

CHAPTER 1

TANE AS A MARKER OF DISCRETENESS IN TURKISH

Emrah GÖRGÜLÜ¹

INTRODUCTION

The word *tane* in Turkish has been investigated in various ways by different researchers over the past a few decades and there have been a number of different analyses about what it actually is and what it does in the structure. In this work, I will review some of the prominent descriptive and theoretical accounts proposed by those such as Lewis (1975), Underhill (1976), Skilliter (1986), Schroeder (1999), Öztürk (2005), Göksel and Gerslake (2005) and Sağ (2019, 2021), among others. Based on a new corpus analysis on the uses of *tane* in the language, I will argue that it is not functioning as an optional classifier and also not a semantic marker that is an overt realization of an otherwise covert cardinal head, as has been proposed in recent theoretical analyses. Instead, I will show that *tane* should be considered to be a pragmatic marker rather than a semantic one. More specifically, *tane* appear within the numeral construction when the referent, whether it is singular or plural, needs to be marked as discrete in the context. This way various uses of it can be uniformly captured in the language.

The structure of this paper is as follows: in the next section, I provide an overview of various accounts that deal with the nature of *tane* and its use in Turkish. This section also includes various

¹ Assoc. Prof. Dr., İstanbul Sabahattin Zaim University, emrah.gorgulu@izu.edu.tr

counterarguments that show that the existing analyses do not fully capture its true behavior. In the section following, based on a corpus work, I propose my own account and argue that *tane* is functioning as a marker of discreteness. In other words, the various uses of *tane* can be uniformly captured as marking discrete entities or individuals. The final section is the conclusion of the paper where I also make certain suggestions for remaining work.

TANE IN TURKISH

It is well-attested that in Turkish numeral constructions, the numeral itself appears with the head noun without any other grammatical elements, as in (1) and (2).²

1. bir öğrenci
one student
'one student'
2. beş elma
five apple
'five apples'

In that respect, Turkish is different from languages such as English in that the head noun does not need to get marked for plurality when the numeral is greater than one, as in (3). It also exhibits differences from languages like Chinese since there is no need to use an additional elements known as classifiers that are used to categorize the referent in terms of shape, animacy or some other inherent characteristics (Aikhenvald 2000, 2006), as in (4),

² The abbreviations used in the glosses are as follows: 1 = first person; 2 = second person; 3 = third person; ABIL = ability; ACC = accusative case; ABL = ablative case; CAUS = causative marker; CLF = classifier; COND = conditional; DAT = dative case; EVID = evidential marker; FOC = focus marker; GEN = genitive case; INST = instrumental case; LOC = locative case; NEG = negation; OBJP = object participle; PASS = passive voice; PAST = past tense; PL = plural; POSS = possessive marker; PROG = progressive aspect; SG = singular; SUBJP = subject participle

even though the head noun also does not get plural marking when the numeral is greater than one in these languages.

3. a. two students / *two student

b. five boxes / *five box

4. wǔ běn shū

five CLF book

'five books'

On the other hand, the Turkish lexical item *tane* that can be translated as 'item', 'piece' or 'unit' can appear in numeral constructions, as in (5) and (6).

5. bir tane öğrenci

one item student

'one student'

6. beş tane elma

two item/piece apple

'two apples'

The question that arises at this point is what is the contribution of this lexical element in the structure. Is it syntactic, semantic or pragmatic in nature? Historically, *tane* is considered to be a lexical item borrowed from Persian which is itself a classifier language (Skilliter 1986, Schroeder 1999).³ In the next section, I will provide a brief overview of earlier descriptive and theoretical analyses on the subject.

PREVIOUS WORK ON *TANE*

Various researchers have pointed out its different uses in the language. For instance, Lewis (1975) treats *tane* as a classifier and focuses on its various functions such as being an anaphor and a

³ *Tane* in its bare form means 'seed' or 'grain' in the language.

pro-form. Similarly, Underhill (1976) notes that it is a counting word and associates its use with similar elements such as *paket* 'pack' as in *üç paket bisküvi* 'three packets of biscuits' and *bardak* 'glass' as in *bir bardak ayran* 'one glass of ayran'. However, it should be noted that the fact that *tane* shares certain characteristics with actual classifiers in languages such as Chinese and Thai does not necessarily mean that it should be called and treated as such in Turkish. Classifiers are obligatory elements in numeral constructions and there is a strong semantic relationship between them and other elements such as numerals and nouns in the structure. In that sense, the idea that it is some sort of a classifier is just an illusion and would lead to wrong predictions.

Schroeder (1999), in his study, notes that *tane* has a number of different uses, one of them being used with a new referent denoting an entity introduced in the discourse as an antecedent for subsequent reference. This usage is positively correlated with a referent having a high degree of pragmatic referentiality or the quantity is either emphasized or contrasted, as in (7). In addition to that, Schroeder notes that *tane* is also used in a wh-question asking about the (sizeable) number of distinct entities, as in (8).

7. Delhi'li bir adam-la tanış-tı-m. **Üç tane**

New Delhi-from a man-INST meet-PAST-1SG three tane

kız-ı var-dı. Hiç oğl-u yok.

girl-POSS-3SG exist-PAST no son-POSS-3SG NEG.exist

Evlen-dir-miş **bir tane-si-ni**.

marry-CAUS-EVID one tane-POSS-ACC

'I met a man from New Delhi. He had three daughters. He had no son. One of them (the daughters) he gave in marriage.'

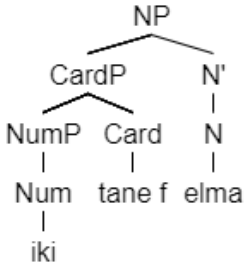
8. Kaç tane yatağ-ınız var?

How many tane bed-POSS.2PL exist

'How many beds have you got?'

On the other hand, in a number of recent theoretical studies, Sağ (2019, 2021) argues that *tane* is a semantic marker that is an overt realization of an otherwise covert head. Following Scontras (2014), Sağ argues that counting is universally ensured by a cardinal head across languages and *tane* is the overt counterpart of this head in Turkish. She also maintains that the morphological contrast in Turkish and English nouns in numeral constructions can be explained by the presence or absence of number agreement, as was proposed by Ionin and Matushansky (2006, 2019). More specifically, *tane* is needed for counting itself and numeral constructions are formed around the cardinal head. This cardinality is what achieves counting in the language. The syntactic and semantic structure of *tane* would then be as in (9) and (10) respectively.

9. NCs with *tane*



10. $((tane_f)) = \lambda n \lambda P_{AT}. f(\lambda x \exists S (\Pi(S)(x) \wedge |S| = n \wedge \forall s \in S P(s)))$

The proposed syntactic structure for numeral constructions with *tane* is given in (9). Basically, Sağ argues that every numeral construction within a noun phrase universally includes a null cardinal head. That cardinal head denotes the cardinality measuring function since measuring (i.e. counting) is achieved by this null head and not by numerals themselves. On the other hand, the numeral construction with *tane* is its overt counterpart, as illustrated in (9). In other words, *tane* is taken to be the overt realization of this measuring function. The only difference is that

when *tane* appears in the sentence, the entire numeral construction is assigned the indefinite reading. The presence of *tane* here is only optional since it does not have a close relationship with nouns as a true classifier does in classifier languages.

However, this line of reasoning does not seem to be so viable for several reasons. First of all, it is rather difficult to see why an overt semantic marker of cardinality would have an effect on the pragmatic characteristics (in-)definiteness of noun phrases in which numeral constructions are embedded. Consider the example in (11).

11. Masa-da üç (tane) elma ve beş (tane) portakal

table-LOC three (piece) apple and five (piece) orange

var-dı. Üç (*tane) elma çürük görün-üyor-du.

exist-PAST three (piece) elma rotten look-PROG-PAST

‘There were three (pieces) of apples and five (pieces) of oranges on the table. The three (*pieces of) apples looked rotten.’

As illustrated by the anaphoric relationship in (11), numeral constructions that do not include *tane* can get a definite interpretation in Turkish. On the other hand, it is well-known since Schroeder (1999) that numeral construction that include *tane* are not able to be an anaphor for an antecedent in the discourse. The presence of *tane* will always indicate indefiniteness in these cases. However, it would be rather difficult to see why this would be the case and to make a logical connection between a semantic notion such as cardinality on the one hand, and a pragmatic notion like (in-)definiteness on the other.

It should also be noted here that the presence of *tane* in numeral constructions is not always required for the entire noun phrase to have an indefinite interpretation. In other words, its absence in the structure does not always indicate definiteness, as exemplified in (12) and (13).

12. On kitap al-mış-tı. Kitap-lar-ın üç
ten book buy-PERF-PAST book-PL-GEN three
tane-si-ni kaybet-ti.
piece-POSS-ACC lose-PAST
'She bought ten books. She lost three of the books.'
13. On kitap al-mış-tı. **Kitap-lar-ın üç-ün-ü**
ten book buy-PERF-PAST book-PL-GEN three-POSS-ACC
kaybet-ti.
lose-PAST
'She bought ten books. She lost three of the books.'

What we see in (12) and (13) is that even though *tane* does not appear in the latter, it has the same (specific) indefinite reading as in the former with *tane*. This clearly indicates that the claim that its presence is not always required for the noun phrase to have an indefinite reading.

Finally, the presence of *tane* is not always allowed in numeral constructions, as shown in (14) and (15).

14. Üç (*tane) hafta geçt-ti.
three item week pass-PAST
'Three weeks passed.'
15. Beş (*tane) dolar borç al-dı.
five item dollar loan take-PAST
'He borrowed five dollars.'
16. On (*tane) metre sıçra-dı.
ten item meter jump-PAST
'He jumped ten meters.'

What the sentences in (14), (15) and (16) illustrate is that if the head noun in the numeral construction refers to a stretch of time,

distance or some currency, *tane* is banned from the structure.⁴ However, if *tane* were really the overt realization of a cardinal head, it would be able to appear in numeral constructions across the board. However, this is clearly not the case. In the next section, I will propose my own account of *tane* and argue that it primarily functions as a marker of discreteness.

AN ANALYSIS OF PRAGMATIC DISCRETENESS

Based on the argumentations made in the previous section, I propose that there is no need to postulate such a semantic cardinal head, whether it is overt or covert, in Turkish.⁵ This is because *tane* seems to have functions other than being an overt head of a counting mechanism. Instead, following Schroeder (1999), Göksel and Kerslake (2005) and Görgülü (2012) who note that *tane* marks distinctness of referents in Turkish, I argue that the use of *tane* is not entirely optional but its presence is required when the reference in the structure is to an entity that is regarded as a discrete object or an individual. This idea of discreteness is also different from the use of numeral classifiers in languages like Chinese and Thai in that classifiers semantically provide some kind of individuation and their presence is obligatory in these languages. However, *tane* should be considered to be more like a pragmatic marker that provides discreteness to the head noun, an idea first entertained by Serzisko (1980). In other words, the referent stands out as something discrete and among others. This can be exemplified by the corpus data collected through the TS Corpus v2 (Sezer and Sezer 2013).

⁴ The asterisk inside the parentheses (i.e. (*xxxxx)) indicates that the presence of the element inside the parentheses is unallowed.

⁵ Note that the arguments made for *tane* can be probably expanded to adet 'item/piece' and 'piece' that are used with non-human nouns in the language. However, I will not pursue that in this paper and leave it for further research.

17. Üstelik de ikinci dünya harb-i-nde adam-lar
moreover FOC second world war-POSS-LOC man-PL
baş-lar-ın-a **iki tane atom bomba-sı** ye-di-ler.
head-PL-POSS-DAT two item atom bomb-POSS get-PST-PL
'Moreover, these men got two atom bombs on their heads in the
second world war.'
18. **On beş tane hakemlik pozisyon** ol-muş. Sen o-nun
fifteen item refereehood position be-PERF you it-GEN
onüç-ün-ü doğru değerlendir-e-me-miş-sin.
thirteen-GEN-ACC right evaluate-ABIL-NEG-EVID-2SG
'There were fifteen positions that called the attention of the
referee. However, you failed to evaluate thirteen of them.'
19. Yukarı-da ad-ı geç-en eser-de Türki
above-LOC name-POSS mention-SBJP work-LOC Turkic
dil-ler-in **otuz dokuz tane** ol-duğ-u tespit
language-PL-GEN thirty nine item be-OBJP establishing
ed-il-miş.
do-PASS-EVID
'In the work mentioned above, it has been established that there
are thirty-nine Turkic languages.'

The examples above illustrate that when *tane* is used in numeral constructions, its main function is to mark the referent of the head noun stand out as a discrete entity. This is the case regardless of whether *tane* is used in between the numeral and the head noun, as in (17) and (18) on the one hand, or whether it appears as a pronominal element, as in (19). Therefore, the conclusion here is that even though the presence of *tane* in numeral construction is not obligatory, it plays an important role of the marker of discreteness when it does appear in them.

This idea of discreteness also manifests itself when *tane* is used as the head of partitive constructions, which Schroeder calls a pro-form in his descriptive work. In this usage, *tane* functions as the head in a genitive-possessive or an ablative-possessive construction. This is exemplified in (20) and (21).

20. Katıl-an atmış dört otomobil-in kırk bir tane-si
participate-SBJP sixty-four car-GEN forty-one item-POSS
benzin-li ve **yirmi üç tane-si** dizel motor-a
gas-with and twenty-three item-POSS diesel motor-DAT
sahip-ti.
possess-PAST

‘Of the sixty-four cars that participated, forty-one of them had gas motors and twenty-three of them had diesel motors.’

21. On yedi kaza-dan sadece **bir tane-si** ciddi-ydi.
seventeen accident-ABL only one item-POSS serious-PAST

‘Of the seventeen accidents, only one of them was serious.’

The examples above indicate that when *tane* is used as the head of a partitive (i.e. as a proform) in the structure, it marks the referent or the referents as discrete entities, standing out among others. In both cases, *tane* would be omitted and the sentences would still be fine since it is also possible to have *sadece bir-i* in (21) where *tane* is not the head of the partitive. However, its presence in the numeral is to mark one entity or multiple entities as being discrete.

Note, however, that I digress from arguments found in analyses such as Schroeder (1999) in that he maintains that *tane* is almost never used in numeral constructions in which the numeral is over twenty. He goes on to say that the numbers range from one and ten in numerals where *tane* appears. However, when we look closely at the data, especially the data in the corpora, we immediately observe that there is in fact no apparent restriction for *tane* to be

used along with any numeral in Turkish. This is exemplified in (22), (23) and (24).

(22) Ders-i **üçyüz elli tane öğrenci** al-ıyor-du.

class-ACC three hundred fifty item student take-PROG-PAST
'Three hundred and fifty students were taking the class.'

(23) **Yüz tane trafik kazası** yap-san-ız nafile.

one hundred item traffic acc.-POSS have-COND-2PL useless
'It is in vain even if you have one hundred traffic accidents.'

(24) Çıkartma-ya **bin tane asker** katıl-dı.

landing-DAT one thousand item soldier participate-PAST
'One thousand soldiers participated in the landing.'

The grammaticality of the sentences above and the appropriateness of the use of *tane* in them clearly shows that *tane* is compatible with any numeral and not just with smaller numbers, unlike what was proposed in prior descriptive studies.

CONCLUSION

In this paper, I investigated various functions of *tane* in Turkish and argued that even though it was a borrowed element from Persian along with its classifier function, it does not have this semantic function as a true classifier as in Chinese or Thai in its modern usage. Nor is it an overt realization of a semantic marker of cardinality, as it was proposed to be functioning as in more recent theoretical work. Instead, I proposed that it should be considered to be a marker of discreteness that has the function of making the referent of the noun stand out as a discrete entity in the numeral construction. Moreover, I also showed that *tane* can be used with any numeral and the bigness or smallness of the numeral does not appear to be a restricting or determining factor. Further work on these issues will shed more light on the topic.

REFERENCES

- Aikhenvald, Y. A. (2000). *Classifiers: A typology of noun categorization devices*. Oxford studies in typology and linguistic theory. Oxford University Press.
- Aikhenvald, Y. A. (2006). Classifiers and noun classes: Semantics. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*. (2nd Edition). pp. 320-325. Elsevier: Oxford.
- Göksel, A. & Kerslake, C. (2005). *Turkish: An essential grammar*. Routledge: London & New York.
- Görgülü, E. (2012). Semantics of nouns and the specification of number in Turkish. Unpublished PhD Dissertation. Simon Fraser University.
- Ionin, T. & Matushansky, O. (2006). The composition of complex cardinals. *Journal of Semantics*, 23 (4), 315-360. Doi: <https://doi.org/10.1093/jos/ffl006>
- Ionin, T. & Matushansky, O. (2019). *Cardinals: The syntax and semantics of cardinal-containing expressions*. Cambridge, MA: MIT Press.
- Lewis, G. L. (1975). *Turkish grammar*. Oxford: Oxford University Press.
- Öztürk, B. (2005). *Case, referentiality and phrase structure*. Linguistics Today. John Benjamins Publishing Company.
- Sağ, Y. (2019). The semantics of number marking: reference to kinds, counting, and optional classifiers. Unpublished PhD Dissertation. Rutgers University.
- Sağ, Y. (2021). On the (non-)optionality of the Turkish classifier *tane*. Ms. Harvard University.
- Schroeder, C. (1999). *The Turkish nominal phrase in spoken discourse*. Turkologica 40. Wiesbaden: Harratssowitz.
- Sezer, T. & Sezer, B. (2013). TS Corpus herkes için Türkçe derlem. *Proceedings of the 27th National Linguistics Conference*. May 3-4 2013. Antalya, Kemer: Hacettepe University, English Linguistics Department. pp. 217-225.
- Scontras, G. (2014). The semantics of measurement. Unpublished PhD Dissertation. Harvard University.
- Serzisko, F. (1980). Sprachen mit zahlklassifikatoren: Analyse und vergleich. Akup 37.
- Skilliter, S. A. (1986). The family of *tane*: Ottoman classifiers and rules governing their usage. Wiener Zeitschrift für die Kunde des Morgenlandes, Vol. 76, Festschrift Andreas Tietze zum 70. 255-261.
- Underhill, R. (1976). *Turkish grammar*. Cambridge, MA: MIT Press.