

On the Mass-Count Distinction in an Obligatory Classifier Language

1. Introduction. Among theories of mass-count distinction, there is a debate as to whether the distinction exists at the lexico-semantic level, e.g. Link 1983, Krifka 1989, or is rather mediated through syntactic structure, e.g. Borer 2005, Acquaviva 2019. One piece of evidence in favor of the latter approach comes from classifier languages, in which all nouns seem to exhibit mass-like behavior in that they cannot combine with a numeral directly but only with the mediation of a classifier. In the present paper, we examine Tashkent Uzbek, an obligatory classifier dialect of Uzbek (Beckwith 1998, 2007). We argue that this language exhibits the mass-count distinction even in the absence of classifiers or, if they are present, in lower positions, before the classifiers are merged (at the NP/nP level). We provide a formal analysis of a range of Uzbek classifiers, showing that they are sensitive to the mass/count distinction (rather than constituting a source thereof.) Thus, we provide evidence that the mass-count distinction is lexically encoded in nouns even in classifier languages (e.g. Cheng & Sybesma 1998, Doetjes 1997, Sudo 2016, forthcoming.)

2. Mass-Count Contrasts. Evidence suggesting that Tashkent Uzbek makes mass-count distinction at the lexico-semantic level comes from the contrasts in the distribution of (non-)counting modifiers, the plural morphology, and different types of classifiers.

(i) Modifiers. First, there exist quantifiers compatible only with notionally count nouns. They also require the presence of the plural morpheme on the noun (1a). Conversely, some quantifiers select for only notionally mass nouns (1b). The mass-count contrasts are also found in modification by numerals that cannot co-occur with classifiers and are only compatible with count nouns, i.e. **approximatives** and **collectives** (1c). The mass-count contrast also arises with non-counting modifiers like distributive adjectives known to only modify individuated units (1d).

- | | |
|--|--|
| <p>(1) a. ba'zi kitob-lar/ *suv(-lar)
 some book-PL water-PL
 'some books (*waters)'</p> | <p>b. ozgina *kitob(-lar)/ suv(-lar)
 small amount book-PL water-PL
 'a small amount of water (*book(s))'</p> |
| <p>c. ming-lab/ uch-ala kitob/ *suv
 thousand-APPROX three-COLL book water
 'thousands of/ all three books (*waters)'</p> | <p>d. katta kitob/ *suv
 big book water
 'a big book (*water)'</p> |

(ii) Plural. While pluralization of count nouns leads to 'more than one' interpretation, pluralization of mass nouns is restricted, resulting in plurality of abundance/subkinds. In certain environments, e.g. partitives, the plural is **obligatory** on count, but **ungrammatical** on mass nouns:

- | | |
|--|--|
| <p>(2) a. kitob-*(lar)-dan oluvdim
 book-PL-ABL take.PST.1SG
 'I took some of the books.'</p> | <p>b. suv-*(lar)-dan ichuvdim
 water-PL-ABL drink.PST.1SG
 Int.: 'I drank some of the water.'</p> |
|--|--|

(iii) Classifiers. In the presence of numerals, the choice of the classifier is determined by whether the noun is notionally count or mass. In other words, **the distribution of the classifiers is sensitive to the mass-count distinction on the NP**. Some classifiers select exclusively count nouns (3a), others are compatible with both count and mass nouns (3b), yet others look for mass nouns (3c), and some look exclusively for aggregates (3d), meaning that the substance/aggregate distinction is linguistically relevant, too (Grimm 2012). The distinction between these classifier expressions is left unexplained if we assume that all nouns are mass before a classifier is attached.

- | | |
|---|--|
| <p>(3) a. ikki dona/ dasta/ juft kitob
 two CL_{item} CL_{pile} CL_{pair} book
 'two (items of)/ piles of/ pairs of books'</p> | <p>b. ikki qop tuz / kitob
 two CL_{sack} salt book
 'two sacks of salt/ books'</p> |
|---|--|

c. **ikki** **litr** **suv**
 two CL_{liter} water
 ‘two liters of water’

d. **ikki** **zarracha** **tuz**
 two CL_{particle} salt
 ‘two particles of salt’

The facts are accounted for straightforwardly if we assume that some nouns in Uzbek denote sets of individuated entities, while others do not, denoting instead non-atomic/ non-disjoint entities.

3. Formal Analysis of Uzbek Classifiers. 3.1 Syntactic Analysis. We propose that Uzbek classifiers do not merge directly with the NP. Rather, they first combine with the numeral, and the resulting expression, in turn, forms a constituent with the NP ([_{DP} [_{MeasP} ikki dona] [_{NP} kitob]]) ‘two CL_{ITEM} book’). Our motivation for this claim is both theoretical and empirical. **Theoretically**, robust evidence in favor of the mass-count distinction in Uzbek suggests that classifiers are not needed in order to “turn the nouns into count”. Therefore, we follow the approach according to which classifiers are needed for numerals, turning them into functions that can combine with an NP without type mismatch (Krifka 1995, Bale & Coon 2014, Sudo 2016, Little & Winarto 2019, Sağ-Parvardeh 2019). **Empirical evidence** can be divided into two parts. First, the particularly widespread and underspecified classifier *-ta* is a suffix which, crucially, attaches to the numeral, rather than the noun (*ikki-ta kitob* ‘two-CL books’, not **ikki kitob-ta*). Second, classifiers (whether suffixational or free morphemes) are incompatible with other counting-related units which get suffixed to the numeral, such as the approximative and the collective suffixes (1c). A combination leads to ungrammaticality, suggesting that classifiers, approximatives and collectives compete for the same syntactic position and are thus all adjoined to Num.

3.2 Semantic Analysis. We propose the following general formula for representing the semantic contribution of classifiers (the formalism partly based on Chierchia 2010): $\lambda n \lambda P \lambda x. \mu_{U, Q^P}(x) = n$. A classifier looks for a numeral, a property and an individual argument, and specifies that the number of individuals is n . The property that these individuals bear, Q^P in the formula, is typically identical to P ; however, we will see that with some classifiers, it is related to P in a different way. Finally, the main difference between classifiers has to do with the nature of the units into which Q^P objects are divided (U). Below, we consider several examples.

-ta is an underspecified classifier, compatible with notionally count and (some) mass nouns. It does not impose any requirements regarding the division of the nominal’s denotation into units. With count nouns, these will be natural units (cf. Krifka 1989, Rothstein 2010), which are built in into the semantics of the N. When such units are absent, e.g. with substance mass nouns, the appropriate division should be provided by the context (e.g. water bottles may serve as units of water). If the context does not provide such a division, *-ta* will be inapplicable.

(4) [[*-ta*]] = $\lambda n \lambda P \lambda x. \mu_{U, P}(x) = n$

nafar ‘person’ is a sortal classifier that is only compatible with count human nouns. If P is not divisible into natural units, *nafar* is unacceptable. It does not allow a contextually triggered division. The underlined part is presupposed.

(5) [[*nafar*]] = $\lambda n \lambda P \lambda x. \mu_{NU, P}(x) = n \ \& \ \underline{P \subset \text{HUMAN}}$

qop ‘sack’ is a mensural (mass) classifier that imposes a particular measure unit. Mass nouns often require such classifiers, because division into units is not built in into their own semantics.

(6) [[*qop*]] = $\lambda n \lambda P \lambda x. \mu_{\text{SACK}, P}(x) = n$

gala ‘flock’ is an example of a group classifier. We propose that such classifiers impose a NU requirement, like *nafar*, but they further create a cluster reading: the natural units that get counted are not P units but rather units of P -clusters (in the sense of Grimm 2012).

(7) [[*gala*]] = $\lambda n \lambda P \lambda x. \mu_{NU, \text{CLUSTER}(P)}(x) = n \ \& \ \underline{P \subset \text{AVIAN}}$