

***The three musketeers: plural marking
in Turkish nominal phrases with cardinal numerals***

Roumyana Pancheva & Yaqing Cao*

Abstract. Numerically-quantified noun phrases in Turkish prohibit plural marking, except in the case of reference to well-known groups such as *The Three Musketeers*. We argue that the observed plural in such expressions is not the regular *additive* plural, the plural that is required in English and that is prohibited in the general case in Turkish in the presence of numerals, but is an instance of a *group* plural, which appears structurally higher and co-occurs with a *proprial* feature in D, responsible for the proper-name-like interpretation of *The Three Musketeers*.

Keywords. numerals; plurals; groups; proper names; quasi-names; Turkish

1. Introduction. This paper is concerned with an apparent case of optionality of nominal number in numerically-quantified noun phrases in Turkish. The Turkish counterpart of the expression *three musketeers* may appear with or without a plural morpheme, but as we will show, this variation is associated with a difference in meaning. We further propose that the different meanings are derived from different structures.

Here is a preview. In the presence of the plural morpheme *lar* (which has a variable surface realization), *üç silahşör-ler* ‘three musketeer-s’, denotes the group known as *The Three Musketeers*, of the Alexandre Dumas’ novel by the same name. We suggest that this meaning is derived from a structure with a covert *proprial determiner* – the determiner that appears overtly with proper names in some languages – and a *group plural*, which is encoded in a Number head that is structurally higher than the numeral. In the absence of the plural marker *lar*, *üç silahşör* ‘three musketeer’, is interpreted as a definite or indefinite description – Turkish lacks an overt definite article – denoting a maximal plurality (definite) or some plurality (indefinite) of three musketeers, who may (in the case of the definite interpretation) be the three musketeers of Dumas’ novel. We posit that this meaning corresponds to a structure with a covert definite / indefinite determiner and a singular-marked Number head that is lower than the numeral.

1.1. LACK OF PLURAL MARKING IN NUMERICALLY-QUANTIFIED NOUN PHRASES. Cardinal numerals in Turkish cannot combine with plural-marked nouns, as can be seen in (1), where the addition of the plural morpheme to the noun in the subject phrase renders the sentence unacceptable. In the absence of a plural morpheme, the subject is still interpreted as plural, as evidenced by its semantic compatibility with the distributive predicate. The semantic plurality of the subject in (1) is not reflected by the verb morphology, as subject-verb number agreement in Turkish is optional with 3rd person subjects (Kornfilt 1997, a.o.).¹ We will refer to nominal phrases with cardinal numerals, as in (1), with the abbreviation C-Ns (for cardinal-nominals).

* Many thanks to the conference organizers and volume editors Young-Hoon Kim and Burak Oney, as well as to the two anonymous reviewers, whose comments improved the paper. Thanks also to Ümit Atlamaz, Duygu Demiray, Yağmur Kiper, Jaklin Kornfilt, Mete Oğuz, and Elif Ulusoy for judgments and discussion of the Turkish data, and to Luismi Toquero-Pérez for discussion of associative plurals and plural last names in Spanish. Authors: Roumyana Pancheva, University of California, Santa Cruz (pancheva@ucsc.edu) & Yaqing Cao, University of California, Santa Cruz (ycao83@ucsc.edu).

¹ There are subtle differences in interpretation between sentences with singular and plural verb agreement, which we will put aside here.

- (1) Üç öğrenci(*-ler) birer birer gel-di.
 three student-PL one.each one.each come-PAST.3SG
 ‘Three students came one by one.’

The obligatory lack of plural marking in Turkish C-Ns is notable, because the plural morpheme is obligatory for plural interpretation of arguments that are not numerically quantified, as can be seen in (2a), where the absence of the plural morpheme in the subject leads to unacceptability.² Plural marking is similarly obligatory in complements to partitive numerals, as illustrated in (2b). The subjects in (2) are both semantically and morpho-syntactically plural, but they too do not obligatorily trigger number agreement on the verb.

- (2) a. Öğrenci*(-ler) birer birer gel-di.
 student-PL one.each one.each come-PAST.3SG
 ‘Students came one by one.’
 b. Öğrenci*(-ler)-den üç-ü birer birer gel-di.
 student-PL-ABL three-AGR one.each one.each come-PAST.3SG
 ‘Three of the students came one by one.’

That C-Ns in Turkish do not admit plural marking is also notable in light of cross-linguistic variation in nominal number in C-Ns: Finnish, Estonian, Hungarian and others also prohibit plurals, while English, the Slavic and Germanic languages, among others, require plurals, and while still others, like Western Armenian, optionally allow plurals (see e.g., Ionin & Matushansky 2018). Understanding this variation is important and relevant for the related issues in (3).

- (3) a. the link between number morphology and interpretation
 b. the semantics of numerals
 c. the position of merge and interpretation of the number feature relative to numerals

With respect to (3a), ideally, morphological number would consistently match with semantic number, yet in Turkish semantic plurality obtains both with unmarked *öğrenci* ‘student’ in (1) and with plural-marked *öğrenciler* ‘students’ in (2). If nouns that are not overtly marked as plural in Turkish are interpreted as number neutral, i.e., as predicates of atomic individuals and their sums, rather than as singular, as they are in English, there will be no mismatch between nominal number morphology and interpretation, and no cross-linguistic variation in the grammar of C-Ns. But if unmarked nouns in Turkish have singular semantics, as in English, then there must be grammatical differences between C-Ns in Turkish and English that determine the absence/presence of plural morphology.

This brings us to the issue in (3b). Semantic theories differ in whether they assign numerals like *three* the semantics of numbers (of type *n*) or of intersective or subsective predicate modifiers. Semantic theories also differ as to whether they posit that numerals combine with semantically plural or singular predicates. Given that numerals share a number-denoting core cross-linguistically, what other grammatical formatives are employed in C-Ns to yield the observed cross-linguistic variation? How does Turkish fit in the cross-linguistic typology?

² Direct objects that are unmarked for case can receive a singular or plural interpretation in the absence of plural marking, arguably as an instance of pseudo incorporation. A reviewer points out that this is the case for some subjects as well; it is possible that these are unaccusative themes. We do not address these issues here; see Sağ 2022.

Issue (3c) is concerned with the way nominal number is encoded in C-Ns. If number features are merged below numerals, as more traditionally thought, then the necessity of a plural marker, as in English C-Ns, or the prohibition against a plural marker, as in Turkish C-Ns, could be accounted for by positing a different semantics for numerals or for covert measure expressions combining numerals and nouns, in the two types of languages. But if number features are merged higher than the numeral, as sometimes suggested, it becomes possible to account for the variation solely by attributing a different semantics to number marking itself.

Given the importance of the issues in (3), both for the proper analysis of C-Ns, and for nominal number more generally, the lack of plural marking in C-Ns in Turkish has received a fair amount of attention in the theoretical literature, including from a cross-linguistic perspective (e.g., Görgülü 2012, Sağ 2018, 2019, 2022; Sağ 2024, Alexiadou 2019, Bale et al. 2011, Ionin & Matushansky 2018, Scontras 2022). A number of different accounts have been offered, with different takes on the issues in (3a), (3b) and (3c) for Turkish and for the wider typology of C-Ns.

Much less attention has been devoted to the possibility of plural marking in C-Ns in Turkish, although it bears on the theoretical questions concerning the analysis of nominal number marking and numerals, both in Turkish and cross-linguistically. We turn to this phenomenon next.

1.2. PLURAL MARKING IN NUMERICALLY-QUANTIFIED NOUN PHRASES. Contrary to the general prohibition against plurals, C-Ns that denote well-known groups regularly admit the plural marker (as briefly noted in Göksel & Kerslake 2005: 148, Görgülü 2012: 70, Sağ 2019: 153, Sağ 2024: ex. (80)). In fact, in the case of the C-Ns in (4), the plural is required for reference to the specific *groups*: the three musketeers Athos, Porthos and Aramis of the eponymous novel by Alexandre Dumas in (4a); the seven dwarfs from the fairy tale *Snow White* by the Brothers Grimm in (4b); and the forty thieves from the folk tale *Ali Baba and the Forty Thieves* in (4c). In the absence of the plural marker, the C-Ns in (4) can refer to definite pluralities, i.e., a maximally unique group of three musketeers, seven dwarfs or forty thieves, (or for that matter, to indefinite pluralities of three musketeers, seven dwarfs or forty thieves). In fact, the plural-less definite descriptions may refer to the very same individuals that the C-Ns with the plural marker refer to, e.g., the plurality of Athos, Porthos and Aramis, etc. This is why we mark the absence of the plural marker with a #, rather than a *. We will come back to this issue.

- (4) a. üç silahşör-#(ler)
 three musketeer-PL
 ‘The Three Musketeers’ (of the novel *The Three Musketeers*)
- b. yedi cüce-#(ler)
 seven dwarf-PL
 ‘The Seven Dwarfs’ (of the *Snow White* fairy tale)
- c. kırk harami-#(ler)
 forty thief-PL
 ‘The Forty Thieves’ (of the folk tale *Ali Baba and the Forty Thieves*)

We will refer to the C-Ns like those in (4) as gC-Ns (for group-denoting nominals with cardinal numerals) and we will aim to provide an analysis that also speaks to the right analysis of C-Ns.

The prior literature has taken gC-Ns to show that the lack of plural marking in C-Ns in Turkish is optional to some limited extent, conditioned by factors such as definiteness. Such optional-

ity is easier to accommodate on some accounts of C-Ns than others, and so gC-Ns could potentially adjudicate between those accounts. Yet, we argue here that gC-Ns employ a different kind of plural from the one prohibited in Turkish C-Ns and allowed in English C-Ns, thereby rendering the lack of plural marking in Turkish C-Ns categorical.

Specifically, we suggest that the plural seen in gC-Ns is a *group* plural, while the plural seen in English C-Ns and prohibited in Turkish C-Ns is an *additive* plural. Both are different from yet another kind of plural, an *associative* plural, that Turkish also has. While we do not offer here an analysis of the precise semantics of group plurals, we identify some distinct aspects of their meaning that justify the addition of this new category of plural.

2. gC-Ns as quasi proper names. We propose that gC-Ns are quasi proper names, i.e., they are semantically similar to proper names but are formed on the basis of common nouns. An illustrative example of a quasi-name is *Mom*, based on the common noun *mom*. gC-Ns are like *Mom*, but plural, though of course they lack the indexicality of *Mom*, which refers to the speaker's mother or to the addressee's mother (e.g., in child-directed speech). Like proper names and quasi-names, and unlike definite descriptions, gC-Ns refer rigidly, resisting bound readings.

2.1. REFERENTIAL AND PREDICATIVE USES OF PROPER NAMES AND QUASI-NAMES. Before we illustrate the difference in the availability of bound readings, we need to clarify that proper names and quasi-names can be used in two ways. Their referential use is illustrated in (5). Here the proper name *Ana* and the quasi-name *Mom* occur as bare singulars, not accompanied by an overt determiner. This is the realization of proper names and quasi-names that doesn't allow bound readings. This is what we suggest that gC-Ns are like: they have the interpretation of referential proper names and quasi-names.

(5) Ana/Mom won her first gold medal in the Paris Olympics tonight.

Proper names and quasi names also have a predicative use. Quasi-names are formed on the basis of common nouns, so naturally, those common nouns appear as predicates. But proper names too may appear as predicates, as in (6), accompanied by a determiner. The definite description in (6c) is similar to the form of *The Three Musketeers*, but it is not a gC-N. Rather, it is a definite C-N. The semantic difference is particularly sharp in the case of *the three moms*, which doesn't have the indexicality component of *Mom*.

- (6) a. the Ana/mom that we both love
b. every Ana/mom that I know
c. the three Anas/moms

Thus, there is a distinction between a referential use and a predicative use of proper names and quasi-names (the latter being just the ordinary common nouns). When we say that gC-Ns are quasi-names, we mean that they are formed on the basis of a common noun, but have the semantics of referential quasi-names, like *Mom* in (5) (except plural), and not that of definite descriptions, like *the three moms* in (6c). The latter is the interpretation of definite C-Ns.

The difference in meaning between gC-Ns (quasi-names) and definite C-Ns is much more subtle than that between *Mom* and *the mom*, but is nevertheless real. It exists in English between *The Three Musketeers* (a quasi-name) and *the three musketeers* (a definite description), although

it is not reflected in the form of the two expressions, but only in the orthography. In Turkish, the distinction is overtly realized by the presence vs. the absence of the plural marker.

- (7) a. üç silahşör-ler
 three musketeer-PL
 ‘the Three Musketeers’ (gC-N: a referential quasi-name)
- b. üç silahşör
 three musketeer
 ‘(the) three musketeers’ (C-N: a(n in)definite description)

2.2. BOUND READINGS. The bare proper noun *Ana* and quasi-name *Mom* behave differently from the definite descriptions *the Ana* and *the mom*. The examples in (8) are from Muñoz (2019).

- (8) a. In every race, Mom/Ana always wins.
 b. In every race, the mom/the Ana always wins.

The interpretation of (8a) is that a single individual – who is named Ana, or who is (typically) the speaker’s mother – wins in all the races. On the other hand, a possible reading of (8b) is that a different person wins in each race and that person is whoever is called Ana or whoever is a mom. In other words, definite descriptions allow bound readings, while referential proper names and quasi-names do not. There is also another reading of (8b), easily accessible for the definite description *the mom*, where the same individual, who is a contextually unique mother, wins in all the races. In the case of the predicative use of the proper name, *the Ana*, the bound reading is the more salient one, and perhaps even the only one possible.

The same contrast is found in Turkish numerically quantified nominals: in (9a), in the presence of the plural marker, we have a gC-N and it does not allow a bound reading. Athos, Porthos and Aramis win in every single race, there’s no variation of musketeers by race. This is in contrast with (9b) where in the absence of a plural we have a C-N and a possible reading is that each race is won by potentially different three musketeers, whoever the three musketeers are in that race. This is the bound reading. In the absence of the ‘whoever they are’, it is also possible to interpret the C-N as a definite plurality of individuals who win all the races (not a bound reading).³

- (9) a. Her yarışta üç silahşör-ler (Dumas-ın romanındaki) daima kazanır.
 each race three musketeer-PL (Dumas-of novel) always win
 ‘In every race, The Three Musketeers (of Dumas’ novel) always win.’
- b. Her yarışta üç silahşör (onlar kim olursa olsun) daima kazanır.
 each race three musketeers (they who if though) always win
 ‘In every race, the three musketeers (whoever they are) always win.’

The contrast in the availability of a bound reading between (9a) and (9b) suggests that gC-Ns in Turkish behave like proper names and quasi-names, instead of definite descriptions.

³ A reviewer finds that the sentences in (9) would be more natural if the material in the parenthesis precedes the gC-N/C-N, and in particular, if in (9b), the pronoun *onlar* ‘they’ is omitted.

The English quasi-name *The Three Musketeers* and definite description *the three musketeers* behave just like their Turkish gC-N and C-N counterparts, as the translations of (9) reveal. Thus, any structural differences that underlie the interpretative differences between definite descriptions and quasi-names must not be tied solely to the presence or absence of the plural marker, nor to the exponence of the determiner. *The* can evidently express whatever features are on D in definite descriptions and in quasi-names. Similarly, in Turkish, in the absence of an overt article, the relevant features on D may remain covert in both definite descriptions and in quasi-names.

2.3. A FURTHER ELABORATION ON DEFINITE C-NS. Notably, the C-N *üç silahşör* ‘three musketeer’, when interpreted as definite, *can* refer to Athos, Porthos and Aramis. The example in (10) is acceptable and natural in the given context. This shows that the gC-N *üç silahşörler* ‘The Three Musketeers’ does not have exclusivity rights on reference to the individuals in the plurality.

(10) *Context: Athos, Porthos and Aramis were all invited to the dinner.*

Üç silahşör birer birer gel-di.
 three musketeer one.each one.each come-PAST.3SG
 ‘The three musketeers came one by one.’

Because both *üç silahşör-ler* and *üç silahşör* can refer to the same individuals, the addition of the plural marker *-LAR* might be considered optional in C-Ns. This might be why the semantic difference between the quasi-name and the definite description has been previously overlooked. And this is why earlier, in the example in (4), we indicated the effect of the absence of the plural marker *-lAr* on the meaning with a # rather than a *.

The English definite description *the three musketeers* too can refer to the individuals covered by the quasi-name *The Three Musketeers*, as the translation of (10) in the given context reveals. The gC-Ns and C-Ns in the two languages behave identically, despite the presence/absence of the plural marker and the overt/covert realization of the D head.

2.4. PROPRIAL D. Given that proper names and quasi names may readily occur as predicates, it is tempting to analyze their referential use as also having a predicative source. A uniform account would posit the same underlying [D NP] structure that would be realized as a definite description with an overt definite article as in (11a), and as a bare proper name or quasi-name when the D is null, as in (11b) (see e.g., Matushansky 2008, Fara 2015).

(11) a. [_{DP} the Ana/mom] (definite description)
 b. [_{DP} \emptyset_{the} Ana/Mom] (proper name/quasi-name)

The semantic difference between proper names and quasi-names, on the one hand, vs. definite descriptions on the other, with respect to the availability of bound readings, casts doubt on the analysis of the former as having the structure in (11b), with an unpronounced definite article. But this is so only if we assume that the overt and covert articles are semantically the same. This need not be the case. Some languages have so-called *proprial* articles that occur with proper names and quasi-names and that are distinct from definite articles, lending support for the idea that proper names and quasi-names are intrinsically predicates that always appear in a DP structure with a determiner. When proper names and quasi-names are used referentially, D is realized

by a proprial article and when they are used as a definite description, D is realized by the definite article. In other words, instead of (11), we would have (12).

- (12) a. [_{DP} the Ana/mom] (definite description)
 b. [_{DP} ∅_{proprial} Ana/Mom] (proper name/quasi-name)

The Māori examples in (13), taken from example sentences in Muñoz (2019), illustrate the proprial article with a personal proper name and with a name of an object, and also one of the definite articles in the language.

- (13) a. a Puhihuia, a Mātaatua
 PROPRIAL Puhihuia PROPRIAL Mātaatua
 ‘Puhihuia’ (a woman’s proper name), ‘Mātaatua’ (the name of a voyaging canoe)
 b. te kotiro o te rangatira
 the girl of the chief
 ‘the daughter of the chief’

The proprial article in Māori occurs with names but not with common nouns.⁴ It may sometimes be omitted from proper names in argument position, but what conditions this omission is unclear.

However, the definite article can also be used to form proper names in Māori. Walker (1969) quotes Buck (1950): “... creating a list of proper names by converting common names such as *kore* (nothing) and *po* (night, darkness) into proper names by prefixing the definite article *te* (the) and so creating Te Kore (the void) and Te Po (the unknown)”. Some other examples from Walker (1969), which we glossed, are in (14). The first example is a toponym (and it also includes an instance of the proprial article), and the last one is a personal proper name. In the terminology used here, both have the form of a quasi-name, as they are derived on the basis of common nouns.

- (14) a. Te Kupenga a Kupe
 the fishing-net PROPRIAL Kupe
 ‘Kupe’s fishing net’ (a place near Jackson’s Head, Queen Charlotte Sound)
 b. Te Waharoa
 the Waharoa
 ‘Waharoa’ (a person’s name, after *waha roa*, the main gateway to a settlement)

We take the lessons from Māori to be that there is a PROPRIAL feature, responsible for the referential properties of proper names and quasi-names, that can be hosted by D and that can be overtly realized as a specialized article, *a*, or remain covert as it is realized together with the regular definite article, *te*. Thus *a Kupe* and *te Waharoa* have the same structure, with a D hosting a DEF and a PROPRIAL feature, as in (15), and the realization of that D as *a* or *te* depends on the class membership of the lexical noun: a (predicative) proper name or a common noun.

⁴ Although Muñoz (2019) notes that it may also occur with some locational nouns and with personal pronouns.

(15) [_{DP} D_[DEF, PROPRIAL] Kupe/Waharoa]

What this means for Turkish and for English is that, putting the issue of plural marking aside, gC-Ns have the structure in (16a), while definite C-Ns have the structure in (16b).

(16) a. [_{DP} D_[DEF, PROPRIAL] three musketeers] (gC-N: a quasi-name)
 b. [_{DP} D_[DEF] three musketeers] (C-N: a definite description)

In Turkish, in the absence of a definite article, both Ds are covertly realized. In English, in the absence of a specialized proprial article, both Ds are realized as *the*. Of course in English, the proprial feature in D may also have a null realization, as with proper names like *Ana*. Like the case of Māori, this reflects sensitivity to the class membership of the lexical noun: D_[DEF, PROPRIAL] is realized as phonologically null in combination with proper names, like *Ana*, and as *the* in combination with common nouns, like *musketeer*.⁵

This account, of course, is also not complete without offering an explicit semantics for the PROPRIAL feature, but we will not do so here; see Muñoz (2019). We just note that the semantics of the proprial feature is responsible for the lack of bound readings.

3. Nominal number in C-Ns and gC-Ns. So far we have proposed that gC-Ns are quasi-names which differ from their definite description C-N counterparts in having a PROPRIAL feature in D. We now turn to the role of number marking. As we have also seen, C-Ns in Turkish do not allow the plural marker *-lar*, while gC-Ns allow it, and in the case of *The Three Musketeers*, among other gC-Ns, in fact require it. We address number marking in Turkish C-Ns first.

3.1. NO PLURAL MARKING IN C-NS. Cardinality measures have three necessary ingredients. One, there need to be roots denoting precise numbers, (17a). Two, there need to be expressions encoding a cardinality measure function, a function (of type $\langle e, n \rangle$) that maps individuals to numbers representing their cardinality (17b). Three, in order for the measure function to output numbers greater than 1, there need to be semantically plural predicates, (17c), where the * is Link (1983)'s operator that pluralizes predicates, adding sums to their denotation.

(17) a. $\llbracket \sqrt{\text{three}} \rrbracket = 3$
 b. $\llbracket \mu_{\text{card}} \rrbracket = \lambda x. |x|$
 c. $* \llbracket \text{student} \rrbracket = \lambda x. x \text{ is a student or students}$

How these three ingredients – precise number morphemes, cardinality measure functions, and the pluralizing operator * – are realized in C-Ns could be subject to cross-linguistic variation, and is still very much under debate. The combinatorial options, when it comes to the lexical semantics of numerals, can be seen in (18).

⁵ Names like *Canada* vs. *the United States*, *the Beatles* conform to this pattern, yet it is unclear why the quasi-name *Mom* surfaces with a null D rather than with a *the*. Furthermore, the picture becomes more complicated once modifiers are introduced, as can be seen in (i), from Fara (2015), where the non-restrictive modifier makes the proper name compatible with *the*. The nature of the modifier is considered relevant because non-restrictive modifiers require the use of *the*. We have to put this issue aside; see Matushansky (2008), Fara (2015), a.o. for more examples and discussion.

(i) (The) brave Philip slew the dragon with just one blow.

- (18) a. $\llbracket \ddot{u}\check{c} \rrbracket = \lambda P_{\langle e,t \rangle} \lambda x_e. P(x) \wedge |x| = 3$
 b. $\llbracket \ddot{u}\check{c} \rrbracket = \lambda P_{\langle e,t \rangle} \lambda x_e. *P(x) \wedge |x| = 3$
 c. $\llbracket \ddot{u}\check{c} \rrbracket = 3$

The number root and the cardinality measure function may both be encoded in numerals, as in (18a). This is a very common view of the literature, perhaps the predominant one (Link 1983, Partee 1987, Landman 2004, Chierchia 2010, Rothstein 2017, a.o.) and a version of it is endorsed by Bale et al. (2011) for Turkish. Numerals with the semantics in (18a) can combine directly with nominal predicates, similarly to adjectives, and need to combine with semantically plural predicates. Bale et al. (2011) suggest that cross-linguistic variation in the semantics of nouns is responsible for the observed differences in whether or not C-Ns allow plural markers. Nouns unmarked for number in Turkish are said to be number neutral, i.e., they are weakly plural, denoting predicates of atoms and their sums, as in (19); this is why no plural marking is needed in C-Ns. Unmarked nouns in English denote predicates of atomic individuals, and this is why a plural marker is needed to make them suitable arguments to numerals.

- (19) $\llbracket \ddot{o}\check{g}r\check{e}n\check{c}i \rrbracket = * \llbracket student \rrbracket = \lambda x. x$ is a student or students

Ionin & Matushansky (2018) go a step further and propose that numerals do not just incorporate the number-denoting root and the cardinality measure function but also the * operator, as in (18b); they don't quite put it this way, but this is what their proposed semantics for numerals effectively is like. Numerals with the meaning in (18b) combine directly with semantically singular predicates. Cross-linguistically, nouns in C-Ns denote singular predicates, as in (20), regardless of their morphology. Cross-linguistic variation in number marking in C-Ns for these authors is the result of a morpho-syntactic agreement rule and has no semantic basis. Sağ (2018) adopts this proposal for Turkish.

- (20) $\llbracket \ddot{o}\check{g}r\check{e}n\check{c}i \rrbracket = \llbracket student \rrbracket = \lambda x. x$ is a student

Alternatively, it might be that numerals only encode number-denoting roots, as in (18c), leaving the measure function and the * operator to be expressed separately. This view, largely informed by work on languages with numeral classifiers, has also been suggested for languages without obligatory classifiers (e.g., Krifka 1995, Zabbal 2005). It is adopted for Turkish in Scontras (2013, 2022), Sağ (2018, 2019); Sağ (2024), and Martí (2020). Numerals with the semantics in (18c) cannot combine directly with nominal predicates but need a covert linker. This linker may in principle incorporate just the cardinality measure function, (21a) or both the cardinality measure function and the * operator, (21b). The two make different demands on their complements: $CARD_1$ can combine with semantically plural predicates, while $CARD_2$ combines only with semantically singular predicates and pluralizes them.

- (21) a. $\llbracket CARD_1 \rrbracket = \lambda P \lambda n \lambda x. P(x) \wedge |x| = n$
 b. $\llbracket CARD_2 \rrbracket = \lambda P : \forall y. P(y) \rightarrow atom(y) \lambda n \lambda x. *P(x) \wedge |x| = n$

Both options for the semantics of the linker expression have been entertained for Turkish. On the one hand, Sağ (2019) adopts a $CARD_2$, as in (21b). The claim is that Turkish nouns in

C-Ns are semantically singular, as are English nouns, as in (20); $CARD_2$ supplies the requisite pluralizer itself (though this is not explicitly stated). The plural morphology seen in English C-Ns has no semantic contribution; it is just the reflex of morpho-syntactic agreement. On the other hand, Scontras (2022) and Martí (2020) adopt a $CARD_1$, as in (21a). For these authors, Turkish nouns that are not marked plural have number neutral semantics, as in (19). Although they lack the plural marker *-lar*, the nouns are semantically (weakly) plural and so are suitable arguments for $CARD_1$. Furthermore, on their analysis, there is a number feature higher than the numeral, which is singular in value in Turkish, but it can compose with its semantically plural complement, because its lexical semantics is expressed in terms of minimality rather than atomicity, as in (22a). The different semantics of English singular number, as in (22b) prevents it from being expressed above the numeral, and so English C-Ns require a high plural number, that then gets expressed on the noun.

- (22) a. $\llbracket SG_{Turkish} \rrbracket = \lambda P \lambda x. P(x) \wedge \neg \exists y. P(y) \wedge y \sqsubset x$
 b. $\llbracket SG_{English} \rrbracket = \lambda P \lambda x. P(x) \wedge atom(x)$

As for our own position, we agree with Sağ (2019) when it comes to Turkish, although we do not adopt the view that plural marking in English C-Ns is the result of morpho-syntactic agreement. We suggest that the cross-linguistic typology is best handled by adopting the view that numerals are number-denoting, as in (18c), and that they compose with nominal predicates via the mediating role of either of the two measure-function-incorporating linkers in (21). Languages like English make use of $CARD_1$ and plural marking delivers the needed semantically plural predicates. Languages like Turkish, Estonian and Finnish have $CARD_2$, which combines with semantically singular predicates, and thus C-Ns in these languages preclude plural marking. And then there are languages which exhibit noun class splits in C-Ns, like Bulgarian, where masculine non-human denoting nouns are marked singular, while all other nouns are marked plural; such languages can be said to have both $CARD_1$ and $CARD_2$ (Pancheva 2023). Figure 1 illustrates the (limited) cross-linguistic typology and (23) shows the details of the semantic interpretation.

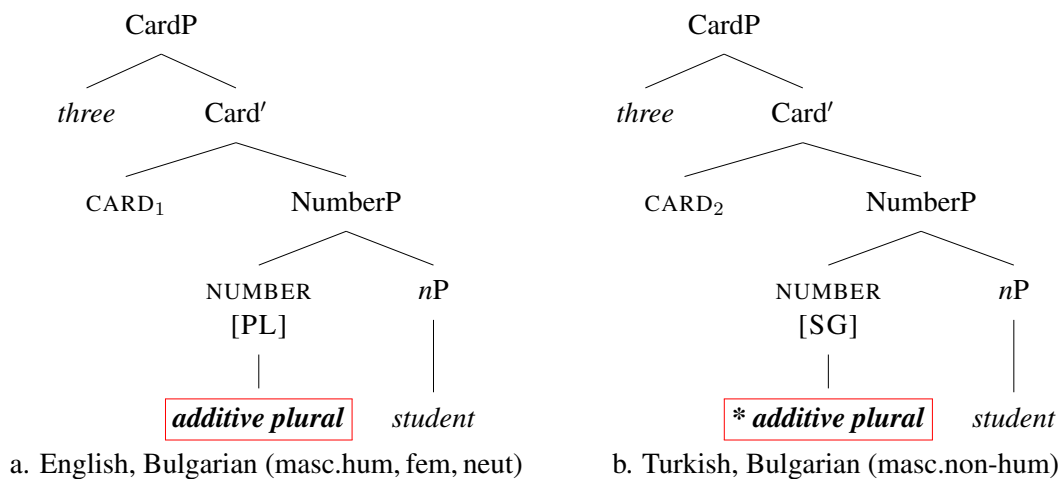


Figure 1: Structures for C-Ns cross-linguistically

The interpretation of C-Ns proceeds as follows. *n*Ps unmarked for number are number neu-

tral, in Turkish, as well as in English, Estonian, Bulgarian, etc., (23a).⁶ The lexical semantics of singular number doesn't vary either, (23b). Its effect on combining with nP is to output a predicate true of atoms only, (23c). This predicate is an appropriate argument of $CARD_2$, which is the measure-function encoding linker in Turkish, (23d). Because $CARD_2$ also encodes the pluralizing * operator, it returns a weakly plural predicate, (23d). In contrast, in English C-Ns, a plural is needed, because the measure-function encoding linker in this language is $CARD_1$. A singular predicate as in (23c) can compose with $CARD_1$ but then could only ever combine with the numeral *one*. Plural number – known as *additive plural* – has the meaning in (23f), an exclusive plural, which upon combining with the nP returns a predicate of sums only, (23g).⁷ This is the right meaning for English $CARD_1$, (23h)-(23i), but it would not be composable with Turkish $CARD_2$. Finally, numerals in the languages under consideration denote numbers, (23j). The resulting interpretation for C-Ns in Turkish and English can be seen in (23k) and (23l).

- (23) a. $\llbracket [_{nP} \text{öğrenci}] \rrbracket = \llbracket [_{nP} \text{student}] \rrbracket = \lambda x. x \text{ is a student or students}$
 $\{a, b, c, a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c, \dots\}$
- b. $\llbracket [SG] \rrbracket = \lambda P \lambda x. P(x) \wedge atom(x)$
- c. $\llbracket [_{NumberP} SG \text{öğrenci}] \rrbracket = \llbracket [_{NumberP} SG \text{student}] \rrbracket = \lambda x. x \text{ is a student}$
 $\{a, b, c, \dots\}$
- d. $\llbracket [CARD_2] \rrbracket = \lambda P : \forall y. P(y) \rightarrow atom(y) \lambda n \lambda x. *P(x) \wedge |x| = n$
- e. $\llbracket [_{Card'} CARD_2 [_{NumberP} SG \text{öğrenci}]] \rrbracket = \lambda n \lambda x. x \text{ is a student or students} \wedge |x| = n$
 $\{a, b, c, a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c, \dots\}$
- f. $\llbracket [PL] \rrbracket = \lambda P \lambda x. P(x) \wedge \neg \exists y. P(y) \wedge atom(y)$
- g. $\llbracket [_{NumberP} PL \text{öğrenci}] \rrbracket = \llbracket [_{NumberP} PL \text{student}] \rrbracket = \lambda x. x \text{ is students}$
 $\{a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c, \dots\}$
- h. $\llbracket [CARD_1] \rrbracket = \lambda P \lambda n \lambda x. P(x) \wedge |x| = n$
- i. $\llbracket [_{Card'} CARD_1 [_{NumberP} PL \text{student}]] \rrbracket = \lambda n \lambda x. x \text{ is students} \wedge |x| = n$
 $\{a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c, \dots\}$
- j. $\llbracket [üç] \rrbracket = \llbracket [three] \rrbracket = 3$
- k. $\llbracket [_{CardP} üç [CARD_2 [_{NumberP} SG \text{öğrenci}]]] \rrbracket = \lambda x. x \text{ is student or students} \wedge |x| = 3$
 $\{a \oplus b \oplus c, \dots\}$
- l. $\llbracket [_{CardP} three [CARD_1 [_{NumberP} PL \text{student}]]] \rrbracket = \lambda x. x \text{ is student or students} \wedge |x| = 3$
 $\{a \oplus b \oplus c, \dots\}$

We endorse the above account, which shares common elements with other approaches though not in their totality, because we think that the alternative approaches face difficulties, empirical or conceptual. Bale et al. (2011)'s account cannot be extended to Estonian and Finnish, where nouns that are not overtly marked plural are interpreted as strictly singular, i.e., as predicates of atomic individuals, just like in English. The account also has problems with Turkish. Sağ (2022) argues that Turkish unmarked nouns in argument positions have singular, not number neutral, semantics and that apparent number neutrality in non-case-marked position arises via singular kind reference. On the other hand, there is a mismatch between morphological number and semantic number on the approach that treats plural marking in C-Ns as morpho-syntactic agreement (Ionin &

⁶ Languages may differ in how their nPs end up denoting predicates of individuals, but this is not of relevance here.

⁷ An inclusive/weak plural semantics may also work; we do not pursue this question here.

Matushansky 2018, Sağ 2018, 2019). Finally, the approach where the semantics of singular number varies (Scontras 2022, Martí 2020) would be challenged by noun class splits like the ones in Bulgarian, as it is specifically designed not to allow plural marking in C-Ns in a language whose singular number has minimality semantics, as in (22a).

3.2. PLURAL MARKING IN gC-NS. Given the account of C-Ns defended above, in order for the plural *-lar* seen in gC-Ns to be an additive plural, merged instead of SG, Turkish would need to have $CARD_1$ in addition to $CARD_2$ in its inventory. If so, we would expect to see variation in noun class, similar to Bulgarian. Some initial evidence may point in this direction. All of our consultants require the plural marker in the gC-Ns in (4), where the lexical nouns are human. Although dwarfs are technically supernatural beings in Germanic folklore, it is not unusual that Happy, Grumpy and the rest of the dwarfs are treated by the grammar as human. On the other hand, the same consultants preferred not to have the plural marker in the examples in (24), which are gC-Ns that also refer to well-known groups, but which are formed on the basis of non-human nouns. *The Three Little Pigs* is a fable, and *The Three Little Wolves and The Big Bad Pig* is a children’s book by Evgenios Trivizas. Both are translated into Turkish without the plural marker.

- (24) a. üç domuzcuk(#-lar)
 three piggy-(PL)
 ‘The Three Little Pigs’ (of the fable)
 b. üç küçük kurt(#-lar)
 three little wolf-(PL)
 ‘The Three Little Wolves’ (of the childrens book)

Yet it is not clear that there is a human-ness restriction on plural marking in gC-Ns in Turkish. A reviewer notes that (25) is an acceptable toponym, both in its bare and plural-marked form.⁸

- (25) üç tepe(-ler)
 three peak-(PL)
 ‘The Three Peaks’

We do not have a good understanding of this variation of plural marking in gC-Ns. Often, the form of toponyms is a historical relic, reflecting grammatical processes that are no longer part of the synchronic grammar. Thus, non-human denoting gC-Ns as in (25) do not justify positing $CARD_1$ in Turkish in addition to $CARD_2$. If this were the right analysis, we would expect to see this variation of plural marking in indefinite and definite C-Ns as well, yet this does not happen.⁹ We suggest that the plural in all gC-Ns is the value of a Number head above CardP, as in Figure 2, a plural we call *group plural*. Up to CardP, gC-Ns and C-Ns are structurally identical.

⁸ Similarly, we find *beş yüz ev-ler* ‘five hundred house-PL’ as the name of a neighborhood, in Sağ (2024).

⁹ The distinction between definite C-Ns and gC-Ns is of course subtle. But a diagnostic exists. A reviewer notes that place names always have non-final stress, even when their common noun counterparts have final stress, e.g., *avcı-lar*, a district in Istanbul vs. *avcı-lar* ‘hunter-s’, where underlining indicates the position of stress. In the presence of a numeral, the plural marker only appears with place names, i.e., with gC-Ns, and never C-Ns, as identified by their stress pattern.

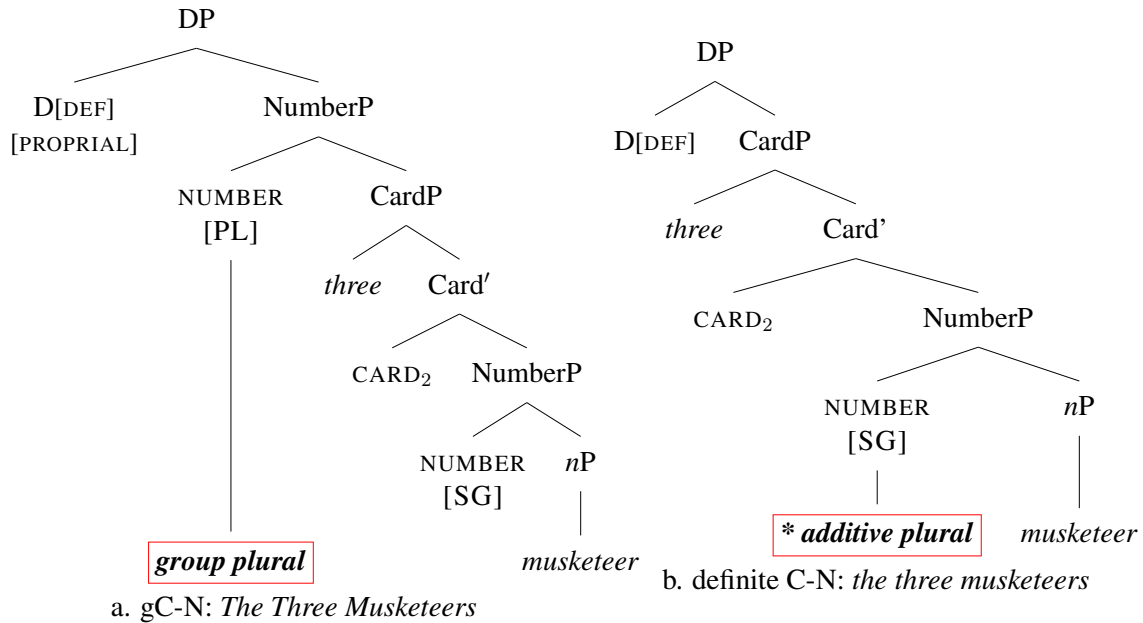


Figure 2: Structures for gC-Ns and C-Ns in Turkish

The higher PL can compose with CardP, given the meaning in (23f), and provided that the numeral is not *one*. Yet its presence is not necessary, if the meaning in (23f) is all that the group plural contributes. There must be an additional component to the meaning of this plural that is responsible for the group interpretation. Landman (1989)'s group-forming operator \uparrow is a good candidate: it conditions the appearance of the group plural.

There are some similarities between the group plural and the *associative* plural, which Turkish also has. Both appear high in the structure and both involve reference to groups. However, associative plural nominals differ from gC-Ns in that the group is associated with a single individual (the focus), whereas gC-Ns lack grounding in a focus. Associative plurals place [PL] even higher, above a singleton-denoting DP (the focus), and a functional head that relates the focus to individuals that are the focus' associates.

The positions of the two plurals can be seen relative to the possessive marker, which is at the left edge of the DP. Example (26), from Görgülü (2011): 72-73, reveals that the additive plural appears low, (26a), whereas the associative plural appears higher than D, (26b).

- (26) a. *abi-ler-im*
 brother-PL-1 SG.POSS
 'my brothers' (additive)
- b. *abi-m-ler*
 brother-1 SG.POSS-PL
 'my brother and his family/associates/friends' (associative)

We cannot offer here a precise semantics for the group and associative plurals, particularly in light of the variation in (24)-(25), leaving this task for another occasion.

4. Some cross-linguistic support. Spanish last names that refer to family groups offer evidence for the co-existence of a high plural with a low singular number, just like in the structure we proposed for Turkish gC-Ns. Mandarin C-Ns prohibit the plural marker *-men* but gC-Ns allow *-men*, lending support for our proposal that the appearance of the plural marker in Turkish C-Ns is restricted to group-denoting expressions.

4.1. SPANISH LAST NAMES. For a last name in Spanish to be group denoting, it must have a plural-marked article but the name itself may not be plural, see (27a). Pluralizing the name as in (27b) yields only an interpretation of a (maximal) plurality of individuals sharing the name, who may or may not belong to the same family. The examples in (27) are from Camacho (2021).

- (27) a. los Velasco
 the.PL Velasco.SG
 ‘The Velasco family’
 b. los Velasco-s
 the.PL Velasco-PL
 ‘The Velasco family members’ or ‘the individuals named Velasco’

The nominal plural seen in (27b) has the meaning that we would expect from an additive plural. The plural on the determiner is best analyzed as the result of concord. The expression *los Velascos* has the properties expected of a definite description, including allowing bound readings, but instead of a common noun it is formed on a predicative name. On the other hand, the plural in (27a) is a high plural as it is realized on the definite article and it is not the result of concord because the nominal itself is not marked plural. We submit that this high plural is a group plural. In line with our analysis of gC-Ns, the definite article further hosts a PROPRIAL feature. The name *los Velasco* does not allow bound readings, as expected.

4.2. MANDARIN GC-NS. In Mandarin regular C-Ns prohibit plural marking, (28a), yet the plural marker *-men* is optional in the case of gC-Ns, (28b). As far as we know, the possibility of *-men* appearing in numerically-quantified noun phrases has not been previously noted.

- (28) a. san-ge nanren-(*men)
 three-CL man-(*PL)
 ‘three men’
 b. san jianke-(men), qi-ge xiaairen-(men), qi-xiannv-(men)
 three musketeer-(PL), seven-CL dwarf-(PL), seven-fairy-(PL)
 ‘The Three Musketeers’ (of Dumas’ novel), ‘The Seven Dwarfs’ of (*Snow White*),
 ‘The Seven Fairies’ (in Chinese folklore)

Though the theoretical status of *-men* in Mandarin is still controversial – some have treated it as a plural morpheme realized in D (Li 1999), and others as a collective marker (Li & Thompson 1989; Zhu 1982), one consensus is that that *-men* alone does not encode an associative plural. Adding *-men* to a proper name in an attempt to express the associative meaning yields unacceptability, (29a). The associative plural requires the presence of the plural pronoun *tamen* ‘they’ after the focus noun, (29b).

- (29) a. xiaozhang-(*men)
 little-Zhang-(*PL)
 attempted: ‘Little Zhang and his associates’
- b. xiaozhang-tamen
 little-Zhang-they
 ‘Little Zhang and his associates’

5. Conclusion. Plural marking in Turkish numerically-quantified noun phrases is restricted to group-denoting expressions like *The Three Musketeers*. We argued that such expressions are quasi proper names rather than definite descriptions. We suggested that their plural marker is not an additive plural but a group plural, merged higher, and co-occurring with a PROPRIAL feature in D. The rest of the structure of group-denoting *The Three Musketeers* is shared with the description (*the*) *three musketeers*: both involve a measure-function encoding CARD that mediates between number-denoting numerals and nouns that denote semantically singular predicates.

References

- Alexiadou, Artemis. 2019. Morphological and semantic markedness revisited: the realization of plurality across languages. *Zeitschrift für Sprachwissenschaft* 38. 123–154. <https://doi.org/10.1515/zfs-2019-0004>.
- Bale, Alan, Michaël Gagnon & Hrayr Khanjian. 2011. Cross-linguistic representations of numerals and number marking. In *Proceedings of SALT XX*, 582–598. CLC Publications. <https://doi.org/10.3765/salt.v20i0.2552>.
- Buck, Sir Peter. 1950. *The coming of the Maori*. Whitcombe and Tombs, Wellington.
- Camacho, Jose. 2021. The structure of plural last names in Spanish and other languages. *Glossa: a journal of general linguistics* 6(1): 140. 1–40. <https://doi.org/10.16995/glossa.5773>.
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149. <https://doi.org/10.1007/s11229-009-9686-6>.
- Fara, Delia Graff. 2015. Names are predicates. *The Philosophical Review* 124(1). 59–117. <https://doi.org/10.1215/00318108-2812660>.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: a comprehensive grammar*. Routledge.
- Görgülü, Emrah. 2011. Plural marking in Turkish: Additive or associative? In *Working Papers of the Linguistics Circle* 1, 70–80. <https://journals.uvic.ca/index.php/WPLC/article/view/5814>.
- Görgülü, Emrah. 2012. *Semantics of nouns and the specification of number in Turkish*. Burnaby, British Columbia: Simon Fraser University PhD thesis. <https://summit.sfu.ca/item/12565>.
- Ionin, Tania & Ora Matushansky. 2018. *Cardinals: the syntax and semantics of cardinal-containing expressions*. MIT Press. <https://doi.org/10.7551/mitpress/8703.001.0001>.
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Krifka, Manfred. 1995. Common nouns: A contrastive analysis of English and Chinese. In Greg Carlson & Jeffrey Pelletier (eds.), *The Generic Book*, 398–411. Chicago: University of Chicago Press.
- Landman, Fred. 1989. Groups, I. *Linguistics and Philosophy* 12(5). 559–605. <https://doi.org/10.1007/BF00627774>.
- Landman, Fred. 2004. *Indefinites and the type of sets*. Blackwell. <https://doi.org/10.1002/9780470759318>.

- Li, Charles & Sandra Thompson. 1989. *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Li, Yen-hui Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8. 75–99. <https://doi.org/10.1023/A:1008306431442>.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use and interpretation of language*, 302–323. Berlin: De Gruyter. <https://doi.org/10.1515/9783110852820>.
- Martí, Luisa. 2020. Numerals and the theory of number. *Semantics and Pragmatics* 13. 1–57. <https://doi.org/10.3765/sp.13.3>.
- Matushansky, Ora. 2008. On the linguistic complexity of proper names. *Linguistics and Philosophy* 31. 573–627. <https://doi.org/10.1007/s10988-008-9050-1>.
- Muñoz, Patrick. 2019. The proprial article and the semantics of names. *Semantics and Pragmatics* 12. 1–36. <https://doi.org/10.3765/sp.12.6>.
- Pancheva, Roumyana. 2023. Morphosyntactic variation in numerically-quantified noun phrases in Bulgarian. *Journal of Slavic Linguistics* 31. 1–18. <https://ojs.ung.si/index.php/JSL/article/view/176>.
- Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds.), *Studies in discourse representation theory and the theory of generalized quantifiers*, chap. 5, 115–141. Dordrecht: Foris. <https://doi.org/10.1515/9783112420027>.
- Rothstein, Susan. 2017. *Semantics for counting and measuring*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9780511734830>.
- Sağ, Yağmur. 2024. On the (non) optionality of the Turkish classifier *tane*. *Natural Language & Linguistic Theory* 1–58. <https://doi.org/10.1007/s11049-024-09626-4>.
- Sağ, Yağmur. 2018. The semantics of Turkish numeral constructions. In Uli Sauerland & Stephanie Solt (eds.), *Sinn und bedeutung* 22, 307–324. <https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/108>.
- Sağ, Yağmur. 2019. *The semantics of number marking: reference to kinds, counting, and optional classifiers*. New Brunswick, New Jersey: Rutgers University PhD thesis. <https://rucore.libraries.rutgers.edu/rutgers-lib/61931/PDF/1/play/>.
- Sağ, Yağmur. 2022. Bare singulars and singularity in Turkish. *Linguistics and Philosophy* 45. 741–793. <https://doi.org/10.1007/s10988-021-09323-0>.
- Scontras, Gregory. 2013. A unified semantics for number marking, numerals, and nominal structure. In Emmanuel Chemla, Vincent Homer & Grégoire Winterstein (eds.), *Proceedings of Sinn und Bedeutung* 17, 545–562. <https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/360>.
- Scontras, Gregory. 2022. On the semantics of number morphology. *Linguistics and Philosophy* 45. 1165–1196. <https://doi.org/10.1007/s10988-021-09345-8>.
- Walker, R. J. 1969. Proper names in Maori myth and tradition. *The Journal of the Polynesian Society* 78(3). 405–416. <http://jstor.org/stable/20704644>.
- Zabbal, Youri. 2005. The syntax of numeral expressions. Unpublished Ms., University of Massachusetts, Amherst. <https://ling.auf.net/lingbuzz/005104>.
- Zhu, Dexi. 1982. *Yufa jiangyi [Lectures on Grammar]*. Shangwu Publishers, Beijing.