On a grammaticized lexical count-mass distinction in classifier languages: Experimental evidence from Tashkent Uzbek

Background Traditionally, nouns in classifier languages (CLs) were claimed to have uniform unindividuated (i.e., mass) semantics (e.g., Sharvy 1978). More recent literature argues that nouns crosslinguistically may be either underspecified (e.g., Borer 2005) or flexible w.r.t. count-mass (e.g., Pelletier 2012). Within these frameworks, then, the count reading is obtained only at the syntactic level, e.g., via classifiers.

Alternatively, others have argued that the count-mass distinction is, in fact, encoded in the semantics of nouns, even in CLs (e.g., Cheng & Sybesma 1998, Chierchia 2010, Rothstein 2010), 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). Importantly, though, even scholars recognizing the non-uniform nature of nouns in CLs assume that the linguistic count-mass distinction in these languages merely aligns with the cognitive object-substance distinction (e.g., Chierchia 2021).

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

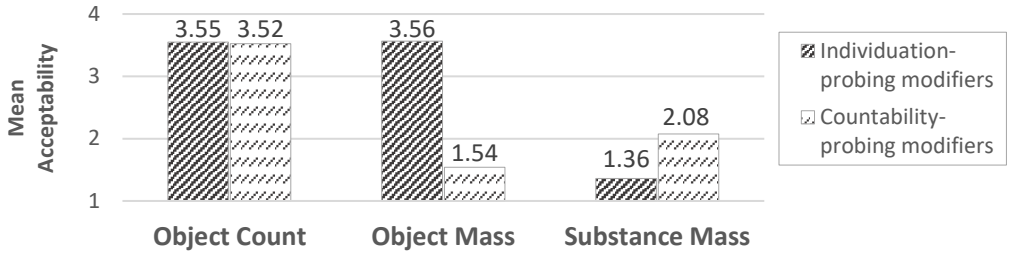
Current study Taking Erbach et al.'s exploratory findings as a starting point, the goal of the current study is to establish the existence of a lexicalized count-mass distinction in Tashkent Uzbek (TU), an obligatory classifier dialect of Uzbek. Specifically, we want to systematically demonstrate that nouns in TU are *not* uniformly unindividuated, and more importantly, that the count-mass distinction in TU – just like in English – *transcends* the cognitive object-substance distinction.

Methods We developed an experimental paradigm to elicit acceptability ratings of sentences with a range of modifier+noun combinations. Three nominal categories were tested: object count (e.g., *xat* 'letter'), substance mass (e.g., *qor* 'snow'), and so-called *object mass* (e.g., *mebel* 'furniture'). 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 followed by either a general classifier *-ta* (i.e., *uchta* 'three.CL') or by a collective suffix *-ala* (i.e., *ikkala* 'both'). The experimental design, along with some example items are 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 condition, for a total of 36 experimental items. Examples from each sentence type are provided below. The task was conducted online via Qualtrics^{xm}. Verbal stimuli were presented as fully randomized audio files. Adult TU speakers (n=40) were asked to determine the likelihood that the test sentences could be produced by a native speaker of TU. Judgments were noted on a 4-point scale, with only the extreme ends explicitly labeled 1= *past* ('low'); 4= *baland* ('high').

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



The graph reveals that acceptability ratings in the object count condition are at near ceiling for both types of modifiers. This is in stark contrast with the results observed in the substance mass condition, where both modifiers receive low ratings. Particularly striking are the results of the object mass condition, in which speakers' judgments are sharply polarized as a function of modifier type. While modification by an individuation-probing modifier essentially mirrors the response pattern in the object count condition, countability-probing modifiers yield judgments that closely pattern with those in the substance mass condition.

A final, minor note concerns 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 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 Our data affirm the existence of two canonical noun classes in TU (object count and substance mass), which is clearly at odds with claims that nouns in CLs have uniform 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 pattern with mass nouns, i.e., they are incompatible with number coding; unlike canonical substance mass nouns, however, object mass nouns refer to individuals (cf. Barner & Snedeker 2005). As such, 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 and the cognitive distinctions fully align, such non-canonical nouns are predicted to be entirely absent in CLs such as TU. This prediction is not borne out by the results of the current study.

In sum, to the best of our knowledge, no existing research to date has been able to offer such clear evidence for three distinct nominal classes (object count, substance mass, and object mass) in a CL. These previously unavailable, systematically controlled, experimental data strongly indicate that a grammaticized lexical count-mass distinction is, in fact, encoded in the semantics of nouns in (at least some) CLs. Hence, our findings pose a serious challenge for 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, H. (2005).** *In name only*. Oxford University Press. | **Cheng, L. & Sybesma, R. (1998).** Yi-wan tang, yi-ge tang: Classifiers and massifiers. *The Tsing Hua Journal of Chinese Studies*, *28*(3), 385–412. | **Chierchia, G. (2010).** Mass Nouns, Vagueness, and Semantic Variation, *Synthese, 174*: 99-149. | **Erbach et al. (2021).** Object Mass Nouns as an Arbiter for the Count–Mass Category. In *Things and Stuff: The Semantics of the Count-Mass Distinction* (pp. 167-192). Cambridge: Cambridge University Press.