:

(42) English # heads:

- a.  $\llbracket \text{SG} \rrbracket = \lambda P : \forall \mu \forall x \in P [QU_{\mu}(P) \rightarrow \mu(x) = 1]. P$
- b.  $\llbracket \text{PL} \rrbracket = \lambda P. P$
- c.  $QU_{\mu}(P) = 1$  iff  $\forall x \forall y [P(x) \wedge P(y) \rightarrow \mu(x) = \mu(y)]$

<sup>23</sup> See also Martí (2020) for a similar analysis where number marking distinctions, concerning also dual, trial, paucal values of grammatical number, are derived by Harbour's (2014) feature system.

<sup>24</sup> Scontras does not call number marking in NCs an agreement phenomenon. This is a contribution made in this paper.

<sup>25</sup> With numerals like *zero* and decimals like *0.5*, we witness plural morphology because the one-ness presupposition of sg is not met in these cases. See also Martí (2022) for plural marking with *zero*. Note that 1.0, which equals 1 as a mathematical fact, triggers the plural marking on the noun as well, e.g., *1.0 dogs*. Based on this, Krifka (1989, 1995, 2003) concludes that plural marking in NCs must be an instance of semantically vacuous syntactic agreement. However, I depart from this view in taking plural agreement in NCs to be semantically contentful. It is possible to conceive 1.0 to be incompatible with singularity since with 1.0, what is at issue is the whole, i.e., the sum of all fractions. Then, the sense of one-ness in decimals might differ from the sense of one-ness in cardinality and measurement with respect to other measure dimensions.

<sup>26</sup> Scontras takes 'of' in MPs like *three kilos of apples* to be syntactically and semantically vacuous and argues that a difference between CARD and other measure terms like *kilo* is that the former operates on the predicative denotation of the noun while the complement of the latter must be bare plurals or mass nouns, which denote at the kind level.

Number morphology surfaces on the closest head to the # head. In the presence of an overt measure head like *kilo*, it is expressed on M, while in the presence of a covert measure head, i.e., CARD, it is expressed on the head of the next closest phrase, the lexical NP.<sup>27</sup>

Scontras gives CARD a restrictive semantics. That is, with numerals other than *one*, CARD restricts a semantically plural predicate (formed via \*) to just those members with cardinality *n*. This is one aspect where my account departs from Scontras’s view. Instead, I follow Ionin and Matushansky (2006, 2019) in taking the lexical NP in NCs to denote sets of atoms cross-linguistically, as stated above. Applying their view of numerals to it, I propose that CARD, is an expression of type  $\langle\langle e, t \rangle, \langle n, \langle e, t \rangle \rangle\rangle$  in English. It takes an atomic property *P* and a number *n*, and returns a set of individuals *x*, where each *x* evaluates to the cardinality *n*, and the atomic parts of each *x* are in *P*, as illustrated below:<sup>28</sup>

- (43) The Semantics of CARD (to be revised):  

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

$$\sqcup C = X$$

Below is the derivation of [two CARD book], where *AT* in *P<sub>AT</sub>* is short for the presuppositional content.

- (44)
  - $\llbracket \text{book} \rrbracket = \lambda x. \text{book}(x) = \{a, b, c\}$
  - $\llbracket \text{CARD} \rrbracket = \lambda P_{AT} \lambda n \lambda x. \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S P(s)]$
  - $\llbracket \text{two} \rrbracket = 2$
  - $\llbracket \text{CARD book} \rrbracket = \lambda n \lambda x. \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S \text{book}(s)]$
  - $\llbracket \text{two CARD book} \rrbracket = \lambda x. \exists S [\prod(S)(x) \wedge |S| = 2 \wedge \forall s \in S \text{book}(s)] = \{a \oplus b, b \oplus c, a \oplus c\}$

Together with this twist, number marking in the presence of numerals proceeds as shown below:

- (45) Number marking with *one*:  
  - $\checkmark \llbracket \text{SG one CARD book} \rrbracket = \{a, b, c\}$
  - $\times \llbracket \text{PL one CARD book} \rrbracket = \text{failure due to Maximize Presupposition}$

(46) Number marking with *two*:  
  - $\times \llbracket \text{SG two CARD book} \rrbracket = \text{presupposition failure}$
  - $\checkmark \llbracket \text{PL two CARD book} \rrbracket = \{a \oplus b, b \oplus c, a \oplus c\}$

To sum up, the lexical NP in NCs is semantically singular though NCs in some languages like English also involve plural number agreement. We have analyzed the semantics of number agreement, locating the semantically contentful SG and PL features on a number head that is situated above NCs.

## 4.2 *tane* as an Overt Form of the Cardinal Head

Scontras’ analysis relies on the presence of a covert cardinal head for a uniform treatment of number marking in constructions with cardinal numerals and measure terms. I propose that the Turkish classifier *tane* is the overt realization of this head. More precisely, Turkish has two forms of CARD: an overt form, which is *tane*, and a covert form, represented as  $\text{CARD}_\emptyset$ . These forms share the same semantics as CARD given in (43), with the exception that they combine with the numeral first, making them  $\langle n, \langle\langle e, t \rangle, \langle e, t \rangle \rangle\rangle$  type expressions:

- (47) The Semantics of CARD in Turkish (to be revised):  

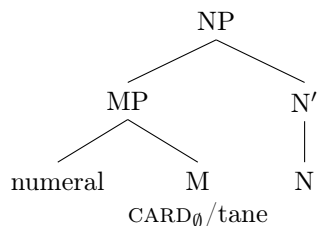
$$\llbracket \text{CARD} \rrbracket = \lambda n \lambda P_{AT} \lambda x. \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S P(s)]$$

<sup>27</sup> Scontras (2022) considers MP with CARD as a ‘cardinal numeral’ structure and MP with measure terms like *kilo* as a ‘measure term’ structure. My understanding is that numerals are not considered independent of the projection of the M head, and thus they are not the closest heads eligible for receiving number marking.

<sup>28</sup> The semantics of CARD differs from Ionin and Matushansky’s semantics of numerals in the following: Instead of their constraint in (39) that ensures the atomicity requirement of simplex numerals and the compositional derivation of complex numerals (see fn 20), I directly impose an atomic property requirement on the cardinal head regardless of simplex and complex numerals. See fn 50 for reasons that relate to complex numerals in Turkish and Farsi.

That CARD combines with the numeral first in Turkish follows from the structural configuration that I adopt for Turkish NCs. Differently from English, MP is represented as a modifier to the noun, which aligns naturally with the head-final characteristics of Turkish, as illustrated below. Unless independent evidence is provided for otherwise, representing a functional projection like M as head-initial in a strictly head-final language would only be a stipulation.<sup>29</sup>

(48) The structure of Turkish NCs

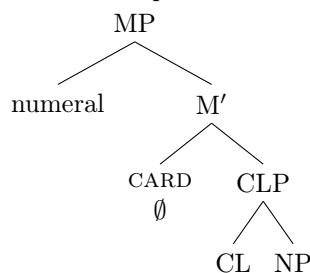


Then, the difference between English and Turkish NCs not only surfaces in the plural number agreement, but also in the form of CARD.<sup>30</sup> In English, the cardinal head is always covert, and in fact, this is the case in many other languages. Turkish departs from these languages also featuring an overt version of this head. This is the optional aspect of *tane* in NCs.

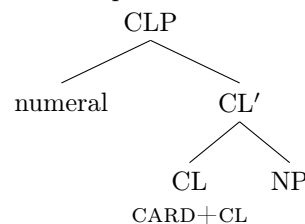
It should be emphasized that optionality is not an inherent property of an overt CARD head. In fact, it would be possible to find languages where CARD is always overt.<sup>31</sup> Furthermore, the overt vs. covert realization of CARD could be vulnerable to some language-specific factors. Scontras (2022) discusses Mayan languages, Chol and Mi'gmaq, as an example where the overt and covert realizations of CARD is dependent on the set of numerals they combine with (see Bale et al 2019 for this generalization).

Before proceeding further, there is one issue remaining that we need to address. As discussed in Section 3.1, obligatory classifiers in Mandarin-like languages take the role of a repair mechanism to ensure counting with nouns that are otherwise unsuitable for counting. Based on the analysis developed in this paper, we then expect a cardinal head besides a classifier in NCs of these languages. There are two routes to take here: CARD might be a separate covert head, as shown in (49a), or its semantics might be intertwined within the denotation of the classifier, as shown in (49b). As we have seen, a view along the lines of the latter has been argued by Krifka (1989, 1995, 2003) where obligatory classifiers are analyzed as functions that take a kind and count the number of specimens of that kind. Scontras (2014) offers a similar analysis.

(49) a. CARD as a separate head



b. CARD as part of CL's denotation



<sup>29</sup> See Sağ (2019) for other reasons that pertain to licensing of ellipsis in NCs and Sağ (2020) for the implications of this view for the interpretation of measure constructions. See also von Heusinger and Kornfilt (2017) for an alternative representation. Note that treating numerals (corresponding to the projection of CARD (MP) in our framework) as modifiers/specifiers of the NP is standard in the literature (e.g., Link 1987, Verkuyl 1993, Carpenter 1998, Landman 2003, Danon 2012, Rothstein 2017, cf. Mittwoch and Sadler 2005 and Ionin and Matushansky 2006). See Giusti (2002) and Alexiadou et al (2007) for an overview.

<sup>30</sup> Number agreement never emerges in constructions with other measure terms in Turkish: *iki kilo elma*, lit.: two kilo apple, meaning 'two kilos of apples'. Notice also that the noun occurs in the unmarked form generally in Turkish measure constructions. Sağ (2020) argues that measurement operates on the simplest form of the property in Turkish, atomic if count, mass otherwise.

<sup>31</sup> Bangla, known as an obligatory classifier language with systematically required plural marking, could be an example of this, where CARD might always be overtly realized as *ta/to* (cf. Dayal 2014, Saha 2023). One reason to analyze *ta/to* as CARD rather than a classifier as in Mandarin is that it can co-occur with partitioning quantizing nouns, e.g., *du(-to) bosta caal* 'two-CL sack rice' (p.c., Ankana Saha). See the discussion to follow for such double occurrences.

As shown in (50a), *tane* does not appear with quantizing nouns that we readily associate with a partitioning function on mass nouns, a role on a par with obligatory classifiers of Mandarin-like languages. This suggests that partitioning and cardinality functions are realized by one lexical item when they co-occur in Turkish, pointing to the structure in (49b) when the interaction of CARD and obligatory classifiers is considered analogously. However, in Western Armenian and Farsi, whose optional classifiers will be analyzed as overt cardinal heads similar to *tane* in Section 5.1, the co-occurrence of the classifier and partitioning words is possible, as seen in (50b) and (50c). Then, these functions can be spelled-out as separate heads, aligning with the structure in (49a). However, further research is required to understand what determines these choices.<sup>32</sup>

- |      |    |                                    |         |
|------|----|------------------------------------|---------|
| (50) | a. | iki (*tane) <i>damla/ şişe</i> kan | Turkish |
|      |    | two CL drop bottle blood           |         |
|      |    | ‘two drops/bottles of blood’       |         |
|      | b. | jergu (had) <i>gatil</i> arujn     | WA      |
|      |    | two CL drop blood                  |         |
|      |    | ‘two drops of blood’               |         |
|      | c. | do (tā) <i>shishe</i> nooshābeh    | Farsi   |
|      |    | two CL bottle coke                 |         |
|      |    | ‘two bottles of coke’              |         |

Until we find evidence favoring one over the other, the two structural configurations are equally possible for NCs in Mandarin-like languages. To leave both alternatives open, I will take these languages to *presumably* involve a covert CARD in their NCs, while assuming the structure in (49a) for simplicity.

Let me summarize the key elements of the analysis laid out so far: I claim that the Turkish optional classifier differs from obligatory classifiers in Mandarin and languages alike in not taking a type-fixing/repairing role to make counting possible with kinds and mass nouns. I propose that *tane* denotes a counting function being the overt counterpart of the cardinal head argued to project typically covertly in NCs cross-linguistically. The nominal argument of the cardinal head is semantically singular, but some languages conceal this, displaying plural number agreement in their NCs.

#### 4.3 Numeral Constructions and (In)definiteness

We have focused on one aspect of *tane*, that it optionally surfaces between the numeral and the noun in Turkish NCs. We are now ready to address the disparity between NCs with and without *tane*: the fact that the realization of *tane* is not an available option when the NC is interpreted as a definite expression.

##### 4.3.1 A Cross-linguistic Glance at the Interpretation of NCs

Since the seminal work of Link (1983), NCs have been widely regarded as inherently predicative expressions (e.g., Partee 1987, Link 1987, Verkuyl 1993, Carpenter 1998, Landman 2003, Ionin and Matushansky 2006, cf. Montague 1974, Bennett 1974, Barwise and Cooper 1981, Scha 1981, van der Does 1992, Dayal 2013, among others). Denoting  $\langle e, t \rangle$  type expressions, NCs can function as arguments of determiners, i.e., definite, demonstrative, and quantificational determiners, and occupy the predicate position:

- |      |    |   |
|------|----|---|
| (51) | a. | These three students didn’t submit their homework.      |
|      | b. | The first gift that I received this year was two books. |

When NCs appear in an argument position without an overt determiner, it is assumed that they default to undergoing existential type-shifting, resulting in a strong indefinite interpretation, which is free in scope-taking abilities:

<sup>32</sup> *tane* can co-occur with container nouns if the container noun forms a compound with the lexical noun: *iki (tane) şişe-su* ‘two bottle-waters’. (In such compounds, the stress falls on the container noun, as opposed to the ungrammatical form given in (50a).) This could mean that quantizing nouns, when used in their lexical sense, cannot spell out CARD, allowing *tane* to surface. As a reviewer hints, the co-occurrence of the classifier with partitioning nouns in (50b) and (50c) could also potentially be attributed to a compound structure. However, this cannot be the case at least for Farsi because the version where the partitioning noun and the noun “coke” form a compound differs in form: do (tā) *nooshābeh-e shisheyi* ‘two bottle-okes’

- (52) Three students are not standing but three are. (Dayal 2013, pg. 22)

This is at odds with the fact that English allows bare nominal arguments only with a narrow scope existential reading. Jiang (2012) further points out that NCs can freely occupy an argument position even in languages that strictly ban the occurrence of bare nominal arguments, e.g., French. In these languages as well, determinerless NCs only yield strong indefinite interpretations.

There are two principles in the neo-Carlsonian approach that regulate the interpretation of bare nominal arguments:

- (53) a. *Blocking Principle* (Chierchia 1998):  
 For any type shifting operation  $\phi$  and for any  $X$ :  $*\phi(X)$  if there is a Determiner  $D$  such that for any set  $X$  in its domain,  $D(X) = \phi(X)$ .  
 b. *Revised Meaning Preservation* (Dayal 2004) :  $\{\cap, \iota\} > \exists$

Revised Meaning Preservation dictates type-shifters to apply in a certain order, as long as the Blocking Principle is respected. The Blocking Principle requires the use of lexical items instead of covert type-shifting operations for reasons of economy. So, based on the Blocking Principle, bare nouns in languages with an overt definite article cannot undergo covert  $\iota$  type-shifting, and hence they disallow definite interpretation. Due to Revised Meaning Preservation, bare nouns cannot be interpreted as strong indefinites because  $\exists$  type-shifting is low-ranked and hence unavailable. The narrow scope existential reading of bare nouns is independently possible through Derived Kind Predication (Chierchia 1998).

The fact that French and English NCs can have strong indefinite interpretations then shows that Revised Meaning Preservation is not applicable to NCs, unlike the case with bare nouns. However, the Blocking Principle still appears to be relevant, given that NCs need the overt determiner or the demonstrative for a definite(-like) interpretation in these languages, as exemplified for French in (54).

- (54) Jean a acheté deux chiens et deux chats. \*(Les)/ \*(ces) deux chats sont coûteux.  
 John has bought two dogs and two cats the/these two cats are costly  
 ‘John bought two dogs and two cats. The/these two cats are very expensive.’ (Jiang 2012, pg. 95)

In articleless languages, the picture is surprisingly similar. Although definite interpretation is available for bare nouns through the covert  $\iota$  operator, this does not seem to be the case with NCs. In Mandarin and Russian, NCs can only be interpreted as a strong indefinite in argument positions. A definite-like behavior is possible if NCs are preceded by a demonstrative, as exemplified below for Russian (see also Dayal 2013 for Hindi):

- (55) Ivan kupil pjat’ sobak i pjat’ koshek, \*(Eti) pjat’ koshek ochen’ dorie.  
 Ivan bought five dogs and five cats these five cats very expensive  
 ‘Ivan bought five dogs and five cats; these five cats are very expensive.’ (Jiang 2012, pg. 95)

A striking confirmation of the fact that NCs do not undergo covert  $\iota$  type-shifting comes from Yi, an obligatory classifier language with a definite article. Yi is special in also allowing bare nouns to receive definite interpretation without the definite article, suggesting that the Blocking Principle does not apply in this language (see Jiang 2018 though). Despite this level of freedom with bare nouns, NCs still require the presence of the article for definite interpretation:

- (56) sse-vo nyip ma \*(su) dza dzu ndzɔ.  
 boy two CL DEF rice eat PROG  
 ‘The two boys are having meal.’ (Jiang 2012, pg. 334)

The interim conclusion based on these languages is as follows: While strong indefiniteness is the default interpretation available to NCs, definiteness of NCs cannot be achieved through covert means, i.e.,  $\iota$  type-shifting, but instead requires an overt marking by a definite article, or alternative markers such as demonstratives, if the language lacks a definite article. The phenomenon of NCs resisting definiteness via  $\iota$  in languages without articles remains puzzling if we consider NCs to be inherently of predicative type. That is, it is unclear why  $\iota$  is not equally applicable to NCs in argument positions in these languages unlike the case with bare nouns.

The situation becomes more intricate when we delve into the interpretation of Turkish NCs. While NCs with *tane* comply with the general restrictedness of NCs to indefiniteness, NCs without *tane* display exceptional behavior in freely allowing both definite and indefinite readings, as detailed next.

#### 4.3.2 Turkish NCs and Indefiniteness

Indefinites differ from quantificational elements in that they have unusual scope behavior, with the ability to take wide scope in unexpected contexts (e.g., Fodor and Sag 1982). Like universal quantifiers, they can take wide scope over a preceding quantifier. Still, unlike universal quantifiers, they can also take scope out of islands such as complex noun phrases and the antecedent of conditionals.

Turkish is a scope rigid language, where scope relations reflect the surface order of quantifiers (Zidani-Eroğlu 1997, Göksel 1997, Aygen-Tosun 1999, Keleşir 2001, among others), as illustrated by the following example from Keleşir (pg. 57):

- (57) Bir öğrenci her kitab-ı oku-du.  
 one student every book-ACC read-PAST  
 ‘A student read every book.’ ( $\exists > \forall$ ,  $*\forall > \exists$ )

In contrast to its English counterpart, (57) is only true in a situation where each book at issue was read by the same single student, and it would be false in a situation where each book was read by different students. This shows that the universal quantifier that is in the object position cannot take wide scope over the indefinite subject. In order to get a narrow scope reading for the indefinite, the universal quantifier needs to be fronted, resulting in different word order:

- (58) Her kitab-ı bir öğrenci oku-du.  
 every book-ACC one student read-PAST  
 ‘Every book is such that a student read it.’

Despite scope rigidity, Turkish indefinites are like indefinites in other languages in showing scope ambiguity and having exceptional scope-taking abilities (Keleşir 2001):<sup>33</sup>

- (59) Öğrenci-ler-in tam olarak yarısı bir kitab-a yorum yaz-dı.  
 student-PL-GEN exactly half one book-DAT comment write-PAST  
 ‘Exactly half of the students wrote comments on a book.’  
 (**exactly half** > **a**, **a** > **exactly half**)

Imagine that there are four students, and two wrote comments on a (possibly different) book. In this situation, (59) is true reflecting the narrow scope reading of the indefinite. Now suppose that out of four, two students wrote comments on the same specific book, while a third student wrote comments on a different book. In this situation, (59) is still true, showing that the indefinite can take wide scope over the quantifier in the subject position. Notice that the narrow scope reading of (59) would be false in this scenario because more than half of the students wrote comments on a book.

Just like regular indefinites, NCs with and without *tane* show scope ambiguity when interacting with other quantifiers. (60) can be true in a situation where three of six students wrote comments on (possibly different) two books, reflecting the narrow scope reading of the NCs. It can also be true in other situations where more than half of the students commented on two books as long as the same two books were chosen by half of the students. Like in English, (60) can also receive a distributive reading in the wide scope interpretation of the NC: There are two books such that exactly half of the students wrote comments for each.

- (60) Öğrenci-ler-in tam olarak yarısı iki (tane) kitab-a yorum yaz-dı.  
 student-PL-GEN exactly half two CL book-DAT comment write-PAST

<sup>33</sup> This is the case for case-marked indefinites. Non-case-marked indefinites always yield narrow scope readings. Keleşir (2001), following Diesing (1992), Kennelly (1994), and Zidani-Eroğlu (1997), argues that non-case-marked indefinites are situated inside the VP, where they are locally  $\exists$ -closed, unlike case-marked indefinites which are outside the VP. Note also that accusative case-marked indefinites always receive a wide scope specific reading, which Keleşir (2001) explains by arguing that accusative case carries a presupposition of existence (cf. Enç 1991). Therefore, in this paper, the behavior of indefinites is shown with other case markers.

‘Exactly half of the students wrote comments on two books.’  
 (exactly half > two, two > exactly half)

Turkish NCs also have exceptional scope-taking abilities, and hence can be interpreted inside or outside of an island. For example, (61) can be felicitous in two contexts: In the first one, it is enough that any two of my projects are selected for me to receive funding, while the other requires the condition that two specific projects of mine be selected. However, contrasting with the case in (60), the wide scope reading requires a collective selection of two projects in line with the facts in English.

- (61) 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 are selected, I will receive funding.’ (if > two, two > if)

One other general aspect of indefinites is that they can take intermediate scope (Ruys 1992, Abusch 1993, Farkas 1981). Turkish indefinites also show this general behavior; accordingly, NCs can take intermediate scope besides the widest and narrowest scope readings. For example, in (62), it is possible that different linguists gave an A to every student if they answered two questions that are fixed for every student of a single professor. (Linguist 1 gave an A to every student who answered Question a and b. Linguist 2 gave an A to every student who answered Question c and d, etc.)

- (62) Çoğu dilbilimci iki (tane) soru-ya yanıt ver-en her öğrenci-ye A ver-di.  
 most linguist two CL question-DAT answer-REL every student-DAT A give-PAST  
 ‘Most linguists gave an A to every student that answered two questions.’

We have seen that NCs behave like regular indefinites in their scope-taking properties. They can also receive specific readings other than the ones induced by scopal interactions, such as partitive specificity and epistemic specificity (see von Stechow 2002). The sentence in (63) shows that both forms of NCs can be partitive specifics. Epistemic specificity, on the other hand, expresses the speaker’s knowledge about the referent of an indefinite. As shown in (64), both NCs can also reflect epistemic specificity.

- (63) Oda-da bir sürü çocuk var-dı. İki (tane) çocuk kart oynu-yor-du. Üç (tane) çocuk  
 room-LOC one many child exist-PAST two CL child card play-IMPERF-PAST three CL child  
 televizyon izli-yor-du.  
 TV watch-IMPERF-PAST  
 ‘There were many children in the room. Two children were playing cards. Three children were watching TV.’
- (64) İki (tane) öğrenci sınav-da kopya çek-ti. Kim ol-duk-larım-ı bil-iyor-um. Zeynep  
 two CL student exam-LOC cheat-PAST who be-NMLZ-3PLPOSS-ACC know-IMPERF-1SG Zeynep  
 ve Merve.  
 and Merve  
 ‘Two students cheated on the exam. I know who they are: Zeynep and Merve.’

In sum, the facts discussed above demonstrate that NCs of Turkish can be indefinite regardless of the absence/presence of *tane*.

#### 4.3.3 Turkish NCs and The Definiteness Puzzle

I will now show that in the absence of *tane*, NCs can also be definite, but in the presence of *tane*, they cannot (Schroeder 1992, Öztürk 2005). As first shown in Section 1.1, unlike NCs with *tane*, NCs without *tane* can receive a definite interpretation, evidenced by their ability to refer back to a unique/maximal entity introduced in the preceding context. However, NCs with *tane* do not show this type of anaphoric behavior, for which they need to be preceded by a demonstrative, as illustrated below. The form with *tane* forces a partitive specific reading or is understood as introducing a new discourse referent in the second sentence when not accompanied by a demonstrative, yielding infelicity in this particular context.<sup>34</sup>

<sup>34</sup> A reviewer provides the example given in (ia), suggesting its acceptability in a context where the engineers introduced in the second conjunct are the same as those introduced in the first conjunct. Building on this observation, the reviewer highlights



- (65) Polis beş (tane) hemşire-nin ölüm-ün-ü araştır-ıyor. Edinilen bilgilere  
 police five CL female.nurse-GEN death-3SGPOSS-ACC investigate-IMPERF gathered information  
 göre, **beş (#tane) kadın-ın/ bu beş (?tane) kadın-ın** ellili yaş-lar-da  
 according.to five CL woman-GEN this five CL woman-GEN fifties age-PL-LOC  
 ol-duğ-u tahmin ed-il-iyor.  
 be-NMLZ-3SGPOSS predict-PASS-IMPERF  
 ‘The police is investigating the death of five nurses. Based on the information gathered, it is predicted  
 that the/these five women were in their fifties.’

The contrast in (65) shows that NCs without *tane* do not support a definite interpretation without a demonstrative. This is also revealed in situational contexts where reference to a unique/maximal entity that is familiar in the common ground can be established. Imagine that Sevgi has three apples only and this is known by both the speaker and the addressee. In this case, the NC with *tane* cannot refer to these three apples, whereas NCs without *tane* can, as shown in (66). In contrast, if Sevgi has four apples, referring to three of them is possible with both forms of NCs, as in (67). This demonstrates once again that while both forms of NCs are compatible with partitive specificity, only NCs without *tane* can yield definiteness.

Context: Sevgi has three apples only, familiar to the interlocutors.

- (66) Sevgi-nin üç (#tane) elma-sın-ı Merve-ye ver-di-m.  
 Sevgi-GEN three CL apple-3SGPOSS-ACC Merve-DAT give-PAST-1SG  
 ‘I gave Sevgi’s three apples to Merve.’

Context: Sevgi has four apples, not necessarily familiar to the addressee.<sup>35</sup>

- (67) Sevgi-nin üç (tane) elma-sın-ı Merve-ye ver-di-m.  
 Sevgi-GEN three CL apple-3SGPOSS-ACC Merve-DAT give-PAST-1SG  
 ‘I gave three of Sevgi’s apples to Merve.’

One other diagnostic showing the contrast between the two forms of NCs in terms of definiteness comes from their occurrence with the particle *dA*, which can be additive or distributive, besides having various other functions (Göksel and Özsoy 2003, Göksel and Kerlake 2005, Kamali and Karvovskaya 2013, and Sağ and Demirok 2023). NCs without *tane* are compatible with both the distributive and additive roles of *dA*, but when *tane* is present, *dA* can only function as an additive particle (see Öztürk 2005):

- (68) Parti-de dört (tane) çocuk kurabiye ye-di.  
 party-LOC four CL kid cookie eat-PAST  
 ‘Four kids ate cookies in the party.’

that NCs with *tane* can potentially have definite interpretations. However, there may be an alternative analysis for such cases. There are instances where the novelty condition of indefinites (Heim 1982) appears to be breached, as illustrated by the English example in (ib). This example implies a surprising scenario —that a child only eight years old could unexpectedly resolve the issue (p.c. Gennaro Chierchia). The example in (ia) might then merely serve as another instance where the novelty condition of an indefinite (the NC with *tane* in the second conjunct) is violated. It is crucial to note that the surprise effect, which the reviewer also points out to be evident in (ia), would not follow from treating NCs with *tane* simply as an anaphoric definite in such examples. Further insights into other cases where indefinites do not exhibit novelty effects can be found in Condoravdi (1994), Krifka (2001), and Chierchia and Heim (2017).

- (i) a. İçeri dört tane mühendis gir-di ve dört tane mühendis şirket-in tüm problem-ler-in-i  
 inside four CL engineer enter-PAST and four CL engineer company-GEN all problem-PL-3SGPOSS-ACC  
 çözdü.  
 solve-PAST  
 Reviewer’s translation: ‘Four engineers entered inside and the four engineers solved all the problems in the company.’  
 b. I was desperate because of a problem on my computer. An eight year-old walked in my office and an eight year-old fixed my problem.

<sup>35</sup> The best way of translating (67) is ‘I gave three of Sevgi’s apples to Merve.’ As highlighted by a reviewer, this involves ‘Sevgi’s apples,’ which is a definite expression, and yet the contextual description states that the identity of the apples that Sevgi owns is not necessarily familiar to the addressee. As evident in (66) and (67), both definite and partitive specific indefinite NCs are within a possessive form, indicating that the Turkish possessive differs from the English one in not being strictly definite.

- a. Dört çocuk **da** bir şişe süt iç-ti.  
 four kid DA one bottle milk drink-PAST  
 ‘Four *additional* kids drank a bottle of milk.’  
 ‘The four kids *each* drank a bottle of milk.’  
 collective  
 distributive
- b. Dört tane çocuk **da** bir şişe süt iç-ti.  
 four CL kid DA one bottle milk drink-PAST  
 ‘Four *additional* kids drank a bottle of milk.’  
 collective

The sentence in (68a), which has a NC without *tane* as its subject, can have a collective reading where *dA* contributes an additive interpretation. That is, the sentence means that in addition to four kids who ate cookies, four other kids collectively drank a bottle of milk. Crucially, (68a) can also have a distributive reading, where the four kids who ate cookies each drank a bottle of milk. In contrast, the distributive reading is not evident in the presence of *tane*, while the additive reading is still available, as shown in (68b).<sup>36</sup>

What matters for our purposes is that the distributive reading is only possible if the NC is interpreted as definite. More precisely, in the distributive reading of (68a), the NC has to refer to the four kids introduced in the previous sentence. This shows that the distributive instance of *dA* imposes some sort of definiteness requirement on its host. NCs with *tane* cannot meet this prerequisite as they are unable to convey definite readings, and hence they are incompatible with the distributive function of *dA*.<sup>37</sup>

In summary, while both forms of NCs can be indefinite, *tane* does not surface in the definite interpretation.<sup>38</sup> This difference highlights that the form with *tane* aligns with the typical cross-linguistic characteristic of NCs, being restricted to indefiniteness. In contrast, NCs without *tane* are extraordinary in also allowing a definite interpretation freely. Our goal is then to account for the obligatory indefiniteness of NCs in the presence of *tane* while this constraint appears to vanish in its absence.

#### 4.4 Associating the Cardinal Head with Indefiniteness

We have seen that NCs are strong indefinites in argument positions without an overt D, irrespective of cross-linguistic variations in the rules governing bare arguments. Jiang (2012) proposes that NCs have a distinct source to form arguments from bare nouns, stemming from a lexical variant of numerals. Numerals, in her view, are lexically ambiguous, acting as modifiers of type  $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$  and modifiers with a built-in choice function variable of type  $\langle\langle e, t \rangle, e\rangle$  à la Reinhart (1997) (cf. Fodor and Sag 1982, Winter 1997, and Kratzer 1998). The first variant of numerals allows NCs to have a predicative denotation that can be an argument to a category at D. In contrast, the second variant yields individual type NCs that require existential closure.

In Reinhart’s theory, the  $\exists$ -closure of the choice function variable is assumed to apply at any compositional level. This explains the intermediate scope readings of indefinites and their ability to take exceptional scope without a need for a mechanism to extract the indefinite from an island. The lack of distributive readings with NCs in exceptional scope cases is also derived in this theory, because the distributive reading can only be obtained by a QR-analysis of indefinites and this account rejects a generalized quantifier analysis of indefinites (cf. Fodor and Sag 1982). Below, I exemplify semi-formally how a NC is interpreted with respect to an island under this view.

<sup>36</sup> It should be noted that both forms of NCs are compatible with distributive, collective, and cumulative readings, and the contrast in terms of distributivity and collectivity discussed above emerges only when they are used with the particle *dA*. For example, both NCs with and without *tane* have a distributive reading when used with dependent numerals, as in the following:

- (i) Dört (tane) çocuk bir-er şişe süt iç-ti.  
 four CL kid one-DIST bottle milk drink-PAST  
 ‘Four kids drank a bottle of milk each.’

<sup>37</sup> Plural definites and plural pronouns are incompatible with the distributive *dA*: *Kızlar da bir şişe süt içti*. ‘The girls, too, drank a bottle of milk.’ Here, *dA* has only an additive and collective reading, which is unexpected because plurals and pronouns can satisfy the definiteness requirement of the distributive *dA*. Sağ (2019) analyzes *dA* as a post-suppositional item associated with universal quantification on a par with Szabolcsi’s (2015) analysis of *mo*, the Japanese kin of *dA*. The unavailability of the distributive reading with plural definites and pronouns is argued to follow from their ‘weak maximality.’ It is well-known that plural definites allow exceptions in their interpretations, as opposed to cardinal definites and universal quantifiers (Kroch 1975). Being associated with universal quantification, *dA* is also sensitive to strong vs. weak maximality potential of its host.

<sup>38</sup> One might wonder if there is any difference between the two forms in the indefinite interpretation. To the best of my knowledge, there is no difference, at least in any visible way.

- (69) If two of my projects are selected, I will receive funding.
- a. Narrow Scope Reading (**if** > **two**):  
 $[\exists f [CH(f) \wedge be.selected(f([\![two\ projects\]]))] \rightarrow funding]$   
 I will get funding if there is a choice function and the two projects that it selects are selected (by the committee).
- b. Wide Scope Reading (**two** > **if**):  
 $\exists f [CH(f) \wedge [be.selected(f([\![two\ projects\]]))] \rightarrow funding]$   
 There is a choice function such that if the two projects that it selects are selected (by the committee), I will get funding.

I argue that the source of indefinite interpretations associated with NCs is *CARD* instead. While the deeper connection between indefiniteness and cardinality awaits future explorations, here I maintain Jiang’s insight to construct a typological framework as a first step to understanding why NCs have a universal tendency to favor indefinite interpretations over definite ones. The twist that I offer is that NCs are cross-linguistically argumental, due to *CARD* bringing with it a choice function variable. Depending on the structure of NCs in a language, *CARD* is either of type  $\langle n, \langle \langle e, t \rangle, e \rangle \rangle$  as in Turkish or type  $\langle \langle e, t \rangle, \langle n, e \rangle \rangle$  as in English, differing only in the order of the number and property arguments. Below, I represent the cross-linguistic semantics of *CARD* based on the English *CARD* (conforming to the generalized structure of NCs given in (72)), but nothing hinges on this choice. The subscript  $f$  on  $CARD_f$  represents the argumental nature of the cardinal head—that it yields argumental type NCs with a choice function.

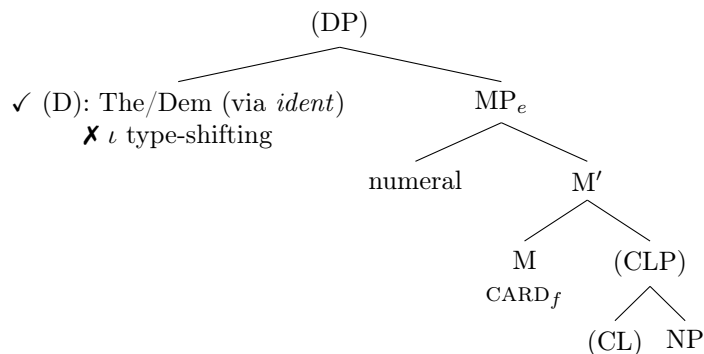
- (70) The cross-linguistic semantics of *CARD* (final)  $\langle \langle e, t \rangle, \langle n, e \rangle \rangle$   
 $[[CARD_f]] = \lambda P_{AT} \lambda n. f(\lambda x \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S P(s)])$

I follow Dayal (2013) in that NCs type-shift to a predicate as a repair operation only occurring in structures requiring the predicative meaning of NCs, i.e., when they serve as arguments to overt determiners or occupy the predicate position. Since NCs are  $e$ -type expressions, I assume the shift to the predicate type occurs via the *ident* operator (Partee 1986), as illustrated below. NCs would not undergo predicative type-shifting to subsequently undergo *iota* type-shifting, given that they are already in the proper type to merge with whatever takes the NC as an argument.<sup>39</sup> That is, there is no impetus for the reparative strategy to activate. In articleless languages, NCs then gain definite-like readings solely through overt means such as demonstratives.

- (71) a. *ident*:  $\lambda x \lambda y. y = x$   
 b. *ident*([two books]) =  $\lambda y. y = f(\lambda x \exists S [\prod(S)(x) \wedge |S| = 2 \wedge \forall s \in S book(s)])$   $\langle e, t \rangle$

In languages like French, English, and Russian, where *CARD* is strictly covert, and in languages like Mandarin and Yi, where *CARD* is presumably covert, NCs have the following construal, represented with a generalized MP structure, ignoring number marking and potential structural variations:<sup>40</sup>

- (72) The Generalized Structure of NCs



<sup>39</sup> I thank an anonymous reviewer for generously bringing this point to my attention.

<sup>40</sup> As discussed in fn 15, Russian has optional classifiers but they seem to have a distinct semantics from the Turkish optional classifier *tane* (Aikhenvald 2000, Khrizman et al 2015, Khrizman 2016). Therefore, in this study, I refrain from analyzing these classifiers as an overt *CARD* head on a par with *tane*, although a thorough comparison demands further investigation.

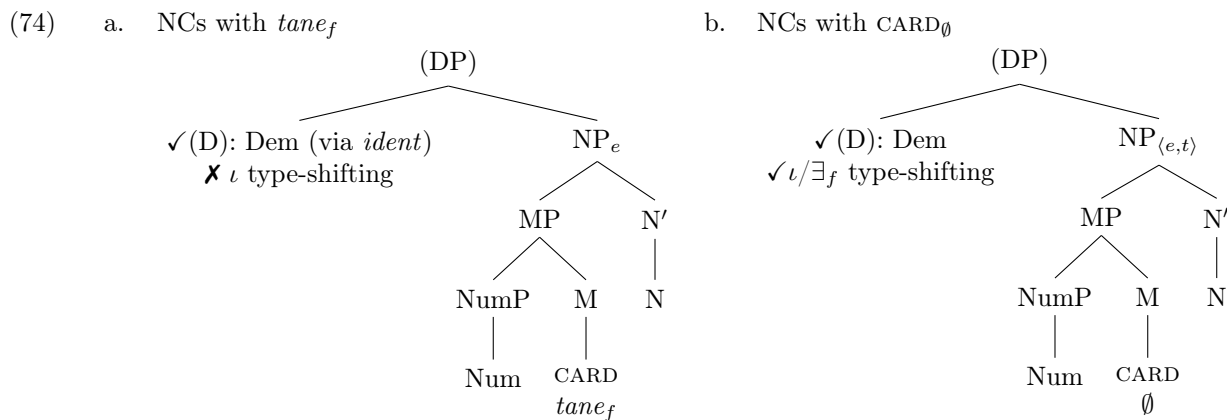
The parentheses around the DP should be interpreted as follows: An argumental type NC can directly occupy an argument position as an MP. Alternatively, it can serve as a complement to a category at D, such as the definite and demonstrative determiners, triggering *ident* type-shifting to resolve the type-mismatch.

In contrast, *iota* type-shifting is unavailable, as explained above. It is essential to highlight that the  $\iota$  operator is represented below the D head for ease of comparison. My account posits  $\iota$  as a covert type-shifting operator and assumes no silent D projection in the absence of an overt definite article. If  $\iota$  were to be inserted under a silent D head, we would not anticipate the covert D head to behave differently from an overt D head—a covert definite determiner would trigger *ident* type-shifting, mirroring the pattern observed with an overt definite determiner, due to the type-mismatch that arises when it takes an  $e$ -type NC as its complement.<sup>41</sup>

Associating the choice function with *CARD* makes it possible to account for the interpretational variation between the two forms of NCs in Turkish. Having identified the cardinal head as the general source of the indefinite force, we can now reduce intra-language variations to its potential variants. More specifically, I propose that in Turkish, one form of NCs continues to retain the general construal in cardinal semantics, but the other form is divorced from the choice function, perhaps as a way to resolve the restriction to indefiniteness. That is, NCs with the covert *CARD* have a predicative denotation.

- (73) a. The Semantics of the overt *CARD* in Turkish  $\langle n, \langle \langle e, t \rangle, e \rangle \rangle$   
 $\llbracket \textit{tane}_f \rrbracket = \lambda n \lambda P_{AT}. f(\lambda x \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S P(s)])$   
 b. The Semantics of the covert *CARD* in Turkish  $\langle n, \langle \langle e, t \rangle, \langle e, t \rangle \rangle \rangle$   
 $\llbracket \textit{CARD}_\emptyset \rrbracket = \lambda n \lambda P_{AT} \lambda x. \exists S [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S P(s)]$

NCs with *CARD*<sub>∅</sub> are then enriched in meaning since, being inherently of  $\langle e, t \rangle$  type, they can not only directly display regular predicative behavior, but also feed into covert type-shifting operators, including  $\iota$  and the choice function. Therefore, they allow both definite and indefinite interpretations. In other words, while NCs with *tane* are just like NCs in other languages reflecting the inherent indefiniteness, NCs with *CARD*<sub>∅</sub> bring a seemingly exceptional status to the Turkish counting system. This is schematized in (74):



As highlighted by the anonymous reviewers, the existential type-shift of NCs with *CARD*<sub>∅</sub> through the choice function would not be hindered by NCs with *tane* since the choice function is introduced at different syntactic levels through distinct sources in these constructions—by a lexical item in one construction and as covert type-shifting operation in the other—thus no blockage effect is anticipated to occur.

To summarize, I have claimed that NCs are cross-linguistically argumental, relating the restrictedness to indefiniteness with a cardinal head that is universally involved in the projection of NCs.<sup>42</sup> I have further proposed that in Turkish, an articleless language that features both an overt and a covert form of *CARD*,

<sup>41</sup> I do not rule out the possibility that there might be articleless languages where  $\iota$  is introduced within a covert DP projection rather than as a type-shifting operation. According to the analysis in this study, definiteness through  $\iota$  is predicted to be available with NCs in such languages. For Turkish, I follow Öztürk (2005) and Bošković and Şener (2014) in that there is no D projection in the absence of an overt definite article (cf. Arslan-Kechriotis 2009, Kornfilt 2005, 2017, von Stechow and Kornfilt 2017).

<sup>42</sup> My proposal should rather be taken as a cross-linguistic tendency. One exception I am aware of is Vietnamese, an articleless obligatory classifier language, where NCs allow both definite and indefinite interpretations (Phan and Chierchia 2022).

the covert CARD leads to predicative NCs and hence allows definiteness without a need for an overt marker, while *tane* is just an ordinary cardinal head hardwired with an indefinite force as in any other language.<sup>43</sup>

As a concluding note, it is crucial to underscore the underlying rationale supporting the claim that NCs are primarily argumental expressions across languages. An inconsistency arises if we consider NCs as uniformly predicative expressions of type  $\langle e, t \rangle$ . This uniformity fails to account for why NCs with *tane* (and those in articleless languages) resist *iota* type-shifting for definite interpretations, unlike NCs without *tane*. On the other hand, we must acknowledge the existence of inherently predicative NCs to account for the availability of definite readings for NCs without *tane* in contrast to the form with *tane*. Consequently, the observed pattern in the two forms of Turkish NCs suggests the presence of two types of NCs: one is the typical argumental type and the other is the predicative type. The latter is a rare phenomenon, possibly coexisting with inherently indefinite NCs within a single language. The discussion of Farsi NCs, as discussed next, will be further illuminating on this point.

## 5 Further Support: Farsi and Western Armenian

With an aim to create a comparative platform and reinforce the cross-linguistic breadth of the analysis, we will now investigate NCs in two more optional classifier languages: Farsi, an articleless language, and Western Armenian, a language that has a definite article. The investigation of these languages proves fruitful in two aspects. Firstly, it sheds light on the variation in number marking in NCs and provides evidence for the agreement-based approach to this variation. Secondly, it strengthens the view of linking indefiniteness of NCs to CARD and shows that our analysis of the interpretation of Turkish NCs is on the right track.

### 5.1 Evidence for the Agreement Approach

We have adopted an agreement approach in the sense of Ionin and Matushansky (2006, 2019) to explain number marking variations in NCs. In short, the cardinal function is universally defined on the semantically singular form of the noun, while languages like English further involve number agreement in NCs, disguising this requirement. While I refer the reader to Ionin and Matushansky (2019) for a more comprehensive list of evidence for this account, one is particularly of interest here: Plural agreement found on predicates, i.e., subject-predicate agreement, and plural marking in NCs are conditioned by similar factors that involve animacy, specificity, and definiteness. Based on this, Ionin and Matushansky conclude that the plural marking in NCs must also be a product of a number agreement phenomenon. They discuss Western Armenian (WA) as one language where we see this correlation, to which I now add Farsi (see also Alexiadou 2019).

WA and Farsi are optional classifier languages with a systematic number marking system, though they differ from Turkish in also allowing plural marking in their NCs. This is possible with a definite or specific reading in WA and a definite reading in Farsi (e.g., Sigler 1996 for WA and Ghomeshi 2003 for Farsi). To see the case of WA first, consider the example in (75). In the presence of the plural marker *-er*, the NC gets a wide scope reading, and the sentence is felicitous in a context where there are two specific elephants that I want to see (de re). Otherwise, the NC receives a narrow scope interpretation.<sup>44</sup>

<sup>43</sup> A reviewer points out that the combination of *every* with an argumental type NC is predicted to yield an odd result because the  $\exists$ -closure of the choice function can only be achieved outside of the NC, and the universal quantification (through *ident*) would be over the individual picked out by the given *f*. NCs are indeed odd in sentences like *I spoke to every three students*. While this aligns with the proposed choice function analysis of NCs, there are also sentences where NCs are used felicitously with *every*, e.g., *There was a policeman every three houses* or *Nine out of every ten patients recommend this toothpaste*. These felicitous cases both involve some sense of partitiveness (p.c. Gennaro Chierchia). The former implies that for every part of the street spanning three houses, there is a policeman. Similarly, in the latter, *every ten patients* is felicitous within a partitive construction. In Turkish, both forms of NCs align with the pattern in English in their combination with *her* ‘every.’ For example, *Her üç (tane) çocukla konuştum* (‘I spoke to every three students’) is peculiar but *Her üç (?tane) binaya bir polis düşüyor* (‘There is a policeman for every three buildings’) is felicitous, with a slight preference for the absence of *tane* similar to the case with demonstratives, as shown in (65) with a question mark. Therefore, the puzzle of universally quantified NCs seems to be independent of whether the NC is inherently argumental or predicative. As the role of partitiveness in the felicitous uses is also yet to be explored, I defer the investigation of the use of NCs with *every* to future research.

<sup>44</sup> WA shows dialectal variation with respect to the co-occurrence of the classifier and the plural marker. Recall that the data discussed here reflect the variety spoken in Beirut (see fn 3). Based on the judgments of a consultant from Istanbul, plural marking does not surface in the presence of the classifier in the variety spoken in Istanbul, though a specific indefinite

- (75) g-uz-em jergu (had) piy(-er) desn-el kazananoṯs-i-n meṯj̄  
 INDC-want-PRES.1SG two CL elephant-PL see-INF zoo-GEN-DEF inside  
 ‘I want to see two elephants in the zoo.’ with PL: (**two** > **want**), without PL: (**want** > **two**)

WA has an overt definite marker, i.e., *-ə*, and both forms of NCs require to be inflected with it to yield definite interpretations. When this is the case, the plural marker is obligatory, as shown in (76). So, we can conclude that WA NCs are compatible with the plural only in specific and definite contexts.

- (76) dup-i-n meṯj̄-ə ga-r jerek had kirk, meg had dedrag, jev jergu had madid.  
 box-GEN-DEF inside-DEF exist-PAST.3SG three CL book one CL notebook and two CL pencil  
**jerek (had) kirk-er-ə** im bəzdig zarmig-i-s dəv-i.  
 three CL book-PL-DEF my little cousin-DAT-1SGPOSS gave-PAST.1SG  
 ‘There were three books, one notebook, two pencils in the gift box. I gave the three books to my little cousin.’

The use of the plural marker in Farsi NCs is more restricted. In the absence of the plural, both forms of NCs can be used as indefinites, including specific interpretations.<sup>45</sup> This is illustrated in (77), which could be paraphrased as follows: (i) If you help any two students in my class, I will increase your grade (if > two). (ii) There are two specific students in my class and if you help both, I will increase your grade (two > if).

- (77) Age tu kelās-e man be do (tā) dāneshju komāk kon-i, be nomra-t ezāfe  
 if in class-EZ I to two CL student help-2SG to grade-2SGPOSS increase  
 mi-kon-am.  
 IMPERF-do-1SG  
 ‘If you help two students in my class, I will increase your grade.’ (**two** > **if** or **if** > **two**)

Farsi lacks an overt definite article as in Turkish, but NCs are interpreted as definite when inflected with the plural marker *-hā*. This is shown in (78) where the *-hā*-marked NC refers to the maximal individual introduced in the preceding sentence, whereas, in the absence of *-hā*, the NC does not have this ability.<sup>46</sup>

- (78) Do (tā) moallem, se (tā) mohandes, va ye doktor vāred-e otāgh shod-an. **Do tā**  
 two CL teacher three CL engineer and a doctor inside-EZ room become-3PL two CL  
**moallem#(-hā)** dar mored-e ye chiz-i sohbat mi-kard-an.  
 teacher-PL about-EZ a thing-INDEF conversation IMPERF-PAST.do-3PL  
 ‘Two teachers, three engineers, and a doctor entered inside the room. The two teachers were talking about something.’

Plural agreement is known to be sensitive to two hierarchies: an animacy hierarchy (i.e., human > animate > inanimate) and a definiteness hierarchy (i.e., definite > specific > non-specific) (Smith-Stark 1974, Corbett 2000, see also Alexiadou 2019). WA and Farsi are languages that display this sensitivity in their subject-predicate agreement mechanism. Sigler (1996) reports that in WA, plural agreement on the predicate is obligatory with definite subjects, though both singular and plural agreement is possible when the subject is an indefinite. This is similar to the pattern in NCs, though with a different degree of sensitivity to the definiteness spine. In Farsi, on the other hand, the subject-predicate plural agreement is restricted with respect to the animacy hierarchy. That is, only animate subjects trigger plural agreement on the predicate

interpretation is still possible (see also Sigler 1996, Borer 2005, Bale and Khanjian 2008, 2014, Khanjian 2013). The definite marker can also inflect both forms of NCs in this variety. While it is not obvious what regulates these variations, I do not take it to be internal to the classifier, considering the distribution in WA spoken in Beirut.

<sup>45</sup> The consultants report that the presence/absence of the classifier in Farsi NCs signals a difference in informal vs. formal registers. While NCs with *tā* are linked to more of a colloquial use, the omission of *tā* is found in formal (often written) contexts. Furthermore, a wide scope interpretation is found harder to get in the absence of the classifier compared to the case with the classifier. However, if the names of two students are mentioned as a follow-up to the conditional, the wide scope reading becomes salient: *Age tu kelas-e man be do daneshju komak koni, be nomrat ezafe mikonam. Amir o Hooman.* ‘If you help two students in my class, I will increase your grade. (These students are) Amir and Hooman.’

<sup>46</sup> The plural marker does not surface in the absence of the classifier. We discuss this in the following section. Note also that Farsi has a so-called uniqueness marker, *-(h)e/a*, which can be mistaken with a definite article. Jasbi (2019) claims that *-(h)e/a* introduces a uniqueness implication on the nominal it attaches to. It can optionally appear with both nominals and indefinites. When it appears on a noun, it results in a definite interpretation, while with indefinites, it leads to specificity.

(Ortmann 2002, Alexiadou 2019, Mahootian 1997, Mache 2012). While this does not fully correlate with the restriction in NCs, it is still viable to consider plural marking in NCs as an instance of agreement. It seems that the two hierarchies are split between NCs and the predicate domain in Farsi.<sup>47</sup>

It should be noted that bare plurals are not constrained by specificity/definiteness in WA and Farsi. They can have a non-specific reading, suggesting that the plural marking on bare nouns and the plural marking in NCs must be of different nature.<sup>48</sup> This is illustrated for WA in (79a) and for Farsi in (79b).

- (79) a. turs-ə manug-ner t̂j-en xay-ar-gor  
 outside-DEF child-PL NEG-PRES.3PL play-NEGPART-PROG  
 ‘Children aren’t playing outside.’ (no children, #some children > not)
- b. In ruz-hā gorbe-hā be bāgh-e-man ne-miāy-and.  
 this day-PL cat-PL to garden-EZ-my NEG-IMPERF.come-3PL  
 ‘These days, cats are not coming to my garden.’ (no cats, #some cats > not)

In English, subject-predicate plural agreement is not subject to interpretational restrictions based on animacy and definiteness hierarchies. Likewise, the plural marking in NCs is not associated with any constraints.

Turkish, like WA and Farsi, exhibits sensitivity to the two hierarchies in subject-predicate agreement. That is, inanimate and non-specific subjects cannot trigger plural marking on the predicate (Kornfilt 1997, Bamyacı et al 2014, see also Alexiadou 2019). Although we take Turkish to completely lack number agreement in NCs for simplicity, it is worth noting that globally unique entities, like well-known fictional characters, involve plural marking on the noun, as shown in (80a). Furthermore, there is a peculiar way of forming place names that involves plural marking in NCs, as demonstrated in (80b). Note that plural marking is not possible with NCs with *tane*, aligning with the restriction to indefiniteness in the presence of the classifier.

- (80) a. yedi cüce-ler  
 seven dwarf-PL  
 ‘the seven dwarfs’
- b. beş yüz ev-ler  
 five hundred house-PL  
 ‘the five hundred houses’ (a name of a neighborhood)

Considering these cases as a sub-type of definiteness, which would then be ranked high in the definiteness hierarchy, these idiosyncratic phenomena might be limited instances of number agreement in Turkish NCs (Alexiadou 2019).

In sum, as primarily observed in WA and Farsi, those factors that govern subject-predicate plural agreement are also involved with the choice of plural marking in NCs, providing support for the agreement-based approach.

A natural move that follows from the discussion above is to analyze WA and Farsi classifiers as the overt form of CARD, analogous to the Turkish classifier.<sup>49</sup> Given that plural number agreement is involved in WA and Farsi NCs, we expect the agreement system of English NCs to apply in these languages as well but with one difference: Agreement is only established in the case of definiteness and specificity in WA, and in the case of definiteness in Farsi. Then, while these languages are similar to Turkish in featuring both an overt and

<sup>47</sup> Ionin and Matushansky (2019) give Chadic language Miya as an example where plural agreement on predicates and in NCs is conditioned by animacy. See Schuh (1998) for this generalization.

<sup>48</sup> A reviewer points out that the plural marker in Farsi NCs could potentially be a plural definite article. The fact that bare plurals are not always definite but can have non-specific existential readings discards this possibility.

<sup>49</sup> Sağ (2019) shows that WA and Farsi pattern with Turkish in number marking semantics and kind reference (cf. Sigler 1996, Borer 2005, Bale and Khanjian 2008, 2014, Bale et al 2010, Khanjian 2013, Ionin and Matushansky 2019, Alexiadou 2019, Martí 2020 and Kalomoiros 2021 for WA, and Ghomeshi 2003, 2016, Gebhardt 2009, Mache 2012, Krifka and Modarresi 2016, and Alexiadou 2019 for Farsi). Farsi is also similar to Turkish in having distributional distinctions in terms of countability (Ghomeshi 2003 and Mache 2012). The count-mass distinction in WA, though, requires further study (see Bale and Gillon 2020). However, crucial for our purposes, *had* and *tā* only co-occur with (ontological) count nouns, as in Turkish. Kalomoiros (2021), building on Sağ’s (2022) pseudo-incorporation analysis, argues that *had* establishes a *belong-to* relation between singular kinds and their object-level members in WA NCs. This is what we have eliminated for *tane* in Section 3.2. If Kalomoiros’s analysis is on the right track, *had* could have both a type-fixing and a counting role in NCs in line with Krifka’s view of classifiers. However, this may be at odds with the fact that *had* can co-occur with partitioning nouns, as in (50b).

a covert cardinal head, they depart from Turkish in displaying a somewhat constrained version of number agreement in their NCs.

## 5.2 Support for the Indefiniteness Analysis of Numeral Constructions

I have proposed that the interpretational variation between NCs with and without *tane* in Turkish is due to NCs with the covert CARD being of predicative type in contrast to the intrinsically indefinite nature of NCs with the overt CARD. Intriguingly, the articleless mate of Turkish, i.e., Farsi, exhibits an opposite pattern compared to Turkish NCs.

While NCs with *tā* can receive a definite interpretation that also triggers plural agreement marking on the noun as we have seen above (see (78)), NCs without *tā* require a demonstrative to display a definite-like behavior, as exemplified in (81) (cf. with the Turkish (65)). Since NCs are incompatible with definiteness in the absence of *tā*, number agreement does not arise, either. Therefore, we do not observe plural marking on the lexical NP in NCs without *tā*.

- (81) Polis dar hāl-e barrasi-e marg-e se (tā) moallem-e zan-(\*hā) ast. Rasāne-hā-ye mahali  
 police in investigation-EZ death-EZ three CL teacher-EZ female-PL is channel-PL-EZ local  
 migooyand se \*(tā) zan-hā/ #(in) se (tā) zan ke dar daheye panjah-e zendegi-e  
 say three CL woman-PL this three CL woman that in fifties-EZ life-EZ  
 khod budand.  
 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.’

This mirror image pattern in interpretation reveals that Turkish is not the only language that allows predicative NCs besides the default argumental one. This discovery also highlights that the choice of which CARD head takes the indefinite/predicative role is language-specific and unrelated to the morphological shape of CARD itself. In Farsi, NCs with the covert CARD represent the argumental form associated with a choice function while NCs with *tā* are devoid of the indefinite force and hence have a predicative denotation.<sup>50</sup>

Our account predicts that the phenomenon of alternating predicative NCs alongside argumental NCs can only be visible in languages without articles. That is, in a language that differs from Turkish and Farsi in featuring a definite article, we expect both forms to be compatible with the definite article and we do not anticipate any form of NCs to freely allow definite interpretations. WA confirms these predictions, where both forms of NCs require the attachment of the definite article for a definite interpretation. Even if one form of CARD yields predicative NCs (represented as  $CARD_{pred}$  below) and the other form yields argumental NCs similar to Turkish and Farsi, both forms remain amenable to co-occurring with the definite article (via *ident* in the case of argumental NCs). This will rule out the application of  $\iota$  on the predicative form due to the Blocking Principle. More precisely, the existence of an overt definite article will never allow  $\iota$  type-shifting to surface even if the NC itself does not inherently prevent it.

Table 2 summarizes how covert *iota* type-shifting is blocked/allowed in WA, Turkish, and Farsi.

<sup>50</sup> Note on complex numerals: As mentioned in fn 20, Ionin and Matushansky (2006, 2019) derive complex numerals compositionally. Their view, when applied to our analysis, would require recurring cardinal heads, where a NC combines with another CARD and a numeral in its specifier, as in [[three CARD] [hundred CARD apples]] (cf. Rothstein 2017). However, this composition is ruled out under our analysis with argumental NCs because CARD requires a property as its argument, rendering it unsuitable for iteration in a complex numeral. This prediction is borne out in argumental NCs in Turkish, where the CARD head *tane* cannot be iterated and only follows the numeral closest to the noun, as in *üç (\*tane) yüz (tane) elma* ‘three hundred apples’. Recurring cardinal heads are not attested in predicative NCs, either, as evidenced by NCs with *tā* in Farsi, where the CARD head cannot be multiplied within a numeral complex: *si (\*tā) sad (tā) sib* ‘three hundred apples’. For this reason, I directly enforced an atomic property requirement on the semantics of CARD instead of Ionin and Matushansky’s constraint in (39), where numerals could combine with properties with individuals of the same cardinality. Given that a predicative NC such as *sad tā sib* ‘hundred apples’ is not an atomic property, it cannot combine with another CARD head under our approach. I suggest that complex numerals are derived by covert arithmetic operators —multiplication and addition—the result of which is a complex number that feeds the argument slot of the cardinality function (see also Turgay 2022 for a syntactic account of this view). E.g., *three hundred* is derived through a covert multiplication operator that takes two numbers and multiplies them.



The type of NCs	with THE	without THE	
	WA	Turkish	Farsi
NCs with $CARD_f$	N/A	N/A	N/A
NCs with $CARD_{pred}$	(if exists) ✗ due to BP	✓	✓

Table 2: The blockage/possibility of *iota* in languages with an overt and a covert CARD

To wrap up the discussion so far, the interpretational possibilities of NCs in the languages discussed in this paper are summarized in Table 3. Recall that the categorization *presumably*  $CARD_\emptyset$  is used for obligatory classifier languages, as stated in Section 4.2.

	$CARD_\emptyset$		Presumably $CARD_\emptyset$	$CARD_\emptyset$ & $CARD_{overt}$	
	French	English	Yi	WA	
<b>Languages with THE</b>					
indefinite NCs	✓	✓	✓	✓	
definite NCs with THE	✓	✓	✓	✓	
definite NCs with $\iota$	✗	✗	✗	✗	
<b>Languages without THE</b>	<b>Russian</b>		<b>Mandarin</b>	<b>Turkish</b>	<b>Farsi</b>
indefinite NCs	✓		✓	✓	✓
definite NCs with Dem	✓		✓	✓	✓
definite NCs with $\iota$	✗		✗	✗ with <i>tane</i> ✓ with $CARD_\emptyset$	✗ with $CARD_\emptyset$ ✓ with $t\bar{a}$

Table 3: The cross-linguistic interpretation of NCs

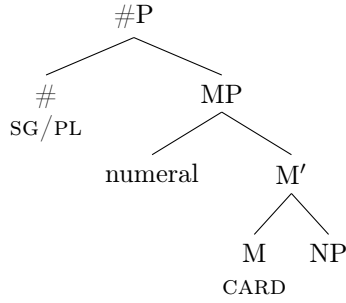
Languages with an overt definite article have the same interpretational distribution regardless of whether they have one form of CARD or two. In the articleless group, though, two languages stand out by allowing definite interpretations via *iota*. These languages are the ones that feature both an overt and a covert CARD in their NCs, highlighting the connection between the indefiniteness of NCs and cardinality. This connection then not only turns out to be useful in capturing intra-language alternations but also sheds light on cross-linguistic variations in the interpretation of NCs.

## 6 Revisiting the Agreement Approach

We have analyzed NCs as argumental expressions departing from the widely accepted view of NCs which treats them as predicates. In Section 4.1, we have discussed an initial take on the semantics of number agreement involved in English NCs, which assumed a predicative semantics for NCs. We now need to revise our account in a way that reflects on the argumental nature of NCs.

Let us recall Scontras’ (2014, 2022) analysis of number marking in English NCs, repeated below. In a nutshell, a number head projects above NCs, i.e., Measure Phrases, either hosting SG or PL features, which are identity functions on properties. SG is defined on the *one-ness* of a quantity-uniform property, which, in the case of cardinal measure phrases, is determined by measurement in terms of cardinality. That is, SG is dependent on whether every member of the set denoted by the MP evaluates to 1 with respect to this measure. PL, on the other hand, only surfaces in cases where the one-ness requirement fails to be met.

(82) Generalized MP



(83) English # heads:

- a.  $\llbracket \text{SG} \rrbracket = \lambda P : \forall \mu \forall x \in P [QU_\mu(P) \rightarrow \mu(x) = 1]. P$
- b.  $\llbracket \text{PL} \rrbracket = \lambda P. P$
- c.  $QU_\mu(P) = 1$  iff  $\forall x \forall y [P(x) \wedge P(y) \rightarrow \mu(x) = \mu(y)]$

To modify the analysis, we simply adjust the type of the number features. Instead of being identity functions on properties, they should take individuals as arguments, as originally proposed in Sauerland (2003), an account which Scontras builds on. Drawing on the insights from the two, I present the following semantics for English # heads:

(84) English # heads:

- a.  $\llbracket \text{SG} \rrbracket = \lambda x : \mu(x) = 1. x$
- b.  $\llbracket \text{PL} \rrbracket = \lambda x. x$

SG takes an  $e$ -type MP and checks whether the individual evaluates to 1 with respect to a measure, which is cardinality in NCs. If not, PL surfaces.

(85) Number marking with *one*:

- a.  $\checkmark \llbracket \text{SG one CARD}_f \text{ book} \rrbracket = f(\{a, b, c\})$
- b.  $\times \llbracket \text{PL one CARD}_f \text{ book} \rrbracket = \text{failure due to Maximize Presupposition}$

(86) Number marking with *two*:

- a.  $\times \llbracket \text{SG two CARD}_f \text{ book} \rrbracket = \text{presupposition failure}$
- b.  $\checkmark \llbracket \text{PL two CARD}_f \text{ book} \rrbracket = f(\{a \oplus b, b \oplus c, a \oplus c\})$

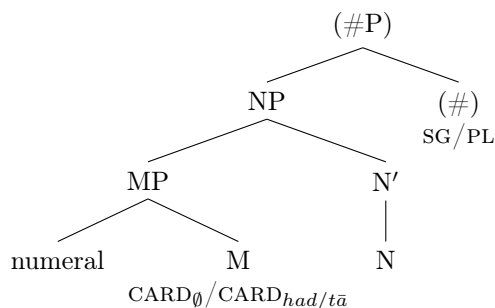
We need to assume that predicative type-shifting of NCs (via *ident*) applies to #P in order to capture number marking of both argumental and predicative uses of NCs. This means that a D head, let it be a definite determiner, demonstrative or quantificational determiner, projects above #P (cf. Sauerland 2003).

While the #-projection always occurs in NCs of English-like languages, agreement is only established in the case of definiteness and specificity in WA, and in the case of definiteness in Farsi, as discussed above. Identifying the exact syntactic mechanism that ensures the restricted realization of number agreement in these languages is beyond the scope of this paper. However, it is necessary to make a note about number marking in Farsi NCs, which is only possible in the presence of  $t\bar{a}$ . Recall that NCs with  $t\bar{a}$  differ from the form with the covert CARD in being predicative, with which definiteness is possible through covert  $\iota$  type-shifting. Number features, being functions on individual type-expressions, should then be able to take a NC with  $t\bar{a}$  as an argument only after its shift to an argumental type is achieved through  $\iota$ .<sup>51</sup>

<sup>51</sup> I have argued that number agreement does not arise in Turkish NCs, excluding the cases in (80). Scontras (2014, 2022) suggests a distinct strategy of number marking in Turkish NCs, where the one-ness presupposition of SG is sensitive to relative atomicity. Martí (2020) applies a similar logic, considering Turkish number marking to be sensitive to [+/-minimal] features. Quantized predicates such as NCs are relatively atomic/[+minimal] because every member in their denotation is a minimal member, and hence an atom, relative to the predicate in question. These accounts analyze WA as a language where both the English and Turkish number marking systems are implemented. This idea, obviously, can be extended to Farsi as well. However, our #-heads are no longer functions on properties, making it challenging to maintain such an account. Furthermore, it would be difficult to explain what triggers a switch from the Turkish system to the English system in case of specificity/definiteness.

The structural construal adopted in Scontras’ analysis, i.e., the structure where the NP is a complement to the M head, shown in (82), predicts number agreement to fall on *had* in WA and *tā* in Farsi. Recall that under this account, number morphology is assumed to surface on the closest overt head to #P. However, plural morphology always appears on the noun in these languages, not on the classifiers when present. This problem disappears if WA and Farsi NCs have a structure similar to what is adopted for Turkish NCs because then the closest head to #P will never be the M head; it will always be the head of the lexical NP. Therefore, I assume the structure in (87) for WA and Farsi NCs, though I leave the exploration of this remedy for future research (see also Mache 2012 for Farsi, cf. Bale and Khanjian 2014 for WA).

(87) The structure of WA and Farsi NCs



Before concluding, I should emphasize that in Sauerland’s and Scontras’s accounts, plural marking on all nominal projections is supplied by the # head. We have seen that plural marking on WA and Farsi bare nouns is not subject to the restrictions observed with the plural marker in their NCs (see (79)). Therefore, I utilize the # head only to account for the semantics of number agreement in NCs (perhaps also for subject-predicate agreement). I take the plural marking on a numeral-less lexical NP in WA, Farsi, and Turkish as the spell-out of the \* operator, leaving the question open whether this is also the case in English.

## 7 Conclusion

The primary goal of this study has been to demonstrate that NCs are argumental expressions with intrinsically indefinite characteristics. Taking the Turkish counting system as the center of the investigation, I have argued that the indefiniteness of NCs across languages is essentially a contribution of a cardinal head, which typically surfaces covertly but also has an overt exponent in some languages.

We have examined four key properties of Turkish NCs: (i) the consistent use of the unmarked noun form with all numerals despite a systematic number marking system, (ii) the optional realization of the classifier *tane*, (iii) the exceptional behavior of NCs without *tane* allowing both definite and indefinite interpretations, in contrast to NCs with *tane*, which are exclusively indefinite, complying with the general behavior of NCs across languages. And finally, (iv) the non-optional aspect of the classifier, i.e., the fact that the presence/absence of *tane* makes a difference in the interpretation of NCs.

We first considered the properties in (i) and (ii) with an aim to understand the optional aspect of the Turkish classifier and cross-linguistic variation in patterns of number marking. Then, shifting our attention to the properties in (iii) and (iv), we examined the (in)definiteness puzzle of Turkish NCs and discussed its consequences for the interpretation of NCs across languages.

The first part of the analysis compared *tane* with obligatory classifiers of Mandarin-like languages, highlighting its distinct characteristics. Drawing on Saĝ’s (2019, 2022) account of the Turkish number marking semantics and examining the count-mass distinction in Turkish, we have concluded that the noun in Turkish NCs denotes a semantically singular property whether the classifier is present or not. To explain the number marking variation in English and Turkish NCs, I have pursued an agreement-based approach, drawing on evidence from two more optional classifier languages, WA and Farsi. I have proposed a semantic account of number agreement in NCs, integrating Ionin and Matushansky’s (2006, 2019) view of numerals into Sauerland’s (2003) and Scontras’s (2014, 2022) analyses of number marking in English. In short, while some languages, including Turkish, morphologically exhibit the semantic singularity of the noun in their NCs, some others, including English, disguise this by employing number agreement in NCs. The agreement

account adopted here posits a cardinal head projection to treat number marking uniformly with cardinal numerals and measure terms. Building on this, I have proposed that *tane* denotes a counting function as an overt spell-out of this head in Turkish. While English and supposedly Mandarin-like languages involve only a covert form of CARD, Turkish, Farsi, and WA have both an overt and a covert exponent of it.

The second half of my analysis has demonstrated a connection between CARD and inherently indefinite characteristics of NCs. Drawing on Jiang's (2012) take on the issue, I have proposed that CARD uniformly comes with a built-in choice function variable, and hence yields individual type NCs. In line with Dayal (2003), I have taken the predicative interpretation of NCs to be derived as a repair strategy, which is invoked only when the structure requires predicative NCs. A particularly intriguing finding is that inherently predicative NCs may exist alongside inherently argumental NCs in languages that allow both overt and covert forms of CARD. Illustrating this point, we identify Turkish and Farsi as languages that demonstrate such coexistence, though in a mirror image pattern. In Turkish, the covert CARD, lacking the choice function, has the role of deriving predicative NCs, while *tane* functions with an indefinite force, similar to the cardinal heads in other languages. Farsi, on the other hand, features *tā* as the predicative form, while the covert CARD corresponds to the canonically indefinite one. Since both languages lack an overt definite determiner, this variation manifests as an enrichment in interpretation, permitting both definite and indefinite readings with the form of CARD that is liberated from the choice function. As a result, this leads to a distinctive pattern for the interpretation of Turkish and Farsi NCs. Our analysis predicts that such enrichment in interpretation cannot be observed in a language with an overt definite article, even if that language employs both intrinsically predicative and indefinite forms of NCs. Our investigation of Western Armenian NCs validates this prediction.

If our discussion is on the right track, then we are one step closer to a general understanding of cardinal semantics. The next task is to explore the overarching question of why counting constructions, specifically cardinality, inherently intertwine with indefiniteness.

## References

- Abusch D (1993) The scope of indefinites. *Natural Language Semantics* 2(2):83–135
- Aikhenvald AY (2000) *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press UK
- Alexiadou A (2019) Morphological and semantic markedness revisited: The realization of plurality across languages. *Zeitschrift für Sprachwissenschaft* 38(1):123–154
- Alexiadou A, Haegeman L, Stavrou M (2007) *Noun Phrase in the Generative Perspective*. de Gruyter Mouton, Berlin & New York
- Arslan-Kechriotis ZC (2009) Referentiality in Turkish: NP/DP. In: Ay S, Aydın Özgür, Ergenç I, Gökmen S, İşsever S, Peçenek D (eds) *Essays on Turkish Linguistics: Proceedings of the 14th International Conference on Turkish Linguistics*, Harrassowitz Verlag, Wiesbaden, pp 83–92
- Aygen-Tosun G (1999) Specificity and subject-object positions/scope interactions in Turkish. In: *Proceedings of the Conference on Turkic Linguistics at Manchester University*
- Bale A, Coon J (2014) Classifier are for numerals, not for nouns: consequences for the mass/count distinction. *Linguistic Inquiry* 45(4):695–707
- Bale A, Gillon B (2020) Re-examining the mass-count distinction. In: Moltmann F (ed) *Mass and Count in Linguistics, Philosophy, and Cognitive Science*, John Benjamins Publishing Company, pp 13–36
- Bale A, Khanjian H (2008) Classifiers and number marking. In: Friedman T, Ito S (eds) *Proceedings of SALT 18*, Cornell University, Ithaca, NY, pp 73–89
- Bale A, Khanjian H (2014) Syntactic complexity and competition: the singular-plural distinction in Western Armenian. *Linguistic Inquiry* 45(1):1–26
- Bale A, Khanjian H, Gagnon M (2010) Cross-linguistic representations of numerals and number marking. In: *Proceedings of SALT 20*, pp 1–15
- Bale A, Coon J, Arcos N (2019) Classifiers, partitions, and measurements: Exploring the syntax and semantics of sortal classifiers. *Glossa: a journal of general linguistics* 4(1)(77):1–30
- Bale AC, Barner D (2009) The interpretation of functional heads: Using comparatives to explore the mass/count distinction: Articles. *Journal of Semantics* 26(3):217–252
- Bamyacı E, Häußler J, Kabak B (2014) The interaction of animacy and number agreement: An experimental investigation. *Lingua* 148:254–277

- Barner D, Snedeker J (2005) Quantity judgments and individuation: evidence that mass nouns count. *Cognition* 97(1):41–66
- Barwise J, Cooper R (1981) Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2):159–219, URL <http://www.jstor.org/stable/25001052>
- Bayırlı IK (2012) On Suffixhood and Verbalness: A Mirror Theoretic Approach. PhD thesis, Bogazici University
- Bayırlı IK (2017) The Universality of Concord. Ph.d. thesis, Massachusetts Institute of Technology
- Bennett MR (1974) Some Extensions of a Montague fragment of English. PhD thesis, UCLA
- Bliss H (2004) The semantics of the bare noun in Turkish. In: Mezhevich I, Dobrovolsky MB (eds) *Calgary Papers in Linguistics*, vol 25, University of Calgary, pp 1–65
- Borer H (2005) *Structuring Sense: Volume 1: In Name Only*. Oxford University Press, Oxford
- Bošković Ž, Şener S (2014) The Turkish NP. In: *Crosslinguistic Studies on Noun Phrase Structure and Reference: Syntax and Semantics: 39*, Brill, Leiden, The Netherlands
- Bunt HC (1985) *Mass terms and model-theoretic semantics*. Cambridge University Press
- Carlson GN (1977) Reference to Kinds in English. PhD thesis, University of Massachusetts, Amherst
- Carpenter B (1998) *Type-Logical Semantics*. MIT Press, Cambridge, MA
- Cheng LLS, Sybesma R (1999) Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4):509–542
- Chierchia G (1998) Reference to Kinds across Language. *Natural Language Semantics* 6(4):339–405
- Chierchia G (2010) Mass nouns, vagueness and semantic variation. *Synthese* 174(1):99–149
- Chierchia G, Heim I (2017) Class notes of ‘Topics in Semantics’ taught in Spring 2017
- Condoravdi C (1994) Descriptions in context. Ph.d. thesis, Yale University
- Corbett GG (2000) *Number*. Cambridge University Press, Cambridge
- Danon G (2012) Two structures for numeral-noun constructions. *Lingua* 122:1282–1307
- Dayal V (2003) Bare nominals: non-specific and contrastive readings under scrambling. In: Karimi S (ed) *Word Order and Scrambling*, Oxford: Blackwell Publishers, pp 67–90
- Dayal V (2004) Number marking and indefiniteness in kind terms. *Linguistics and Philosophy* 27:393–450
- Dayal V (2013) On the existential force of bare plurals across languages. In: Caponigro I, Cecchetto C (eds) *From Grammar to Meaning: The Spontaneous Logicality of Language*, Cambridge University Press, pp 49–80, DOI 10.1017/CBO9781139519328.006
- Dayal V (2014) Bangla plural classifiers. *Language and Linguistics* 15(1)
- Deal AR (2017) Countability distinctions and semantic variation. *Natural Language Semantics* 25(2):125–171
- Diesing M (1992) Indefinites. No. Vol. 20 in *Linguistic Inquiry Monographs*, MIT Press
- van der Does J (1992) *Applied Quantifier Logics*. Ph.d. thesis, University of Amsterdam
- Doetjes J (1997) *Quantifiers and Selection*. Ph.d. thesis, University of Leiden
- Doetjes J (2019) 2. count/mass distinctions across languages. In: Heusinger K, Maienborn C, Portner P (eds) *Semantics - Typology, Diachrony and Processing*, De Gruyter Mouton, Berlin, Boston, pp 29–56
- Enç M (1991) The semantics of specificity. *Linguistic Inquiry* 22(1):1–25, DOI 10.2307/4178706
- Farkas D, de Swart H (2010) The semantics and pragmatics of plurals. *Semantics & Pragmatics* 3(6):1–54
- Farkas DF (1981) Quantifier scope and syntactic islands. In: *Proceedings of the 17th Chicago Linguistics Society (CLS 17)*, pp 59–66
- Fodor JD, Sag IA (1982) Referential and quantificational indefinites. *Linguistics and Philosophy* 5(3):355–398
- Gebhardt L (2009) *Numeral Classifiers and the Structure of DP*. Ph.d. thesis, Northwestern University
- Ghomeshi J (2003) Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica* 57(2):47–74
- Ghomeshi J (2016) Markedness and bare nouns in Persian. In: Karimi S, Samiian V, Stilo DL (eds) *Aspects of Iranian Linguistics*, Cambridge Scholars Publishing, pp 85–111
- Giusti G (2002) The functional structure of noun phrases: A bare phrase structure approach. In: Cinque G (ed) *Functional structure in DP and IP*, Oxford University Press, Oxford, pp 54–90
- Göksel A (1997) Linearity, Focus and Postverbal Position in Turkish. In: Csato E, Johanson L (eds) *Proceedings of the 7th International Conference on Turkish Linguistics*, Wiesbaden:Harrassowitz Verlag
- Göksel A, Kerslake C (2005) *Turkish: A Comprehensive Grammar*. Routledge, London NY
- Göksel A, Özsoy S (2003) dA: a focus/topic associated clitic in Turkish. *Lingua* 113:1143–1167
- Görgülü E (2010) On the Number System and the Count/Mass Distinction in Turkish. In: *Proceedings of the 2010 annual conference of the Canadian Linguistic Association*, pp 1–15

- Görgülü E (2012) *Semantics of Nouns and the Specification of Number in Turkish*. Ph.d. thesis, Simon Fraser University
- Greenberg JH (1990) Numeral classifiers and substantival number. In: Denning K, Kemmer S (eds) *On language: Selected writings of Joseph H. Greenberg*, Stanford University Press, pp 166–193
- Grimm S (2012) *Number and individuation*. Ph.d. thesis, Stanford University
- Guisti G (1996) Is there a FocusP and a TopicP in the Noun Phrase structure? *Working Papers in Linguistics* 6(2):105–128
- Harbour D (2014) Paucity, abundance and the theory of number. *Language* 90:158–229
- Heim I (1982) *The Semantics of Definite and Indefinite Noun Phrases*. Ph.d. thesis, University of Massachusetts, Amherst
- Heim I (1991) Artikel und definitheit. In: Stechow AV, Wunderlich D (eds) *Semantics: An International Handbook of Contemporary Research*, de Gruyter, Berlin, pp 487–535
- Ionin T, Matushansky O (2006) The composition of complex cardinals. *Journal of Semantics* 23(4):315–360
- Ionin T, Matushansky O (2019) *Cardinals: The Syntax and Semantics of Cardinal-containing Expressions*. MIT Press, Cambridge, MA
- Jasbi M (2019) The suffix that makes Farsi nouns unique. To appear in *Advances in Iranian Linguistics*
- Jiang L (2012) *Nominal arguments and language variation*. PhD thesis, Harvard University
- Jiang LJ (2018) *Definiteness in nuosu yi and the theory of argument formation*. *Linguistics and Philosophy* 41(1):1–39
- Kalomoiros A (2021) Bare Singulars and Pseudo-incorporation in Western Armenian. In: *Proceedings of SALT 31*, pp 365–384
- Kamali B, Karvovskaya L (2013) “Also” in Turkish and Ishkashimi. In: *Proceedings of Workshop on Altaic Formal Linguistics (WAFL) 8*, MITWPL
- Kelepir M (2001) *Topics in Turkish syntax: Clausal Structure and Scope*. PhD thesis, Massachusetts Institute of Technology
- Kelepir M (2003) Olmak, değil, var ve yok. In: *Proceedings of the XVIth Dilbilim Kurultayı (National Linguistics Conference)*, Hacettepe University, Ankara
- Kennelly SBD (1994) Caseless existential DPs in Turkish. In: *Paris 8 Working Papers in Linguistics*
- Khanjian H (2013) *(Negative) Concord and Head Directionality In Western Armenian*. Ph.d. thesis, Massachusetts Institute of Technology
- Khrizman K (2016) Functional unit classifiers in (non)-classifier Russian. *Baltic International Yearbook of Cognition, Logic and Communication* 11:1–48
- Khrizman K, Landman F, Lima S, Rothstein S, Brigitta R Schvarcz (2015) Portion Readings are Count Readings, not Measure Readings. In: *Proceedings of the 20th Amsterdam Colloquium*
- Kim J (2009) *Plural Marking in Classifier Languages*. VDM Verlag, Saarbrücken
- Kornfilt J (1996) On some copular clitics in Turkish. In: Alexiadou A, Fuhrhop N, Law P, Loehken S (eds) *ZAS Papers in Linguistics, Zentrum für Allgemeine Sprachwissenschaft 6*, Berlin, pp 96–114
- Kornfilt J (1997) *Turkish*. Routledge, London
- Kornfilt J (2005) Agreement and its placement in Turkic nonsubject relative clauses. In: Cinque G, Kayne R (eds) *The Oxford Handbook of Comparative Syntax*, Oxford University Press, New York, pp 513–542
- Kornfilt J (2017) DP versus NP: a Cross-linguistic typology? In: *Studies in Japanese and Korean Historical and Theoretical Linguistics and Beyond*, Brill, Leiden, The Netherlands
- Kouneli M (2019) Plural marking on mass nouns: Evidence from greek. *Gender and noun classification* pp 233–49
- Kratzer A (1998) Scope or pseudo-Scope: are there wide-scope indefinites? In: Rothstein S (ed) *Events in Grammar*, Kluwer, Dordrecht, pp 163–196
- Krifka M (1989) Nominal reference, temporal constitution and quantification in event semantics. In: Bartsch JvBR, Boas PvE (eds) *Semantics and Contextual Expression*, Foris Publication, Dordrecht, pp 75–115
- Krifka M (1995) Common nouns: a contrastive analysis of Chinese and English. In: *The Generic Book*, University of Chicago Press, pp 398–411
- Krifka M (2001) Non-novel indefinites in adverbial quantification. In: Smith C (ed) *Workshop proceedings: Non-narrative Discourse*. Texas Linguistics Forum, Dept. of Linguistics
- Krifka M (2003) Bare NPs: kind-referring, indefinites, both, or neither? In: *Proceedings of SALT 13*, pp 180–203

- Krifka M, Modarresi F (2016) Number neutrality and anaphoric update of pseudo-incorporated nominals in Persian (and weak definites in English). In: Proceedings of SALT 26, pp 874–891
- Kroch A (1975) The Semantics of Scope of English. Ph.d. thesis, Massachusetts Institute of Technology
- Kurafuji T (1999) Japanese Pronouns in Dynamic Semantics: The Null/overt Contrast. PhD thesis, Rutgers University
- Landau I (2016) Dp-internal semantic agreement: A configurational analysis. *Natural Language & Linguistic Theory* 34(3):975–1020
- Landman F (2003) Predicate-argument mismatches and the adjectival theory of indefinites. In: Coene M, D’hulst Y (eds) *From NP to DP, vol. 1, The Syntax and Semantics of Noun Phrases*, John Benjamins, Amsterdam, pp 211–237
- Landman F (2004) *Indefiniteness and the Type of Sets*. Oxford: Blackwell.
- Landman F (2011) Count nouns, mass nouns, neat nouns, mess nouns. In: Partee B, Glanzberg M, Skilters J (eds) *Formal Semantics and Pragmatics. Discourse, Context and Models*, New Prairie Press, Manhattan, KS
- Lewis GL (2000) *Turkish Grammar*, 2nd edn. Oxford University Press
- Li YHA (1999) Plurality in a classifier language. *Journal of East Asian Linguistics* 8(1):75–99
- Link G (1983) The logical analysis of plurals and mass terms: a lattice-theoretic approach. In: Portner P, Partee BH (eds) *Formal Semantics - the Essential Readings*, Blackwell, pp 127–147
- Link G (1987) Generalized quantifiers and plurals. In: *Generalized Quantifiers: Linguistic and Logical Approaches*, *Studies in Ling. & Philos.*: 31, Reidel, pp 151–180
- Mache A (2012) *Numeral Classifiers in Persian*. Lincom Europa
- Mahootian S (1997) *Persian*. Routledge, London and New York
- Mahowald K, Graff P, Hartman J, Gibson E (2016) Snap judgments: A small n acceptability paradigm (snap) for linguistic acceptability judgments. *Language* 92(3):619–635
- Martí L (2020) Numerals and the theory of number. *Semantics & Pragmatics*
- Martí L (2022) Zero n: Number features and  $\perp$ . *Natural Language Semantics* 30(2):215–237
- Massam D (2001) Pseudo noun incorporation in Niuean. *Natural Language & Linguistic Theory* 19(1):153–197
- ter Meulen A (1981) An intensional logic for mass terms. *Philosophical Studies* 40:105–125
- Mittendorf I, Sadler L (2005) Numerals, nouns and number in Welsh NPs. . In: Butt M, King TH (eds) *Proceedings of the LFG ’05 Conference*, CSLI Publications, Stanford: CA, pp 294–312
- Montague R (1974) *Formal Philosophy; Selected Papers of Richard Montague*. New Haven: Yale University Press
- Ortmann A (2002) Economy-based splits, constraints, and lexical representations. In: Kaufmann I, Stiebels B (eds) *More than words: A festschrift for Dieter Wunderlich*, Berlin: Akademie Verlag, pp 147–178
- Ortmann A (2019) Where plural refuses to agree: feature unification and morphological economy. *Acta Linguistica Hungarica* 47(1-4):249–288
- Öztürk B (2005) Case, Referentiality, and Phrase Structure. *Linguistics today*, J. Benjamins Publishing Company
- Partee BH (1987) Noun phrase interpretation and type-shifting principles. In: *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Groningen-Amsterdam Studies in Semantics: 8, Foris, pp 115–143, URL <https://login.proxy.libraries.rutgers.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=mzh&AN=1988011697&site=eds-live>
- Pelletier FJ, Schubert LK (1989) *Mass Expressions*, Springer Netherlands, Dordrecht, pp 327–407
- Phan T, Chierchia G (2022) Identifying (in)definiteness in vietnamese noun phrase. *Journal of the Southeast Asian Linguistics Society* 15(2):27–49
- Pires de Oliveira R, Rothstein S (2011) Bare singular noun phrases are mass in brazilian portuguese. *Lingua* 121(15):2153–2175
- Quine WVO (1960) *Word and Object*. MIT Press
- Reinhart T (1997) Quantifier scope: how labor is divided between or and Choice Function. *Linguistics and Philosophy* 20:335–397
- Renans A, Tsoulas G, Folli R, Ketrez N, Tieu L, de Vries H, Romoli J (2017) Turkish plural nouns are number-neutral: experimental data. In: *The Proceedings of the Amsterdam Colloquium 2017*

- Renans A, Sağ Y, Ketrez FN, Tieu L, Tsoulas G, Folli R, de Vries H, Romoli J (2020) Plurality and Cross-linguistic Variation: An Experimental Investigation of the Turkish Plural. *Natural Language Semantics* 28:307–342
- Rothstein S (2010) Counting and the mass/count distinction. *Journal of Semantics* 27(3):343–397
- Rothstein S (2017) *Semantics for Counting and Measuring. Key Topics in Semantics and Pragmatics*, Cambridge University Press, DOI 10.1017/9780511734830
- Ruys E (1992) *The scope of indefinites*. PhD thesis, Utrecht University
- Sağ Y (2018) *The Semantics of Turkish Numeral Constructions*. In: *Proceedings of Sinn und Bedeutung 22*
- Sağ Y (2019) *The Semantics of Number Marking: Reference to Kinds, Counting, and Optional Classifiers*. Ph.d. thesis, Rutgers University
- Sağ Y (2020) *The Curious Case of Measure Semantics*. In: *Proceedings of CLS 55*, pp 351–364
- Sağ Y (2022) Bare singulars and singularity in Turkish. *Linguistics and Philosophy* 45(4):741–793
- Sağ Y, Demirok Ö (2023) Getting even without “even” in Turkish’. In: *The Proceedings of the 53rd Annual Meeting of the North East Linguistic Society (NELS 53)*
- Saha A (2023) Reference to kinds: the perspective from Bangla. In: *Proceedings of West Coast Conference on Formal Linguistics (WCCFL 41)*
- Sanches M, Slobin L (1973) Numeral classifiers and plural marking: An implicational universal. *Working Papers in Language Universals* 11:1–22
- Sauerland U (2003) A new semantics for number. In: *Proceedings of SALT 13*, pp 258–275
- Sauerland U, Anderssen J, Yatsushiro K (2005) The plural is semantically unmarked. *Linguistic evidence* 4:409–430
- Scha R (1981) Distributive, collective and cumulative quantification. In: Groenendijk JAG, Janssen TMV, Stokhof MBJ (eds) *Formal Methods in the Study of Language, Part 2*, Mathematisch Centrum, pp 483–512
- Schroeder C (1992) On the discourse function of the so-called numeral classifier TANE and on the count-mass distinction in Turkish. In: Konrot A (ed) *Modern Studies in Turkish Linguistics*, Anadolu University, pp 259–283
- Schuh RG (1998) *A Grammar of Miya*. University of California Press, Berkeley
- Schvarcz BR, Rothstein S (2017) Hungarian classifier constructions, plurality and the mass-count distinction. In: van der Hulst H, Lipták A (eds) *Approaches to Hungarian 15: Papers from the 2015 Leiden Conference*, John Benjamins, Amsterdam, pp 157–182
- Scontras G (2014) *The Semantics of Measurement*. Ph.d. thesis, Harvard University.
- Scontras G (2022) On the semantics of number morphology. *Linguistics and Philosophy*
- Sigler M (1996) *Specificity and Agreement in Standard Western Armenian*. Ph.d. thesis, Massachusetts Institute of Technology
- Smith-Stark TC (1974) The plurality split. In: La Galy MW, Fox RA, Bruck A (eds) *Papers from the 10th regional meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago: IL, pp 657–671
- Spector B (2007) Aspects of the pragmatics of plural morphology: on higher-order implicatures. In: Sauerland U, Stateva P (eds) *Presupposition and Implicature in Compositional Semantics*, Palgrave Macmillan UK, London, pp 243–281
- de Swart HD, Winter Y, Zwarts J (2007) Bare nominals and reference to capacities. *Natural Language & Linguistic Theory* 25:195–222, DOI 10.1007/s11049-006-9007-4
- Szabolcsi A (2015) What do quantifier particles do? *Linguistics and Philosophy* 38(2):159–204, DOI 10.1007/s10988-015-9166-z
- Tat D (2011) APs as reduced relatives: The case of Bir in some varieties of Turkic. In: *Proceedings of the 7th Workshop on Altaic Formal Linguistics (WAFL7)*, pp 73–89
- Tsoulas G (2009) On the grammar of number and mass terms in Greek. In: Halpert C, Hartman J, Hill D (eds) *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT*. MIT Working Papers in Linguistics 57, MITWPL, pp 131–146
- Turgay T (2022) A minimalist account of numerals. *Dilbilim Araştırmaları Dergisi* 33(2):111–144
- Underhill R (1976) *Turkish Grammar*. MIT Press
- Verkuyl H (1993) *A Theory of Aspectuality: The Interaction between Temporal and Atemporal Structure*. Cambridge University Press, Cambridge
- von Stechow K (2002) Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 19(3):245–274



- von Heusinger K, Kornfilt J (2017) Partitivity and case marking in Turkish and related languages. *Glossa: A Journal of General Linguistics* 2(1:20):1–40
- Wągiel M (2018) Subatomic quantification. Ph.d. thesis, Masaryk University
- Wechsler S, Zlatić L (2003) *The Many Faces of Agreement*. CSLI Publications, Stanford: CA
- Winter Y (1997) Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467
- Yang R (2001) Common nouns, classifiers, and quantification in Chinese. PhD thesis, Rutgers University
- Zabbal Y (2005) The syntax of numeral expressions. Ms., University of Massachusetts at Amherst
- Zidani-Eroğlu L (1997) Indefinite Noun Phrases in Turkish. PhD thesis, University of Wisconsin-Madison
- Zweig E (2009) Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32(4):353–407

## APPENDIX

According to the native speakers that an anonymous reviewer consulted (including the reviewer), the continuation in the second sentence in (5) is unacceptable, regardless of the form of the NC. In this appendix, I delve into this issue and include the mean ratings (ranging from 1 to 7) assigned by 15 native speakers for the sentence in (5). For the continuation with NCs with *tane* the mean rating is 1.63, while for the version with NCs without *tane* the mean rating is 6. Below, I also provide further examples along with their mean ratings and detail the informal judgment seeking task I conducted to collect the acceptability ratings.

The examples presented in the paper, illustrating the contrast between the two forms of NCs in terms of anaphoric definiteness, feature NCs in the subject position in the continuation sentence, one in null nominative form (5) and one as an embedded subject with genitive case (65). Additionally, I tested one other example where the NC appeared in the subject position (in null nominative form) and two examples where the NC appeared in the object position with accusative case. To mitigate potential effects of the form of NCs introduced in the first sentence (with or without *tane*) on the choice of the form of the NC in the continuation sentence, I introduced the NCs in different forms in the first sentences. I tested the examples in a randomly mixed order, ensuring that examples with the same type of NC in the continuation sentence were not presented consecutively. I provided the following instructions for each example:

**Turkish:** Şimdi sana iki cümleden oluşan minik bir hikaye vereceğim. İki cümlede de bazı X’lerden bahsedeceğim (X filled with yazarlardan/şarkıcılardan/begesellerden for the instruction of each narrative). Senden ilk cümlede bahsettiğim X’lerin ikinci cümlede bahsettiklerimle aynı olduğunu farz etmeni istiyorum. Bu durumda hikayenin doğallığını 1’den 7’ye kadar derecelendirmeni istiyorum. 1 hiç doğal değil, 7 çok doğal.

**English translation:** I will now present a brief narrative comprising two sentences. In both sentences, I will mention some Xs (X filled with authors/singers/documentaries for the instruction of each narrative). I want you to assume that the Xs that I mention in the first sentence are the same as the ones that I mention in the second sentence. Based on this scenario, I would like you to give me a rating for the naturalness of the narrative ranging from 1 to 7. 1 means very unnatural, 7 means very natural.

I began with a training trial, first uttering the following narrative:

- (88) Sınıf-a bir kız ve bir oğlan gir-di. **Kız-1** dün bizim kafe-de gör-müş-tü-m.  
 class-DAT a girl and a boy enter-PAST girl-ACC yesterday our cafe-LOC see-PERF-PAST-1SG  
 ‘A girl and a boy entered the classroom. I had seen **the girl** at our coffee shop yesterday.’

The instruction clearly stated that the girl that I mentioned in the second sentence was assumed to be the same girl as the one who entered the classroom. We mutually agreed that, in this context, this narrative should be rated as 7. Then, I uttered the following:

- (89) Sınıf-a bir kız ve bir oğlan gir-di. **Bir kız-ı** dün bizim kafe-de  
 class-DAT a girl and a boy enter-PAST a girl-ACC yesterday our cafe-LOC  
 gör-müş-tü-m.  
 see-PERF-PAST-1SG  
 ‘A girl and a boy entered the classroom. I had seen **a girl** at our coffee shop yesterday.’

Here, we also both agreed that the girl that I mentioned in the second sentence is highly unlikely to be the same girl that I mentioned in the first sentence. So, this narrative should be rated as 1. I then continued with the target narratives, which are shown below, together with the mean ratings for each:

- (90) Bugün parti-ye Kerim-le birlikte üç tane yazar ve bir gazeteci gel-di. **Üç**  
 today party-DAT Kerim-with together three CL author and a journalist come-PAST three  
**(tane) yazar-ı** geçen hafta bir fuar-da gör-müş-tü-m.  
 CL author-ACC last week a convention-LOC see-PERF-PAST-1SG  
 ‘Today, three authors and a journalist came to the party with Kerim. I had seen the three authors at a convention last week.’

Mean ratings for the use of the NC in the continuation sentence:

NC with *tane*: 2 —NC without *tane*: 6.13

Comment from a consultant for (90) (translated to English): If I used *tane* in the second sentence, I would add *da* ‘also’ in the second sentence because these are three different authors here.

- (91) Sevgi müzik festival-in-de iki şarkıcı ve bir gitarist-le tanış-tı. **İki (tane) şarkıcı**  
 Sevgi music festival-COMP-LOC two singer and a guitarist-with meet-PAST two CL singer  
 önümüzdeki hafta Taksim-de konser ver-ecek-miş.  
 next week Taksim-LOC concert give-FUT-EVID  
 ‘Sevgi met with two singers and a guitarist in the music festival. Apparently, the two singers will give a concert next week in Taksim.’

Mean ratings for the use of the NC in the continuation sentence:

NC with *tane*: 1.96 —NC without *tane*: 6.06

- (92) Merve dün bütün gün TV baş-ın-da-y-dı çünkü dört tane belgesel ve  
 Merve yesterday all day TV front-3SGPOSS-LOC-COP-PAST because four CL documentary and  
 iki tane Türk dizi-si izle-di. Merve-ye **dört (tane) belgesel-i** daha önce  
 two CL Turkish series-COMP watch-PAST Merve-DAT four CL documentary-ACC earlier  
 abi-si tavsiye et-miş-ti.  
 brother-3SGPOSS recommend-PERF-PAST  
 ‘Yesterday, Merve was in front of the TV the whole day because she watched four documentaries and two Turkish series. Merve’s brother had recommended the four documentaries earlier.’

Mean ratings for the use of the NC in the continuation sentence:

NC with *tane*: 2.1 —NC without *tane*: 6.36

Comment from a consultant for all narratives (translated to English): I rated the versions without *tane* in the second sentence as 6 because I guess saying just the noun (e.g., şarkıcı-lar ‘the singers’) is better in these cases.

In summary, the collective results (together with (5)) —1.92 for NCs with *tane* and 6.13 for NCs without *tane* —consistently demonstrate a sharp contrast between the two forms of NCs in their potential to be used as anaphoric definites. In addressing the reviewer’s concern regarding (5), it is crucial to exercise caution in drawing conclusions about systematic dialectical variation in the use of NCs without *tane* as definite expressions based solely on this particular example. Notably, the reviewer did not express objections to (65), (66), and (68a), as discussed in Section 4.3.3, wherein a contrast between the two forms of NCs regarding definiteness is evident —an observation also reported in Schroeder (1992) and Öztürk (2005).

---

It is essential to highlight that if a subset of speakers consistently deems NCs without *tane* as unacceptable in the context of definiteness, it would suggest that the form without *tane* in their variety does not require a separate analysis and could be considered consistent with the general characteristics of NCs across languages. However, the primary focus of this study remains on the existence of a variety, where NCs without *tane* deviate from the established cross-linguistic pattern.