

From approximation to probability: *bi'* + numeral constructions in Turkish

Nihal Yıldız Gökçe *

Abstract. This paper provides a semantic analysis of *bi'* + numeral constructions in Turkish (e.g., *Markete varmak için bi' 20 metre daha yürüyeceksin*). Although there are similarities between *bi'*+ numeral constructions in Turkish and *some*+ numeral constructions in English, there are unexplained situations about the *bi'* structure in Turkish. We claim that the presence of the indefinite determiner *bi'* 'some' in such constructions is licensed by the presence of the probability operator PROB at the sentence level.

Keywords. formal semantics; probability; numeral; indefinite

1. Introduction. One construction in Turkish that is curious from a semantic point of view is the *bi'*+ numeral construction, an example of which is given in (1).

- (1) *Market-e var-mak için bi' 20 metre daha yürü-(y)ecek-sin.*
 Market-DAT reach-INF for some 20 meter more walk-FUT-2SG
 'You will walk more some 20 meters to reach the market.'

As we shall review shortly, there are some similarities between *bi'*+ numeral constructions in Turkish and *some*+ numeral constructions in English. This raises the possibility that the semantic analysis of *some*+ numeral constructions in English can provide a model for the semantic analysis of *bi'*+ numeral in Turkish.

Some researchers (see. Sauerland and Stateva (2007), Anderson (2014)) have claimed that *some* + numeral constructions give an approximate meaning. On the other hand, Solt and Stevens (2018) brought a new perspective to the semantic analysis of the *some*+ numerical structures, arguing that it would be inadequate to analyze only approximation for *some* + numerical structures.

While Solt and Stevens (2018) provided evidence that *some* can be used with non-round numbers, there are some generalizations about Turkish that their analysis fails to account for, as we shall shortly see.

Besides, there has been no semantic analysis of *bi'* + numeral structures in Turkish. This paper aims to examine *bi'* + numeral constructions in Turkish and provide a novel approach to this expression. We claim that the presence of the indefinite determiner *bi'* 'some' in such constructions is licensed by the probability operator PROB at the sentence level.

This outline of the paper is as follows. Section 2 explains the similarities between *bi'* + numeral constructions in Turkish and *some* + numeral constructions in English. Section 3 presents previous analyses of *some* + numeral expressions in English. Section 4 focuses on the use of *some* + numeral and *bi'*+ numeral constructions with non-round numbers. Section 5 discusses some differences between *some* + numeral constructions in English and *bi'* + numeral constructions in Turkish. In Section 6, I present my proposal in which *bi'* + numeral expressions come with the probability operator PROB.

* Nihal Yıldız Gökçe, TOBB University of Economics and Technology (nihalyildizgokce@gmail.com).

2. Similarities between some + numeral construction and *bi'* + numeral. There are some similarities between *bi'* + numeral constructions in Turkish and *some* + numeral constructions in English. Solt and Stevens (2018) observe that *some* in English may be used with measures in the mass domain (2) and with temporal (3) and spatial (4) extents. However, *some* cannot be used with exact temporal (5) and spatial points (6).

- (2) Anna bought some 5 ounces of gold.
- (3) a. Anna sang for some 45 minutes.
b. Anne moved to Berlin some 20 years ago.
- (4) The tree is some 10 meters from the house.
- (5) *Anna started singing at some 3 o'clock.
- (6) *Anna moved to Berlin in some 1990.

The distribution of *bi'* + numeral constructions in Turkish is quite similar to what we find in English (see 7,8, 9, 10, and 11 below).

- (7) Hasan *bi'* 5 kilo elma satın.al-dı.
Hasan some 5 kilogram apple buy- PST
'Hasan bought some four kilograms apple.'
- (8) a. Nisa *bi'* 2 saat ders anlat-acak.
Nisa some 2 hour lesson teaches- FUT
'Nisa will teach a lesson for some 2 hours'
b. Meryem *bi'* 10 yıl önce Ankara'dan taşın-dı.
Meryem some 10 year ago Ankara- ABL move-PST
'Meryem moved to Ankara some 10 years ago.'
- (9) Ağaç okul-dan *bi'* 30 metre uzaklık-ta.
tree school-ABL some 30 meter away-LOC.
'The tree is some 30 meters from the school.'
- (10) *Bölüm toplantı-sı *bi'* 09.00'da yap-ıl-dı.
Department meeting-POSS some 09.00-LOC do- PASS- PST
'The department meeting was held at 09.00.'
- (11) *Bu ev *bi'* 2013'te inşa.et-il-miş-tir.
This house some 2013+ LOC build-PASS-PST-EPIS
'This house was built in 2013.'

The examples that are given above show that *about* or *approximately* can be used instead of *bi'* and *some* constructions in these expressions.

3. Previous research about *some*+ numeral construction. Sauerland & Stateva (2007) claim that the semantic analysis of *some* + numeral expressions in English is equivalent to the semantic analysis of *about*. They claim that both *some* and *about* give rise to approximative readings, as in (12) and (13). Sauerland & Stateva (2007) analyzed *some* and *about* with the granularity function (*gran*), which maps each point of a scale to an interval that contains it.

- (12) $\|about D\|^{gran} = coarsest(gran)(\|D\|)$
- (13) $\|some D\|^{gran} = coarsest(gran)(\|D\|)$

Moreover, Sauerland & Stateva (2007) show that on the coarsest scale, 49 and 50 belong to the same range. They suggest that one cannot use a non-round number like 49 to indicate this

range, instead 50 should be used. According to Sauerland and Stateva (2007), non-round numbers cannot be used with approximating operators.

(14) *What John cooked was approximately/some/about 49 tapas.

(15) What John cooked was approximately/some/about 50 tapas.

Anderson (2014) discusses the analysis of *some* + numeral expressions with Lasersohn's (1999) halo effect. Lasersohn (1999) uses halo to explain the uncertainty about the exact value in approximating constructions. For instance, in various contexts, *3:00 pm*, in the pragmatic halo of *2:58 pm*, can be used instead of *2:58 pm*.

Anderson (2014) suggests the combination of the precise representation of the number and the pragmatic halo effect of the number with a selection function named *f* in the use of *some* with numerical expressions (16).

(16) $\|twenty\|^c = \|twenty\| \cup halo_c(\|twenty\|)$

Anderson's semantic analysis of *some* also suggests that some are in the sense of approximation.

4. *Some* + numeral and *bi'* + numeral constructions with the non-round numbers. Contrary to the analyses of Sauerland & Stateva (2007) and Anderson (2014), that consider *some* as an approximation, Solt & Stevens (2018) argue that it is mistaken to analyze *some* as only having an approximative reading. Solt and Stevens claim that *some* have other non-approximative uses in the context of non-round numbers as in:

(17) Of some 206 students who responded to the survey, 52% were female.

In Turkish, the *bi'* + numeral constructions can also be used with non-round numbers (18,19).

(18) Günün sonu-n-da bi' 47 çuval findık topla-n-mış-tır.
day-GEN end- POSS-LOC some 47 sack hazelnut collect-PASS-PST-EPIS
'(Probably) at the end of the day, 47 sacks of hazelnuts were collected.'

(19) Okul gezi-si-n-e bi' 32 öğretmen gelecek.
School trip- POSS-DAT some 32 teacher come+ FUT
'32 teachers will come to the school trip.'

We have previously said that we can use *about* or *approximately* instead of *some* in the examples given above. However, when we consider (17), we cannot replace *some* of these structures because structures such as *about* and *approximately* seem odd to the speaker when used with non-round numbers.

Another example indicating that *some* and *about* are not semantically equivalent to each other is that there are cases where replacing *about* with *some* leads to unacceptability.

(20) a. The meeting took place at (about/*some) three o'clock.
b. The tree was planted in (about/*some) 1989.

A similar observation can be made about *bi'* + numeral expressions in Turkish (21).

(21) a. Kahvaltı 10.00 gibi başlar.
Breakfast 10.00 about start- PRS
'Breakfast starts at about 10:00 am.'

- b. *Kahvaltı bi' 10.00 başlar.
 Breakfast some 10.00 start- PRS
 'Breakfast starts at some 10:00 am.'

Solt and Stevens (2018) claim that *some* in English when used with numerals, can be analyzed as a function (f) from sets to sets (22). In its non-approximating uses, some combine with a numeral and give us a subset of the set denoted by the numeral (23). In its approximating uses, the combination of some with a numeral provides a superset of the set denoted by the numeral (24). Below is an example of the non-approximating use of some (25).

- (22) $\| \text{some}_{\langle \alpha t, \alpha t \rangle} \| = \lambda P_{\langle \alpha t \rangle} \lambda x_{\alpha}. f(P)(x)$
 (23) Non-approximating (basic) some = $f(A) \subseteq A$
 (24) Approximating some = $A \subseteq f(A)$
 (25) Some 206 students
 $\| 206 \| = \{ Y: |Y| = 206 \}$
 $\| \text{some } 206 \| = \lambda x. x \in f(\{ Y: |Y| = 206 \})$
 $\| \text{some } 206 \text{ students} \| = \lambda x. \text{students}(x) \ \& \ x \in f(\{ Y: |Y| = 206 \})$,
 where $f(\{ Y: |Y| = 206 \}) \subseteq \{ Y: |Y| = 206 \}$

In this section, we indicated that the use of *some*+ numeral expressions in English with non-round numbers is also accurate in *bi*'+ numeral expressions in Turkish. In the next section, we will consider the dissimilarities between some and *bi*' structures.

5. Some differences between *some* + numeral in English and *bi*' + numeral in Turkish. It is to be noted that there are some differences between *bi*' + numeral constructions in Turkish and some + numeral constructions in English. For instance, Solt and Stevens (2018: 349) observe that non-approximative uses of some are possible in simple indicative past sentences (27). However, as seen in (28), a similar sentence does not appear acceptable in Turkish.

- (26) The storm caused some 103 houses to lose power.
 (27) #Fırtına-da bi' 103 ev-in elektrik-i kes-il-di
 storm-LOC some 103 house-GEN electricity-POSS cut.off-PASS-PST
 'In the storm, some 103 houses lost electricity.'

(27) is acceptable in Turkish only if we add the epistemic marker *-Dir* so that the sentence expresses a guess/prediction by the speaker (28):

- (28) Fırtına-da bi' 103 ev-in elektrik-i kes-il-miş-tir
 storm-LOC some 103 house-GEN electricity-POSS cut.off-PASS-PST-EPIS
 'In the storm, probably some 103 houses have lost electricity.'

The difference between (27) and (28) is that *-DI* in (27) indicates that the speaker has direct evidence about the subject, while *-miş* in (28) suggests that the speaker has indirect evidence or speaker. The speaker about the power outage in (27) means, "I saw that the storm caused some 103 houses to lose power", while in (28) means, "I heard that the storm caused some 103 houses to lose power". In the case of direct evidence, the speaker is certain of the statement it has given, while in the case of indirect evidence, the speaker is not certain of the statement it used. In this

respect, if we recall the use of *bi'* in (19), it is not completely sure what will be in the future; it makes a guess or prediction. To summarize, -Dİr can be used in statements where the speaker is uncertain of itself, such as the future tense (-acak/-ecek) and -mİş. However, -Dİr cannot be used in statements with evidence of the speaker, such as the past tense suffix -DI.

In Turkish, *bi'* + numeral constructions cannot be used when the speaker is certain about the relevant number. The analysis we develop below accounts for the obligatoriness of probability interpretations in the context of the *bi'* + numeral constructions as in (29)

- (29) Toplantı-ya *bi'* 206 kişi katıl-mış-tır.
 meeting-DAT some 206 person attend-PST-EPIS
 '(Probably) some 206 people have attended the meeting.'

The presence of the epistemic marker -Dİr in (29) suggests that the speaker is not certain that there are exactly 206 people who attended the meeting. Instead, the speaker finds it likely that 206 people attended the meeting but cannot rule out the possibility that less than 206 or more than 206 people have attended the meeting.

The analysis by Solt and Stevens (2018) falls short of explaining why it is inappropriate to use *bi'* + numeral expressions in Turkish when indicative of past tense and when the speaker is certain about the subject. For this reason, we present a new proposal for the semantic analysis of *bi'* + numeral expressions in Turkish.

To sum up, we have shown that *bi'* + numeral expressions in Turkish are not used in sentences containing the meaning of certainty, unlike *some* + numeral expressions. In the next section, we will present our proposal for analyzing *bi'* + numeral constructions.

6. Proposal. In Turkish, we assume that *bi'* '*some*' denotes the identity function over predicates. However, the presence of *bi'* inside the nominal phrase makes using the probability operator, PROB, obligatory. This operator takes propositions as an argument and asserts that the probability of the proposition *p* given *O*, which is the total information (a set of propositions) available to an agent, is more than 0.5 (Lassiter 2017). An equivalent way of stating this is to say that, for any proposition *p* and information state *O*, PROB asserts that the probability that *p* is true given *O* is higher than the probability that *p* is false given *O*.

- (30) a. $\| \text{bi}' \| = \lambda f_{\text{et.f}}$
 b. $\| \text{PROB} \| = \lambda p_{\text{st.}}.P(p|O) > 0.5$

In (27), the indicative past tense marker suggests that the speaker is certain of the truth of the relevant proposition. However, as seen in (31), probability operators in natural languages (here, *probably*) cannot be used with propositions whose truth is certain.

- (31) #Probably 2 plus 2 equals 4.

For a probability operator to be felicitous, there must be alternative propositions that have some chance of being true. Below, we indicate this as a presupposition of PROB.

(32) $\|PROB\| = \lambda p_{st}: \exists q \in O \text{ distinct from } p \text{ such that } \text{PROB}(q|O) \neq 0. \text{PROB}(p|O) > 0.5$

This means that *probably p* entails that the speaker is not entirely certain that p holds. Therefore, the use of *bi'* 'some' in past indicative sentences, as in (34), which express certainty on the part of the speaker, is not acceptable.

- (33) Bi' yedi öğrenci sınavdan yüksek not al-mış-tır.
Some seven student exam-ABL high mark get- PST- EPIS
'Some seven students scored high marks from the exam.'
- (34) *Bi' yedi öğrenci sınav-dan yüksek not al-dı.
Some seven student exam-ABL high mark get- PST
'Some seven students scored high marks from the exam.'
- (35) Okul-a bi' 187 öğrenci git-miş-tir.
School- DAT some 187 student go- PST- EPIS
'(Probably) Some 187 students have gone to school.'

Below is an example in which we can see that PROB that the proposition in (36a) as an argument (S_w is the set of students in w , and G_w is the set of entities who went to school).

- (36) a. $\llbracket \text{Okula bi' 187 öğrenci gitmiştir} \rrbracket = \lambda w. |S_w \cap G_w| = 187$
b. $\text{PROB}_o(\lambda w |S_w \cap G_w| = 187) = 1$ iff
 $\text{PROB}(\lambda w |S_w \cap G_w| = 187|O) > 0.5$ assuming $\exists q \in O$ such that $\text{prob}(q|O) \neq 0$

The analysis we have developed above raises the question of the function of *bi'* in the first place. One observation that can be made about sentences containing *bi'* is that the numeral used with it must be focus-marked. That is, the use of *bi'* with a numeral indicates the presence of the focus feature on the numeral. In the alternative semantics of Rooth (1992), this means that propositions obtained by replacing the numeral with its alternatives (i.e., other numerals) are relevant to the truth value of *bi'*+ numeral constructions. This set is given below as the set C, which stands for contextual variable.

- (37) $C = \left(\begin{array}{l} \text{Toplantıya 204 kişi katıldı.} \\ \text{Toplantıya 205 kişi katıldı.} \\ \text{Toplantıya 206 kişi katıldı.} \\ \text{Toplantıya 207 kişi katıldı.} \\ \text{Toplantıya 208 kişi katıldı.} \\ \dots \end{array} \right)$

Yalcin (2010) claims that "probability sentences are evaluated relative not just to a probability space but to a select class of proposition from the space," which we assume is the set of alternative propositions given in C. Given this analysis, we suggest that PROB asserts that its prejacent is more likely than alternatives.

- (38) $\|PROB_C(p)\| = 1$ iff $\forall q \in C$ such that $p \neq q$
 $\text{PROB}(p|C) > \text{PROB}(q|C)$ (Yalcin, 2010)

Yalcin (2010) explains the analysis of the probability operator with an example. We imagine that there is a lottery with 1000 tickets, which we assume is fair. Among the possible possibilities in this lottery, we consider two options. In the first case, Bloggs owns 420 tickets, while the remaining 580 tickets belong to someone else. In the second case, Bloggs still has 420 tickets, while the remaining 580 tickets belong to 580 different people. Even though the number of tickets on Bloggs is the same, the probability of winning between the two situations has changed. In this respect, the probability operator refers to the situation where the probability of the preposition to which it is connected is higher than the other prepositions (Yalcin, 2010).

We prefer to use Yalcin (2010)'s probability operator analysis when examining *bi'* + numeric expressions. When we look at the statement of 206 people to the meeting in (39), the speaker says that the probability of 206 people coming to the meeting is higher than the number of people mentioned in other possible situations.

To give a concrete example, consider (39) with the denotation in (40). (A_w is the set of people who have attended the meeting)

- (39) Toplantı-ya *bi'* 206 kişi katıl-mış-tır.
meeting-DAT some 206 person attend-PST-EPIS
'(Probably) some 206 people have attended the meeting.'

- (40) $\|\text{PROB}_C(\lambda w. |S_w \cap A_w| = 206)\| = 1$ iff $\forall q \in C$ such that $\lambda w. |S_w \cap A_w| = 206 \neq q$

$\text{PROB}(\lambda w. |S_w \cap A_w| = 206 | C) > \text{PROB}(q | C)$ (Yalcin, 2010)

Where C is defined as above.

We can explain why (34) is odd to speaker with the new probability operator analysis. The probability operator arises when the speaker is uncertain of itself and cannot access direct evidence. In sentence 34, since the speaker is speaking in a past indicative sentence with direct proof, it would be wrong to use the probability operator here.

We have already mentioned that *bi'*+numeral constructions, when used with exact temporal and spatial points, give rise to unacceptability.

- (41) * Bölüm toplantı-sı *bi'* 09.00'da yap-ıl-dı.
Department meeting- POSS some 09.00-LOC do- PASS- PST
'The department meeting was held at some 09.00.'

- (42) * Bu ev *bi'* 2013'te inşa.et-il-miş-tir.
This house some 2013+ LOC build-PASS-PST- CER
'This house was built in 2013.'

We have already seen that *bi'* in Turkish denotes the identity function over predicates. Exact temporal and spatial points such as 9 o'clock and 2013 are not sets but points on a scale. Therefore, these expressions cannot serve as the argument of predicates. This explains the unacceptability of (41) and (42).

In Turkish, *bi'*+ numeral constructions are used with sentences containing precise meaning (i.e., simple indicative tenses). The analysis we have done with the probability operator PROB explains why *bi'*+ numeral constructions are not used in sentences with the meaning of certainty.

7. Conclusion. To sum up, *bi'*+ numeral constructions have remarkable semantic properties in Turkish. Although *bi*+ numeral constructions in Turkish have many similarities with *some*+ numeral constructions in English, *bi'* is not used with simple indicative tenses, unlike the structure of *some*. For this reason, we examine *bi'*+ numeral structures with the probability operator in Turkish. As a result of using the probability operator, we explained why *bi'*+ numeral constructions in Turkish could be used with precise prediction sentences when they are not used with sentences that declare certainty.

References

- Anderson, Curt. 2014. Approximation of complex cardinals using some. In *Proceedings of WECOL 2013*. 131–143.
- Lassiter, Daniel. 2017. Probabilistic Language in Indicative and counterfactual Conditionals. *SALT* 27. 525 – 546. <https://doi.org/10.3765/salt.v27i0.4188>
- Lasnik, Howard. 1999. Pragmatic halos. *Language* 75(3). 522–551. <https://doi.org/10.2307/417059>
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75- 116. <https://doi.org/10.1007/BF02342617>
- Sauerland, Uli & Penka Stateva. 2007. Scalar vs. epistemic vagueness: evidence from approximators. *SALT* 17. 228–245. <https://doi.org/10.3765/salt.v17i0.3091>
- Stevens, Jon & Stephanie Solt. 2018. The semantics and pragmatics of “some 27 arrests”. *University of Pennsylvania Working Papers in Linguistics* 24(1). 179–188. <https://repository.upenn.edu/pwpl/vol24/iss1/21>
- Solt, Stephanie & Stevens, Jon. 2018. Some three students: towards a unified account of some. *SALT* 28, 345-365. <https://doi.org/10.3765/salt.v28i0.4436>
- Winter, Yoad. 2016. *Elements of formal semantics: An introduction to the mathematical theory of meaning in natural language (Edinburgh Advanced Textbooks in Linguistics)*. Edinburgh: Edinburgh University Press.
- Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11), 916–937. <https://doi.org/10.1111/j.1747-9991.2010.00360.x>