

**Beyond the object-substance distinction in a classifier language:
Experimental evidence from Tashkent Uzbek**

Background In the literature, it has been commonly assumed that nouns in classifier languages (CLs) have uniform unindividuated semantics, such that they are all invariably mass (e.g., Ritchie 1971, Sharvy 1978). It has also been claimed that nouns in *all* languages may be either underspecified for count-mass (e.g., Borer 2005) or flexible w.r.t. the two values (e.g., Pelletier 2012). Within these frameworks, then, classifiers in CLs are necessary for making uncountable nouns countable.

Others have argued that individuation is, in fact, encoded in the semantics of nouns, even in CLs (e.g., Doetjes 1997, Cheng & Sybesma 1998, Chierchia 2010, Rothstein 2010, Sudo 2016), a position supported by experimental data showing that despite the absence of count syntax, speakers of CLs have access to the core non-uniform semantics of nouns (e.g., Barner et al. 2009, Li et al. 2009, Cheung et al. 2012, Inagaki & Barner 2009). Importantly, though, even scholars recognizing the non-uniform nature of nouns in CLs assume that the grammatical count-mass distinction in these languages merely aligns with the cognitive object-substance distinction (see Chierchia’s 2021 “*cognitively count vs. mass*”).

A recent exception is Erbach et al. (2021), who present preliminary empirical evidence from Japanese, a CL, suggesting that while there is considerable overlap between the linguistic categories *count-mass* and the cognitive categories *object-substance*, the two are not fully aligned.

Current study Taking Erbach et al.’s exploratory findings as a starting point, the goal of the current study is to establish the existence of a lexicalized count-mass distinction in Tashkent Uzbek (TU). TU is an influential dialect of Uzbek that has a sizable inventory of numeral classifiers, including *sortal*, *mensural*, *group*, and *granular*, which are obligatory in numeral constructions, with both prototypically count nouns like *kitob* ‘book’ (1a) and prototypically mass nouns like *ter* ‘sweat’ (1b) (e.g., Levy-Forsythe & Kagan 2022).

- | | |
|---|---|
| (1) a. <i>ikki dona kitob</i>
two CLitem book
‘two books’ | b. <i>ikki tomchi ter</i>
two CLdrop sweat
‘two drops of sweat’ |
|---|---|

We want to systematically demonstrate that nouns in TU are not uniformly unindividuated and, more importantly, that the count-mass distinction in TU – just like in English-type (Chierchia’s 2021 “Type I”) languages – *transcends* the cognitive object-substance distinction.

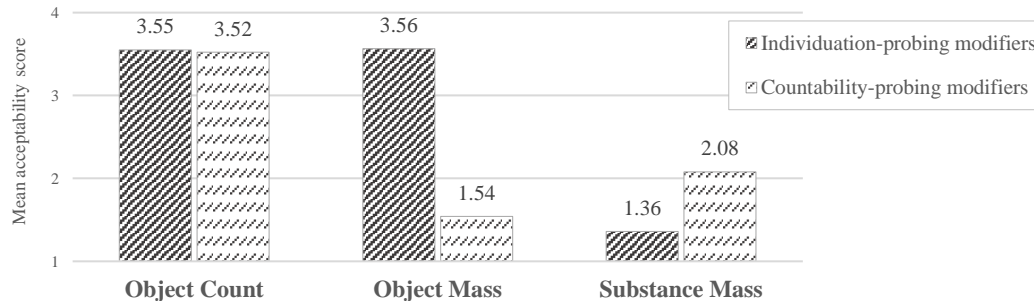
Methods We developed a judgment paradigm to elicit acceptability ratings of sentences with a range of modifier+noun combinations. We tested two canonical noun categories: object-denoting (count) nouns (e.g., *xat* ‘letter’), and substance-denoting (mass) nouns (e.g., *qor* ‘snow’). Crucially, we also tested a non-canonical nominal category: so-called *object mass* nouns (e.g., *mebel* ‘furniture’). Nouns in this category refer to clearly identified units but have mass syntax. The modifiers were of two types: a) those sensitive to notional (un)individuation, and b) modifiers sensitive to morphosyntactic countability. Individuation-probing modifiers included an adjective of size *katta* ‘big’ and a reciprocal *bir-biriga o’xshash* ‘similar to each other’. Countability-probing modifiers included a cardinal numeral *uch* ‘three’ followed by a general classifier *-ta* (i.e., *uchta* ‘three.CL’) and a cardinal numeral *ikki* ‘two’ followed by a collective suffix *-ala* (i.e., *ikkala* ‘both’). The experimental design is presented in the table below.

		Modifier Type	
		Individuation-Probing	Countability-Probing
Noun type	Object Count	Xonada katta televizor o’rnatildi. Room.LOC big TV installed.PSV ‘A big TV was installed in the room.’	Vazirlikda ikkala xat imzolandi. Ministry.LOC two.COLL letter signed.PSV ‘Both letters were signed at the ministry.’
	Object Mass	Zavodda katta mebel ishlab chiqarildi Factory.LOC big furniture produced.PSV ‘Big furniture was produced in the factory.’	Yo’lda ikkala pochta yo’qoldi. Road.LOC two.COLL mail lost.PSV ‘Both mails were lost on the road’
	Substance Mass	Rasmda katta qor chizildi Picture.LOC big snow drew.PSV ‘Big snow was drawn in the picture.’	Laboratoriyada ikkala gaz suyultirildi Lab.LOC two.COLL gas liquefied.PSV ‘Both gases were liquefied at the lab.’

There were 6 items in each experimental condition, for a total of 36 critical items. All nouns appeared in their bare form, i.e., without the plural suffix. The experiment was conducted online via

Qualtrics^{xm}. Verbal stimuli were presented as fully randomized audio files. Adult speakers of TU (N=40) were asked to provide acceptability judgments on a 4-point scale, with only the extreme ends explicitly labeled 1= *past* ('low'); 4= *baland* ('high'). The midpoint was excluded to create a forced-choice scale (cf. Chyung et al. 2018)

Results and analysis A summary of the results is plotted in the graph below, presenting the mean scores for each modifier type across conditions.



As can be seen from the graph, acceptability scores in the object count condition are near ceiling for both types of modifiers. This is in stark contrast to the results of the substance mass condition, in which both modifiers receive low ratings. Most interestingly for the current study are the data in the object mass condition. Here, speakers' judgments greatly diverge. When the noun is preceded by an individuation-probing modifier, the ratings are essentially identical to those in the object count condition; in contrast, nouns preceded by countability-probing modifiers yield judgments that closely pattern with those in the substance mass condition. Lastly, it is worth noting the slightly elevated ratings of sentences with countability-probing modifiers in the substance mass condition. We attribute this to the contextual mass-to-count shift enabled by the availability of the 'standard packaging' and the '(sub)kinds' reading, typical for substance mass nouns.

To statistically analyze the significance of the findings, we performed a Paired-Samples T Test. We found a main effect of Noun Type ($p < 0.001$). Additionally, a significant interaction of Noun Type and Modifier Type was found in the Object Mass and the Substance Mass conditions ($p < 0.001$), but not in the Object Count condition ($p = 0.5567$).

Discussion The current study conclusively affirms the existence of two canonical noun classes in TU: object count (e.g., *xat* 'letter', and substance mass nouns, e.g., *qor* 'snow'. Such findings are clearly at odds with claims that nouns in CLs have uniform unindividuated semantics. Most notably, our study also provides robust evidence for the existence of an additional, non-canonical nominal class, namely, *object mass nouns*. Morphosyntactically, object mass nouns (e.g., *mebel* 'furniture') pattern with mass nouns in that they are incompatible with number coding; unlike canonical substance mass nouns, however, object mass nouns refer to individuals (Barner & Snedeker 2005). In this sense, object mass nouns represent a dissociation between the linguistic count-mass distinction and the cognitive object-substance distinction (cf., Carey & Spelke 1996). Accordingly, under the view that in CLs, the linguistic distinction fully aligns with the cognitive distinction, such non-canonical nouns are predicted to be entirely absent in CLs such as TU (Chierchia 2021: 23). This prediction is not borne out by the results. Instead, the study provides robust experimental evidence that in TU, modification by classifiers (and other modifiers) is selectively restricted by the quantificational properties of the modified NP.

In sum, we have been able to offer previously unavailable, clear evidence for three distinct nominal classes in a CL. This challenges the prevailing typology of noun semantics, which assumes a fundamental distinction between number-marking languages such as English and CLs like TU.

Selected references: Barner et al. (2009). Language, thought, and real nouns. *Cognition* 111. 329–344. | Borer (2005). *In name only*. OUP. | Carey & Spelke (1996). Science and core knowledge. *Philos. Sci.* 63, 515–533. | Cheng & Sybesma (1998). Yi-wan tang, yi-ge tang: Classifiers and massifiers. *J. Chin. Stud.* 28, 385–412. | Chierchia (2010). Mass Nouns, Vagueness, and Semantic Variation, *Synthese* 174, 99–149. | Chierchia (2021). Mass vs. count: Where do we stand? Outline of a theory of semantic variation. In Kiss et al. (Eds.), *Things and Stuff: The Semantics of the Count-Mass Distinction* (21–54). CUP. | Doetjes (1997). Quantifiers and Selection: On the Distribution of Quantifying Expressions in French, Dutch and English (*PhD dissertation*), Leiden U. | Erbach et al. (2021). Object Mass Nouns as an Arbiter for the Count–Mass Category. In Kiss et al. (Eds.), *Things and Stuff: The Semantics of the Count-Mass Distinction* (167–192). CUP. | Levy-Forsythe & Kagan (2022). Classifiers and the mass-count distinction in Uzbek. In *Semantics and Linguistic Theory*, 685–704. | Pelletier (2012). Lexical Nouns are both +MASS and +COUNT but they are neither +MASS nor +COUNT. In Massam (Ed.), *Count and Mass across Languages*, 9–26. OUP. | Rothstein (2010). Counting and the mass/count distinction. *J.Semantics* 27. 343–397. | Sudo (2016). The Semantic Role of Classifiers in Japanese. *Balt. Int. Yearb. Cogn. Log. Commun* 11, 1–15. |