Vagueness, Overlap, and Countability[1]

Peter R. Sutton — Heinrich Heine Universität Düsseldorf
Hana Filip — Heinrich Heine Universität Düsseldorf

Abstract. We propose a novel semantic analysis of the mass/count distinction, within a new framework combining the theory of mereology with Probabilistic Type Theory with Records, Prob-TTR (Cooper et al. 2014)). While the notions akin to VAGUENESS (Chierchia 2010) and OVERLAP (Landman 2011) are needed to ground this distinction, neither on its own is sufficient to accommodate the whole range of data, especially the puzzling intra- and crosslinguistic variation in count vs. mass encoding. This variation becomes tractable, if we generally treat the grammatical differences between mass and count nouns as following from the interaction of two notions: namely, VAGUENESS sharpened in terms of graded (probabilistic) type judgements, and DISJOINTNESS relative to a probability threshold. As a result, in the form-denotation mappings, this leads us to a novel semantic classification of nouns into four classes. The mass/count distinction is a bipartite grammatical distinction manifested in the standard diagnostics like a direct combination with numerals, the indefinite article and quantifiers like *every, much*, among others.

Keywords: mereology, mass/count distinction, probabilisitic semantics, vagueness.

1. Introduction

A major challenge for any semantic account of the mass/count distinction in nouns is to account for intra- and crosslinguistic variation in grammatical mass/count encoding. For languages with a grammatically encoded mass/count distinction, some nouns are fairly universally encoded as either MASS or COUNT. Mass nouns of this sort tend to be prototypical substance nouns (*air*, *water*, *mud*). Count nouns of this sort tend to be prototypical object nouns (*chair*, *car*, *girl*, *cat*). However, there are a very large number of nouns for which variation in mass/count encoding is rife. For example, we have the approximate synonyms in (1)–(4):[2] .

(1) *furniture*$_{-C}$; *huonekalu-t*$_{+C,PL}$ (Finnish); *meubel-s*$_{+C,PL}$, *meubilair*$_{-C}$ (Dutch).
(2) *kitchenware*$_{-C}$; *Küchengerät-e*$_{+C,PL}$ (German, lit. kitchen device-s).
(3) *lentil-s*$_{+C,PL}$; *linse-n*$_{+C,PL}$ (German); *lešta*$_{-C}$ (Bulgarian); *čočka*$_{-C}$ (Czech).
(4) *oat-s*$_{+C,PL}$, *oatmeal*$_{-C}$; *kaura*$_{-C}$, *kaurahiutale-et*$_{+C,PL}$ (Finnish, lit. oat.flake-s).

Focusing on such data, in this paper, we examine two influential analyses of the mass/count distinction. One offered by Chierchia (2010), which takes it to be a matter of vagueness. We will

[1]This research was funded by the German Research Foundation (DFG) CRC991, project C09. We would like to thank Robin Cooper, Zsofia Gyarmathy, Todor Koev, Fred Landman, Susan Rothstein, Karoly Varasdi, and the participants of Sinn und Bedeutung 20, and of the HHU Semantics and Pragmatics Exchange (SemPrE) colloquium.

[2]Subscripts $+C$ and $-C$ indicate COUNT and MASS respectively.

show that on Chierchia's vagueness based account, we are forced to give two disparate explanations for the variation in "fake" mass nouns, such as those in (1)–(2), and "granular" nouns, such as those in (3)–(4). Largely inspired by an attempt to improve on Chierchia (2010), the other account is that of Landman (2011). It relies on the notion of overlap, and while it provides a better account of "fake" mass nouns, it still lacks any proposal for why we see variation in "granular" nouns, such as those in (3)–(4).

Although the notions akin to VAGUENESS (Chierchia 2010) and OVERLAP (Landman 2011) are needed to ground the mass/count distinction, neither on its own is sufficient to accommodate the whole range of puzzling intra- and crosslinguistic variation in count vs. mass encoding. This variation becomes tractable, as we argue, if we generally treat the grammatical differences between mass and count nouns as motivated by the interaction of two notions: namely, VAGUENESS sharpened in terms of graded (probabilistic) type judgements, and DISJOINTNESS relative to a probability threshold. This effectively amounts to the innovative claim that the mass/count distinction is dual-source, rather than just mono-source, as is proposed in previous mereologically-based accounts. Assuming this DUAL-SOURCE HYPOTHESIS, we propose a novel semantic account of the mass/count distinction: namely, just as Chierchia (2010) and Landman (2011), it relies on the theory of mereology, but in departure from previous mereologically-based approaches, it enriches mereology with certain assumptions from Probabilistic Type Theory with Records Prob-TTR (Cooper et al. 2014).

In §§2–3, we will introduce the vagueness based theory of Chierchia (2010) and the overlap based theory of Landman (2011). In §4 we argue that both fail to capture intra- and crosslinguistic variation in mass/count encoding and motivate a dual-source hypothesis which predicts four semantic classes of nouns. It turns out that two of these show little or no mass/count variation and two display a large amount. The distinct advantage of our dual-source framework is a much wider data coverage than previous accounts like Chierchia (2010) or Landman (2011), among others, can offer. In §5, we will (very briefly) introduce Type Theory with Records, and its probabilistic variant. In §6 we outline a mereological enrichment of probabilistic Type Theory with Records, detail how vagueness and disjointness are represented in the system, and describe the different semantic properties of the four entity classes. In §7, we summarize these results and suggest some future extensions for our probabilistic mereological approach.

2. Vagueness

Chierchia's (2010) main claim is that mass nouns are vague in a way that count nouns are not. The denotations of count nouns are generated from "stable atoms": It is clear, non-vague, what to count. The denotations of mass nouns are generated from "unstable" individuals: it is vague what the suitable minimal elements are for counting. Taking a simple example, for a count noun like *cat*, we have a reasonably clear idea of what qualifies as a cat atom. Even though *cat* may be vague (when does a cat embryo become a cat?), there are some individuals that will be atoms in the denotation of *cat* no matter how this vagueness is resolved. In contrast, mass nouns like *water* or *rice* are vague in another way in so far as there is no systematic basis for deciding which water/rice amounts qualify as water/rice atoms. Chierchia attributes this vagueness to a variation in what is semantically an atom in the denotation of mass nouns across contexts:

> "A spoonful of rice is rice. What about a single grain of rice? In many contexts we would not consider a single grain of rice to be enough to reach the threshold of significance. To a child saying she has finished her rice, no parents in their right mind would reply 'no you have not' upon detecting a single grain. Yet for some other purposes we might consider a single grain of rice, rice. But then, that applies to half grains as well. And to quarters of grains." (Chierchia 2010: p. 117)

Chierchia proposes to enrich the mereological structure with a supervaluationist semantics to model this vagueness. Supervaluationism interprets vague NL predicates as including a vagueness band. Relative to a world and a context,[3] a predicate P has a positive extension ^+P, the set of things which count as P, no matter how one might precisify P, and a negative extension ^-P, the set of things which do not count as P, no matter how one might precisify P. The context relative to which ^+P and ^-P is defined is called the *ground context*. If P is vague, then ^+P and ^-P do not form a total partition on the domain. There are elements that sometimes do and sometimes do not count as P depending on the way P is precisified.

Here is a simplified example. Relative to a world and a ground context, the intension of $rice$ will denote a set of objects that are of sufficient quantity to be clear/indisputable cases of rice. For a domain $\mathbf{D} = \{a, b, c, a \cup b, a \cup c, b \cup c, a \cup b \cup c, d\}$, where '$\cup$' is a mereological sum, assume that $^+rice = \{a \cup b \cup c\}$. Further, assume that the negative extension of $rice$ is d (some non-rice): $^-rice = \{d\}$. Since $rice$ is vague, there are ways to precisify its meaning. Precisifications are total partitions of the domain. They form a partial order in that each precisification is either a positive or a neutral extension of the denotation of ^+rice. In the current case, for the total precisification contexts c_0, c_1, c_2, if:

(5)
$$\begin{array}{rcr} rice_{c_0} & = & \{a \cup b \cup c\} \\ rice_{c_1} & = & \{a \cup b, a \cup c, b \cup c, a \cup b \cup c\} \\ rice_{c_2} & = & \{a, b, c, a \cup b, a \cup c, b \cup c, a \cup b \cup c\} \end{array}$$

then c stands in the order $c_0 \propto c_1 \propto c_2$. Chierchia adapts the standard notion of an atom relative to a predicate (6), to an atom relative to a predicate at a ground context (7), and defines stable atoms for a predicate (8) in terms of a definitely operator (9). Atoms relative to each total precisification are given in (10) which means that the set of stable atoms for *rice* is empty. *Rice* has only unstable individuals, not stable atoms.

(6)
$$AT(P) = \{x \in P : \forall y \in P(y \subseteq x \rightarrow y = x)\}$$

[3] Chierchia employs a form of Data Semantics. As such, his 'contexts' play an equivalent role to 'completions of a partial model' in other forms of supervaluationism.

(7) $$AT_c(P) = \{x \in P^+ : \forall y \in P^+(y \subseteq x \rightarrow y = x)\}$$

(8) $$\mathbf{AT}(P) = \lambda x.D[AT(P)(x)]$$

(9) $$[[D\phi]]_c = 1 \text{ iff for all total precisifications } c' \text{of } c, [[D\phi]]_c = 1$$

(10) $$\begin{array}{rcl} AT_{c_0}(rice) & = & \{a \cup b \cup c\} \\ AT_{c_1}(rice) & = & \{a \cup b, a \cup c, b \cup c\} \\ AT_{c_2}(rice) & = & \{a, b, c\} \end{array}$$

Put another way, the intersection of the sets in (10) is empty: there are no stable atoms for $rice$. Counting is counting stable atoms, so $rice$ is MASS.

3. Overlap

A central concept in Landman (2011) is that of a *generator*. The set of generators, $\mathbf{gen}(X)$, of the regular set *X* is the set of *semantic building blocks*, which are either "the things that we would want to count as one" (Landman 2011: p. 26), relative to a context, or are contextually determined minimal parts.[4] If the elements in the generator set are non-overlapping, as in the case of count nouns, then counting is sanctioned: Counting is counting of generators and there is only one way to count. However, if generators overlap, as in the case of mass nouns, counting goes wrong. One of Landman's innovations is to provide a new delimitation of the two cases when this happens, and hence two subcategories of mass nouns: Mess-Mass nouns like *mud* and Neat-Mass nouns like *furniture* (aka "aggregate mass terms" in Payne and Huddleston (2002)). A mass noun is *Neat* if its intension at every world specifies a regular set whose set of minimal elements is non-overlapping. A noun is a *Mess-Mass* noun if its intension at every world specifies a regular set whose set of minimal elements is overlapping.

An example of Mess-Mass noun is *mud*. Relative to a context, there are assumed to be minimal *mud* elements. For the purposes of presentation, let us assume that mud is, roughly, wet dirt, and any minimal element of *mud* must have at least one water component (w_i) and at least one dirt component (d_i), put in the simplest terms. In Figure 1, the minimal elements for mud are highlighted (anything with one dirt and one water component). In this example, the set of minimal elements could equal the set of generators or some superset could, but notice that in either case, the elements of the generator set overlap ($(d_1 \cup w_1) \cap (d_1 \cup w_2) \cap (d_2 \cup w_1) \cap (d_2 \cup w_2) \neq \varnothing$), so *mud* is mass. The minimal elements also overlap, so *mud* is mess. As far as counting is concerned, simultaneously, we have two *variants* of mud of two non-overlapping building blocks: (i) $(d_1 \cup w_1)$ and $(d_2 \cup w_2)$, (ii) $(d_1 \cup w_2)$ and $(d_2 \cup w_1)$. However, it is equally appropriate to regard the mud as being built from $(d_1 \cup w_1)$ and $(d_2 \cup w_2)$ or from $(d_1 \cup w_2)$ and $(d_2 \cup w_1)$. If we count generators we would

[4] We will discuss this difference in more detail below.

count four entities, if we count variants, we count only two. This clash of answers is the reason why *mud* cannot be counted.

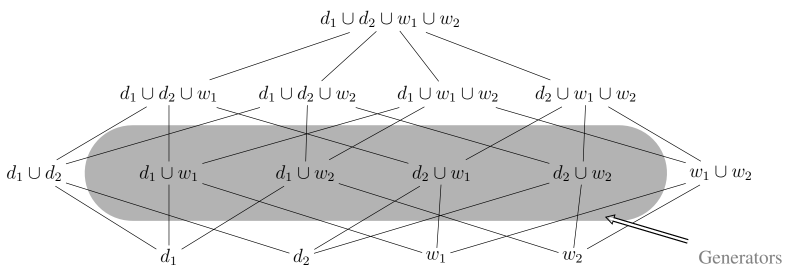

Figure 1: Generators and minimal elements for *mud*

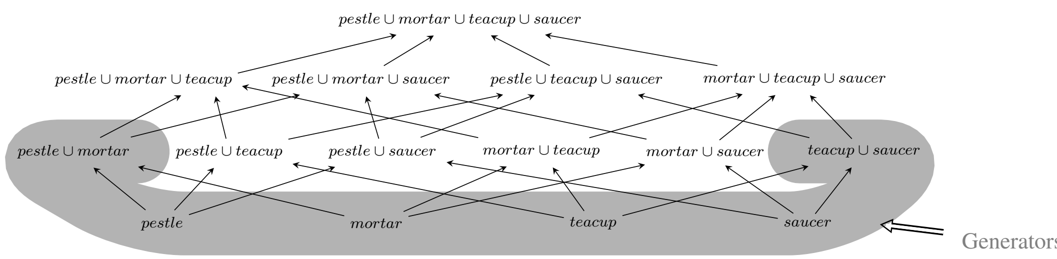

Figure 2: Generators for *kitchenware*

Neat-Mass nouns have a different pattern, illustrated in Figure 2, where the minimal elements of *kitchenware* are a pestle, mortar, teacup and saucer. These are the *non-overlapping minimal generators*. For many purposes, a pestle and a mortar together, and a teacup and a saucer together naturally count as single items of *kitchenware*, in an appropriate context, and so these building blocks are also in the generator set (albeit *not* in the *minimal* generator set). The minimal elements (minimal generators) of *kitchenware* are thus a subset of the generator set. The minimal generators do not overlap ($mortar \cap pestle \cap teacup \cap saucer = \varnothing$), which makes *kitchenware* "neat", but the generator set does, which makes it MASS. Counting goes wrong precisely because the set of generators, i.e., the semantic building blocks that we intuitively want to count as one, includes more than just the minimal elements (generators), namely generators that overlap, as we see in the highlighted area in Figure 2. When we have a pestle, mortar, teacup and saucer, do we have four, three or two items of kitchenware? This clash of answers, due to the fact that both singularities and pluralities can be counted as one simultaneously in the same context, is the reason why *kitchenware* cannot be counted.

4. Vagueness, Overlap and Mass/Count Variation.

Both Chierchia's vagueness-based approach and Landman's overlap-based approach can account for some classes of nouns that were previously not well accounted for in countability research.

Chierchia (2010) is able to explain why nouns which have perceptually salient minimal parts, but are granular in nature (e.g. *rice*) are encoded as MASS. Namely, the perceptually salient parts (the grains) are not stable atoms, but unstable individuals. Landman (2011) is able to address why superordinate/aggregate nouns with clearly salient minimal entities (e.g. *furniture, footwear*) are encoded as MASS. Namely, the perceptual and/or functional parts (items of furniture/footwear) overlap, and so are not defined on the counting function. Despite these clear marks of progress, both accounts face considerable challenges from cross- and intralinguistic mass/count data.

In this section, we will discuss a variety of cross- and intralinguistic data that are problematic for any account defined purely in terms of either vagueness (alone) or overlap (alone). As we explain in §§4.1-4.2 using vagueness as a criterion for count/mass encoding alone or using overlap as a criterion for count/mass encoding alone insufficiently captures the range of relevant data.

4.1. Vagueness is Insufficient

There are vague nouns that are COUNT*:* If *čočka*$_{-C}$ ('lentil', Czech) is vague because, across contexts, what counts as having *čočka*$_{-C}$ varies from lentil dust (severe allergy contexts) up to some larger amount (making lentil soup contexts), then *čočka*$_{-C}$ is MASS. The same should be true for *lentils*, but it is not. *Lentils* is plural COUNT. In other words, the same criteria for count/mass encoding cannot be applied across languages, because there are near synonyms across different languages which are vague in the sense of Chierchia (2010), but that are COUNT in some languages and MASS in others.

Intralinguistically, this single criterion is also insufficient. In British English, (*porridge*) *oats* and *oatmeal* are frequently used as to mean the stuff one buys to make porridge from. These nouns are vague in the sense of Chierchia (2010) because, across contexts, what counts as having *oats/oatmeal* varies from oat dust (severe celiac contexts), up to less than around a cupful (making porridge for breakfast contexts), yet one is MASS and the other is COUNT.

Chierchia (2010) briefly addresses the issue of vague nouns being count:

> "What this suggests is that standardized partitions for the relevant substances are more readily available in such languages/dialects. This type of variation is a consequence of the fact that vagueness comes in degrees: some nouns may well be less vague than others, in the sense that a usable notion of 'smallest sample' can more readily be devised." (Chierchia 2010: p.140)

Although an appeal to degrees of vagueness might explain why one finds crosslinguistic variation in mass/count encoding for vague nouns (different languages have different standards for standardized partitions), it does not explain the intralinguistic cases. It remains unexplained, on Chierchia's account, why the standards for partitioning shift between uses of, for example, *hair*$_{+C}$ and *hair*$_{-C}$ or of *(porridge) oats* and *oatmeal*. In addition, it is not as though usable notions of 'smallest sample' are hard to devise for many of these nouns. For example, in Russian, *kartoška* (potato) is MASS, but potatoes come out of the ground in clearly packaged units. What must be explained is why, despite coming in standardized units, vague nouns may nonetheless be MASS or COUNT.

There are non-vague nouns that are MASS*:* If the Finnish *huonekalu-t*$_{+C}$ ('furniture') is not vague and therefore COUNT, then the English *furniture* should be COUNT too, but it is MASS. Equally, if vagueness is the only factor in mass/count encoding, we should not expect to find mass/count pairs such as the Dutch *meubel-s*$_{+C}$, *meubilair*$_{-C}$ ('furniture').

Chierchia (2010) is aware of this complexity, and suggests that number marking languages sometimes undergo a copycat process in which a potentially count noun is listed as a singleton property in the lexicon. Although this could be what is behind mass/count variation in non-vague superordinate nouns, we worry that as an explanation it is slightly *ad hoc*. In particular, we should expect to find other instances of 'copycatting' between lexical entires outside of the mass/count distinction, but Chierchia (2010) does not provide any such instances.

More concretely, however, Chierchia's proposal makes a prediction with respect to classifier languages that may be false. Chierchia's account predicts that copycatting should only occur in number marking languages, and that classifier languages should not have "fake" mass nouns. Although classifier languages such as Mandarin Chinese have been argued not to display a mass/count distinction in nouns, but in the classifier system, there is some evidence that superordinates such as the cognate of *furniture* display behavior with classifiers that is distinct from either prototypical count classifiers or prototypical mass classifiers. Cheng (2012) observes that the classifier noun pair *jiàn jiājù* ('piece furniture' Mandarin) behaves by one test more like a count noun and by another more like a mass noun. On the adjective test, *jiàn jiājù* can be modified by *dà* ('big') which patterns with mass constructions. However, it is ungrammatical to include *de* between the classifier and noun which is the pattern of a count construction. Although a more careful analysis of putative "fake" mass nouns in Mandarin and other classifier languages needs to be made, we suggest that at least for *jiàn jiājù* ('piece furniture' Mandarin), there is reason not simply to assume this classifier noun construction is straightforwardly COUNT or MASS. If, as Chierchia's account predicts, copycatting is not possible in classifier languages, at the very least, we would need an extra explanation for this difference which may suggest the possibility of a more parsimonious explanation.

4.2. Overlap is Insufficient

There are overlapping nouns that are COUNT*:* Nouns such as *furniture, kitchenware* have overlapping generators and so are encoded as MASS. However, if such nouns have overlapping entities that can simultaneously and in the same context 'count as one' it is highly puzzling why cognates *huonekalu-t*$_{+\mathrm{C}}$ ('furniture' Finnish) and *Küchengerät-e*$_{+\mathrm{C}}$ ('kitchenware', German) should be COUNT. Landman (2011) gives some details for the intralinguistic pair in Dutch, *meubel-s*$_{+\mathrm{C}}$ and *meubilair*$_{-\mathrm{C}}$ ('furniture'). Where MEUBEL is a disjoint set of items of furniture, and the first in the ordered pair is the generated set, and the second in the pair is the generator set:

$$\begin{array}{lll} \textit{meubel} & \rightarrow & \langle \text{MEUBEL}, \text{MEUBEL} \rangle \\ \textit{meubels} & \rightarrow & \langle {}^{*}\text{MEUBEL}, \text{MEUBEL} \rangle \\ \textit{meubilair} & \rightarrow & \langle {}^{*}\text{MEUBEL}, {}^{*}\text{MEUBEL} \rangle \quad \text{(Landman 2011: p. 35)} \end{array}$$

The neat mass noun *meubilair* has an overlapping generator set, but the single (plural) *meubel(s)* does not. Landman does not discuss what licenses such variation within his system, however, a reason can be given for why this kind of variation is possible. Perhaps cross- and intralinguistic variation occurs for neat mass nouns, but not mess mass nouns. In the above, the count noun *meubel* has non-overlapping *minimal* generators some of which can form the non-overlapping generator set. However, mess mass nouns have overlapping minimal generators, so provide no such basis for enabling a COUNT counterpart. Below, we will build on this thought.

There are non-overlapping nouns that are MASS*:* More problematic for Landman's (2011) account is the following. Presumably, count nouns such as *lentil, bean* must have non-overlapping generators. This is *prima facie* plausible given that the denotations of these granular nouns have clearly perceptible units, namely, individual lentils and beans. However, this makes it highly puzzling why we should find *čočka*$_{-C}$ ('lentil', Czech) and *bob*$_{-C}$ ('bean', Bulgarian). In the neat mass noun case, it was fairly intuitive to think of, say, a cup and saucer sum counting as one item of kitchenware, but with granular nouns is is hard to find any intuitive sense in which, say, two beans or half a bean could count as one bean-item simultaneously and in the same context as a whole bean. Alternatively, one could argue that *čočka*$_{-C}$ ('lentil', Czech) and *bob*$_{-C}$ ('bean', Bulgarian) are mess mass nouns and so have overlapping minimal generators. However, to do so would be to lose the ability to explain mass/count variation as a feature of neat nouns with overlapping generators. We, therefore, consider it better to understand foodstuff granular nouns *rice, lentils, beans* etc., to be neat, but only sometimes MASS. However, this suggests that there may be more to mass encoding than the property of having overlapping generators.

4.3. Summary of Mass/Count Variation Data and a Dual-Source Hypothesis

Table 1 and Table 2 collate some of the data for crosslinguistic and intralinguistic data, respectively. The fields Vague/Not Vague and Overlapping/Not Overlapping should be understood as they are defined in Chierchia (2010) and Landman (2011).

Given the problem cases outlined in §§4.1-4.2, and in considering the way these data group in Tables 1 and 2, a striking pattern emerges.

Table 1: Crosslinguistic Data

	Not Overlapping	Overlapping
Not Vague	*chair*$_{+C}$ = *tuoli*$_{+C}$ (Finnish) = *Stuhl*$_{+C}$ (German)	*furniture*$_{-C}$ = *huonekalu-t*$_{+C}$ (Finnish) = *Meubles*$_{-C}$ (French)
	car$_{+C}$ = *auto*$_{+C}$ (Finnish) = *Auto*$_{+C}$ (German)	*kitchenware*$_{-C}$ = keittiöväline-et$_{+C}$ (Finnish) = *Küchengerät-e*$_{+C}$ (German)
	girl$_{+C}$ = *tyttö*$_{+C}$ (Finnish) = *Mädchen*$_{+C}$ (German)	*footwear*$_{-C}$ = *jalkine-et*$_{+C}$ (Finnish) = *Schuhwerk* $_{-C}$ (German)
	cat$_{+C}$ = *kissa*$_{+C}$ (Finnish) = *Katze*$_{+C}$ (German)	*cutlery*$_{-C}$ = ruokailuväline-et$_{+C}$ (Finnish) = *Besteck*$_{-C}$ (German)
Vague	*rice*$_{-C}$ = *riisi*$_{-C}$ (Finnish) = *rýže*$_{-C}$ (Czech)	*mud*$_{-C}$ = *muta*$_{-C}$ (Finnish) = *Schlamm*$_{-C}$ (German)
	lentil-s$_{+C}$ = *linssi-t*$_{+C}$ (Finnish) = *čočka*$_{-C}$ (Czech)	*air*$_{-C}$ = *lenta*$_{-C}$ (Finnish) = *Luft*$_{-C}$ (German)
	oat-s$_{+C}$ = *kaura*$_{-C}$ (Finnish) = *vločky*$_{+C}$ (Czech)	*blood*$_{-C}$ = *veri*$_{-C}$ (Finnish) = *Blut*$_{-C}$ (German)
	bean-s$_{+C}$ = *pavu-t*$_{+C}$ (Finnish) = *bob*$_{-C}$ (Bulgarian)	*oil*$_{-C}$ = *öljy*$_{-C}$ (Finnish) = *Öl*$_{-C}$ (German)

Table 2: Intralinguistic Data

	Not Overlapping	Overlapping
Not Vague	*chair*$_{+C}$ vs. *seat*$_{+C}$ vs. *stool*$_{+C}$ *bus*$_{+C}$ vs. *coach*$_{+C}$	*meubel-s*$_{+C}$ vs. *meubilair*$_{-C}$ ('furniture' Dutch)
Vague	*(porridge) oat-s*$_{+C}$ vs. *oatmeal*$_{-C}$ (British English) *hair*$_{+C}$ vs. *hair*$_{-C}$	*mud*$_{-C}$ vs. *dirt*$_{-C}$ vs. *grime*$_{-C}$ *oil*$_{-C}$ vs. *grease*$_{-C}$ vs. *lubricant*$_{-C}$

(i) Those nouns which are vague and show intra- and crosslinguistic variation are the nouns which are non-overlapping.

(ii) Those nouns which are not vague and show intra- and crosslinguistic variation are the nouns which are overlapping.

(iii) Those nouns which are vague and show little or no intra- and crosslinguistic variation are the nouns which are also overlapping.

(iv) Those nouns which are not vague and show little or no intra- and crosslinguistic variation are the nouns that are also not overlapping.

In general terms, the above strongly suggests that although vagueness or overlap alone may enable mass encoding, a single feature alone also allows for count encoding. However, when BOTH vagueness and overlap are present, mass encoding is virtually enforced.

We hypothesize that there is more than mere coincidence to this pattern, and that there is some kind of interaction between vagueness and overlap that combines to block counting. Specifically, we predict that there will be some form of flexibility in the way the denotations of nouns are conceived when only ONE source (vagueness or overlap) is present, but no such flexibility when either BOTH vagueness and overlap are present, or when NEITHER vagueness nor overlap are present.

To test this hypothesis, we develop a formalism that can represent both vagueness and overlap.

However, we will also incorporate some changes to Chierchia's and Landman's accounts. In particular, we will include a broader and richer notion of what 'counts as one' into our account, beyond that of formal atomicity assumed by Chierchia (2010), and we will give a more intuitive formulation of in what way substance mass nouns are overlapping that does not rely on identifying entities that are minimal in context.

5. Type Theory with Records and Probabilistic Type Theory with Records

Type Theory with Records (TTR, Cooper 2012) is a richly typed semantic platform that combines both the rich expressivity of lexical semantic frames (Fillmore 1976) with a compositional semantics in the Montagovian tradition. However, TTR is also helpfully understood as a development of a situation theoretic semantics at least in so far as the truth makers of propositions in TTR are Records which, in application to natural language, should be understood as situations (which are partial) as opposed to possible worlds which are total structures.

TTR includes basic types such as Ind for individual, but also predicates which in the Frege-Montague tradition are n-place functions. For example, the predicate $cat(x)$ is short form for a function $\langle \lambda v.cat(v), \langle x \rangle \rangle$ which takes the value of some label, x, and returns the type of situation in which that individual is a cat. For example if $felix$ is the value for x in a situation, the resulting type would be $cat(felix)$, the type of situation in which felix is a cat.

TTR interprets propositions as types, in particular Record Types such as the one in (11). Propositions have proofs (things which make them true), which in application to natural language semantics are situations or events. The proposition/Record Type in (11) which is the type of situation in which some individual is a cat. Labels x, s_{cat} are provided values by the situation/Record which one is judges to be/not to be of the (Record) Type. As agents, we form judgements as to whether situations are of certain types. So, for example, if a situation, r contains Felix the cat, an agent may judge (truly) that r is of the type in (11). This is expressed in (12).

(11) $$\begin{bmatrix} x & : & Ind \\ s_{cat} & : & cat(x) \end{bmatrix}$$

(12) $$r : \begin{bmatrix} x & : & Ind \\ s_{cat} & : & cat(x) \end{bmatrix}$$

In probabilistic TTR, (Cooper et al. 2014, 2015), judgements are probabilistic. Where T is a type, probabilistic judgements are of the form $p(a : T) = k$ where $k \in [0, 1]$. Below we assume Cooper et al.'s implementation of the probabilistic variant of TTR into a simple Bayesian learning model. An agent A records a set of probabilistic judgements $\mathfrak{J}$ from her learning data and calculates the probabilities of new judgements from the information she has in her judgement set. Judgement sets evolve as the agent is exposed to new judgements made, for example by competent speakers. The value k in (13) will represent the prior probability an agent A has for some individual being a cat, given her judgement set $\mathfrak{J}$. Conditional probabilities are then computed as in (14) using a type theoretic version of Bayes' Rule where $||T||_{\mathfrak{J}}$ is the sum of all probabilities associated with T in $\mathfrak{J}$.

(13) $$p_{A,\mathfrak{J}}(r : \left[\begin{array}{lll} x & : & Ind \\ s_{cat} & : & cat(x) \end{array}\right]) = k$$ (14) $$p_{A,\mathfrak{J}}(s : T_1 | s : T_2) = \frac{||T_1 \wedge T_2||_{\mathfrak{J}}}{||T_2||_{\mathfrak{J}}}$$

6. Probabilistic Mereological Type Theory with Records (probM-TTR)

The simple enrichment of prob-TTR we make is replace the type Ind, the basic type for individuals, with the basic type $^*Ind$ which has a domain of 'stuff' which may include substances, individuals, their parts and sums thereof. Formally, the domain of $^*Ind$ will be represented as a whole Boolean semilattice closed under sum.[5] A learner's task will be to establish what, if anything, the individuals denoted by a particular predicate are. For example, given a world full of stuff, a learner of the predicate cat must learn which portions of stuff are individual cats. The type of individual for a predicate P will be represented Ind_P.

For languages that have any true mass nouns at all, we take this process of individuation to be a necessary but not a sufficient condition for countability. Individuation is not sufficient because (i) granular mass nouns, such as *rice*, denote stuff that comes in perceptually salient and identifiable units, namely grains; and (ii) aggregate mass nouns like *furniture* also denote stuff that comes in clearly individuable units (functionally if not also perceptually), namely tables, chairs etc. However, individuation is necessary. Substance and liquid nouns denote stuff that does not come in perceptually or functionally individuable units and it is precisely in these nouns we find uniformity in mass encoding for languages that display any form of mass/count distinction at all. We now sketch how the property of being individuable with respect to a predicate can be represented within probM-TTR.

Following some suggestions in Krifka (1995), we want to distinguish between a qualitative criterion of application and a quantitative criterion of application for predicates. Qualitative criteria include, *inter alia*, functional features, color, size, shape, separatedness, and perceptible granularity.[6] We will leave a more thorough investigation of these features for further research, but this large array of non-necessary but highly informative features makes the frame-based aspect of TTR a highly appropriate modeling tool. Quantitative criteria will be represented with a "quantitative" function, which operates on the values of the qualitative criteria of a noun and outputs a natural number (a function of type $(RecType \rightarrow NatNum)$). Given the size, shape, functionality, etc. of some object, stuff, or sum of objects, we can make a judgement about how many individuable entities there are with respect to a predicate. The special case will be where there is one (the output of the quantitative function is 1). When this occurs, the record type represents the type of situation in which there are objects that are individuals with respect to P. In other words, this record type may be abbreviated as the type of P-Individuals (Ind_P).

[5] This could equally be achieved using sets. For a set of formal atoms $\{a, b, c\}$, the domain of Ind entities would be $\{a, b, c, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$

[6] Some of these features are clearly mereotopological in the sense of Grimm (2012). Just how the insights of Grimm's work could be combined within our formalism is a project we intend to pursue.

For simplicity and brevity, for each predicate, we compress all of the qualitative features into one predicate P_{Qual}. The Record Type for the qualitative and quantitative criteria for applying a predicate P will contain: a record type containing P_{qual}; the specification of a quantitative function labelled as $f_{P_{quant}}$; and a condition stating that the function, when applied to the record type yields a natural number value i. A schema for such a type is given in (15), and an example assuming a numerical measure of 1 is given for the predicate $rice$ in (16)

(15) $$\begin{bmatrix} s_{p_{stuff}} & : & \begin{bmatrix} x & : & {}^{*}Ind \\ s_{p_{qual}} & : & P_{Qual}(x) \end{bmatrix} \\ f_{p_{quant}} & : & (\begin{bmatrix} x & : & {}^{*}Ind \\ s_{p_{qual}} & : & P_{Qual}(x) \end{bmatrix} \rightarrow \mathbb{N}) \\ i & : & \mathbb{N} \\ s_{p_{quant}} & : & f_{p_{quant}}(s_{p_{stuff}}) = i \end{bmatrix}$$

(16) $$\begin{bmatrix} s_{rice_{stuff}} & : & \begin{bmatrix} x & : & {}^{*}Ind \\ s_{rice_{qual}} & : & rice_{Qual}(x) \end{bmatrix} \\ f_{rice_{quant}} & : & (\begin{bmatrix} x & : & {}^{*}Ind \\ s_{rice_{qual}} & : & rice_{Qual}(x) \end{bmatrix} \rightarrow \mathbb{N}) \\ s_{rice_{quant}} & : & f_{rice_{quant}}(s_{rice_{stuff}}) = 1 \end{bmatrix}$$

The special case where the value of the quantitative function with respect to a predicate is 1 yields the type for an individual of that predicate. In other words the record type in (16) is the type for single grains of rice. This type can be abbreviated as $[x : Ind_{rice}]$.

These Ind_P types already allow us to mark a difference between substance/liquid nouns and all other concrete nouns. Perceptual and functional features of the denotations of nouns, we assume, allow one to infer what ‘counts as one’ in the denotation of the relevant noun. On the perceptual side, this, in part, involves identifying bounded units (what Grimm (2012) refers to as *Maximally Strongly Self Connected* entities). These include whole cats and apples, as well as, for example, whole lentils or rice grains. On the functional side, one might identify multiple bounded entities that jointly perform some function. For example, a pestle and mortar could be inferred to count as one item of kitchenware. However, for substance nouns such as *mud, blood* there are none of the consistent perceptual or functional cues to infer what counts as one that there are with prototypical count nouns, granular nouns, or aggregate nouns. From our probabilistic learning perspective, this will translate as high degrees of uncertainty as to what would be of the type Ind_P when P is a substance predicate. For example, for mud, there will be no a such that an agent would judge a high value for $p(a : Ind_{mud})$.

6.1. Contextual Variation, Vagueness, and Uncertainty in probM-TTR

In the supervaluationism of Chierchia (2010), contexts play the role of precisifications in other forms of supervaluationism. One begins with a ground context and an extension gap, and contexts create complete classical extensions of partial (vague) predicates. Some problems with supervaluationist approaches are well known. For one, it is unclear why ground contexts should be non-vague when expressions are vague. This is the problem of higher-order vagueness. If it is vague what falls into the extension of $rice$, it should also be vague what falls into the extension of that which is definitely rice. However, on Chierchia’s analysis, there is a sharp line between that which is always in the extension of e.g. $rice$, and that which is in the extension gap (relative to some ground context).

From a situation theoretic perspective, contexts are themselves situations, or in TTR terminology, records. In prob-TTR, judgments about contexts are recorded in agents' judgement sets. Following, approximately, the account of vagueness presented in Sutton (2013, 2015), we will now show how varied learning data (across situations/contexts) can yield uncertainty as to whether to apply a predicate. This form of metalinguistic uncertainty captures vagueness from a probabilistic perspective. As we shall argue, in some cases (such as with 'granular' predicates such as $rice$), even if one is completely clear as to what the individual units relative to a predicate are, one may still be uncertain about applying the predicate to small quantities of these units.

The way vagueness arises for granular nouns is very close to the way informally described in Chierchia (2010), however, we shift his informal observations into a learning setting. A learner is presented with competent speaker judgements with respect to situations, but sometimes the information they get is inconsistent. Here is a simple case:

Situation/context 1: Child learner has eaten all but around 10 or so grains of rice on her plate. Parent says: "You have eaten all of your rice."
Situation/context 2: Child learner spills 10 or so grains of rice from her plate. Parent says: "Oops, you spilled some rice."

The quantity (and potentially quality) of the rice is the same in each case, but the learner gets conflicting information. In Situation 1, she learns that ten or so grains does not count as rice. In situation/context 2, she learns that 10 or so grains does count as rice. From a Bayesian learning perspective, giving equal weight to the parent's assertions, she should assign a probability value of 0.5 that a competent speaker would judge ten or so grains to be rice. Within the probM-TTR framework, when the learner next comes to judge this quantity of rice, she applies Bayes' rule to her judgement set, and is left with a high degree of uncertainty whether or not to judge this quantity to be $rice$. The outcome of this calculation is shown in (17):[7]

(17) $$p_{A,\mathfrak{J}}(r : \begin{bmatrix} x & : & {}^{*}Ind \\ s_{rice} & : & rice(x) \end{bmatrix} \mid r : \begin{bmatrix} s_{rice_{stuff}} & : & \begin{bmatrix} x & : & {}^{*}Ind \\ s_{rice_{qual}} & : & rice_{Qual}(x) \end{bmatrix} \\ f_{rice_{quant}} & : & (\begin{bmatrix} x & : & {}^{*}Ind \\ s_{rice_{qual}} & : & rice_{Qual}(x) \end{bmatrix} \rightarrow \mathbb{N}) \\ i & : & \mathbb{N} \\ s_{rice_{quant}} & : & f_{rice_{quant}}(s_{rice_{stuff}}) = 10 \end{bmatrix}) = 0.5$$

Larger quantities of rice almost always get judged to be rice, smaller quantities will much less frequently do so. This will lead to a graded slope and an individuation challenge. Larger collections of grains will have a high degree of certainty of being judged to be rice, but larger collections of

[7]The value, 10, of the quantity function in (17) need not require there to be exactly 10 grains of rice. Larger measure values could be increasingly course grained in this respect.

grains do not provide stable bases for counting, since, at the very least, we are not clearly able to discern larger collections of grains from slightly smaller ones, nor to discern how many multiples of larger numbers of grains are in, say a big bowl of rice. So, in searching for a quantity that is sufficiently certain to be judged $rice$, one is left with something one cannot count. Most importantly, we conclude that this is one way in which contextual variation and conflicting learning data can give rise to mass encoding.

We have made no appeal to ground contexts on this account of vagueness. Furthermore, near certainty as to whether some entities are of some type will gradually fade into ever increasing amounts of uncertainty (for example, for smaller amounts of rice). In contrast to supervaluationism, this effectively removes sharp boundaries for ground contexts, and so alleviates problems associated with higher order vagueness, because totally clear cases of, for example, *rice* will not mark a sharp boundary, but rather the edge of a very gradual slope which eventually leads to the unclear cases.

However, the way of representing granular noun denotations such as in (17) yields two subtly different ways of conceiving of the referents. One can make classifications using, for example, the $rice(x)$ predicate. This leads to a 'de-emphasizing' of the type Ind_{rice}, since the presence of a single grain is not a always a good basis for judging something to be rice. However, alternatively, one may also judge whether some stuff is of the type Ind_P (or in the upward closure of this type under sum). For example, *lentil-s* in English are just as contextually variant as *rice* is with respect to conflicting learning data and so *lentil-s* could be treated the same as rice (as in the Czech *čočka*). However, in the case of *lentils*, English has 'chosen' to use the ready made type we have as learners which can form counting base, namely the type for individual, perceptually salient grains/lentils: Ind_{lentil}. Indeed, languages in general have made choices for various granular nouns on the basis of two alternative strategies. They could emphasize the number neutral predicate (e.g. $rice(x)$). This leads to non-countability in the face of uncertainty generated by contextual variation, since it is highly uncertain that one has rice when one has only one grain in certain situations. Alternatively, languages could emphasize the predicate indexed Ind_P type (e.g. Ind_{lentil}), in which case a lexical item would come to denote only individual grains and be straightforwardly countable. It is because both of these strategies are available that helps to explain, as we propose, why we see such common crosslinguistic variation in the mass/count encoding of granular nouns. Indeed, since both strategies may be useful, this also accounts for why we should expect to find intralinguistic COUNT/MASS pairs for granular nouns such as *(porridge) oats* and *oatmeal*.

Prototypical count nouns are not influenced by contexts in this way. It would be rare in the extreme for a learning data provided by competent speakers to conflict with respect to what are and what are not (whole) cats. Therefore, a learner will have little or no uncertainty about making corresponding cat judgements. In other words, as shown in (18) the probability that a single cat quantity (i.e. a single cat) will be judged to be a cat is high or close to 1:

(18) $$p_{A,\mathfrak{J}}(r : \begin{bmatrix} x & : & {}^{*}Ind \\ s_{cat} & : & cat(x) \end{bmatrix} \mid r : [x : Ind_{cat}]) \approx 1$$

So, whether or not a language 'emphasizes' the type cat or the type Ind_{cat}, the result will be the same, namely, a type with a clear counting base of individual cats.

In contrast, although substance nouns such as *mud* are context sensitive and vague in the same way as granular nouns like *rice*, our model can also explain why we so rarely find languages that encode substance nouns as count. For nouns such as *rice, lentils*, two options were available: emphasize the predicate type $P(x)$, or emphasize the Ind_P type. In the former case, one is left with uncertainty of what to count since the stuff which one is sufficiently certain of being rice comes in fairly large, hard to identify portions. In the latter case, one can count, because one can be highly certain that, say, individual lentils are of the type Ind_{lentil}. No such alternative strategy is available for substance nouns such as *mud*, however. As we argued above, nothing can be clearly judged to be an individuated unit of the relevant Ind_P type when P is a substance predicate. Hence, in contrast to vague granular nouns, no countable type is available.

6.2. Modeling Overlap in probM-TTR

Given our mereological enrichment of prob-TTR, defining overlap is fairly straightforward. In standard mereology, disjointness is a higher order property of predicate denotations. This can be interpreted as a type of types in TTR. To integrate this notion into a probabilistic system, we also introduce the threshold θ as a minimal degree of certainty for making a declarative type judgement. Types can then be judged to be disjoint relative to some, possibly context-sensitive threshold of certainty.

A type T is disjoint relative to a probability threshold θ ($Disj_\theta$):

IF	there is at least some a such that $p(a : T) \geq \theta$,
THEN	$T : Disj_\theta$ iff, for all a, b such that $p(a : T) \geq \theta$ and $p(b : T) \geq \theta,$ if $a \neq b$, then $a \cap b = \varnothing$,
ELSE	Undefined.

Since non-disjointness gives rise to multiple and inconsistent counting results, we suggest that numeral phrases are semantically restricted to applying to types that are disjoint.[8]

For prototypical count nouns such as *cat*, the related Ind type (Ind_{cat}) will be disjoint, and so countable. For granular nouns in languages that emphasize the Ind type for that noun (such as *lentil-s* and Ind_{lentil} in English), this type is also disjoint and so countable.

For neat mass ("fake" mass) nouns, the story is more complex. Recall that types of the form Ind_P are abbreviations for more complex types which include *inter alia* types of the form P_{Qual} and a quantitative function. Importantly, the quantitative function is a function, and so it should not be

[8] We do not have the space for a formal description of the semantics of numerals in probM-TTR here.

possible, across situations, for qualitatively identical entities to receive multiple quantity values as the output of the function. However, this is precisely what seems to be required for neat mass nouns. This is because, for example, a pestle and mortar sum should be able to count both as one item of kitchenware and as two. The implication of this is that there is no single quantitative function that can deliver this result. However, when a learner is learning to individuate furniture and kitchenware etc., she will be faced with such conflicting data. The strategy left available to resolve this conflict is to begin to track different quantitative functions for the same predicate such that on one, a pestle and mortar sum will receive a value of 1, and on the other a value of 2. In other words, given the way we see "fake" mass nouns (or their count counterparts) being used, we are forced to learn multiple, inconsistent schemes of individuation where each such scheme will be represented by a different quantitative function. Interestingly, this multiple quantitative function approach seems to describe something akin to 'overlapping simultaneously in the same context' in Landman (2011), and "counting contexts" as described in Rothstein (2010). We leave further investigation of this parallelism to further research.

For "fake" mass nouns, then, rather than being a single type Ind_P, there will be multiple types Ind_{P_i}, each one corresponding to a different individuation scheme. Given that each of these Ind_{P_i} types will relate to a single individuation scheme, then each will be disjoint. Therefore the countability of nouns such as the German *Küchengerät-e*$_{+\mathrm{C}}$ ('kitchenware') or the Finnish *huonekalu-t*$_{+\mathrm{C}}$ ('furniture') is gained merely by counting relative to a single type of the form Ind_{P_i}. However, allowing multiple Ind types for a single predicate is a matter of lexical choice, a language may encode a generalized Ind type formed as the join of all other Ind types for that predicate:

(19) $$Ind_{P_{generalized}} = Ind_{P_i} \vee Ind_{P_{i+1}} \vee ... \vee Ind_{P_n}$$

Generalizing comes at a cost however, since the generalized type will no longer be disjoint. As such, one cannot grammatically count the entities of this type due to the disjointness type restriction on the interpretation of number phrases. It is precisely this that we suggest models why the English *kitchenware*$_{-\mathrm{C}}$ and *furniture*$_{-\mathrm{C}}$ are MASS.

Substance nouns such as *mud, blood* are not countable for a slightly different reason. Given that we cannot be certain of what the individual entities for substance nouns are, substance Ind types are undefined for disjointness and so, for example, even a single Ind_{mud_i} type will not be disjoint. Of course, one way to boost the probability of something being of a type is to form a more general join type since the probability of an entity being of a join type increases with the number of joins (disjuncts). However, even if it were the case that for some a, $p(a : Ind_{mud_i} \vee Ind_{mud_{i+1}} \vee ... \vee Ind_{mud_n}) \geq \theta$, the resulting type would no longer be disjoint. Hence, substance nouns are not likely to be encoded as COUNT.

This strategy of forming join types and thereby allowing in multiple individuation schemas is perhaps an approximation of what Landman (2011) means by mess mass nouns having overlapping minimal generators. However, on our account we do not need to assume anything like a portion of mud, blood etc., that is minimal at a context.

7. Summary and Conclusions

By pursuing this approach we have made some headway into gaining a greater coverage of the puzzling variation one can observe in intra- and crosslinguistic mass/count variation. We have argued that one should adopt a dual-source approach to semantically modeling the mass/count distinction. In doing so, we have arrived at a formal characterization of four classes of nouns, each of which has distinct semantic properties. These are:

Noun type	Exemplars	Semantic Properties
Individual Object	*cat*, *chair*	Denote entities that are clearly (non-vaguely) classified under the predicate. There are no pressures from conflicting learning data to infer multiple quantitative schemas and so counting is felicitous for the associated Ind_P type.
Aggregates	*furniture*, *huone-kalu-t*$_{+\mathrm{C}}$ ('furniture' Finnish)	Denote entities that are clearly (non-vaguely) classified under the predicate. There is pressure to infer multiple quantitative schemas. A language may (i) lexically encode these nouns to a single Ind_{P_i} type at every context in which case, counting is felicitous for each associated Ind_{P_i} type; or (ii) form a generalized join type which is not disjoint in which case the counting is not felicitous.
Granular	*rice*, *lentils*	Denote entities that are not clearly (vaguely) classified under the predicate. There are no pressures to infer multiple quantitative schemas, but entities of the relevant Ind_P type are not clear cases of Ps. A language may (i) Lexically encode only the Ind_P type in which case counting is felicitous for this type; or (ii) Lexically encode the predicate $P(x)$. In this case, entities of the Ind_P type are not sufficiently clear cases of Ps and so cannot form the counting base.
Substance	*mud*, *blood*	Denote entities that are not clearly (vaguely) classified under the predicate. No entities are clearly of any of the relevant Ind_{P_i} types. There is nothing to count for any particular Ind_{P_i} type, and if anything is clearly of the more generalized join type, this type is not disjoint and so not defined for counting.

In effect, we have outlined strategies for ignoring vagueness and ignoring overlap. For Individual Object nouns, there is no overlap or vagueness to ignore and so these nouns are straightforwardly countable. For Aggregates, if overlap is ignored, the lexical item is countable. If overlap is not ignored, the lexical item will not be countable. For Granular nouns, if vagueness is ignored, the lexical item is countable. If vagueness is not ignored, the lexical item will not be countable. For Substance nouns, neither strategy for ignoring vagueness/overlap succeeds. This is due, mainly, to the difficulty in inferring any clear entities as individuals for these substances. Hence, our learning driven account derives four semantic classes of nouns. For two of these we expect stable count/mass encoding, for the other two, we expect to find a large amount of cross- and intralinguistic mass/count variation.

We plan to investigate the ways in which our probabilistic mereological approach could be extended. For example, we have restricted our discussion only to concrete nouns, and we shall investigate whether concepts such as vagueness and overlap can be translated into the abstract domain, and how. Furthermore, while argue for four semantic classes of nouns, we tacitly assume a binary morphosyntactic distinction between COUNT and MASS. However, a further line of research would be to investigate whether our four semantic classes form a correspondence with the four classes proposed in Grimm's (2012) scale of individuation.

References

Cheng, L. (2012). Counting and classifiers. In D. Massam (Ed.), *Count and Mass Across Languages*, pp. 199–219. OUP.

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese 174*, 99–149.

Cooper, R. (2012). Type theory and semantics in flux. In R. Kempson, T. Fernando, and N. Asher (Eds.), *Philosophy of Linguistics, Handbook of the Philosophy of Science*, pp. 271–323. Elsevier.

Cooper, R., S. Dobnik, S. Lappin, and S. Larsson (2014). A probabilistic rich type theory for semantic interpretation. *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics*.

Cooper, R., S. Dobnik, S. Larsson, and S. Lappin (2015). Probabilistic type theory and natural language semantics. *Linguistic Issues in Language Technology 10*(4), 1–43.

Fillmore, C. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences 280*(1), 20–32.

Grimm, S. (2012). *Number and Individuation*. PhD Dissertation, Stanford University.

Krifka, M. (1995). Common nouns: A contrastive analysis of english and chinese. In G. Carlson and F. J. Pelletier (Eds.), *The Generic Book*, pp. 398–411. Chicago University Press.

Landman, F. (2011). Count nouns – mass nouns – neat nouns – mess nouns. *The Baltic International Yearbook of Cognition 6*, 1–67.

Payne, J. and R. Huddleston (2002). Nouns and noun phrases. In G. Pullum and R. Huddleston (Eds.), *The Cambridge grammar of the English language*, pp. 323–524. Cambridge University Press.

Rothstein, S. (2010). Counting and the mass/count distinction. *Journal of Semantics 27*(3), 343–397.

Sutton, P. R. (2013). *Vagueness, Communication, and Semantic Information*. Ph. D. thesis, King's College London.

Sutton, P. R. (2015). Towards a probabilistic semantics for vague adjectives. In H.-C. Schmitz and H. Zeevat (Eds.), *Language, Cognition, and Mind*, pp. 221–246. Springer.