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

Yağmur Sağ

Abstract Numeral constructions (NCs) display a robust pattern of strong indefiniteness. While they can also be definite via a definite determiner, in languages lacking articles, this typically relies on alternative markers like demonstratives. Conversely, bare nouns in articleless languages can be freely definite, attributed to the covert *iota* operator in the neo-Carlsonian framework. The prevalent view treats NCs as predicative expressions of type $\langle e, t \rangle$, defaulting to existential interpretation in argument positions in the absence of an overt determiner, but it remains unclear why *iota* does not apply to NCs in articleless languages. This paper seeks to unravel this puzzle by analyzing the counting system in Turkish, an articleless language with an optional classifier, *tane*. Turkish NCs stand out by freely conveying both definiteness and indefiniteness without *tane*, while *tane* renders them exclusively indefinite. Drawing from this contrast, I argue that NCs crosslinguistically function as inherently argumental expressions of type e, with indefiniteness (via a choice function) originating from a cardinal head residing within their structure. Taking the predicative use of NCs as derived only when structurally necessary and relying on *iota* as a type-shifting operator, I attribute the incompatibility of NCs with *iota* to the absence of a type mismatch in argument positions. Turkish, however, accommodates inherently predicative NCs alongside default argumental NCs by featuring both covert and overt cardinal heads. While *tane* spells out the form with indefinite force, the covert head yields predicative NCs, making definiteness possible via *iota* type shifting. The analysis finds support from two more optional classifier languages, Farsi and Western Armenian, reinforcing the link 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.

In articleless languages such as Mandarin, Russian, and Hindi, definiteness is not freely available for NCs, which typically require demonstratives instead to display a definite-like behavior (Jiang 2012, Dayal 2013). The inability of NCs to convey definiteness in these languages is particularly puzzling because bare nominals freely allow definite interpretations in articleless languages. In the neo-Carlsonian approach, this is standardly achieved through the covert *iota* operator, with theories differing on whether it acts as a covert type-shifting operator or is inserted under a silent D head.

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

S. Author
second address

A generally accepted view due to Link (1983) treats NCs as inherently predicative expressions, which allows them to be arguments of determiners or occupy the predicate position (e.g., Partee 1986, Link 1987, Verkuyl 1993, Krifka 1999, Winter 2001, Landman 2003, Ionin and Matushansky 2006, cf. Montague 1974, Bennett 1974, Barwise and Cooper 1981, Scha 1981, van der Does 1992, Dayal 2013). When they occupy an argument position of a verb without an accompanying overt determiner, NCs are assumed to gain an existential force by default, for instance through $\exists$ type shifting, and hence the indefiniteness of NCs. However, the resistance of NCs to definiteness in articleless languages is perplexing if NCs are inherently $\langle e, t \rangle$ type expressions. That is, it is not obvious why the covert *iota* is not equally applicable to 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 determiner.

The key contribution of this study is to show that NCs are inherently argumental expressions of type e. I argue that the source of indefiniteness in NCs crosslinguistically lies in the projection of a cardinal head that may surface covertly and/or overtly, 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 (via *ident*) only in structures requiring a predicative meaning —for instance, when they serve as arguments to determiners (cf. Dayal 2013). Adhering to views where *iota* is a type-shifting operator, rather than functioning as a covert D head, the incompatibility of NCs with definiteness in articleless languages is reduced to the absence of type mismatch in the argument position of a verb. As a result, definiteness is only attainable for NCs through alternative overt markers. However, I further illustrate that inherently predicative NCs may coexist with default argumental NCs in languages 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, and Japanese, 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, among others, for Mandarin, Japanese, and Korean).

(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 a complementarity between systematic number marking in a language and the presence of an obligatory classifier in NCs with count nouns (Sanches and Slobin 1973, Greenberg 1990). English-like languages systematically distinguish between the singular and plural forms of nouns. 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 'one' while the plural is used with other numerals. In contrast, Mandarin-like languages do not exhibit such a fine distinction. Unmarked nouns consistently yield number neutral readings, while plural marking is possible only under certain conditions. Accordingly, these languages use unmarked nouns with all numerals but require the mediation of a classifier.

And yet, a third group of languages, including Turkish, presents an intriguing puzzle to the crosslinguistic semantics of counting, sharing features with these two types of languages. Like English, Turkish systematically

differentiates between unmarked and plural forms of nouns. However, the noun in Turkish NCs appears in the unmarked form with all numerals. Turkish NCs also feature an optional item between numerals and count nouns. 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):[1]

(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 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 as *tane* is not entirely optional. As stated above, NCs are freely indefinite in argument positions, with definiteness achieved through overt marking —by the definite determiner in languages with articles, or by other means like demonstratives in articleless languages. In Turkish, which lacks a definite article, only NCs with *tane* are restricted to indefiniteness. NCs without *tane* can be both definite and indefinite, exhibiting an exceptional behavior within the crosslinguistic picture (Schroeder 1992, Öztürk 2005). I show this in (5), delaying the discussion of the empirical facts for now. While both forms of NCs can introduce new discourse referents, as evidenced in the first sentence, only NCs without *tane* can refer to a unique/maximal entity from the preceding context, hence the contrast in the second sentence. Let us call this the (in)definiteness puzzle of Turkish NCs.[2]

(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:

i. The noun is unmarked for all numerals despite a systematic number marking system in the language.
ii. An intervening item, i.e., *tane*, optionally surfaces between a numeral and a count noun.
iii. NCs with *tane* are indefinite in argument positions 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 primary challenge posed by these properties is internal to Turkish and the group of languages it belongs to, in general. We need to understand how the Turkish counting system works and what role *tane* plays

---

[1] *tane* is compatible with all sorts of count nouns. There is another classifier in Turkish, i.e., *adet*, which is only compatible with non-human count nouns. In this paper, I only discuss *tane* as the distribution of the two classifiers is the same.

[2] 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.

in this. A key question is how the current accounts of relatively better-understood obligatory classifiers shape our understanding of seemingly optional classifier-like words such as *tane*, and vice versa. A further question concerns the consequences for a more general crosslinguistic 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 or 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 crosslinguistic implications.

I start by illustrating that *tane* is distinct from obligatory classifiers of Mandarin-like languages, often seen as repair mechanisms 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), where NCs are taken to universally involve the projection of a cardinal head, I propose that *tane* denotes a cardinality function, 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 crosslinguistic variation in number marking in NCs. Building on Ionin and Matushansky's (2006, 2019) view of numerals, I argue that the cardinal head necessitates a semantically singular form of the noun, a requirement 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 presents the core analysis, associating the universal restriction of NCs 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 *tane* as the indefinite form, whereas NCs with the covert variant realize the predicative form.

While the discussion primarily revolves around Turkish NCs, it also involves two other 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 specificity and definiteness, form the central motivation behind an agreement-based approach to number marking variation 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 as its NCs display the mirror image of Turkish NCs in their interpretation. In Farsi, it is the covert cardinal head that reflects exclusively indefinite characteristics, while the overt form can be freely definite and indefinite. This finding, which reinforces the link between the indefiniteness of NCs and the cardinal head, affirms the validity of the analysis offered for Turkish NCs. Additionally, we predict that the effects of featuring both argumental and predicative NCs can only be evident in articleless languages. NCs in Western Armenian, a language with a definite article, align with 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 lays the foundation of the account developed here. Section 3 compares *tane* with obligatory

classifiers. Section 4 is dedicated to the analysis. I first present my account of *tane*, adopting an initial semantics for number marking in NCs. Subsequently, I address the (in)definiteness puzzle and provide the core of the analysis. Section 5 discusses Western Armenian and Farsi data. Section 6 revises the number marking analysis adopted earlier. Section 7 concludes.[3]

2 Turkish Number Marking Semantics

The initial step 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.　$[\![kitap]\!] = \{a,\ b,\ c\}$
　　b.　$[\![kitap + PL]\!] = \{a,\ b,\ c,\ a \oplus b,\ a \oplus c,\ b \oplus c,\ a \oplus b \oplus c\}$

The number neutrality of Turkish plurals is evidenced in downward entailing environments 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)　Eğer **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)

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.[4] 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.'

---

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

[4] English unmarked nouns do not always receive a strictly singular reading, as seen 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.

Despite the dual nature of unmarked nouns, there is evidence for identifying them as strictly singular terms.[5] 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 singular and plural subjects. 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 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 taxonomic 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 be definite through covert *iota* type shifting.[6] 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 unique book familiar in the common ground via *iota* type shifting. This explains its strictly singular and definite interpretation.

(11) a. $[\![kitap]\!] = \lambda x.\ book(x) = \{a\}$
      b. ι: $\lambda P.\ \iota x.\ P(x)$
      c. $\iota([\![kitap]\!]) = \iota x.\ book(x) = a$
      d. $[\![(8c)]\!] = read(Ali, \iota x.\ book(x))$

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

---

[5] 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-debt 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.

[6] The *iota* operator is hindered by *the* in English due to the Blocking Principle, which prioritizes the use of overt determiners over covert type-shifting operations. Furthermore, type-shifting operators are assumed to be regulated by Revised Meaning Preservation, which bans ∃, and hence strong indefinite interpretation of bare nouns, in both English and Turkish (Dayal 2004). We revisit these principles in Section 4.3.1.

as *come from different regions*, as illustrated in (12b). This contrast also holds for English as evident in the translations, which 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 evrimleş-miş-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 worlds/situations s to the maximal entity satisfying that property in that world/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
$$^\cap P = \begin{cases} \lambda s.\ \iota x.\ P_s(x),\ \text{if } \lambda s.\ \iota x.\ P_s(x) \text{ is in } K, \text{the set of kinds} \\ \text{undefined, otherwise} \end{cases}$$
      b. $[\![(12a) \text{ with plural}]\!] = evolved(\lambda s.\ \iota x.\ dinosaur_s(x))$

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

In episodic contexts, as in (12b), Derived Kind Predication (DKP) provides sort-adjustment through $\exists$-quantification over the instantiations of the kind provided by *pred* in a given situation, resulting in an existential reading.[7] The ability of plural kind terms to be type shifted to sets of object-level entities makes them compatible with distributive predicates that 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) $[\![(12b) \text{ with plural}]\!] = \exists x\ [^{\cup\cap}bear(x) \wedge \forall y, z\ [[y < x \wedge z < x \wedge y \neq z]\ \rightarrow \iota r_1\ [region(r_1) \wedge came.to.zoo.from(r_1)(y)] \neq \iota r_2\ [region(r_2) \wedge came.to.zoo.from(r_2)(z)]]]$

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) a. $[\![dinozor]\!] = \lambda x^K.\ dinosaur^K(x^K)$
      b. $\iota([\![dinozor]\!]) = \iota x^K.\ dinosaur^K(x^K)$
      c. $[\![(12a) \text{ with unmarked}]\!] = evolved(\iota x^K.\ dinosaur^K(x^K))$

Dayal (2004) claims that singular kind terms denote grammatically atomic entities, similar to group terms like *team*. They do not allow type shifting to sets of object-level entities associated with kinds. Hence, the

---

[7] 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*.

derivation fails when they combine with a distributive predicate, as in (12b). Additionally, 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, 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 grammatically atomic nature, akin to the membership relation between an atomic group term and its members. Sağ calls this relation *belong-to*, as defined below:

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

As stated above, non-case-marked unmarked nouns have been previously analyzed as pseudo-incorporated arguments. Sağ argues that incorporation takes place with an incorporating thematic function, which establishes a *belong-to* relation between singular kinds and their object-level members, resulting in number neutrality.[8] The restriction in modification follows from the fact that taxonomic kinds can only be modified by adjectives denoting a classificatory/sub-kind forming property, such as *religious*, as in (9a).[9] 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) $\exists e \exists y \, [belong\text{-}to(y, \iota x^K. \, book^K(x^K)) \wedge read(e) \wedge Th(e) = y \wedge Ag(e) = Ali]$

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. In contrast, *doktor* in (10b) denotes the doctor kind, and hence it is only compatible with taxonomic adjectives like *practitioner*.

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 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. $[\![COP]\!] = \lambda x^K \lambda y. \, belong\text{-}to(y, x^K)$
        b. $[\![Ali \; child]\!] = belong\text{-}to(Ali, \iota x^K. \, child^K(x^K))$
        c. $[\![Ali \; and \; Merve \; child]\!] = belong\text{-}to(Ali \oplus Merve, \; \iota x^K. \, child^K(x^K))$

To sum up, Turkish and English nominals are similar in their ordinary object and kind-level denotations. 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 key difference between English and obligatory classifier languages lies in the fact that, while only mass nouns in English cannot directly combine with a numeral (excluding 'packaging/sorting' coercions), all nouns

---

[8] See Sağ (2018) and Martí (2020) for other accounts dedicating the number neutrality of unmarked nouns to the semantics of incorporation.

[9] 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-hand book store. This is known to be the *name-worthiness* requirement of incorporation.

in obligatory classifier languages, even those typically categorized as count nouns, require an intervening item. This 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. Consequently, classifiers are seen 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 they allow bare nominal 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
         without PL: 'I will go find a child/children/the child/the children.'
         with PL: 'I will go find the children.'

Unmarked nouns can receive kind-level, existential, and definite readings in these languages, as exemplified in (21a) and (21b) for Mandarin (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 in their number neutrality and kind-level interpretations, which align with plural kind reference (Yang 2001). The generally accepted view due to Chierchia (1998) posits that unmarked nouns of Mandarin-like languages are inherently kind terms of type $\langle s, e \rangle$, which can be type shifted to sets of object-level entities via *pred*. The availability of this type shifting also makes unmarked nouns 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 the 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. $[\![zhi]\!] = \lambda k \lambda x \, [^{\cup}k(x) \wedge AT(x)]$
        b. $[\![bi]\!] = pen^k$
        c. $[\![zhi \; bi]\!] = \lambda x \, [^{\cup}pen^k(x) \wedge AT(x)]$

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).[10] 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 instances, OU corresponds to the measurement of the number of atoms or 'an object unit' of an entity:[11]

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

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. $[\![gallon]\!] = \lambda y^k \lambda n \lambda w \lambda x \, [R_w(x, y^k) \wedge GALLON_w(y^k)(x) = n]$
        b. $[\![ben]\!] = \lambda y^k \lambda n \lambda w \lambda x \, [R_w(x, y^k) \wedge OU_w(y^k)(x) = n]$

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.[12] 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.

---

[10] 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).

[11] In Krifka's account, count nouns in English NCs are singular kind terms. The plural marking in NCs is an instance of syntactic agreement with the numeral, absent in languages like Turkish and Hungarian. In the absence of a numeral, 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 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.

[12] Other studies do not take the role of kind reference as the key for the semantics of classifiers in NCs. But in these accounts, classifiers are also given some function that 'fixes' 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 languages.

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 through the covert *iota* operator.

Given that Turkish NCs require unmarked nouns, there are two possibilities for the noun denotation: it can either be a singular kind or an atomic property of object-level individuals. If the former, a transformation from the kind-level to object-level denotation is needed for 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, precluding type shifting to sets of object-level instances, and thus 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 the noun still denote a singular kind in NCs with *tane*? One analysis worth considering aligns 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 be attributed to *tane*, let us revisit the *kind specification* phenomenon discussed 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 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:

(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*, akin to the copula, could be establishing a *belong-to* relation between a singular kind and object-level entities associated with it, as illustrated in (26). In this case, we would expect the same modificational restrictions to surface in NCs with *tane*, but this is not attested. Unmarked nouns in NCs can undergo object-level modification regardless of the classifier, as shown in (27). Therefore, we can conclude that NCs do not involve kind specification with singular kind terms.[13]

(26) $[\![tane]\!] = \lambda x^K \lambda y.\ belong\text{-}to(y, x^K)$

(27)   Sevgi iki (tane) *yakışıklı* doktor-a   mesaj at-tı.
        Sevgi two CL   handsome doctor-DAT text   send-PAST
        'Sevgi texted two handsome doctors.'

Given the problem of covert type shifting with singular kind terms and the absence of kind specification in NCs, *tane* cannot be ascribed an atomizer semantics in the sense of Chierchia (1998) either. This is because a function of this sort would demand access to object-level entities to extract atoms. Such semantics would be plausible 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.

---

[13] 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 (Giusti 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'

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

3.3 *Tane* and Mass Nouns

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

So far, our focus has been on ontological count nouns in Turkish. However, it is crucial to see how Turkish compares to English and Mandarin-like languages regarding the count-mass distinction. While English-like languages grammatically distinguish between the count and mass senses of nouns, Mandarin-like languages reflect this distinction only through the choice of the classifier (Cheng and Sybesma 1999).

In Section 1.1, I demonstrated that Turkish numerals, as in 'two waters' in English, cannot directly combine with a mass noun unless the noun is coerced into a count interpretation through covert universal packaging/sorting (see (3) and (4)). This coercion is also necessary for *tane* to accompany mass nouns.[14] The differing abilities of count and mass nouns to directly combine with numerals and *tane* suggest a more English-like characteristic of Turkish nominals, indicating that *tane* selectively pairs with the count sense of nouns, just like numerals. However, before reaching this conclusion, further evidence is required to confirm the count-mass distinction in Turkish. Moreover, 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 an analysis along the lines of Krifka's (2003), assuming that a count noun inherently involves a number argument. The hypothesis is then as follows: In NCs without *tane*, the noun is count, with a denotation as in (28a), enabling direct combination with a numeral. In NCs with *tane*, the noun is mass, with a denotation as in (29a), hence requiring a classifier for counting.[15]

(28)   *üç köpek* 'three dog':
        a. $[\![dog_{count}]\!] = \lambda n \lambda w \lambda x \, [OU_w(dog)(x) = n]$

---

[14] 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 arasından, 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.'

[15] A similar analysis 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* is 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 counting expressions. This is particularly evident in agreement facts and their inability to be antecedents of reciprocals. Below, I discuss why Turkish NCs should not be analyzed analogously. Nevertheless, it is crucial to note that NCs with or without *tane* fundamentally differ by being compatible with reciprocals:

(i)   İki (tane) öğrenci birbirin-e   aşık ol-du.
       two CL   student each.other-DAT love be-PAST
       'Two students fell in love with each other.'

      b. $[\![three \ dog]\!] = \lambda w \lambda x \ [OU_w(dog)(x) = 3]$

(29)   *üç tane köpek* 'three CL dog':
      a. $[\![dog_{mass}]\!] = \lambda w \lambda x \ [dog_w(x)]$
      b. $[\![tane]\!] = \lambda n \lambda P \lambda w \lambda x \ [OU_w(P)(x) = n]$
      c. $[\![three \ tane \ dog]\!] = \lambda w \lambda x \ [OU_w(dog)(x) = 3]$

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 (e.g., 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 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.

kedi 'cat' vs. kan 'blood' We know that direct combination 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*-type nouns are compatible with *bir kaç* 'a few', *blood*-type nouns require *biraz* 'a little', as shown in (30) (see Göksel and Kerslake 2005 and Görgülü 2010).[16] Critically, while *tane* is an option with the count quantifier, it is incompatible with the mass quantifier, as illustrated 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 imparts an *abundance* or *unorderly scattered* interpretation, as in (31a) and (31b), a phenomenon also found in Greek (see Tsoulas 2009 and Kouneli 2019 for Greek). The pluralization of mass nouns in Turkish, therefore, contrasts with that of count nouns.

(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

---

[16] 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.'

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:[17]

(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.

(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.'

In summary, these diagnostics reveal that Turkish count and mass nouns exhibit both distributional and semantic distinctions.

mobilya 'furniture' Nouns like *mobilya* 'furniture' and *bagaj* 'luggage' exhibit characteristics of both count and mass nouns, thus form a third variety in Turkish, as in English, albeit with some less clear-cut differences. These nouns are degraded, if not fully ungrammatical, when paired with a numeral or a count quantifier directly. The intervention of *parça* 'piece' resolves the issue, but the degradedness persists when *tane* is used instead, as illustrated in (34a). It is crucial to note that *parça* is incompatible with count nouns unless the noun is coerced into a mass interpretation via universal grinding. For example, (34b) could not mean 'three chickens', mirroring the situation in the English equivalent.

(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, it can become acceptable if the context emphasizes the volume of the furniture. Suppose that we need to rent a truck for moving. The following would be felicitous in a conversation with the rental agent while assessing the truck size required.

(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:

---

[17] Mass nouns also differ from unmarked 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.'

(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, displaying traits of both count and mass nouns. While (37) naturally compares quantities of furniture based on the number of pieces, it can also compare their volume if followed by an utterance such as 'She will need a larger moving truck.'

(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).

	kedi	mobilya	kan
direct combination with numerals and the count quantifier	✓	??	*
combination with *tane*	✓	??	*
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

In conclusion, Turkish nouns display clear distributional and semantic distinctions in countability, aligning Turkish more closely with English-like languages despite some puzzling variations, such as the pluralization of mass nouns. While these variations merit independent research, our goal has been to establish that Turkish fundamentally differs from obligatory classifier languages by having a three-way count-mass distinction.

3.3.2 Interim Discussion

What mainly concerns us in this comparison is that *tane* is only compatible with count nouns like *kedi* '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 *parça* 'piece', similar to the case with mass nouns. This mediation is also not (readily) provided by *tane*.

   The facts also illustrate that analyzing Turkish unmarked nouns as flexible nouns is problematic. As shown above, there are cases where *furniture*-type nouns have count behavior, but ontological count nouns do not display any mass behavior (modulo coercion).[18] 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*.

---

[18] 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.)

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.[19]

By eliminating the possibility that Turkish count nouns have a flexible denotation, we find ample evidence against analyzing *tane* as a repair mechanism with kind terms or mass nouns, akin to obligatory classifiers in Mandarin-like languages. Consequently, I argue that unmarked count nouns in Turkish NCs uniformly denote atomic properties of object-level individuals, thereby obviating the need for a mediator in counting.

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 crosslinguistic 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* manifests as the overt form of this head, contrasting with English, where the cardinal head is only realized covertly.

We will then shift our focus to the non-optional aspect of *tane*, i.e., the (in)definiteness puzzle, and elaborate on the crosslinguistic interpretation of counting constructions. I explain the general restriction of NCs to indefiniteness by associating the cardinal head with a choice function variable in the sense of Reinhart (1997) (cf. Jiang 2012). In essence, I analyze NCs to be inherently argumental expressions with a built-in indefinite force. I further propose that the covert form of the cardinal head in Turkish is exempt from this force, and thus is enriched in meaning, whereas *tane* embodies the typical inherently indefinite cardinal head.

4.1 Counting with Atoms and Number Agreement in NCs

We 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 nouns in NCs are semantically singular crosslinguistically, despite their appearance in languages like English. They argue that counting is feasible only for individuals of identical cardinality. Specifically, numerals, as modifiers of type $\langle\langle e, t\rangle, \langle e, t\rangle\rangle$, require atomic properties as their argument since members of a plural property do not necessarily evaluate to the same cardinality.[20] 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).[21]

(38) $[\![two]\!] = \lambda P \lambda x. \exists S [\prod(S)(x) \wedge |S| = 2 \wedge \forall s \in S\ P(s)]$
   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) $[\![two]\!](P)(x)$ is defined iff $\exists n \forall z [P(z) \rightarrow |z| = n]$

Then, $[\![two\ books]\!]$ 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

---

[19] We observed that *furniture*-type nouns are not entirely ungrammatical when directly combined with numerals and *tane*. However, their degradedness should not be interpreted as evidence that *tane* operates on *furniture*-type mass nouns. If it did, we would expect *tane* to be perfectly compatible with these nouns, as is the case with count nouns. It instead prompts a question about the extent to which *furniture*-type nouns are count in nature, allowing for some degree of compatibility with *tane*.

[20] 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 identical cardinality.

[21] $\prod$ 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., $a \leq_i a \oplus b$.

this view at first glance as the noun occurs in the plural form. Ionin and Matushansky argue that the plural marking on the lexical NP is number agreement reflecting the semantic plurality of NCs. For instance, *books* in *two books* is lexically singular, denoting a set of atomic individuals, but the NC as a whole denotes a set of plural individuals, with the *-s* marking signaling this plurality.[22] This view then dedicates the contrast in the noun form between English and Turkish NCs to the presence or absence of number agreement.

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.

Similar to Ionin and Matushansky's view, in Scontras's analysis, the plural in English NCs 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 is a number head, i.e., #, that projects above NCs and and hosts singular (SG) and plural (PL) features.[23] 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.[24]

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 one, 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 sensitive to cardinality, but also to a measure specified by the head of a measure phrase. For instance, in *three kilos of apples*, the number morphology is determined by the specific measure provided by *kilo*. In *one kilo of apples*, each member of the set evaluates to one with respect to the kilo measure, resulting in the singular form of the measure term. However, with a different number, such as *three*, the plural form *kilos* surfaces due to the presupposition failure.[25]

For a uniform account of number marking in constructions with cardinals and measure terms, Scontras analyzes numerals as individual-denoting expressions of type n referring to natural numbers, which serve as arguments to a functional measuring element, 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, as shown below:

(41)   Generalized MP

---
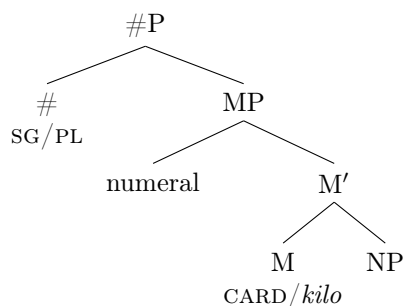
[22] A syntactic implementation of this approach can be found in Alexiadou (2019) (building on Landau 2016 and Wechsler and Zlatić 2003) and Bayırlı (2017). See also Krifka (1989, 1995, 2003), Ortmann (2000), Farkas and de Swart (2010), Sağ (2018) for studies adopting an agreement-based approach to number marking in NCs.

[23] 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.

[24] Scontras does not call number marking in NCs an agreement phenomenon. This is a contribution made in this paper.

[25] 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.

An MP denotes a quantity-uniform (QU) property relative to the measure expressed by the M head.[26] Every member of the set denoted by the MP maps to the same number in terms of this measure and the one-ness presupposition of SG involves this information. With CARD, number marking is sensitive to the quantity-uniform measure μ_{CARD}, while with measure terms, it is sensitive to the quantity-uniform measure intrinsic to the measure term, e.g., μ_{kg}. Building on this, the semantics of English # heads is given below:

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

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, namely CARD, it is expressed on the head of the next closest phrase, the lexical NP.[27]

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 crosslinguistically, 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:[28]

(43)   The Semantics of CARD (to be revised):
   $\llbracket\text{CARD}\rrbracket = \lambda P\lambda n\lambda x : \forall y\ [P(y) \to AT(y)].\ \exists S\ [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S\ P(s)]$
   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:
        $\sqcup C = X$

Below is the derivation of [two CARD book], where AT in P_{AT} is short for the presuppositional content.

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

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

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

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

---

[26] 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.

[27] Scontras 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.

[28] 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 51 for reasons that relate to complex numerals in Turkish and Farsi.

To sum up, the lexical NP in NCs is semantically singular although in some languages, like English, NCs 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's 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):
$\llbracket \text{CARD} \rrbracket = \lambda n \lambda P_{AT} \lambda x. \exists S \ [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S \ P(s)]$

That CARD combines with the numeral first in Turkish follows from the structural configuration I adopt for Turkish NCs. Unlike in English, MP is represented as a modifier to the noun, which aligns with the head-final characteristics of Turkish, as illustrated below. Representing a functional projection as head-initial in a strictly head-final language, without independent evidence to the contrary, would merely be a stipulation.[29]

(48)   The structure of Turkish NCs
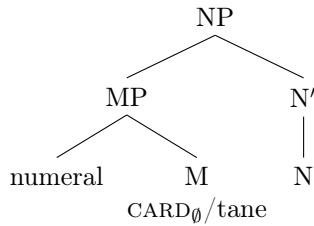
The difference between English and Turkish NCs then not only surfaces in the plural number agreement, but also in the form of CARD.[30] In English, the cardinal head is always covert, and in fact, this is the case in many other languages. Turkish departs from these languages by also featuring an overt version of this head. This is the optional aspect of *tane* in NCs.

It is important to emphasize that it could be possible to find languages where CARD is always overt.[31] 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 cases where the overt and covert realizations of CARD depend on the numerals they combine with (see Bale et al 2019 for this generalization).

Before proceeding further, there is one remaining issue that needs to be addressed. As discussed in Section 3.1, obligatory classifiers in Mandarin-like languages take the role of a repair mechanism to ensure counting with nouns 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

---

[29] 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 analysis) as modifiers/specifiers in the extended projection of NP is standard in the literature (e.g., Link 1987, Verkuyl 1993, Krifka 1999, Landman 2003, Rothstein 2017, cf. Giusti 1991, Ritter 1991, Zamparelli 1995, Danon 2012, Mittendorf and Sadler 2005, Ionin and Matushansky (2006)).

[30] 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.

[31] Bangla, an obligatory classifier language with systematic plural marking, serves as a potential example 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 Mandarin-type classifier is its ability to 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.

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
                         MP   CLP
              ┌──────────┴──────────┐   ┌──────────────┴──────────────┐
          numeral   M′   numeral   CL′
                         ┌──────┴──────┐   ┌──────┴──────┐
                      CARD   CLP   CL   NP
                        ∅   ┌───┴───┐   CARD+CL
                               CL   NP

As illustrated in (50a), *tane* does not appear with quantizing nouns that denote 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, which points 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 nouns is possible, as shown 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.[32]

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

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

    To summarize the analysis so far: I claim that the Turkish optional classifier differs from obligatory classifiers in Mandarin and similar languages in that it does not serve a type-fixing/repairing role to enable counting with kinds and mass nouns. Instead, *tane* denotes a counting function by overtly realizing the cardinal head, which is typically covert in NCs across languages. 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.

---

[32] *tane* can co-occur with container nouns that form a compound with the lexical noun, as in *iki (tane) şişe-su* 'two bottle-waters', where the stress falls on the container noun, contrasting with (50a). This could mean that quantizing nouns, when used in their lexical sense, cannot spell out CARD, allowing *tane* to surface. While a reviewer suggests that the grammatical forms in (50b) and (50c) might also result from compounding, this cannot be the case at least for Farsi, as the version where the partitioning noun and 'coke' form a compound differs in structure: do (tā) *nooshābeh-e shisheyi* 'two bottle-cokes'

4.3.1 A Crosslinguistic 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 1986, Link 1987, Verkuyl 1993, Krifka 1999, Winter 2001, Landman 2003, Ionin and Matushansky 2006, cf. Montague 1974, Bennett 1974, Barwise and Cooper 1981, Scha 1981, van der Does 1992, Dayal 2013). 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 occupy an argument position of a verb without an overt determiner, they are presumed to attain an indefinite interpretation by default. One commonly proposed mechanism for this is existential type shifting that leads to a strong indefinite interpretation with free scope-taking abilities:

(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* (Chiercha 1998):
              For any type shifting operation ϕ 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 a specific order for type shifters to apply, adhering to the Blocking Principle, which mandates the use of overt determiners over covert type-shifting operations for economy reasons. Thus, in languages with an overt definite determiner, bare nouns cannot undergo covert ι type shifting, ruling out definite interpretations. Furthermore, bare nouns cannot be interpreted as strong indefinites, as $\exists$ type shifting is low-ranked and thus not an option. The narrow scope existential reading of bare nouns is independently possible through Derived Kind Predication (Chiercha 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 a 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 bare nouns can be definite through the covert ι operator, NCs are invariably interpreted as strong indefinites in argument positions, indicating that they do not undergo ι type shifting. Jiang (2012) illustrates this with Mandarin and Russian NCs. Likewise, Dayal (2013) makes a similar observation for Hindi NCs. However, NCs in these languages can exhibit definite-like characteristics when paired with a demonstrative, as exemplified for Hindi below:[33]

(55)   do bacce kamre meN the. *(ve) do bacce khel rahe the
        two kids room in   were those two kids play PROG PAST

---

[33] Hindi NCs can alternatively be definite when the numeral is inflected with the particle *-no*, as in *do-no bacce khel rahe the* 'The two kids were playing.' Although the precise nature of *-no* is unclear, its use is restricted to certain numerals and does not extend to bare nouns, ruling out its categorization as a definite determiner (p.c. Veneeta Dayal). Furthermore, in some articleless languages, like Bangla, the definiteness of NCs is argued to be possible through the syntactic movement of the noun, placing it before the numeral+classifier combination (Dayal 2014).

'Two kids were in the room. Those two kids were playing.'   (p.c. Veneeta Dayal)

A striking confirmation of the fact that NCs do not undergo covert ι type shifting comes from Yi, an obligatory classifier language with a definite article. Yi allows bare nouns to be definite without the definite article, implying that the Blocking Principle may not be applicable in this language (see Jiang 2018 though). Despite this flexibility with bare nouns, NCs still require the presence of the article for definite interpretations:

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

The interim conclusion drawn from these languages is as follows: While NCs default to a strong indefinite interpretation, their definiteness cannot be achieved covertly through ι type shifting. Instead, it requires overt marking with a definite article or alternatives like demonstratives if the language lacks a definite article. The resistance of NCs to definiteness via ι in articleless languages is perplexing, if NCs are inherently predicative. That is, it is unclear why ι is not equally applicable to NCs in these languages, in contrast to bare nouns.

The situation becomes more intricate when we delve into the interpretation of Turkish NCs. While NCs with *tane* adhere to the general trend of being restricted to indefiniteness, NCs without *tane* display exceptional behavior, 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, Kelepir 2001, Demirok 2021, among others), as illustrated by the following example from Kelepir (2001: 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 was read by the same single student, and it would be false if each book was read by different students. This shows that the universal quantifier in the object position cannot take wide scope over the indefinite subject. For a narrow scope reading of 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 pattern with indefinites in other languages by exhibiting scope ambiguity and exceptional scope-taking abilities (Kelepir 2001). This is exemplified below:[34]

(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** $> \exists$, $\exists >$ **exactly half**)

Suppose 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

---

[34] This is the case for case-marked indefinites. Non-case-marked indefinites always yield narrow scope readings. Kelepir (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, introduced outside the VP. Furthermore, accusative-marked indefinites always receive a specific indefinite interpretation, for which Kelepir (2001) argues 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.

commented on the same specific book, while a third student commented on a different book. Here, (59) is still true, showing that the indefinite can take wide scope over the quantifier in the subject position. However, the narrow scope reading of (59) would be false as more than half the students commented on a book.

Just like regular indefinites, NCs with and without *tane* exhibit 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 NC. It can also be true in a situation 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
      '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 outside of an island. For example, (61) is felicitous in two contexts: In the first, funding is secured if any two of my projects are selected, while in the second, funding hinges on the selection of two specific projects. However, unlike the case in (60), the wide scope interpretation necessitates the collective selection of two projects, as 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). This behavior is also exhibited by Turkish indefinites, allowing NCs to take intermediate scope in addition to 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 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 independent of scopal interactions, such as partitive and epistemic specificity (see von Heusinger 2002). The sentence in (63) shows that both forms of NCs can be partitive specific. 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-ların-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 or presence of *tane*.

4.3.3 Turkish NCs and The Definiteness Puzzle

As first demonstrated in Section 1.1, NCs without *tane* can receive a definite interpretation, unlike NCs with *tane*. This is evidenced by their ability to refer to a unique/maximal entity introduced in the preceding context. However, NCs with *tane* do not exhibit this type of anaphoric behavior (Schroeder 1992, Öztürk 2005). Instead, they require a demonstrative, as illustrated below. When not accompanied by a demonstrative, the form with *tane* either forces a partitive specific reading or is understood as introducing a new discourse referent in the second sentence, yielding infelicity in this particular context.[35]

(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 inability of NCs without *tane* to freely convey definiteness is further revealed in situational contexts where it is possible to refer to a unique/maximal entity familiar in the common ground. Suppose Sevgi has only three apples, known to 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, reference to any three of them is feasible with either form of NCs, as in (67). This reaffirms that while both forms of NCs align 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.[36]

(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.'

Another indicator of the contrast in definiteness between the two forms of NCs involves the particle *dA*, which can serve additive or distributive roles, among other functions (Göksel and Özsoy 2003, Göksel and Kerslake

---

[35] A reviewer provides the example in (i), 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, the reviewer highlights that NCs with *tane* can potentially be definite. 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 shown by the English example in (ii). 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 (i) might then merely serve as another instance where the novelty condition of an indefinite is violated. Notice that the surprise effect, which the reviewer also points out to be evident in (i), 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)   İçeri dört tane mühendis gir-di   ve dört tane mühendis şirket-in   tüm problem-ler-in-i   çöz-dü.
      inside four CL   engineer enter-PAST and four CL   engineer company-GEN all problem-PL-3SGPOSS-ACC solve-PAST
      Reviewer's translation: 'Four engineers entered inside and the four engineers solved all the problems in the company.'

(ii) 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.

[36] 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 owned by Sevgi is not necessarily familiar to the addressee. As evident in (66) and (67), both definite and partitive specific indefinite NCs are used within the same possessive form, indicating that the Turkish possessive is not strictly definite.

2005, Kamali and Karvovskaya 2013, Sağ and Demirok 2023). NCs without *tane* are compatible with both distributive and additive uses of *dA*, but with *tane*, *dA* is restricted to only an additive role (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.'

    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.'   collective
        'The four kids *each* drank a bottle of milk.'   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 an NC without *tane* as its subject, can have a collective reading, where *dA* contributes an additive interpretation, as follows: Besides four kids eating cookies, four other kids collectively drank a bottle of milk. Crucially, (68a) also allows a distributive reading, where each of the four kids, who ate cookies, drank a bottle of milk. However, in the presence of *tane*, this distributive reading is not apparent, though the additive interpretation remains possible, as shown in (68b).[37]

What matters for our purposes is that the distributive reading is only possible if the NC is interpreted as definite. Specifically, in (68a)'s distributive reading, the NC must refer to the four kids mentioned earlier. This shows that the distributive instance of *dA* imposes a definiteness requirement on its host. Since NCs with *tane* cannot convey definite readings, they are incompatible with the distributive function of *dA*.[38]

In summary, NCs with *tane* align with the crosslinguistic trait of NCs, being restricted to indefiniteness, whereas NCs without *tane* are extraordinary in also conveying definiteness freely.[39] Our aim is to explain why *tane* necessitates indefiniteness in NCs, 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, regardless of crosslinguistic variations in the rules governing bare arguments. Jiang (2012) posits that NCs have a distinct source from bare nouns to form arguments, viewing numerals as lexically ambiguous. The first variant, a modifier of type $\langle\langle e, t\rangle, \langle e, t\rangle\rangle$, enables NCs to denote predicates that can be an argument to a category at D. The second variant, a modifier of type $\langle\langle e, t\rangle, e\rangle$ with a built-in choice function variable à la Reinhart (1997) (cf. Fodor and Sag 1982, Winter 1997, and Kratzer 1998), yields individual type NCs requiring ∃-closure.

In Reinhart's theory, the ∃-closure of the choice function variable can occur at any compositional level. This elucidates the intermediate scope readings of indefinites and their capacity for exceptional scope without postulating a mechanism that extracts indefinites from an island. This theory also explains the lack of distributive readings with NCs in exceptional scope cases as distributive readings can only be obtained through QR. However, this account rejects a generalized quantifier analysis of indefinites (cf. Fodor and Sag 1982). Below, I exemplify semi-formally how an NC is interpreted with respect to an island under this view.

(69)   If two of my projects are selected, I will receive funding.

---

[37] Note that both forms of NCs support distributive, collective, and cumulative readings. The contrast discussed above only arises with the particle *dA*. For instance, both NCs with and without *tane* have a distributive reading with dependent numerals:

(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.'

[38] Despite meeting the definiteness requirement, plural definites and pronouns are incompatible with distributive *dA*, e.g., *Kızlar da bir şişe süt içti.* 'The girls, too, drank a bottle of milk.' Here, *dA* has only an additive and collective reading. Sağ (2019) analyzes *dA* as a post-suppositional item associated with universal quantification following Szabolcsi's (2015) analysis of *mo*, the Japanese kin of *dA*. The lack of distributive reading with plural definites and pronouns is attributed to their 'weak maximality.' Unlike cardinal definites and universal quantifiers, plural definites allow exceptions in their interpretations (Kroch 1975). As *dA* is linked to universal quantification, it is also sensitive to strong vs. weak maximality potential of its host.

[39] One might question if there is any difference in indefinite readings between the two forms. To my knowledge, none exists.

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 indefiniteness in NCs is CARD instead. While the deeper connection between indefiniteness and cardinality awaits future exploration, here I maintain Jiang's insight to construct an initial typology, as a first step in understanding why NCs universally tend to favor indefinite interpretations over definite ones. The twist I offer is that NCs are crosslinguistically 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 arguments. Below, I represent the crosslinguistic 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$_f$ signifies that the cardinal head is argumental, meaning it yields argumental type NCs with a choice function.

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

I follow Dayal (2013) in that NCs type shift to a predicate as a repair operation only in structures requiring the predicative meaning of NCs, i.e., when they are arguments to overt determiners or occupy the predicate position. Since NCs are e-type expressions, I assume the shift to predicate type occurs via the *ident* operator (Partee 1986), as shown below. NCs would not undergo predicative type shifting followed by *iota* type shifting in argument positions of verbs since they are already of the suitable type —there is no impetus for the reparative strategy to activate when NCs merge with an element looking for an e-type argument.[40] In articleless languages, NCs then gain definite-like readings solely through overt markers like 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, Russian, and Hindi, 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:[41]

(72)   The Generalized Structure of NCs
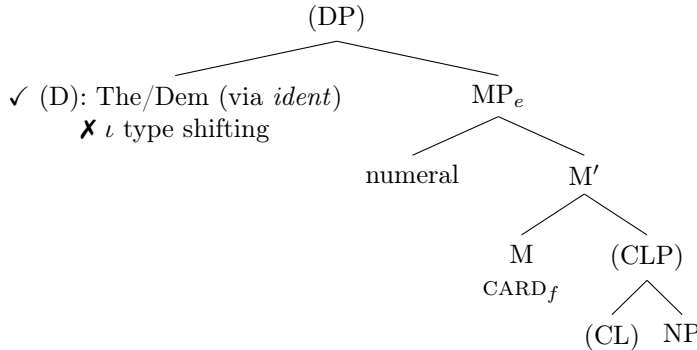
The parentheses around the DP should be interpreted as follows: An argumental type NC can directly occupy an argument position of a verb as an MP. Alternatively, it can serve as a complement to a category at D,

---

[40] I am grateful to an anonymous reviewer for generously bringing this point to my attention.

[41] 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.

such as definite and demonstrative determiners, or occupy the predicate position as a complement to the copular verb, triggering *ident* type shifting to resolve the type mismatch.

In contrast, ι type shifting is unavailable, as explained above. It is essential to highlight that the ι operator is represented below the D head for ease of comparison. My account posits ι as a covert type-shifting operator and assumes no silent D projection in the absence of an overt definite determiner. If ι 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.[42]

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 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 severed from the choice function, perhaps as a way to resolve the restriction to indefiniteness. That is, while NCs with *tane* are of argumental type, NCs with the covert CARD have a predicative denotation:

(73)   a.   The Semantics of the overt CARD in Turkish $\langle n, \langle\langle e, t\rangle, e\rangle\rangle$
                $[\![tane_f]\!] = \lambda n \lambda P_{AT}.\ f(\lambda x\ \exists S\ [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S\ P(s)])$

       b.   The Semantics of the covert CARD in Turkish $\langle n, \langle\langle e, t\rangle, \langle e, t\rangle\rangle\rangle$
                $[\![\text{CARD}_\emptyset]\!] = \lambda n \lambda P_{AT} \lambda x.\ \exists S\ [\prod(S)(x) \wedge |S| = n \wedge \forall s \in S\ P(s)]$

NCs with CARD$_\emptyset$ 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 ι 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$_\emptyset$ bring a seemingly exceptional status to the Turkish counting system. This is schematized in (74):

(74)   a.   NCs with *tane_f*   b.   NCs with CARD$_\emptyset$
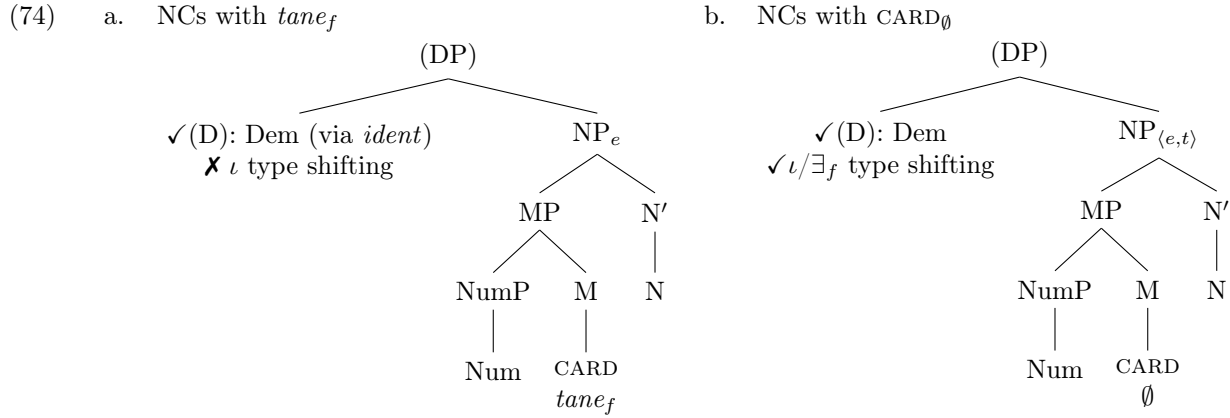
As highlighted by the anonymous reviewers, the existential type shift of NCs with CARD$_\emptyset$ through the choice function would not be hindered by NCs with *tane*. This stems from the choice function being introduced at different syntactic levels in these constructions —by a lexical item within the NC in one case and as covert type-shifting operation applying above the NC in the other.

To summarize, I have argued that NCs are argumental across languages, linking their restrictedness to indefiniteness to a cardinal head that is universally involved in the projection of NCs.[43] I have further proposed that in Turkish, an articleless language that features both overt and covert forms of CARD, the

---

[42] I do not rule out the possibility that there might be articleless languages introducing ι within a covert DP projection rather than via type shifting. Our analysis predicts that definiteness through ι should 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 determiner (cf. Arslan-Kechriotis 2009, Kornfilt 2005, 2017, von Heusinger and Kornfilt 2017).

[43] My proposal should rather be taken as a crosslinguistic 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).

covert CARD leads to predicative NCs, allowing definiteness without a need for an overt marker. In contrast, *tane* is just an ordinary cardinal head hardwired with an indefinite force as in any other language.[44]

As a concluding note, it is crucial to underscore the rationale behind the claim that NCs are primarily argumental expressions across languages. Considering NCs uniformly as predicative expressions of type $\langle e, t \rangle$ introduces an inconsistency because it fails to account for why NCs with *tane* (and those in articleless languages) resist *iota* type shifting for definite interpretations, unlike NCs without *tane*. However, we must also acknowledge inherently predicative NCs to explain the availability of definite readings for NCs without *tane*, in contrast to the form with *tane*. Consequently, the observed pattern in 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 crosslinguistic 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 atomic properties, requiring the semantically singular form of the noun, whereas 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. This leads Ionin and Matushansky to conclude that plural marking in NCs is also a result of number agreement phenomena. 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. However, they differ from Turkish by also allowing plural marking in their NCs, which yields a definite or specific reading in WA and a definite reading in Farsi (e.g., Sigler 1996 for WA and Ghomeshi 2003 for Farsi). For WA, consider the example in (75). With the plural marker *-er*, the NC gets a wide scope reading, felicitous when referring to two specific elephants desired to be seen (de re), otherwise, it receives a narrow scope interpretation.[45]

---

[44] A reviewer notes that combining *every* with argumental type NCs should yield an odd result, as the ∃-closure of the choice function occurs outside the NC, and universal quantification (through *ident*) would be over the individual chosen by f. This prediction aligns with observed oddities in sentences like *I spoke to every three students*, supporting the choice function analysis for NCs. However, NCs are felicitous with *every* in sentences involving some sense of partitiveness, e.g., *There was a policeman every three houses* or *Nine out of every ten patients recommend this toothpaste* (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, similar patterns are observed with *her* 'every', irrespective of the form of NCs. For example, *Her üç (tane) çocukla konuştum* ('I spoke to every three students') sounds odd, but *Her üç (?tane) binaya bir polis düşüyor* ('There is a policeman for every three buildings') is felicitous, with a slight preference for the form without *tane*, similar to the case with demonstratives, as shown in (65) with a question mark. Thus, the compatibility of NCs with *every* does not seem to hinge on whether NCs are inherently argumental or predicative. Given that the role of partitiveness in felicitous uses has yet to be explored, I defer the investigation of universally quantified NCs to future research.

[45] WA shows dialectical variation regarding the co-occurrence of the classifier and the plural marker. Data discussed here reflect the variety spoken in Beirut (see fn 3). According to judgments of a consultant from Istanbul, plural marking does not occur with the classifier in the Istanbul variety, though a specific indefinite reading is still possible (see also Sigler 1996, Borer 2005, Bale and Khanjian 2008, 2014, Khanjian 2013). Moreover, the definite marker can inflect both forms of NCs. While it is not obvious what regulates these variations, I do not attribute it to the classifier, due to the distribution in the Beirut variety.

(75) g-uz-em   jergu (had) piɣ(-er)   desn-el kazananot͡s-i-n met͡ʃ
      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 convey definite interpretations. When this is the case, the plural marker is obligatory, as shown in (76). Thus, we can conclude that WA NCs are compatible with the plural only in specific indefinite and definite contexts.

(76) dup-i-n   met͡ʃ-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 indefinite, including specific interpretations.[46] 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 to 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 does not have a definite article, similar to 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.[47]

(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 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, Croft 2002, see also Alexiadou 2019). For example, both WA and Farsi exhibit this sensitivity in subject-predicate agreement. According to Sigler (1996), WA requires plural agreement on the predicate with definite subjects, while both singular and plural agreement are possible with indefinite subjects. This mirrors the pattern observed in NCs, albeit with distinct degrees of sensitivity to the definiteness spine, suggesting that the mechanisms governing both are fundamentally of the same nature. In contrast, Farsi restricts subject-predicate plural agreement to the animacy hierarchy, with plural agreement triggered only by animate subjects (Ortmann 2002, Alexiadou 2019, Mahootian 1997, Mache 2012). Although this does not align perfectly with the

---

[46] Consultants note that NCs with *tā* are colloquially used, while omission of *tā* is common in formal (often written) contexts. Furthermore, a wide scope interpretation is found harder to get without the classifier compared to when it is present. However, if two student names are mentioned following 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.'

[47] 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 confused with a definite article. Jasbi (2020) claims that *-(h)e/a* can optionally appear with both nominals and indefinites, introducing a uniqueness implication. When applied to a noun, it yields a definite interpretation; with indefinites, it conveys specificity.

constraints observed in NCs, it is still plausible to consider plural marking in NCs as a form of agreement. It seems that the two hierarchies are split between NCs and the predicate domain in Farsi.[48]

Note that plural nouns in WA and Farsi are not constrained by specificity or definiteness. They can convey a non-specific existential interpretation, as illustrated for WA in (79a) and for Farsi in (79b). This suggests that the plural marking on bare nouns and the plural marking in NCs must be of a distinct nature.[49]

(79) a. turs-ə   manug-ner t͡ʃ-en   xaɣ-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 restricted by animacy or definiteness hierarchies. Similarly, 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. Inanimate and non-specific subjects cannot trigger plural marking on the predicate (Kornfilt 1997, Bamyacı et al 2014, see also Alexiadou 2019). While we take Turkish to entirely 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 proper names, such as neighborhood names, that involves plural marking in NCs, as demonstrated in (80b). 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' (the name of a neighborhood)

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

In summary, primarily observed in WA and Farsi, the factors governing subject-predicate plural agreement are also involved in 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, akin to the Turkish classifier.[50] 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, albeit with a distinction: Agreement is contingent upon definiteness and specificity in WA, and solely on definiteness in Farsi. Therefore, while these languages resemble Turkish in featuring both an overt and a covert cardinal head, they diverge from Turkish by exhibiting a somewhat restricted version of number agreement in NCs.

---

[48] Ionin and Matushansky (2019) present the Chadic language Miya as an example where plural agreement on predicates and in NCs is conditioned by animacy. See Schuh (1998) for this generalization.

[49] A reviewer suggest that the plural marker in Farsi NCs could potentially be a plural definite article. However, this possibility is dismissed by the fact that bare plural nouns, while they can be definite, can also allow non-specific existential readings. The latter aligns with the DKP-based existential readings of bare plurals in English and Turkish (see Section 2).

[50] 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, akin to Turkish, also exhibits a count-mass distinction (Ghomeshi 2003 and Mache 2012). The countability 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 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 shown in (50b).

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 predicative, 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.

While NCs with tā can be definite, also triggerring plural agreement marking on the noun, as shown in (78), NCs without tā require a demonstrative to convey definiteness, as exemplified in (81) (cf. with the Turkish (65)). Since NCs are incompatible with definiteness in the absence of *tā*, number agreement does not occur, 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 reveals that Turkish is not the only language that features predicative NCs besides the default argumental form. It also underscores that the choice of which CARD head takes on the indefinite or predicative role is language-specific and unrelated to the morphological form of CARD. In Farsi, NCs with the covert CARD are the argumental type associated with a choice function, precluding covert ι type shifting. However, they can co-occur with demonstratives via *ident*, activated by the type mismatch arising when the NC merges with the D head. Conversely, NCs with tā have a predicative denotation, allowing for covert type shifting through ι or the choice function, as well as enabling combination with overt determiners.[51]

Our account predicts that the alternation between predicative and argumental NCs is observable only in languages without articles. In a language distinct from Turkish and Farsi in featuring a definite article, we expect both forms of NCs to require the article for definite interpretations. WA corroborates these predictions. Even if one form of CARD leads to predicative NCs (represented as CARD$_{pred}$ below) and the other form to argumental NCs, as in Turkish and Farsi, both are amenable to co-occurrence with the definite article —via *ident* for argumental NCs. The presence of an overt definite determiner will rule out the application of ι on the predicative form due to the Blocking Principle, even though the NC does not intrinsically impede it. Table 2 summarizes how covert *iota* type shifting is blocked/allowed in WA, Turkish, and Farsi.

	with THE	without THE	
The type of NCs	**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 overt and covert CARD forms

---

[51] Note on complex numerals: As mentioned in fn 20, Ionin and Matushansky (2006, 2019) propose a compositional derivation of complex numerals (cf. Rothstein 2017). When applied to our analysis, their view entails recurring cardinal heads, as in [[three CARD] [hundred CARD apples]], but this composition is ruled out with argumental NCs, as CARD requires a property as its argument. Activating *ident* type-shifting does not resolve the issue, as it returns a singleton set with its member chosen by f. However, [*three* CARD] would require a non-singleton set to return a set of individuals divisible into two non-overlapping individuals, each with 100 non-overlapping parts. As predicted, *tane* only follows the closest numeral to the noun, e.g., *üç (*tane) yüz (tane) elma* 'three hundred apples.' Predicative NCs also lack recurring cardinal heads, as evidenced by NCs with tā in Farsi, where tā cannot be iterated within a numeral complex, e.g., *si (*tā) sad (tā) sib* three hundred apples.' To address this, I replace Ionin and Matushansky's constraint in (39) with an atomic property requirement on the semantics of CARD, preventing predicative NCs from composing with another CARD. 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 example, *three hundred* is derived through a covert multiplication operator that takes two numbers and multiplies them.

To conclude, Table 3 summarizes the interpretations of NCs in the languages addressed in this paper. Recall that *presumably* CARD$_\emptyset$ categorization is utilized for obligatory classifier languages, as stated in Section 4.2.

	CARD$_\emptyset$		*presumably* CARD$_\emptyset$	CARD$_\emptyset$ & CARD$_{overt}$	
Languages with THE	**French**	**English**	**Yi**	**WA**	
indefinite NCs	✓	✓	✓	✓	
definite NCs with THE	✓	✓	✓	✓	
definite NCs with ι	✗	✗	✗	✗	
Languages without THE	**Russian**	**Hindi**	**Mandarin**	**Turkish**	**Farsi**
indefinite NCs	✓	✓	✓	✓	✓
definite NCs with Dem	✓	✓	✓	✓	✓
definite NCs with ι	✗	✗	✗	✗ with *tane* ✓ with CARD$_\emptyset$	✗ with CARD$_\emptyset$ ✓ with t$\bar{a}$

Table 3: The crosslinguistic interpretation of NCs

In languages with a definite article, NCs have the same interpretational distribution regardless of having one form of CARD or two. However, among articleless languages, two stand out by allowing definite interpretations in NCs via *iota*. These languages feature both overt and covert CARD in their NCs, highlighting the connection between the indefiniteness of NCs and cardinality. This connection then not only helps capture intra-language alternations but also sheds light on the crosslinguistic variations in the interpretation of NCs.

6 Revisiting the Agreement Approach

We have analyzed NCs as argumental expressions departing from the common view of NCs, which treats them primarily as predicates. In Section 4.1, we discussed number agreement in English NCs, initially assuming a predicative semantics for NCs. We now need to revise our account to reflect the argumental nature of NCs.

Recall Scontras's (2014, 2022) analysis of number marking in English NCs, repeated below. Essentially, 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, determined by cardinality in cardinal MPs. That is, SG checks whether every member of the set denoted by the MP evaluates to 1 with respect to cardinality. Conversely, PL surfaces if the one-ness requirement fails to be met.

(82)   Generalized MP
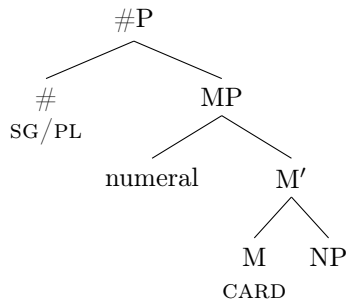
(83)   English # heads:
     a. $[\![\text{SG}]\!] = \lambda P : \forall \mu \forall x \in P \ [QU_\mu(P) \rightarrow \mu(x) = 1]. \ P$
     b. $[\![\text{PL}]\!] = \lambda P. \ P$
     c. $QU_\mu(P) = 1$ iff $\forall x \forall y \ [P(x) \wedge P(y) \rightarrow \mu(x) = \mu(y)]$

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 both studies, I present the following semantics for the 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 it denotes evaluates to 1 with respect to a measure, which is cardinality in NCs. If not, PL surfaces.

(85) Number marking with *one*:
   a. ✓ $\llbracket \text{SG } one \text{ CARD}_f \ book \rrbracket = f(\{a, b, c\})$
   b. ✗ $\llbracket \text{PL } one \text{ CARD}_f \ book \rrbracket = $ *failure due to Maximize Presupposition*

(86) Number marking with *two*:
   a. ✗ $\llbracket \text{SG } two \text{ CARD}_f \ book \rrbracket = $ *presupposition failure*
   b. ✓ $\llbracket \text{PL } two \text{ CARD}_f \ book \rrbracket = f(\{a \oplus b, b \oplus c, a \oplus c\})$

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 consistently occurs in English NCs, agreement is manifested solely in cases of definiteness and specificity in WA, and for definiteness in Farsi, as previously discussed. Detailing the syntactic mechanism for this restricted realization of number agreement exceeds the scope of this paper. Nonetheless, it is pertinent to make a note about number marking in Farsi NCs, which necessitates the presence of *tā*. Recall that NCs with *tā* differ from those with the covert CARD, as they are predicative and allow for definiteness through covert ι type shifting. Hence, number features, as functions on individual type expressions, should be able to take an NC with *tā* as an argument only after the NC undergoes ι type shifting.[52]

In Scontras's analysis, 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, yet plural morphology in these languages appears on the noun, not the classifier. This issue is resolved if WA and Farsi NCs have a structure akin to that of Turkish NCs, ensuring the closest head to #P is always the head of the lexical NP, not the M head. Therefore, I adopt the structure in (87) for WA and Farsi NCs, while leaving 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
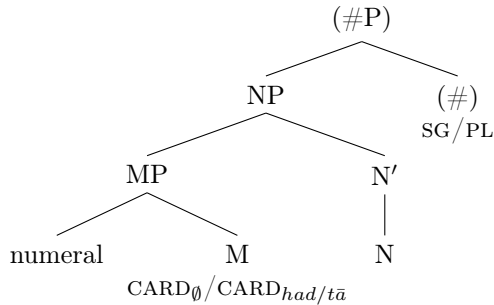
Before concluding, I should emphasize that in Sauerland's and Scontras's accounts, plural marking on all nominal projections is supplied by the # head. We have seen that plural marking on WA and Farsi bare nouns is not subject to the restrictions observed with the plural marker in their NCs (see (79)). Therefore, I utilize

---

[52] I have argued that Turkish NCs lack number agreement, except in cases noted in (80). Scontras (2014, 2022) proposes a distinct strategy for number marking in Turkish NCs, linking the one-ness presupposition of SG to relative atomicity. Martí (2020) adopts a similar logic, viewing Turkish number marking as 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 could naturally extend to Farsi as well. However, our #-heads are no longer functions on properties, making it challenging to maintain such an account. Additionally, explaining what triggers the switch from the Turkish system to the English system in case of specificity and definiteness could be challenging.

the # head solely to explain the semantics of number agreement in NCs (and possibly for subject-predicate agreement as well). I take the plural marking on a numeral-less lexical NP in WA, Farsi, and Turkish as the spell-out of the * operator, while leaving open the question of whether the same applies to 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 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 examined four key properties of Turkish NCs: (i) consistent use of unmarked nouns with all numerals, despite a systematic number marking system; (ii) optional use of the classifier *tane*; (iii) exceptional behavior of NCs without *tane*, which permit both definite and indefinite interpretations —unlike NCs with *tane* that are exclusively indefinite, following the general pattern of NCs across languages; and (iv) non-optional aspect of the classifier, i.e., how the presence or absence of *tane* influences the interpretation of NCs.

We first focused on the properties in (i) and (ii) to understand the optional aspect of the Turkish classifier and crosslinguistic variation in number marking patterns. Then, shifting our attention to the properties in (iii) and (iv), we examined the (in)definiteness puzzle of Turkish NCs and discussed the implications for the crosslinguistic interpretation of NCs.

The initial analysis compared *tane* with obligatory classifiers of Mandarin-like languages, highlighting its distinct nature. Drawing on Sağ's (2019, 2022) account of number marking semantics in Turkish and further examining the count-mass distinction, we concluded that the noun in Turkish NCs denotes a semantically singular property, the classifier notwithstanding. To explain number marking variations between English and Turkish NCs, I pursued an agreement-based approach, drawing on evidence from two more optional classifier languages, WA and Farsi. I 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 like Turkish morphologically exhibit the semantic singularity of the noun in their NCs, others, including English, obscure this through number agreement in NCs. The agreement account adopted here posits a cardinal head projection for uniform treatment of number marking with cardinal numerals and measure terms. I proposed that *tane* in Turkish is an overt spell-out of this head and denotes a counting function. While English-like and supposedly Mandarin-like languages involve only a covert CARD, Turkish, Farsi, and WA utilize both overt and covert exponents of it.

The second half of my analysis demonstrated a connection between CARD and inherently indefinite characteristics of NCs. Drawing on Jiang (2012), I proposed that CARD has a built-in choice function variable, thus yielding individual type NCs. With Dayal (2003), I take predicative interpretations of NCs to be a repair, invoked only for structural necessities. Intriguingly, inherently predicative and inherently argumental NCs can coexist in languages with both overt and covert CARD. We identified Turkish and Farsi as exhibiting this coexistence, albeit in mirror image patterns. In Turkish, the covert CARD, lacking the choice function, derives predicative NCs, while *tane* has an indefinite force, akin to cardinal heads in other languages. Farsi, by contrast, utilizes NCs with *tā* as the predicative form, while NCs with the covert CARD are canonically indefinite. Since both languages lack overt definite determiners, this variation manifests as an enrichment in interpretation, permitting both definite and indefinite readings with the form of CARD severed from the choice function. Our analysis predicts such enrichment to be absent from languages with articles, even if they employ both intrinsically predicative and indefinite NCs. Western Armenian validates this prediction.

If our discussion is on the right track, we are one step closer to a general understanding of cardinal semantics. Our 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

Arslan-Kechriotis ZC (2009) Referentiality in Turkish: NP/DP. In: Ay S, Aydın Özgür, Ergenç I, Gökmen S, Işsever 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, Barner D (2009) The interpretation of functional heads: Using comparatives to explore the mass/count distinction: Articles. Journal of Semantics 26(3):217–252

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: Proceedings of SALT 18, 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

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

Bayırlı IK (2017) The Universality of Concord. PhD 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

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. PhD thesis, Yale University

Corbett GG (2000) Number. Cambridge University Press, Cambridge

Croft W (2002) Typology and Universals, 2nd edn. Cambridge Textbooks in Linguistics, Cambridge University Press

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

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

Demirok Ö (2021) Intervention effects follow from scope rigidity in Turkish. In: Proceedings of SALT 31, pp 82–103

Diesing M (1992) Indefinites. No. Vol. 20 in Linguistic Inquiry Monographs, MIT Press

van der Does J (1992) Applied Quantifier Logics. PhD thesis, University of Amsterdam

Doetjes J (1997) Quantifiers and Selection. PhD 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

Farkas D (1981) Quantifier scope and syntactic islands. In: Proceedings of the 17th Chicago Linguistics Society (CLS 17), pp 59–66

Farkas D, de Swart H (2010) The semantics and pragmatics of plurals. Semantics & Pragmatics 3(6):1–54

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. PhD 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 (1991) The categorial status of quantified nominals. Linguistische Berichte (136):438–452

Giusti G (1996) Is there a FocusP and a TopicP in the Noun Phrase structure? Working Papers in Linguistics 6(2):105–128

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. PhD 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. PhD thesis, Stanford University

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. PhD 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 (2020) The suffix that makes Persian nouns unique, Advances in Iranian Linguistics: Current Issues in Linguistic Theory, vol 351, pp 107–118

Jiang LJ (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. PhD 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, Saarbrüken

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 R, van Benthem J, van Emde Boas P (eds) Semantics and Contextual Expression, De Gruyter Mouton, Berlin, Boston, 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 (1999) At least some determiners aren't determiners. In: Turner K (ed) The Semantics/Pragmatics Interface From Different Points of View, Elsevier, pp 1–257

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. PhD 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 Prarie 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 13(3):1–57

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 (2000) Where plural refuses to agree: feature unification and morphological economy. Acta Linguistica Hungarica 47(1-4):249–288

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

Öztürk B (2005) Case, Referentiality, and Phrase Structure. Linguistics today, J. Benjamins Publishing Company

Partee BH (1986) Noun Phrase Interpretation and Type-Shifting Principles, De Gruyter, Berlin, Boston, pp 115–144

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

Ritter E (1991) Two functional categories in noun phrases: Evidence from Modern Hebrew, Academic Press, New York, pp 37–62

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

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, pp 307–324

Sağ Y (2019) The Semantics of Number Marking: Reference to Kinds, Counting, and Optional Classifiers. PhD 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. PhD thesis, Harvard University.

Scontras G (2022) On the semantics of number morphology. Linguistics and Philosophy 45(5):1165–1196

Sigler M (1996) Specificity and Agreement in Standard Western Armenian. PhD 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

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 Heusinger 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. PhD 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

Winter Y (2001) Flexibility Principles in Boolean Semantics: Coordination, Plurality and Scope in Natural Language. MIT Press, Cambridge, MA

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

Zamparelli R (1995) Layers in the determiner phrase. PhD thesis, University of Rochester

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 not at all natural, 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-ı**   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., şarkıcılar '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 crosslinguistic pattern.