

TWO INDEFINITE ARTICLES IN UZBEK

KLAUS VON HEUSINGER, UDO KLEIN
University of Stuttgart

1. Introduction^{*}

It has sometimes been claimed that Uzbek has “no definite or indefinite articles” (Bodrogligeti 2001, p. 55). Contrary to these claims, we show that Uzbek has not just one but two indefinite articles, namely *bir* and *bitta*. Uzbek thus shares this crosslinguistically rather unusual property with languages such as Lakhota, which distinguishes between two “quasi-indefinite determiners” (Lyons 1999), Moroccan Arabic, which employs a “potential” and a “concretizing” (Harrell 1962), and Maori, a more well-known example of a language where we also find two indefinite articles (Bauer 1993).

In section 2 we show that *bir* and *bitta* are indeed used as indefinite articles. First we present Heine’s theory of the development of indefinite articles from numerals, then we discuss the relevant aspects of the numeral and classifier system of Uzbek, and finally we provide textual evidence that both *bir* and *bitta* occur in various usages as indefinite articles. In section 3 we compare the occurrences of these two indefinite articles in two novels, the first one published in 1926 and the second one published in 2001, and show that while *bir* has reached the last development stage already by 1926, the use of *bitta* is a more recent development (at least in written Uzbek). The second important conclusion from this corpus study is that the use of *bitta* appears to have spread to the expense of *bir* not just in written but also in colloquial Uzbek. Due to the inherent limitations of corpus studies, some important questions could not be answered by this methodology. In section 4 we present the results of an online questionnaire, which was designed to elicit acceptability judgements in order to address some of these questions. First, the absence of certain usages of *bitta* (from our corpus) is, of course, not evidence of the absence of

^{*} We would like to thank our informant Dildora Niyazmetowa for annotating the Uzbek corpus which we used and for her help with the questionnaire, as well as Jaklin Kornfilt for discussion of an earlier version of this paper. All remaining errors are, of course, our own. The present work has been carried out as part of the project C2 “Case and referential properties” of the collaborative research centre SFB 732 “Incremental specification in Context” of the German Science Foundation, whose financial support we would like to acknowledge.

this usage. Acceptability judgements can help settle this issue. Secondly, we wanted to know whether, and if so in which contexts of use, in modern spoken Uzbek the use of *bitta* is judged better than the use of *bir*, because if this were the case then this would corroborate the hypothesis that *bitta* has spread to the expense of *bir*. The main result is that when it is used as an indefinite article introducing human referents, *bitta* is at least as good as *bir* (except in predicative constructions), and indeed better if used to introduce referents whose identity neither the speaker nor the hearer knows or cares about.

2. Numerals, classifiers and indefinite articles in Uzbek

Bodrogligeti (2001, p. 55) claims that in Uzbek “[t]here is no special morphological or grammatical means specifically designed to distinguish between definite and indefinite nouns: No definite or indefinite articles”. In this first section we aim to show that contrary to Bodrogligeti’s claim there is not just one indefinite article in Uzbek, but two, namely *bir* and *bitta*. Both of them developed from the numeral *bir* ‘one’. Since this is an instance of a general pattern of diachronic development of indefinite articles from numerals, we briefly present the five-stage model developed in Heine (1997) to account for this development. Next we turn to the basic ingredients for the development of these indefinite articles, namely the numeral *bir* ‘one’ and the classifier system of Uzbek. With these important preliminaries clarified, we turn to the two items *bir* and *bitta*, and indicate what stage they have reached in their development into indefinite articles.

2.1 Heine’s theory of development of indefinite articles from numerals

Heine (1997, p. 71) claims that “in the vast majority of languages that have developed an indefinite article it is the numeral ‘one’ that was recruited for this purpose.” He then proposes a five-stage model in order to account for the diachronic evolution of an indefinite article from a numeral. In the first stage, “there is an item for ‘one’ which functions exclusively as a numeral” (p. 72). Such an item reaches the second stage in its development into an indefinite article when it “introduces a new participant presumed to be unknown to the hearer and this participant is taken up as definite in subsequent discourse” (p. 72). In the third stage, the use of this item “is extended typically to any participant in discourse known to the speaker but presumed to be unknown to the hearer, irrespective of whether or not the participant concerned is expected to be taken up in subsequent discourse” (p. 72f). The fourth stage is reached if the item can “be used when a participant is introduced whose referential identity neither the hearer nor the speaker knows or cares to know” so that the article is “no longer confined to marking specific reference at this stage” (p. 73). In the fifth stage, “the article can be expected to occur on all types of nouns, even if there may remain a number of exceptions. Heine emphasises that “[t]he various stages must not be viewed as discrete entities; rather the evolution from stage I to V is continuous and involves overlaps of various kinds.”

Heine’s five stage model can be viewed as an elaboration of the three stage model proposed in Givon (1981). Heine’s first stage corresponds to Givon’s quantification category, Heine’s second, third and fourth stages can be viewed as a subclassification of Givon’s referentiality category, while Heine’s fifth stage corresponds to Givon’s genericity category.

2.2 Numerals and classifiers in Uzbek

Since in Uzbek, too, the indefinite articles have their root in the numeral *bir* ‘one’, in this subsection we will briefly present the main property of numerals in Uzbek (which sets Uzbek apart from other Turkic languages), namely that they require what Beckwith (1998) calls a specifier if they occur with a count noun. We will use the terminology used in Beckwith (1998), since this is the most detailed and theoretically informed paper on Uzbek classifiers that we are aware of. At the end of this subsection we will provide what we think are the corresponding notions in Aikhenvalds terminology.

Beckwith (1998) distinguishes two main classes of specifiers (i.e. lexical items which must occur with numerals plus count nouns) in Uzbek, namely measures and classifiers. These two types of specifiers are syntactically similar but semantically different. Measures differ from classifiers in that “the features common to the set of expressions using them are inherent not to the nouns [as is the case with classifiers, vH&K] but to the measures, which subdivide and qualify the nouns.” (Beckwith 1998, p. 129). To illustrate this distinction, note that what the meanings of expressions of the form *piyala N*, e.g. (1), have in common is due to the meaning contribution of *piyala*, whereas what expressions of the form *nafar N* have in common, e.g. in (2), is due to the meaning contribution of the noun *N*.

- | | | |
|-----|--|------------|
| (1) | <i>ikki piyala cay</i> two cup tea ‘two cups of tea’ | Measure |
| (2) | <i>ikki nafar oquwci</i> two CL:human student ‘two students’ | Classifier |

Having introduced this notion of specifier, Beckwith claims that in Uzbek “a specifier – including as an alternative the nonclassifying specifier suffix *-ta*, which is unique to Uzbek among all Turkic languages – is obligatory in numeral expressions for all count nouns” (127). To give some examples:

- | | | | | | |
|-----|---|----|---|----|--|
| (3) | <table border="0"> <tr> <td style="vertical-align: top;">a.</td> <td style="vertical-align: top;"><i>ikki kosa olma</i> two bowl apple ‘two bowls of apples’</td> </tr> <tr> <td style="vertical-align: top;">b.</td> <td style="vertical-align: top;">*<i>ikki olma</i> two apple Int.: ‘two bowls of apples’</td> </tr> </table> | a. | <i>ikki kosa olma</i> two bowl apple ‘two bowls of apples’ | b. | * <i>ikki olma</i> two apple Int.: ‘two bowls of apples’ |
| a. | <i>ikki kosa olma</i> two bowl apple ‘two bowls of apples’ | | | | |
| b. | * <i>ikki olma</i> two apple Int.: ‘two bowls of apples’ | | | | |
| (4) | <table border="0"> <tr> <td style="vertical-align: top;">a.</td> <td style="vertical-align: top;"><i>ikki nafar uq'ituvch'i</i> two CL:HUMAN teacher ‘two teachers’</td> </tr> <tr> <td style="vertical-align: top;">b.</td> <td style="vertical-align: top;">*<i>ikki uq'ituvch'i</i> two teacher Int.: ‘two teachers’</td> </tr> </table> | a. | <i>ikki nafar uq'ituvch'i</i> two CL:HUMAN teacher ‘two teachers’ | b. | * <i>ikki uq'ituvch'i</i> two teacher Int.: ‘two teachers’ |
| a. | <i>ikki nafar uq'ituvch'i</i> two CL:HUMAN teacher ‘two teachers’ | | | | |
| b. | * <i>ikki uq'ituvch'i</i> two teacher Int.: ‘two teachers’ | | | | |

- (5) a. *ikki-ta hona*
two-SPEC room
'two rooms'
- b. **ikki hona*
two room
Int.: 'two rooms'

As the examples above show, expressions containing a numeral and a count noun are ungrammatical if they do not contain a measure as in (3), a classifier as in (4), or the nonclassifying specifier suffix *-ta* as in (5).

The only exception to this generalisation is the numeral *bir* 'one', which may also occur without a specifier: "some speakers in Uzbekistan, especially in the cities, generally drop the suffix with the numeral *bir* 'one'" (Beckwith 1998, p. 127). That is, with the numeral *bir* it is not necessary to use a specifier (either a measure, a classifier or the nonclassifying specifier suffix *-ta*), so that *bir* (6a) can be used instead of *bitta* (6b), which as suggested by Beckwith is the result of the assimilation of the numeral *bir* to the nonclassifying specifier suffix *-ta*.

- (6) a. *bir q'iz*
one girl
'one/a girl'
- b. *bitta q'iz*
one:SPEC girl
'one/a girl'

According to Beckwith, this "may have developed due to influence from the use of *bir* as an indefinite article, in which usage no specifier occurs" (127). Note that the way in which Beckwith formulates this exception to the rule that numerals require a specifier implies that *bir* can occur without a specifier not only when it is used as an indefinite article, but also when it is used as a numeral. In the next subsection we will illustrate the distribution of *bir* as an indefinite article, in order to show which stage in the development it has reached.

A brief terminological note: Beckwith's class of measures corresponds to Aikhenvald (2000)'s class of mensural classifiers, whereas Beckwith's classifiers correspond to Aikhenvald's notion of sortal classifier. See Aikhenvald (2000, 114-120) for criteria distinguishing these types of classifiers.

2.3 The indefinite article *bir*

In this subsection we show that, contrary to claims made in the literature according to which Uzbek does not have indefinite articles, the numeral *bir* can indeed be used as an indefinite article. In section 2.4 we will show that *bir* was used as indefinite article already in 1926.

In example (7) *bir* is used to introduce a new participant which (i) is presumed to be unknown to the hearer and which (ii) is the subject of further specification in the subsequent discourse.

- (7) *Bir zamon-lar bir schoh bu'l-ib u't-gan ekan...*
A time-PL a(=one) king be-GER be-PTCP apparently
'Once upon a time there was a king. . .'

According to Heine (1997, 66-82), this is the second stage in the development of an indefinite article from a numeral. Sentence (8) below, which similar to the following sentences, is from Tohir Malik's novel *Shaytanat*, illustrates that *bir* can also be used to introduce a new referent without picking this referent up in the following discourse (stage 3); sentence (9) illustrates that *bir* can be used without introducing a specific referent (stage 4); and sentence (10) shows that *bir* can be used in predicative constructions (stage 5). The following sentences are from two novels (see next section).

- (8) *Eshpo'latov ichkari bir erkak-ni ko'r-gan. Hayron*
 Eschpolatov inside-LOC a man-ACC see-PRF wonder
bo'l-ib zavod-ga bor-gan.
 be-GER workshop-DAT walk-PRFT

'Eshpolatov saw a man inside. Bewildered he walked into the workshop.'

- (9) *Uch-av-lar-i ham bir narsa-ning xayol-i-ni*
 three-together-PL-AGR also a thing-GEN think-AGR-ACC
sur-gan-dek ko'r-i-nar di-lar.
 do-PTCP-as.if see-AGR-PROG be-PL

'The three looked as if they were thinking about something.'

- (10) *Hurmat-i zo'r bo'l-sa boshqa-lar-ga zo'r, sen bilan*
 honor-AGR strong be-COND other-PL-DAT strong 2SG with
men-ga u bor-yo'g'i bir oshna.
 1SG-DAT 3SG be-just a friend

'If he is very honorable, then only for the others. For you and for me he is just a friend.'

We therefore conclude that the Uzbek numeral *bir* has reached the fifth stage of development into an indefinite article. In the next subsection we turn to the development of *bitta* into an indefinite article.

2.4 The indefinite article *bitta*

Uzbek differs from other Turkic languages, e.g. Turkish which also uses the word *bir* as an indefinite article (Kornfilt, 1997, p. 106), in that in addition to *bir* it has developed a second indefinite article, namely *bitta*. The following sentences, found again in the Tohir Malik's novel *Shaytanat*, illustrate that it has reached at least the fourth stage of development into an indefinite article. In (11) it is used to introduce a new referent which is then picked up again in the following discourse (stage 2), in (12) it is used to introduce a new referent which is not picked up again (stage 3), and in (13) it is used without introducing a specific referent (stage 4). In the contemporary novel which we analysed we found no use of *bitta* in a predicative construction (see next section):

- (11) *Ammo bitta shart-i-m bo'l-a-di: singl-i-m-ni bun-*
 but a condition-AGR-1SG be-FUT-3SG little.sister-AGR-1SG-ACC now-
dan buyon qimor-ga tik-may-san.
 ABL from game-DAT put-NEG-2SG:FUT
 'But there will be a condition; from now onwards you will not bet my little sister.'
- (12) *Bitta shahar-da ikki-ta imom-domla bo'l-gan ekan. It bilan*
 a city-LOC two-SPEC imam-priests be-PRF EVID dog with
mushuk murosa qil-sa qil-arkan-ki, bu-lar ittifoq
 cat agreement make-COND make-EVID-CONJ DEM-PL agreed
bo'la ol-ish-mas ekan-lar.
 be take-NOM-NEG EVID-PL
 'In a city there were two imams. Even dogs and cats could agree, but these two couldn't.'
- (13) *Katta-ng-ga bor-ib ayt: men-ga bitta samolyot*
 big-2SG:AGR-DAT go-GER say.IMP 1SG-DAT a plane
bilan bir million pul kerak.
 with a million money need
 'Go to your boss and say: I need a plane and a/one million.'

2.5 Semantic contrast of indefinite articles

In some contexts, the choice of one indefinite article over the other may give rise to semantic contrasts. For example, the availability of two indefinite articles appears to be exploited in order to distinguish between neutral and non-specific uses of indefinite NPs. In (14) there appears to be a strong preference for interpreting the direct object NP non-specifically (i.e. as introducing a referent whose identity the speaker and hearer neither know or care about), whereas in (15) the indefinite can but does not have to be so interpreted.¹ In section 4 we will return to this contrast and hint at what may be the reason for it.

- (14) *Professor bir student-ni tekschir-moq'-chi.*
 professor a(=one) student-ACC examine-want-3SG
 'The professor wants to examine a student.' [non-specific]
- (15) *Professor bitta student-ni tekschir-moq'-chi.*
 professor a(=one:SPEC) student-ACC examine-want-3SG
 'The professor wants to examine a student.' [specific, non-specific, numeral]

The existence of two indefinite articles in Uzbek raises a two questions: first, how did these two indefinite articles develop, and secondly how are the articles used nowadays in modern colloquial Uzbek? To investigate the first question we compared the use of *bir* and *bitta* in a novel from 1926 with a novel from 2001, the results of which are presented in the next section.

¹ For the interaction between accusative case marking and the two indefinite articles see Niyazmetowa (in preparation).

3. Corpus study

To throw some light on the first question, we initially analysed the use of *bir* and *bitta* in the novel *Otgan kunlar* by Abdulla Qodiriy, which was published in 1926, and then we compared this usage with the usage of *bir* and *bitta* in Tohir Malik's novel *Shaytanat*, published in 2001.

In the novel from 1926, we annotated the first 113 occurrences of *bir* and *bitta* as indefinite articles with respect to the kind of usage and with respect to the animacy of the referent. As table 1 shows *bir* has been employed in all four types of usages, from introducing a referent which is picked up again (stage 2) to occurrences in predicative constructions (stage 5). However, in the older novel we did not find *bitta* used as an article (we only found one use of *bitta* as a numeral).

| Article | Usage | Occurrences | human | inanimate |
|--------------|-------|-------------|----------|-----------|
| <i>bir</i> | 2 | 29 | 14 | 15 |
| <i>bir</i> | 3 | 29 | 16 | 13 |
| <i>bir</i> | 4 | 41 | 8 | 33 |
| <i>bir</i> | 5 | 14 | 11 | 3 |
| total | | 113 | 49 (43%) | 64 (57%) |
| <i>bitta</i> | 2 | 0 | 0 | 0 |
| <i>bitta</i> | 3 | 0 | 0 | 0 |
| <i>bitta</i> | 4 | 0 | 0 | 0 |
| <i>bitta</i> | 5 | 0 | 0 | 0 |
| total | | 0 | 0 | 0 |

Table 1: Use of *bir/bitta* in the novel from 1926

There are two (logically) possible reasons for why *bitta* is lacking in this novel. Either, *bitta* was not yet in use as an indefinite article in 1926, or it was already in use, but not yet in the written register. Assuming that the written register lags behind the spoken register, it is plausible to conclude that *bitta* has developed into an indefinite article after *bir*.

The lack of *bitta* in the novel published in 1926 stands in sharp contrast to the presence of *bitta* in the novel from 2001, in which we annotated the first 125 occurrences of *bir* and *bitta* as indefinite articles. As illustrated in table 2, approximately two thirds (81) were occurrences of *bir*, and approximately one third (44) were occurrences of *bitta*. Given that in the first novel the lack of *bitta* may have been due to differences in register between written and colloquial language, we also annotated whether the indefinite articles *bir/bitta* occurred in the author's narrative, which is likely to be closer to the written register, or whether it occurs in direct speech, which is likely to be closer to the colloquial register. As illustrated in table 3, out of 66 occurrences of indefinite articles in the speaker's narrative, 62 were occurrences of *bir*, and only 4 were occurrences of *bitta*. From this we can conclude that there is a strong preference for *bir* in the speaker's narrative. On the other hand, out of 59 occurrences of indefinite articles in direct speech, only 19 were occurrences of *bir*, compared to 40 occurrences of *bitta*. This indicates that in direct speech there is a preference for using *bitta*.

| Article | Usage | Occurrences | human | inanimate |
|--------------|-------|-------------|----------|-----------|
| <i>bir</i> | 2 | 16 | 2 | 14 |
| <i>bir</i> | 3 | 40 | 8 | 32 |
| <i>bir</i> | 4 | 16 | 3 | 13 |
| <i>bir</i> | 5 | 9 | 2 | 7 |
| total | | 81 | 15 (12%) | 66 (53%) |
| <i>bitta</i> | 2 | 12 | 2 | 10 |
| <i>bitta</i> | 3 | 20 | 7 | 13 |
| <i>bitta</i> | 4 | 12 | 4 | 8 |
| <i>bitta</i> | 5 | 0 | 0 | 0 |
| total | | 44 | 13 (10%) | 31 (25%) |

Table 2: Use of *bir/bitta* in the novel from 2001

| Article | Usage | Occurrences | Narrative | Direct speech |
|--------------|-------|-------------|-----------|---------------|
| <i>bir</i> | 2 | 16 | 13 | 3 |
| <i>bir</i> | 3 | 40 | 33 | 7 |
| <i>bir</i> | 4 | 16 | 11 | 5 |
| <i>bir</i> | 5 | 9 | 5 | 4 |
| total | | 81 | 62 (50%) | 19 (15%) |
| <i>bitta</i> | 2 | 12 | 1 | 11 |
| <i>bitta</i> | 3 | 20 | 0 | 20 |
| <i>bitta</i> | 4 | 12 | 3 | 9 |
| <i>bitta</i> | 5 | 0 | 0 | 0 |
| total | | 44 | 4 (3%) | 40 (32%) |

Table 3: Use of *bir/bitta* in the novel from 1926, anoted by narrative versus direct speech

Regarding *bir*, we can therefore conclude that its use as an indefinite article had already spread to Heine's fifth stage by 1926, as the examples from the 1926 novel indicate, in which *bir* is used in predicative constructions.

- (16) *Xon qizig'a loyiq bir yigit.*
 khan very worthy a(=one) young.man

'The khan is a very worth young man.'

Regarding *bitta*, we have evidence that in the written register, the use of *bitta* as an indefinite article has increased from the 1926 novel, where we found no occurrence of *bitta* as an indefinite article in the first 113 occurrences of *bir* and *bitta*, to the 2001 novel, where we found 44 occurrences in the first 125 occurrences of *bir* and *bitta* as indefinite articles. Moreover, as the example in (13), repeated below as (17), illustrates, *bitta* may be used to introduce a participant "whose referential identity neither the hearer nor the speaker knows or cares to know", and has thus reached at least stage IV of development into an indefinite article.

- (17) *Katta-ng-ga* *bor-ib* *ayt:* *men-ga* *bitta* *samolyot*
 big-2SG:AGR-DAT go-GER say.IMP 1SG-DAT a plane
- bilan bir million pul kerak.*
 with a million money need

‘Go to your boss and say: I need a plane and a million.’

The frequency of occurrence of the indefinite articles *bir* and *bitta* in the novel from 2001 correlates with the distinction between literary and colloquial language: *bir* is significantly more frequent in the author's narrative, whereas *bitta* is significantly more frequent in direct speech.

We can therefore conclude that at least in the written register, there is an obvious development of the use of *bitta* as an indefinite article. The fact that our corpus was limited both in size as well as in kind precludes any claims about the use of *bitta* as an indefinite article in spoken Uzbek around 1926. We simply cannot tell, based on our corpus, whether *bitta* was not used in written or spoken Uzbek around 1926 or whether *bitta* was used, but only in spoken Uzbek. What is plausible, however, is that *bitta* has developed into an indefinite article after *bir*. The first indication that this might be the case is that the written register usually lags behind the spoken register, and if *bitta* had developed before *bir* we would have expected at least as many occurrences of *bitta* as of *bir* in the novel from 1926. The second indicator that *bitta* has developed after *bir* is that we found 9 occurrences of *bir* in a predicative construction in the novel from 2001, while *bitta* has not occurred in predicative constructions in this novel. To the extent that we have reason to believe that the item which develops first into an indefinite article also reaches the last stage of development first, we also have reason to believe that *bir* developed first into an indefinite article, since it appears to have reached the last stage of development first. Moreover, due to the fact that in the novel from 2001 three quarters of the occurrences of *bir* are in the author's narrative, whereas nine tenths of the occurrences of *bitta* are in the direct speech of the characters, it is also plausible to hypothesise a significant difference between the written and spoken Uzbek: whereas in written Uzbek *bir* is more widely used as an indefinite article, in spoken Uzbek it is *bitta* which is more widely used as an indefinite article.

Putting together the hypothesis that *bir* has started developing into an indefinite article before *bitta* with the hypothesis that in modern spoken Uzbek *bitta* is more widespread than *bir*, leads to the conclusion that the use of *bitta* as an indefinite article has spread in spoken Uzbek to the expense of *bir*.

In the next section we provide additional evidence from an online questionnaire bearing on this issue. First we wanted to know whether, and if so in which contexts of use, in modern spoken Uzbek the use of *bitta* is judged better than the use *bir*. Secondly, since the absence of evidence from our corpus that *bitta* has reached stage V in the development from numeral to indefinite article is not evidence that it has not reached this stage, the question arises how good sentences are judged in which the indefinite article in a predicative construction is not *bir* but *bitta*.

4. Questionnaire

To answer these questions we designed and carried out a questionnaire in order to elicit acceptability judgements about sentences containing the two articles in the four different usages. The independent factors of our experiment were first the indefinite article, namely *bir* or *bitta*, secondly the animacy of the referent (human or inanimate), and thirdly the four different usages

(stage 2 to 5). Crossing the values of these three factors with each other, we got 16 different types of sentences which were instantiated in the questionnaire by means of 3 different lexicalisations. The total of 48 items was divided onto 6 actual questionnaires, so that each participant judged only 8 items which were randomly mixed with 20 filler sentences. The task of the 45 participants in the questionnaire was to judge the acceptability of the sentences (“how good the sentence sounds”) on a scale from 2, meaning very bad, to 5, meaning very good.

The judgements were collected in part online (using the WEBEXP2 software developed in Edinburgh) and in part by means of written questionnaires, and were then evaluated by means of an analysis of variance (ANOVA). The basic idea behind an ANOVA is to estimate how likely or unlikely it is for a given difference in judgements to be due to chance. If it is unlikely that the difference is due to chance, then the difference is called statistically significant. This estimation is based on comparing the variance between two (or more) sets of judgements (e.g. the set of judgements of sentences whose indefinite article is *bir* with the set of judgements whose indefinite article is *bitta*) with the variance within these sets of judgements. If the variance between the sets of judgements is high while the variance within the sets of judgements is low, then the difference is likely to be statistically significant. If on the other hand, the variance between the sets of judgements is low, while the variance within the sets of judgements is high then the difference is likely to be due to chance.

Interaction between usage type and article, split by animacy

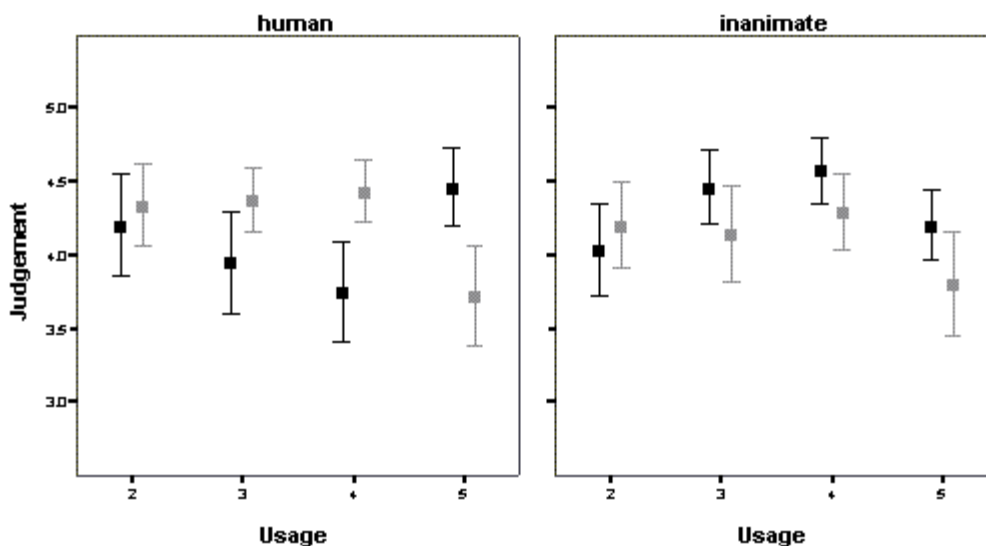


Table 4: Interaction between usage and article, split by animacy.
(Black error bars = *bir*, grey error bars = *bitta*)

First, we found a significant interaction (with $F(3,561) = 4.7$; $p < 0.005$) between the three factors usage, article and animacy, as the two charts in table 4 show. The error bars indicate that if we repeated the experiment with other participants the likelihood of the mean of the judgements being between the upper and lower bounds of the error bars is 95%. Therefore, if two error bars do not overlap, then the likelihood is very high that the difference is not due to chance but statistically significant. The black error bars indicate the mean judgements for *bir*, whereas the grey error bars indicate the mean judgements for *bitta*. The horizontal axis represents the four different types of usages (2 = introduction of referent which is picked up again, 3 = introduction

of referent which is not picked up again, 4 = referential identity unknown by speaker, 5 = predicative construction), and the vertical axis represents the judgements from 2 (very bad) to 5 (very good).

First of all, note that except for the use of *bitta* in predicative constructions, all other uses of *bitta* as an indefinite article have a mean judgement of over 4 on a scale from 2 to 5, irrespective of the animacy of the referent. The judgement mean for sentences containing *bitta* in predicative constructions is around 3.75, and is thus clearly above the mean for ungrammatical filler sentences, which was around 2.5. So the hypothesis from the previous section, namely that *bitta* is more widespread in spoken Uzbek than it is in written Uzbek, is corroborated by the fact that with the exception of stage V uses it is also consistently judged to be as acceptable as the grammatical filler sentences. Moreover, the fact that the stage V use of *bitta* is clearly more acceptable than the ungrammatical sentences, but not as good as *bitta* in its other uses as an indefinite article indicates that the use of *bitta* is in the process of spreading to the fifth stage of development into an indefinite article. The fact that for both humans and inanimates *bir* is judged better than *bitta* in predicative constructions can be taken as evidence that *bir* has reached the fifth stage of development before *bitta* has.

Next note that the acceptability of *bir* with human referents decreases with increasing stages of development. While it makes no statistically significant difference whether *bir* or *bitta* is used to introduce a human referent which is then picked up again in discourse (stage 2), it makes a significant difference if a human referent is to be introduced, whose referential identity is unknown to the speaker (stage 4) – for this usage *bitta* is significantly better than *bir*. We interpret this as evidence that the use of *bitta* with human referents has begun to impact on the use of *bir*, to the effect that *bitta* is preferred to *bir*, resulting in the use of *bir* for human referents starting to be less acceptable than it used to be (at least in some contexts of use).

For inanimate referents there appear to be no significant differences between the use of *bir* or *bitta*. This means that while the use of *bitta* as an indefinite article has spread to both human and inanimate referents, it is only for human referents (which are not in predicative constructions) that the use of *bitta* as an indefinite article is preferred over *bir*. The contrast between the usage of the indefinite articles for stages 4 and 5 shows that *bitta* is preferred for the referential stage 4, while *bir* is preferred for the non-referential stage 5. This may provide the motivation for the semantic contrast observed in examples (14) and (15) in section 2.5, where the use of *bir* implies a non-specific reading, whereas the use of *bitta* is unmarked.

5. Conclusion

In this paper we have provided converging evidence both from a corpus study as well as from an online questionnaire that modern Uzbek has two indefinite articles *bir* and *bitta*. The corpus study revealed a significant difference in usage between *bir* and *bitta*, to the effect that *bir* was strongly preferred in the author's narrative, whereas *bitta* was preferred in direct speech. On the basis of this we hypothesised that *bitta* has spread in modern colloquial Uzbek to the expense of *bir*. This hypothesis was confirmed by the evaluation of an online questionnaire, which indicates that in certain contexts of use *bitta* is judged more acceptable than *bir*. Moreover, the acceptability judgements on sentences containing *bitta* in predicative constructions show that the use of *bitta* is in the process of spreading to the last stage of development into an indefinite article. As we hinted at in the previous section, the development of *bitta* to the expense of *bir*

may be the reason for the emergence of semantic contrasts between the use of *bir* and *bitta*. The detailed investigation of this and related semantic contrasts had to be left for future work.

References

- Aikhenvald, Alexandra Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press, Oxford.
- Bauer, Winfred 1993. *Maori*. Routledge, London and New York.
- Beckwith, Christopher 1998. Noun Specification and Classification in Uzbek. *Anthropological Linguistics*, 40:124–140.
- Bodrogligeti, András J. E. 2001. *An Academic Reference Grammar of Modern Literary Uzbek*. Volume 1. Lincom Europa, München.
- Givón, Talmy 1981. On the development of the numeral ‘one’ as an indefinite marker”. In *Folia Linguistica Historica*, 2:35-53.
- Harrell, Richard S. 1962. *A short reference grammar of Moroccan Arabic*. Georgetown University Press, Washington D.C..
- Heine, Bernd 1997. *Cognitive Foundations of Grammar*. Oxford University Press, Oxford.
- Kornfilt, Jaklin 1997. *Turkish*. Descriptive Grammars. Routledge, London.
- Lyons, Christopher 1999. *Definiteness*. Cambridge University Press, Cambridge, 1999.
- Malik, Tohir 2001. *Shaytanat* (part 4).
- Niyazmetowa, Dildora (in preparation). *Differential object marking in Uzbek*. Master thesis, University of Stuttgart.
- Qodiriy, Abdulla 1926. *Otgan kunlar*.