The Baltic International Yearbook of
Cognition, Logic and Communication

December 2016 Volume 11: *Number: Cognitive, Semantic and Crosslinguistic Approaches*
pages 1-45 DOI: http://dx.doi.org/10.4148/1944-3676.1110

PETER R. SUTTON
Heinrich Heine University Düsseldorf, Germany

HANA FILIP
Heinrich Heine University Düsseldorf, Germany

MASS/COUNT VARIATION: A MEREOLOGICAL, TWO-DIMENSIONAL SEMANTICS

ABSTRACT: We argue that two types of context are central to grounding the semantics for the mass/count distinction. We combine and develop the accounts of Rothstein (2010) and Landman (2011), which emphasize (non-)overlap at a context. We also adopt some parts of Chierchia's (2010) account which uses precisifying contexts. We unite these strands in a two-dimensional semantics that covers a wide range of the puzzling variation data in mass/count lexicalization. Most importantly, it predicts where we should expect to find such variation for some classes of nouns but not for others, and also explains why.

1. INTRODUCTION

In this paper, we identify two context indices that underpin the semantics of the mass/count distinction. One index is a counting context (Rothstein 2010), that either ensures the set of entities that count as 'one' in a noun's denotation is disjoint and suitable for counting, or, allows overlap leading to mass encoding (inspired by Landman 2011).

The other index is a precisification index (inspired by Chierchia 2010). At some precisifications, the set of entities that count as 'one' can be excluded from a noun's denotation leading to mass encoding. We combine these indices into a two dimensional semantics that not only models the semantics of a number of classes of nouns, but can also explain and predict the puzzling phenomenon of mass/count variation. Specifically, our account can explain why mass/count distribution patterns are uniform for some noun classes, but highly varied cross- and intralinguistically.

In Section 2, we introduce the theories of Rothstein (2010), Chierchia (2010), and Landman (2011) and analyze the role of context in each of them. We outline how Chierchia's account uses contexts as precisifications of vague expressions. We then present a perhaps surprising way of combining Rothstein's and Landman's approach into a single account, based on (non-)overlap at a context, which allows us then to offer greater coverage of the data.

In Section 3, we present the puzzle of mass/count variation. We identify five classes of nouns: PROTOTYPICAL OBJECTS, as exemplified by English nouns like *boy*, *chair*; SUBSTANCES, LIQUIDS, AND GASSES encoded by mass nouns like *mud*, *blood* in English; COLLECTIVE ARTIFACTS such as the mass noun *furniture* in English, but also the count noun like *huonekalu* ('(item of) furniture') in Finnish; HOMOGENOUS OBJECTS encoded, for instance, by the English count like *fence* and the mass noun *fencing*; and finally GRANULARS, which in English are exemplified by the mass noun *rice* or the count noun *lentils*. The first two classes contain nouns with strong tendencies for either count or mass lexicalization. The other three display cross- and intralinguistic variation in the encoding as count or mass. We argue that neither a vagueness based account, along the lines of Chierchia (2010), nor a synthesized Rothstein-Landman (non-)overlap based account can adequately account for all of the data.

In Section 4, we argue that a dual-source account that employs indices for contexts that govern overlap/non-overlap as well as precisifications allows to motivate the puzzling mass/count variation in the lexical encoding that we systematically find in the domain of the three classes of noun concepts identified above. In Section 5, we lay out our formal model for the dual source account, a two-dimensional semantics.

In Section 6, we show how our model can account for the distribution of count and mass nouns cross- and intralinguistically, and provide the lexical entries for paradigmatic nouns representing each noun class. This account, just like most other current accounts of the mass/count distinction, relies on a pretheoretical notion of individuation, i.e., what counts as 'one'. We will discuss this shortcoming in Section 7, and briefly mention how it may be overcome in a richer representational language such as that proposed in Sutton & Filip (2016c,d). The final section, Section 8, then provides an overall summary.

2. VAGUENESS, OVERLAP AND CONTEXT

2.1. *Context Sensitivity and Vagueness: Chierchia (2010)*

Chierchia (2010) claims that mass nouns are vague in a way in which count nouns are not. Count nouns have "stable atoms" in their denotation, that is, they have entities in their denotation that are atoms in every context. Mass noun denotations lack stable atoms. If a noun lacks stable atoms, there is no entity that is an atom in the denotation of the predicate at all contexts. In this sense then, mass nouns have only unstable individuals in their denotation. But counting is counting of stable atoms only. This motivates why mass nouns cannot be counted, i.e., straightforwardly occur with numericals, as in *?three muds* (unless they undergo a shift into a count interpretation).

Chierchia enriches mereological semantics with a form of supervaluationism wherein vague nouns interpreted at *ground contexts* have extension gaps (vagueness bands). Contexts then play the role of classical completions of a partial model in other supervaluationist formalisms such that at every (total) context, a nominal predicate is a total function on the domain.

Contexts stand in a partial order to one another such that if c' precisifies c ($c \propto c'$), then the denotation of a predicate P at c and at a world w is a (possibly not proper) subset of P at c' and w. That is to say that for an interpretation function $\mathcal{F}$:

$$\mathcal{F}(P)(c)(w) \subseteq \mathcal{F}(P)(c')(w) \qquad (1)$$

Let us briefly illustrate Chierchia's supervaluationist account with his

paradigm example of a mass noun *rice*. It is vague in the following way. It is not the case that, across all contexts, for example, a few grains, single grains, half grains, and rice flour dust, fall under the denotation of the predicate *rice*. But this means that such various quantities of rice are all in the vagueness band of rice, they fall in and out of the denotation of rice depending on the context. There may be some total precisifications of the ground context c, in which single grains are rice atoms. There may also be some c' such that $c \propto c'$, where half grains are rice atoms. There may also be some c'' such that $c' \propto c''$ in which rice dust particles are rice atoms. Most importantly, there is, therefore, no entity that is a rice atom at every total precisification of *rice*. In this sense, the denotation of *rice* lacks stable atoms, and so is vague. Counting is counting stable atoms, on Chierchia's account, but *rice* has no stable atoms in its denotation, what it denotes cannot be counted, which motivates its grammatical property.

2.2. *Counting Context Sensitivity: Rothstein (2010)*

Rothstein (2010) builds on Krifka (1989, 1995) in so far as she rejects Link's (1983) two-domain (atomic and non-atomic) ontology from which nouns take their denotation. However, she rejects Krifka's idea that the meanings of all (concrete) count nouns are to be analyzed in terms of the NU ("natural unit") extensive measure function. On Krifka's account, the entry for *cow* is as in (2):

$$\lambda n.\lambda x.[\mathbf{COW}(x) \wedge \mathbf{NU}(\mathbf{COW})(x) = n] \qquad (2)$$

In Krifka's theory, singular count nominal predicate meanings are derived with extensive measure functions (like **NU**) from cumulative predicate denotations (like $\mathbf{COW}(x)$), and they are *quantized*: entities they denote have no proper parts falling under the same predicate. A proper part of what is a cow, say, just a tail, is not describable by *a cow*. However, it was pointed out, at least since Zucchi & White (1996, 2001) (and Partee, p.c., among others) that there are count nouns that are not quantized, they do not come in natural units: e.g., *line*, *twig*, *fence*, *wall*. For instance, *fence* fails to be quantized, because it denotes entities which may have proper parts that also fall under the same predicate *fence*. Rothstein (2010) calls such nouns "homogenous" and provides a unified semantics that accommodates both such puzzling count nouns as well as prototypical count nouns like *cow*, *apple*, *boy*, which come in natural units, i.e., they are "naturally atomic", and therefore quantized, and hence are unproblematic for Krifka's theory.

Central to Rothstein's account is the notion of a *counting context*. Formally, counting contexts are subsets of the domain that are intersected with the denotation of a noun, but for "default contextual interpretations", contexts operate to remove overlap and yield a set of entities that count as 'one'. Rothstein introduces a typal distinction between count and mass nouns. Mass nouns are of type $\langle e, t\rangle$ (predicates of individuals). They may or may not be naturally atomic. Count nouns are indexed to counting contexts and are of type $\langle e \times k, t\rangle$ (predicates of indexed individuals), they are "semantically atomic". I.e., they denote individuals indexed for the context k in which they count as one, they are disjoint in a given context, and hence countable.

This means that "homogenous" count nouns like *fence* can denote non-overlapping, and so countable, "semantic atoms" at each context, even if these semantic atoms vary from context to context. For example, Rothstein's semantics captures that fencing enclosing a square field could plausibly count as four fences in one context and one fence in another.

Suppose the denotation of *fence* is the upward closure of $\{f_1, f_2, f_3, f_4\}$, and there are two contexts $k_1 = \{f_1, f_2, f_3, f_4\}$ and $k_2 = \{f_1 \sqcup f_2 \sqcup f_3 \sqcup f_4\}$. Applying these contexts to *fence* would yield the following sets of ordered pairs:

$$[\![\text{fence}]\!]_{k_1} = \{\langle f_1, k_1\rangle, \langle f_2, k_1\rangle, \langle f_3, k_1\rangle, \langle f_4, k_1\rangle\} \quad (3)$$

$$[\![\text{fence}]\!]_{k_2} = \{f_1 \sqcup f_2 \sqcup f_3 \sqcup f_4, k_2\} \quad (4)$$

This means that there would be four fences in context k_1 and one fence in context k_2. Although all count nouns, as a matter of semantic type, are indexed to counting contexts, not all count nouns are counting context sensitive (i.e. not all count nouns change their denotation from counting context to counting context). For example, a naturally atomic predicate such as *cow* would denote the same set (the set of individual cows) across all counting contexts.[1]

2.3. Overlap in a Context: Landman (2011, this volume)

The notion of what counts as 'one' is formalized in Landman (2011) as the members of a generator set for a predicate. A set of generators, **gen**(X), of the regular set X is a set of semantic building blocks for the predicate, namely, those entities that generate the denotation of a number neutral predicate via complete closure under sum $\sqcup$. There may be multiple sets that are generator sets. For example, if $X = \{a, b, a \sqcup b\}$, then two possible generator sets are $\{a, b\}$ and $\{a, b, a \sqcup b\}$.

For practical purposes, possibly relative to a context, a single generator set will be assigned to a noun predicate. For example, if the (number neutral) set CAT $= {}^*\{\text{tibbles}, \text{felix}, \text{garfield}\}$,[2] then the selected generator set will be $\{\text{tibbles}, \text{felix}, \text{garfield}\}$, and not, for example, $\{\text{tibbles}, \text{felix}, \text{garfield}, \text{tibbles} \sqcup \text{felix}\}$. In this sense, the generator set reflects the set of entities that count as 'one' with respect to a predicate.

If the elements in the (chosen) 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 a disjoint set. However, if generators overlap, as in the case of mass nouns, counting goes wrong, because more than one simultaneous counting outcome is possible. Landman provides 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*. A mass noun is neat if its intension at every world specifies a regular set whose set of minimal elements is non-overlapping. The entities in the generator set of neat mass nouns are what we would wish to count as 'one', such as single items of furniture. However, entities that count as one item of furniture, such as a vanity (consisting of a dressing table, a mirror, and possibly also a stool) may be sums of entities that each individually, and also their various sums, also count as one item of furniture "simultaneously in the same context" (Landman 2011, pp. 33-4). A noun is a mess mass noun if its intension at every world specifies a regular set whose set of minimal elements is overlapping. Landman suggests that what counts as minimal for mess mass nouns may vary with context (see his *salt*, *water* and *meat* examples).

There are many similarities, at least in spirit, between Landman 2011 and Landman's work in this volume. There are some important differences, however, which we briefly discuss here. The notion of a generator set is still central to Landman's new account. Lexical entries for nouns are represented, at least in part, as *i-sets*:

$$X = \langle \mathbf{body}(X), \mathbf{base}(X) \rangle : \mathbf{body}(X) \subseteq {}^*\mathbf{base}(X)$$

The central ideas are: bases generate bodies under $\sqcup$; bases form the notional set of entities for counting; counting is only possible for i-sets with disjoint bases:

X is *count* iff: $\mathbf{base}(X)$ is disjoint,
otherwise X is mass.

Landman (this volume) adds to this the definition of *Neat-mass i-sets*:

X is *neat* iff: $\mathbf{min}(\mathbf{base}(X))$ is disjoint, and
$\mathbf{min}(\mathbf{base}(X))$ generates $\mathbf{base}(X)$ under $\sqcup$,
otherwise X is *mess*.

This allows the definition of neat mass:

X is *neat mass* iff: $\mathbf{base}(X)$ is not disjoint, and
$\mathbf{min}(\mathbf{base}(X))$ is disjoint

An important feature of Landman's contribution to this volume is to move from what he refers to as *Mountain Semantics* to *Iceberg Semantics*. In much formal work on the mass/count distinction, it is assumed that there is a predicate-independent domain of atoms (Link 1983; Rothstein 2010; Chierchia 2010). Count denotations are 'built', mountain-like, from a set of atoms at the bottom of a domain. As discussed by Landman (this volume), this leads to the need for mapping atoms, that are also atoms with respect to a count predicate, to the entities which are parts of the entities represented by the atoms. (For example, from the atom a denoting Alex, to entities $b, c, d, \ldots$ denoting the parts of Alex.) In *Iceberg Semantics*, there is no assumed domain of predicate independent atoms. The same predicate e.g. *CAT* could be assigned a disjoint base set or an overlapping base set relative to the same domain of stuff.

An interesting question, one which we do not have space to elaborate on here, is to what extent iceberg semantics is connected to other formalisms that do not assume that sets of atoms in a universal domain

play a relevant part in defining the mass/count distinction such as the work of Krifka (1989, 1995, among others) and some work of our own (Sutton & Filip 2016b,c).

Iceberg semantics addresses an issue that is a little puzzling in Landman (2011), namely, the claim that mess mass nouns have minimal generators that overlap (albeit relative to a context). This involved the claim that, for nouns like *meat*, there are, in any context, entities that are minimal in the denotation of *meat*. This stipulation is no longer made (or required) in Landman's latest work (this volume) which is compatible with bases defined over atomless domains. (Such bases are provably overlapping.)

Landman (this volume) also emphasizes the connection to previous work in the verbal domain (Landman 1992, appx.). What counts as 'one', be it regarding entities or events, should be understood as a pragmatic notion that is sensitive to perspective. Landman further makes more explicit the connection between his theory and that of Rothstein (2010) in that bases can be disjoint or overlapping *relative to a context*. We make the connection between Landman's and Rothstein's work more precise in Section 2.4 where we propose a synthesis of Rothstein (2010) and Landman (2011).

2.4. *Overlap or Disjointness at a Context: A Synthesis of Rothstein (2010) and Landman (2011)*

In Rothstein's (2010) analysis, count nouns, which are counting context sensitive ("semantically atomic"), denote, in default contexts, non-overlapping entities that count as 'one'. In Landman's (2011) analysis, neat mass nouns like *kitchenware* and *furniture* denote an overlapping set of entities that count as 'one' (although their minimal generators do not overlap). Crucially, however, as (5-a)-(5-d) show, neat mass nouns are precisely a class of nouns that have cross- and intralinguistic count counterparts. Subscripts [+C] and [-C] indicate count and mass, respectively:

(5) a. furniture$_{[-C]}$; huonekalu-t$_{[+C,PL]}$ ('furniture', Finnish); meubels$_{[+C,PL]}$ ('furniture', Dutch); meubilair$_{[-C]}$ ('furniture', Dutch)

b. kitchenware$_{[-C]}$; keittiöväline-et$_{[+C,PL]}$ ('kitchenware', Finnish); Küchengerät-e$_{[+C,PL]}$ ('kitchenware', German)

c. footwear$_{[-C]}$; jalkine-et$_{[+C,PL]}$ ('footwear', Finnish); Schuhwerk$_{[-C]}$ ('footwear', German)
d. jewelry$_{[-C]}$; koru-t$_{[+C,PL]}$ ('jewelry', Finnish)

As (6-a)-(6-c) illustrate, context-sensitive count nouns like *fence* in English are precisely a class of nouns that have cross- and intralinguistic mass counterparts.

(6) a. fence$_{[+C]}$; fencing$_{[-C]}$
b. wall$_{[+C]}$; Mauer$_{[+C]}$ ('wall', German); walling$_{[-C]}$; Gemäuer$_{[-C]}$ ('walling/ masonry', German)
c. bush$_{[+C]}$; Busch$_{[+C]}$ ('bush', German); shrubbery$_{[-C]}$; Gebüsch$_{[-C]}$ ('shrubbery', German)

Let us consider an example of each kind: *furniture*$_{[-C]}$ versus *huonekalu-t*$_{[+C]}$ ('furniture' Finnish) and *fence*$_{[+C]}$ versus *fencing*$_{[-C]}$. One might argue that if you have a neat mass noun (*furniture*$_{[-C]}$, *kitchenware*$_{[-C]}$) in Landman's terms that has a count counterpart, then that count counterpart is analyzable as "semantically atomic", or a counting context sensitive count noun, on Rothstein's theory (*huonekalu-t*$_{[+C]}$). Vice versa, if a counting context sensitive count noun *fence* in Rothstein's terms has a mass counterpart *fencing*, then that mass counterpart will be analyzable as a neat mass noun on Landman's theory. Indeed Landman has claimed that nouns such as *fencing* are neat mass nouns (Landman p.c.). This general argumentation strategy is shown in Figure 1.

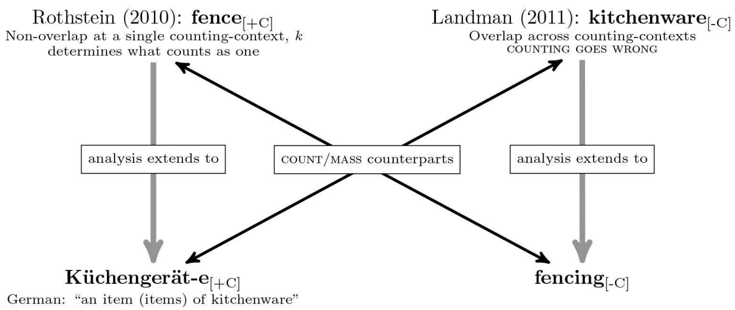


Figure 1: Connecting Rothstein (2010) and Landman (2011)

furniture versus *huonekalu-t*: On Landman's analysis, for *furniture* we get overlapping entities that count as 'one' simultaneously in the same counting situation, i.e., we have many different possible partitions of the domain and different ways of counting. For example, a vanity, but also the table and mirror that compose it, may each count as one item of furniture. However, in Finnish, it is not the case that *only* the vanity counts as 'one' *or only* the mirror and table each count as 'one'. What one counts for *huonekalut* is a context dependent matter. From the perspective of Rothstein's analysis, the count counterparts of neat mass nouns are counting context sensitive, and hence analyzable as semantically atomic.

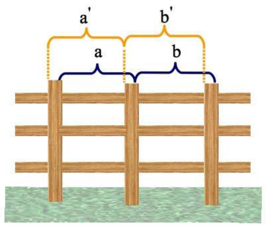
fence versus *fencing:* On Rothstein's analysis, *fence* is counting context sensitive (Section 2.2). That means that if one takes all possible contexts, then there is overlap between the members of the set of entities that count as 'one' in one context, and the members of the set of entities that count as 'one' in another context. However, on the assumption that *fence* is the number neutral property for both *fencing* and *fence*,[3] then two conditions must hold for *fencing* to be a neat mass noun in Landman's (2011) terms: (i) some of the entities in the denotation of *fence/fencing* are minimal (i.e. minimal generators); (ii) the same entities are non-overlapping.

Let us now assess conditions (i) and (ii). With respect to (i), on the face of it, it might seem that there are no clear minimal fence entities. If a 2m stretch of fence counts as fence, then surely 1cm can't make a difference, so a 1.99m stretch also counts as *fence*, therefore *fence* is sorites susceptible and so, arguably, vague. However, if *fence* is vague, then there are no clear minimal *fence* entities. The question is whether vagueness, characterized as sorites sensitivity, really does mean that *fence/fencing*-like nouns do not have minimal entities in their denotations. The reason for thinking that vagueness, in this sorites susceptible sense, is not relevant, is that many highly prototypical count nouns such as *girl* can be pushed down a soritical forced march. For example, take a clear case for the predicate *GIRL*, call her Billie. One single human cell can't ever make a difference between applying *GIRL* and applying $\neg$*GIRL*, and so, *Billie minus one cell* must still be a girl. Absurd conclusion: a single cell, any single cell that was originally part of Billie, is also a girl. In this sense, sorites susceptibility is a problem for Chierchia's (2010) account. By playing with precisifications, one can upset the stable atomicity of virtually any basic lexical predicate (excepting, perhaps, scientific terms of art or mathematical expressions).

The relevant 'vagueness' for nouns such as *rice*, we will argue, is that across precisifications, entities that we would intuitively count as 'one' (the individual grains of rice) are not in the denotation of *RICE* across all precisifications. This is not so for *girl*. Our prototypical clear case for *GIRL*, Billie, will be in the denotation of *GIRL* across all precisifications, hence *girl* is not like *rice*.[4]

Now take *fence* and *fencing*. We know from Rothstein 2010 that these nouns are sensitive to a different kinds of context, namely *counting contexts*. If we keep the counting context fixed, then *fence* and *fencing* behave like *girl* and not like *rice*. Relative to a specific counting context, there will be clear cases of entities that fall under the predicate *FENCE*. Now, even if we modulate the precisification context, *those* entities will not cease to be in the extension of *FENCE* just as Billie would not cease to be in the extension of *GIRL*. Nouns like *fence* and *fencing* have minimal entities to the same extent that nouns such as *girl* do. Therefore condition (i) holds.

Let us now examine condition (ii). We have just argued that there is some set of minimal fence entities. The question now is, would these units be non-overlapping? Here we hit a problem. As Figure 2 shows, this condition on a minimal fence unit might at different counting contexts yield an overlapping set of entities. Suppose, as in Figure 2, that $\{a, b\}$ are minimal at c_1, but $\{a', b'\}$ are minimal at c_2. The trouble is that a overlaps with b' and a' with b with respect to the middle upright stake.



Portion of fencing divided into minimal entities at two different contexts.

Figure 2: Overlapping minimal fencing entities

However, notice that at each counting context (or "variant" in Landman's sense), the number of semantic atoms (minimal generators) is the same: For the case in Figure 2, there are two fences in c_1 and two fences in c_2. This means that *fencing* is not, properly speaking a neat mass noun, since it does not have non-overlapping minimal generators. All is not lost, however. We now show that, using Rothstein's notion of a counting context (or Landman's notion of a variant), we can capture the spirit of Landman's account albeit in slightly different terms.

Recall that, for Landman, what makes counting go wrong is that the number of entities at each variant of the generator set may differ. With no single answer to the question 'How many?', counting goes wrong. However, one way of understanding and redefining Landman's notion of *neat* is to say that a mass noun is neat iff the set of *minimal generators* (non-overlapping minimal entities) returns the same cardinality at all variants. This would allow fencing to be neat, but would also allow both *fencing* and nouns like *furniture* to be uncountable, and hence mass, since, although the minimal generator set would return the same cardinality at all variants, the generator set would not. Put another way, the set of entities that count as 'one' varies across counting contexts, not only in membership, but also in cardinality.

In summary, given a slight adjustment of the definition of *neat mass* that is nonetheless in the spirit of Landman (2011), we have argued that count counterparts of neat mass nouns are counting context sensitive and that mass counterparts of counting context sensitive count nouns are neat mass nouns. Thus we have proposed a unified Rothstein-Landman account for the analysis of both neat mass nouns and count nouns that are counting context sensitive. The theme that unites the accounts of Landman and Rothstein is (non-)overlap at a context. Rothstein's (default) counting contexts remove overlap and yield count nouns. Landman's contexts (involving a multiplicity of variants (i.e. a multiplicity of counting contexts in Rothstein's sense)) allow overlap and so yield mass nouns. In Section 5, we formally define these two notions of context in terms of each other.

3. THE CHALLENGE OF MASS/COUNT VARIATION

In Section 2.4, we provided some data that neat mass nouns have counting context sensitive count counterparts within a single language or across different languages, and that counting context sensitive count nouns have neat mass counterparts. Another group of nouns which have been observed to display mass/count variation are what we dub GRANULARS. This group, in English, includes *oats*, *lentils*, *rice*, *beans*. Examples of cross- and intralinguistic variation are given in (7-a)-(7-c).

(7) a. lentil-s$_{[+C,PL]}$; linssi-t$_{[+C,PL]}$ ('lentils', Finnish); lešta$_{[-C]}$ ('lentil', Bulgarian);
čočka$_{[-C]}$ ('lentil', Czech)
b. oat-s$_{[+C,PL]}$; oatmeal$_{[-C]}$; kaura$_{[-C]}$ ('oat', Finnish); kaurahiutale-et$_{[+C,PL]}$ ('oatmeal' lit. oat.flake, Finnish)
c. bean-s$_{[+C,PL]}$; pavu-t$_{[+C,PL]}$('beans', Finnish); bob$_{[-C]}$('bean', Bulgarian)

Other nouns display considerable uniformity with respect to whether they are count or mass within a single language and across languages, however. On the one hand, prototypical objects, examples of which in English are *cat*, *chair*, *table*, *car*, are highly probably lexicalized as count nouns, whereas substances, English examples of which are *mud*, *blood*, *water*, *air*, are highly probably lexicalized as mass nouns intra- and crosslinguistically. These prototypical cases are not problematic for any of the theories we have discussed. For example, entities like cats are not overlapping (Landman 2011), are indexed to counting context (Rothstein 2010), and are not vague with respect to their countable units (Chierchia 2010). Substances are overlapping (mess mass) (Landman 2011), not indexed to counting context (Rothstein 2010), and vague (Chierchia 2010). We summarize the mass/count variation data in Table 1.

We will now argue that variation in mass/count encoding poses a major challenge for a vagueness-based account such as Chierchia's (2010), and for disjointness/overlap based accounts such as Rothstein (2010) and Landman (2011). Then in Section 4, we put forward a dual-source account which overcomes many of these difficulties.

Table 1: Mass/count variation data grouped into classes

Noun Class	Examples
Prototypical Objects	chair$_{[+C]}$; tuoli$_{[+C]}$ ('chair' Finnish); Stuhl$_{[+C]}$ ('chair' German) dog$_{[+C]}$; koira$_{[+C]}$ ('dog' Finnish); Hund$_{[+C]}$ ('dog' German) boy$_{[+C]}$; poika$_{[+C]}$ ('boy' Finnish); Junge$_{[+C]}$ ('boy' German)
Collective Artifacts	furniture[-C]; huonekalu-t$_{[+C,PL]}$ ('furniture', Finnish); meubel-s$_{[+C,PL]}$ ('furniture', Dutch); meubilair$_{[-C]}$ ('furniture', Dutch) kitchenware[-C]; keittiöväline-et$_{[+C,PL]}$ ('kitchenware', Finnish); Küchengerät-e$_{[+C,PL]}$ ('kitchenware', German) footwear[-C]; jalkine-et$_{[+C,PL]}$ ('footwear', Finnish); Schuhwerk$_{[-C]}$ ('footwear', German) jewelry$_{[-C]}$; koru-t$_{[+C,PL]}$ ('jewelry', Finnish)
Homogenous Objects	fence$_{[+C]}$ vs. fencing$_{[-C]}$ wall$_{[+C]}$; Mauer$_{[+C]}$ ('wall', German); walling[-C]; Gemäuer$_{[-C]}$ ('walling/masonry', German) bush$_{[+C]}$; Busch$_{[+C]}$ ('bush', German); shrubbery$_{[-C]}$; Gebüsch $_{[-C]}$ ('shrubbery', German)
Granulars	lentil-s$_{[+C,PL]}$; linssi-t$_{[+C,PL]}$ ('lentils', Finnish); lešta$_{[-C]}$ ('lentil', Bulgarian); čočka$_{[-C]}$ ('lentil', Czech) oat-s[+C,PL]; oatmeal[-C]; kaura$_{[-C]}$ ('oat', Finnish); kaurahiutale-et$_{[+C,PL]}$ ('oatmeal' lit. oat.flake, Finnish) bean-s$_{[+C,PL]}$; pavu-t$_{[+C]}$ ('beans', Finnish); bob$_{[-C]}$ ('bean', Bulgarian)
Substances, Liquids & Gasses	mud$_{[-C]}$; muta$_{[-C]}$ ('mud' Finnish); Schlamm$_{[-C]}$ ('mud' German) blood$_{[-C]}$; veri$_{[-C]}$ ('blood' Finnish); Blut$_{[-C]}$ ('blood' German) air$_{[-C]}$; lenta$_{[-C]}$ ('air' Finnish); Luft$_{[-C]}$ ('air' German)

3.1. Vagueness-based accounts and mass/count variation

It should be acknowledged that Chierchia (2010) states the need for more than one source for the mass/count distinction. The primary evidence for this are fake mass nouns which are not vague, but are nonetheless mass nouns. In fact, Chierchia (2010) offers a number of different explanations for the mass/count lexicalization of different nouns. We argue in Section 4 for a more parsimonious explanation than that given by Chierchia (2010).

We call neat mass nouns such as *furniture* and count counterparts such as *huonekalu* ('(item of) furniture', Finnish) collective artifacts. As Chierchia suggests, collective artifacts are problematic for a vagueness-based account of the mass/count distinction. This is because they are not vague in the sense of having unstable individuals, since, for example, items of furniture such as chairs are in the denotation of *furniture* at all precisifications. Chierchia's explanation for why collective artifacts can be mass is that:

> "fake mass nouns arise as a 'copy cat' effect from the way in which number marking languages react to unstably atomic nouns. Since listing a potentially count noun as a singleton property is essentially a matter of lexical choice, we expect there to be variation, even across closely related languages or language families on this score, which has, in fact, been often observed in connection with fake mass nouns." (Chierchia 2010, p. 139)

However, even if were to accept this mode of explanation, and were to accept that we should expect variation, even across closely related languages or language families, this does not necessarily explain why we should find variation within a single language as with the Dutch *meubel*$_{+C}$ and *meubilair*$_{-C}$ ('furniture'). We may ask, why should not other, prototypical, count nouns be similarly *copy catted*? An explanation Chierchia gives is that copycatting may be driven by a "lack of interest in the atoms" (Chierchia 2010, p. 139, fn. 48) which "might account for the superordinate flavor of nouns like *furniture*, as for the fact they tend to be 'superordinate' category" (Chierchia 2010, p. 139, fn. 48). However, plural count nouns can be collective too (*vehicles, vegetables*), and it is not clear that they display less 'lack of interest' in their atoms. Finally, Chierchia's copycat addendum predicts that the phenomena of fake mass nouns should not be found in classifier languages. However, as Sutton & Filip (2016b) argue, based on evidence from Cheng (2012), this does not appear to be borne out for *jiājù* ('furniture', Mandarin).

Furthermore, similar concerns arise for mass/count counterparts such as *fencing* and *fence*. Chierchia (2010, pp. 112-3) claims that count nouns such as *heap*, *mountain*, and *fence* are not vague in the relevant sense. Furthermore, nouns such as *fence* do not display a similar 'lack of interest' in atoms, so we should not expect mass counterparts to be generated by copy catting any more than for prototypical count nouns such as *cat*. Nonetheless, we do find mass terms such as *fencing*, *hedging*, but do not tend to find mass counterparts for nouns such as *cat*.[5]

A more parsimonious account would be one in which, in addition to vagueness, a second semantic factor could explain mass/count variation both in collective artifacts such as *furniture* and in *fence*-like nouns. We give such an account informally in Section 4, and formalize it in Sections 5 and 6. A second challenge to a vagueness-based account is that there are vague nouns that are not mass nouns. Examples of these are (7-a)-(7-c). If *lentil* is not-vague and therefore count, then *lešta*$_{[-C]}$ ('lentil', Bulgarian) and *čočka*$_{[-C]}$ ('lentil', Czech) should be count too. Alternatively, if *lešta*$_{[-C]}$ ('lentil', Bulgarian) and *čočka*$_{[-C]}$ ('lentil', Czech) are vague and therefore mass, then *lentil* should be mass too. Chierchia's response to this variation is:

> "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. Clearly, for example, defining smallest samples for (non commercially packaged) liquids is harder than for granular substances." (Chierchia 2010, p. 140)

Whatever the merits of this tentative explanation turn out to be for crosslinguistic mass/count variation, it does not account for intralinguistic cases. Take the case of *oat-s*$_{[+C,PL]}$, *kaura*$_{[-C]}$ ('oat', Finnish), *oatmeal*$_{[-C]}$, *kaurahiutale-et*$_{[+C,PL]}$ (lit. 'oat.flake', Finnish). In British English, at least, *oatmeal* is commonly used to refer to the rolled oats used for making porridge, likewise with *kaurahiutaleet* in Finnish. As these data show, the mass/count encoding of the basic oat grain as opposed to the processed product (the meal/flakes) is switched between English and Finnish. This allows an argument against a vagueness-only view. Either the oats and the flakes/meal are, as a 'smallest sample', as available as each other, or one is more available than the other. If the former, then we should not expect mass/count variation in either English or Finnish. If the latter, we should expect mass/count variation in either English or Finnish, but not both. Whether the same or different, Chierchia's vagueness-based approach plus story of crosslinguistic "smallest available sample" variation is insufficient to account for the data.

In conclusion, there are nouns with unstable entities in their denotations that are sometimes lexicalized as either mass or count cross- and intralinguistically. There are also nouns with stable entities in their denotations that are sometimes lexicalized as either mass or count cross- and intralinguistically. Together, these data can be used to form compelling arguments that an account based on vagueness alone is insufficient to form a basis for the mass/count distinction.

3.2. Disjointness/Overlap-based accounts and mass/count variation

In Section 3.1 we argued that a vagueness-only based account of the mass/count distinction faces two challenges given mass/count variation data. First, homogenous objects and collective artifacts are not-vague but are mass as well as count. Second, granulars are vague nouns that are count as well as mass.

On the face of it, a hybrid Landman-Rothstein account based on (non-)overlap in context (Section 2.4) fairs better than vagueness-based approaches. Homogenous objects (fence) and collective artifacts can overlap in context (and be mass), or be disjoint in context (and be count).

In Section 2.4, we argued that the count counterparts of Landman's neat mass nouns can be interpreted as counting context-sensitive nouns on Rothstein's account, and that mass counterparts of counting context sensitive nouns can be interpreted on Landman's account as neat mass. However, to establish that nouns like *fencing* are in the same broad class as nouns like *furniture*, we had to make a minor adjustment to Landman's definition of neat mass, namely that the set of minimal generators has the same cardinality at all variants.

For a semantic unification of these accounts, one must define Landman's notion of context in which the set of entities that count as 'one' overlap *simultaneously in the same context* differently from Rothstein's default counting contexts which return disjoint sets. Our proposal in Section 5 does this by defining Landman's notion of context in terms Rothstein's default contexts. (In brief, the interpretation of a predicate at a Landman context will be the sum of the interpretation of that predicate over all Rothstein contexts.)

Assuming one has two notions of context (one in the Landman sense and one in the Rothstein sense), then, on a (non-)overlap based ac-

count, collective artifacts and homogenous objects may be classed as mass when interpreted relative to Landman contexts and as count when interpreted relative to Rothstein contexts. Thus a synthesized Rothstein-Landman account could accommodate the mass/count variation data for homogenous objects and collective artifacts from Table 1. This is a clear advantage over a vagueness-only based account for which nouns in these classes were problematic.

Nonetheless, nouns in the granulars class still pose a challenge for even a hybrid Rothstein-Landman account. The difficulty is that no matter how one looks at it, the set of entities that count as 'one' for granulars does not overlap. Overlap was present for homogenous objects, because if one has two (relevantly) similar homogenous objects that count as one (such as two portions of fence), then one can append one to the other and get a sum entity that also (in at least some contexts) counts as one fence too. Overlap was present for collective artifacts, because if one has two objects that count as 'one', such as a pestle and a mortar for kitchenware, such that these two objects have a joint function qua item of kitchenware (holding and grinding things), then the sum of these entities could well also count as 'one' with respect to kitchenware.

The same does not hold for granulars. For example, single lentils count as one with respect to the predicate *lentils*, however, there are no contexts in which sums of lentils count as *one lentil*. Hence, there cannot be overlap going 'upwards' in the lattice from single lentils as we found with both homogenous objects and collective artifacts. However, nor can one locate overlap going either 'downwards' or 'horizontally' in the lattice. Parts of single lentils do not count as 'one' lentil, so downwards overlap is ruled out. Furthermore, single lentils do not themselves overlap, hence there cannot be 'horizontal' overlap either.

Granulars can be lexicalized as count (e.g. *lentils*$_{[+C,+PL]}$) or as mass (e.g. *čočka*$_{[-C]}$ ('lentil', Czech)), but, as we have seen, the set of entities that count as 'one' for such nouns is non-overlapping at all contexts. Non-overlap within and between contexts means that granulars should all be count nouns. However, we often find mass nouns in this class, therefore (non-) overlap at a context cannot be the whole story when it comes to the mass/count distinction.

In summary, neither a vagueness-only account nor an account only based on (non-)overlap at a context can accommodate the full range of cross- and intralinguistic data, therefore, either an entirely new source for the distinction may be at play, or something more complex than (non-)overlap or vagueness alone may be at work. We opt not to throw the baby out with the bathwater and so, choose the latter. In Section 4 we argue that these data can be accommodated with a dual-source account (vagueness *and* (non-)overlap at a context).

4. A DUAL SOURCE ACCOUNT

In this section, we argue that two different indices of evaluation affect mass/count encoding and so should feature in a semantics for concrete nouns. One index is the set of counting contexts in the sense of Rothstein (2010), the other is the set of permissible precisifications in the sense of Chierchia (2010). We will show how all of the above cross- and intralinguistic mass/count variation data can be accommodated by an account that uses both of these indices. We formally present these ideas in Section 6.

Given that, in Section 2, we argued that homogenous objects (*fence*) and collective artifacts (*furniture*) can be subsumed under the same broad class of nouns, we now have four main classes to analyze: (i) prototypical objects; (ii) homogenous objects & collective artifacts; (iii) granulars; (iv) substances, liquids and gasses. In the following Sections (4.1-4.4) we discuss the concept of counting as 'one' for each of these classes. We assess each class in terms of sensitivity or insensitivity to counting context, and to precisification context.

We show that when a set of entities that count as 'one' for a predicate is insensitive to both indices, the predicate noun is stably count. When there is no clear set of entities that count as 'one', it is not determinate whether the predicate is sensitive or insensitive to counting context. This leads to predicates being stably mass. We further show that when a set of entities that count as 'one' for a predicate is insensitive to one, but not both indices, then we should expect mass/count variation.

4.1. Prototypical Objects

Nouns in this class tend to be 'naturally atomic' in the sense of Rothstein (2010), or come in 'natural units' in the sense of Krifka (1989). The set

of entities that count as 'one' are these natural units: single chairs, boys, cats for *chair*, *boy*, *cat*, respectively. The set of entities that count as 'one' also does not greatly vary depending on the precise interpretation of, for example, *cat*. That is to say that, for some set of entities (such as the clear cases of cats), these individuals will be in the set of things that count as *one cat* at every index of interpretation for *cat*.

Notice that this way of thinking differs from Chierchia's (2010). In Chierchia (2010), there is no concept of what counts as 'one' independently of the set of entities that are atoms in a predicate's denotation across all permissible precisifications (the stable atoms). We go a slightly different route. We take what counts as 'one' as a more pre-theoretical notion, and then assess whether or not different precisifications include or exclude individuals that count as 'one' from the noun's denotation.

For prototypical objects, there is a set of countable entities on any and across all precisifications. This characteristic of a predicate is what we define as insensitivity to precisification contexts with respect to the entities in the predicate's denotation that count as 'one'.

Prototypical objects are insensitive to counting contexts too. Relative to some precisification, there is no change in the number of countable entities as the counting context varies. So, even on an account that integrates both variation in interpretation (precisifications) and variation in the way one individuates (counting contexts), nouns in this class are countable insofar as at any (or all) precisifications and at any (or all) counting contexts, there is a non-empty set of non-overlapping, and thereby countable set of entities available.

4.2. Collective Artifacts and Homogenous Objects

Here we adopt the viewpoint of the synthesized Rothstein-Landman framework we outlined in Section 2. The set of entities that counts as 'one' in this class is sensitive to counting context. For example, depending on the counting context, a square of fencing may count as *one fence*, or as, say, *four fences* (Rothstein 2010). Also, depending on the counting context, what counts as *one meubel* ('(item of) furniture', Dutch) or what counts as *one item of furniture* will vary from context to context. For example, a vanity may count as 'one' in some contexts, but as two (a table and a mirror) in others.

Just like prototypical objects, the set of entities that count as 'one' for nouns in this class are not sensitive to precisification contexts in the sense that there are some entities (fences, pieces of furniture) that count as 'one' no matter how one varies the precisification of *fence/furniture* (the set of entities that count as 'one' is not empty).

Also like prototypical objects, the relative lack of vagueness for nouns in this class means that at any or across all precisifications, there is a set of entities that form a notional counting base. However, the sensitivity of these nouns to counting contexts means that whether or not one is left with a non-overlapping countable set is determined by whether one interprets the nouns relative to a Rothstein-type non-overlapping context or a Landman-type overlapping context. Therefore, we should expect possible variation in the mass/count encoding of these nouns. If they are indexed to counting contexts *à la* Rothstein, nouns in this class will be lexicalized as count. If indexed to contexts *à la* Landman, nouns will be lexicalized as mass.

4.3. Granulars

Nouns in this class denote stuff made up of grains, granules, flakes and the like. Normally, that their denotations are so constituted, is clearly perceptible and well known for those who have grasped the sense of these nouns. Furthermore, for count nouns in this class, the single grains, granules, flakes etc., are what are denoted and explicitly available for counting. We take this as a strong reason to identify the individual grains, granules, flakes etc. as the entities that count as 'one' for nouns in this class. For example, for both *lentils*$_{[+C,PL]}$ and *čočka*$_{[-C]}$ ('lentil', Czech), what counts as 'one' are single lentils.

Nouns in this class are precisification context sensitive; their denotation changes depending on the precisification. Crucially, the context in which these predicates are precisified can sometimes exclude the single grains, granules or flakes etc. from their denotations. For example, on less precise interpretations of *rice*, and *lentils*, single rice grains/lentils will be of too small a quantity to be in the denotation of *rice*, *lentils*, respectively. In more precise contexts, they will be in the denotation. That means that the set of entities that intuitively count as one can be excluded from the denotations on some precisifications. This is represented graphically in Figure 3, where four possible precisifications of

rice are displayed ($\pi_1, \pi_2, \pi_3, \pi_4$).

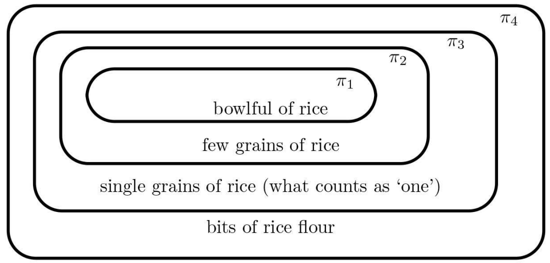


Figure 3: Precisifications of *rice*. Not all precisifications include the entities that count as 'one'

As we argued in Section 3.2, just like prototypical objects, nouns in this class are not counting context sensitive. That is to say that, single lentils, rice grains etc. in the denotation of rice, lentils, respectively, count as 'one' at every counting context. This means that whether they are interpreted relative to any counting context, including a Landman-type context (which allows a multiplicity of simultaneous variants), there will be a non-empty set of non-overlapping countable entities, provided these entities are in the denotation of the noun at all.

Whether these entities are in the denotation of the noun at all is where precisification context sensitivity comes in. For example, there are precisifications on which rice contains single grains in its denotation (π_3 and π_4 in Figure 3) and precisifications that exclude the single grains (π_1 and π_2 in Figure 3). A similar thing could be said of *lentils* (*pace* Chierchia 2010). When predicates such as *RICE* and *LENTIL* are interpreted relative to the intersection of interpretations across all contexts (what we will call the *null precisification context*), the set of entities that count as 'one' is empty, because the single grains, lentils etc. are excluded from the denotations of these predicates at this context.

This means that, depending on how strictly the noun is interpreted by default, countability is affected. If interpreted on all admissible precisifications, the individual grains, flakes etc. will be excluded. This results in nouns being lexicalized as mass because the single grains are made unavailable as a counting base. If, by default, a noun is interpreted in a more precise manner, then the individual grains, flakes etc., may be included, these grains, flakes etc. will form the disjoint counting base for the noun, and so the noun can be lexicalized as count.

4.4. Substances, liquids and gasses

Nouns in this class do not, intuitively, have any entities that clearly (or even, perhaps, at all) count as 'one'. There is nothing that we can identify either on perceptual or functional grounds that forms individuals in their denotations. If we take this basic intuition seriously, then when we come to represent the semantics of these nouns we will leave the set of entities that count as 'one' for e.g., *mud*, *blood*, and *air* as undefined. This is a departure from Landman (2011) who argues that mess mass nouns have contextually provided (overlapping) minimal entities.

With respect to what demarcates substances, liquids and gasses from other (concrete) nouns, we invoke a distinction from Soja et al. (1991) who found that pre-linguistic infants can distinguish substances from objects. We assume in the following that this prelinguistic conceptual distinction determines whether or not a concrete noun predicate/concept has a defined individuation schema. A lack of such an individuation schema implies a single, stable way to conceptualise these referents, namely as mass.

What we propose is that, when there is no defined individuation schema for a predicate, the counting base in the corresponding lexical entry will be the same as the number neutral predicate.[6] As we will show, this makes such nouns undefined for counting context and precisification context sensitivity.

In Section 5, we develop our two-dimensional formal framework and in Section 6, we formally represent the above classes and their sensitivity/insensitivity to precisification contexts and counting contexts in order to account for patterns that have been observed in mass/count variation.

5. FORMAL MODEL: A TWO-DIMENSIONAL SEMANTICS

In this section we outline our formal representation for the above informal observations. That is to say, we give a model that can represent

both Rothstein's counting contexts that enforce disjointness, Landman's contexts that allow overlap, and a range of precisifications.

The model is two-dimensional, since expressions are evaluated along two indices: counting contexts (including both Rothstein's and Landman's contexts) and precisification contexts.

5.1. Model

Our model will allow evaluations across two dimensions of contexts, counting contexts and contexts that determine precisifications. Models are tuples $\langle \mathcal{D}, \mathcal{I}, \mathcal{W}, \mathcal{C}, \Pi \rangle$. $\mathcal{D}$ is the domain (entities and truth values). $\mathcal{D}_{\langle e \rangle}$ the domain of entities is structured as a complete Boolean algebra minus the bottom element. $\mathcal{I}$ is the interpretation function (we will mostly use $[\![\ .\ ]\!]$). $\mathcal{W}$ is the set of worlds, however, for simplicity we will suppress intensions here, and describe the semantics of expressions in purely extensional terms. $\mathcal{C}$ is the set of counting contexts (details below). Π is the set of precisifying contexts (details below). The standard definition of disjoint and overlapping (not disjoint) sets is:

$$DISJ(X) \leftrightarrow \forall x, y \in X[x \neq y \rightarrow x \sqcap y = 0] \tag{8}$$

$$OVERLAP(X) \leftrightarrow \exists x, y \in X[x \neq y \wedge x \sqcap y \neq 0] \tag{9}$$

Counting contexts ($c_i \in \mathcal{C}$) map sets onto maximally disjoint subsets. For a set X, there is a constraint on counting contexts:

$$X_{c_i} = \{Y : Y \subseteq X, DISJ(Y), \forall x \in X\ \exists y \in Y[x \sqcap y \neq 0]\} \tag{10}$$

This means that, when applied to disjoint sets, counting contexts are the identity function. When applied to non-disjoint sets, each counting context will yield a maximally disjoint subset of that set.

This means that contexts $c_i \in \mathcal{C}$ will behave like Rothstein's *default contexts* (and Landman's *variants*). Interestingly, we can then define Landman's contexts (that allow overlap) in terms of a sub-valuation on the set of contexts. Subvaluations are applied here to capture Landman's (2011) idea that furniture-like nouns *overspecify* the set of entities that count as one. Subvaluating is a means of overspecifying. Formally speaking, subvaluations can be defined in terms of unions:

$$X_{c_0} = \bigcup_{c_i \in \mathcal{C}} X_{c_i} \tag{11}$$

This approximates Landman's (2011) notion of context, since it will include the union of all maximally disjoint subsets of X. For a set X, X_0 will not be disjoint unless X was already disjoint. This means that the null counting context, c_0, will be the identity function on sets of entities that count as 'one'. Disjoint sets map to the same disjoint set. Overlapping sets map to the same overlapping set.

Precisifying contexts ($\pi_i \in \Pi$) map sets that do not completely partition the domain (has an extension gap) into sets that do completely partition the domain. Like Chierchia (2010) and other supervaluationists, precisifying contexts form a partial order. $\pi \propto \pi'$ (π' precisifies π), relative to a set X:

$$\pi \propto_X \pi' \leftrightarrow X_\pi \subseteq X_{\pi'} \tag{12}$$

This means that precisifying does not 'shrink' the membership of a set, but may expand it. We follow a characterization of vagueness-based accounts made by Landman (2011), namely as involving *under-specification* as opposed to overspecification. Indeed, this is what supervaluations provide. Supervaluations capture the idea that nouns like *rice* underspecify their denotations insofar as there are amounts of rice, such as a single grain, that do not count as *rice* in every precisifying context. Formally speaking, supervaluations can be defined in terms of intersections:

$$X_{\pi_0} = \bigcap_{\pi_i \in \Pi} X_{\pi_i} \tag{13}$$

Predicates will then be interpreted relative to counting contexts and precisifications. However, in some cases (to be outlined below) the counting and precisifying contexts will be the null counting context c_0 or the null precisification context π_0.

5.2. *The IND Function*

Crucial to our account is the notion of what counts as 'one'. We treat this as a pretheoretical notion, and represent it as a function that maps sets of entities to the sets of entities that count as 'one' (we discuss the merits of this 'pretheoreticality' assumption in Section 7). We allow, as in Landman (2011), that the set of entities that count as 'one' may be disjoint or overlapping. For example, *IND*(*CAT*) will be the set of disjoint single cats. *IND*(*FURN*) will be the set of single items of furniture

(that may overlap such as with a table, a mirror and their sum which forms a vanity). *IND*(*RICE*) and *IND*(*LENTIL*) will be the sets of disjoint single grains of rice and single lentils respectively. *IND*(*MUD*) will be undefined (*IND* is a partial function on predicates). This models the intuition that substances, liquids and gasses lack identifiable individuals independently of the provision of an explicitly or contextually provided measure or package.[7]

However, the *IND* function also introduces a requirement for a counting context. Predicates, P, we assume, are already typed with a precisification context argument, so are of type $\langle \pi, \langle e, t\rangle\rangle$. Applying *IND* (which is of type $\langle\langle \pi, \langle e, t\rangle\rangle, \langle c, \langle \pi, \langle e, t\rangle\rangle\rangle\rangle$) to yield an expression of type $\langle c, \langle \pi, \langle e, t\rangle\rangle\rangle$. The intuitive idea here is that it yields a set of entities that count as 'one'. This set can then be interpreted relative to either a specific counting context ($c_{i>0}$) or the null counting context (c_0) and relative to a specific precisification context ($\pi_{i>0}$) or the null precisification context (π_0).

5.3. *Precisifying context sensitivity (Π-Sensitivity)*

A predicate P is Π-sensitive relative to a counting context c_i:

$$\begin{array}{lcl} \text{Presupposition:} & & \exists \pi_i \exists c_i \exists x[\mathit{IND}(P)(c_i)(\pi_i)(x)] \\ \Pi\text{-sensitive}(P)_{c_i} & \leftrightarrow & \neg\exists x[\mathit{IND}(P)(c_i)(\pi_0)(x)] \end{array} \tag{14}$$

If a predicate is such that, over precisification contexts, the set of entities that count as 'one' are excluded, then there will be no entities that count as 'one' on all precisifications (at π_0). For example, if single grains are in the denotation of *RICE* at some, but not all $\pi_i \in \Pi$, then single rice grains will not be in the denotation of *RICE* at π_0. Nouns in the substances, liquids and gasses class such as *mud* do not satisfy the presupposition in (14), so we assume that it is undefined whether such nouns are Π-sensitive. The definition in (14) means that predicates such as *CAT*, *FURNITURE*, and *FENCE* will not be Π-sensitive relative to at least some counting contexts c_i since there will be entities that count as individuals even at the null precisification context π_0.

5.4. *Counting context sensitivity ($\mathcal{C}$-sensitivity)*

We can define $\mathcal{C}$-insensitivity as the denotation of a predicate being disjoint with respect to the entities that count as 'one', relative to a precisification π_i, by using counting contexts $c_i \in \mathcal{C}$, and specifically using the subvaluated index c_0:

$$\begin{array}{lcl} \text{Presupposition:} & & \exists \pi_i \exists c_i \exists x[IND(P)(c_i)(\pi_i)(x)] \\ \mathcal{C}\text{-insensitive}(P)_{\pi_i} & \leftrightarrow & DISJ(IND(P)(c_0)(\pi_i)) \end{array} \tag{15}$$

Recall that c_0 takes the union of the interpretations of the formula across all counting contexts. $\mathcal{C}$-insensitivity entails that, presupposing there are entities that count as 'one' and the interpretation of $IND(P)$ at some contexts, $IND(P)$ is disjoint at c_0. This will capture the idea that nouns in the prototypical objects class such as *cat* will be disjoint at c_0, since $IND(CAT)$ is the same disjoint set at all counting contexts.

$\mathcal{C}$-sensitivity, defined in terms of overlap, entails that presupposing that there are entities that count as 'one' and the interpretation of $IND(P)$ at some contexts, $IND(P)$ is not disjoint at c_0.

$$\begin{array}{lcl} \text{Presupposition:} & & \exists \pi_i \exists c_i \exists x[IND(P)(c_i)(\pi_i)(x)] \\ \mathcal{C}\text{-sensitive}(P)_{\pi_i} & \leftrightarrow & OVERLAP(IND(P)(c_0)(\pi_i)) \end{array} \tag{16}$$

Nouns in the substances, liquids and gasses class are not defined for $\mathcal{C}$-sensitivity or $\mathcal{C}$-insensitivity since they fail the presupposition in that there are no entities that (clearly) count as 'one' at any precisification or counting contexts with respect to the relevant predicates.

Nouns in the granulars class such as *rice*, *lentils* come out as disjoint when assessed at precisifications that include the entities that count as 'one', since the set of single grains of rice/single lentils is anyway disjoint, so does not change across counting contexts. At these contexts, such nouns will be countable (we can count the grains, granules, flakes etc).

In precisifying contexts where the entities that count as 'one' are excluded from the denotations of granulars, then, at those precisifying contexts, they are (trivially) disjoint. However, this does not mean they are still countable. That the single grains are unavailable at these

precisification contexts is what prevents them from having a counting base.

Nouns in the collective artifacts and homogenous objects class will not be disjoint. Take for example, *kitchenware*. In some counting contexts, a pestle and mortar will count as two items of kitchenware, but in other contexts it will count as one. This means that at c_0, the *IND* set of *K_WARE* will not be disjoint, since it will contain pestles, mortars and pestle and mortar sums.

6. MASS/COUNT LEXICALIZATION PATTERNS, LEXICAL ENTRIES AND FURTHER PREDICTIONS

6.1. *Distribution of Mass/Count Lexicalization Across and Within Languages*

Having defined Π-sensitivity and $\mathcal{C}$-sensitivity, we can derive predictions for the likely encoding of lexical items as count or mass cross- and intralinguistically. This is summarized in Table 2.

Table 2: Vagueness, Overlap, and Lexicalization Patterns

Properties		Example Class	Mass/Count Distribution
Π-sens.	$\mathcal{C}$-sens.		
No	No	Prototypical Objects	Strong tendency as count
Yes	No	Granulars	Variation
No	Yes	Homogenous Objects & Collective Artfacts	Variation
Undefined	Undefined	Substances, Liquids, gasses	Strong tendency as mass

By using these two properties as two sources for the grounding of the mass/count distinction, we are able to predict where one should expect to find variation in mass/count lexicalization patterns. This is something that the other accounts we considered above do not manage to do. The reason why should be clear from Table 2. With just one single source (either Π-sensitivity or $\mathcal{C}$-sensitivity) one would be forced to make a binary decision with respect to mass/count lexicalization leaving no room for variation. With two sources, there is room to motivate how displaying one source, but not the other, can give rise to a tension, the resolution of which leads to either count or mass encoding.

As a brief but important digression, let us observe that our four-way distinction displays parallels with Grimm's (2012) work on the *Scale of Individuation*:

Figure 4: Grimm's (2012) Scale of Individuation

Grimm (2012) did not address artifacts. However, substances align roughly with our class of *substances, liquids, gasses*, and Grimm's *individual entities* aligns with our *prototypical objects*. It is less clear how the intermediate classes of Grimm relate to our *granulars*. Grimm tends to place what we call 'granulars' in either the *granular aggregates* category (for more fine-grained entities), or in the *collectives* category for more coarse-grained entities (Grimm p.c.).

From the ordering in Grimm's (2012) scale of individuation, one can predict patterns in mass/count encoding for specific languages. For example, for a language such as English which arguably has two countability classes, count and mass, it should be ruled out, for example, that substances are lexicalized as count, but collectives as mass. It should also be ruled out that individual entities should be lexicalized as mass, but granular aggregates as count.

With respect to our classes, we doubt there is an ordering relation between homogenous objects and collective artifacts on the one hand and granulars on the other. That is to say that the count or mass encoding of, say, collective artifacts in a language does not predict the mass/count encoding of granulars.

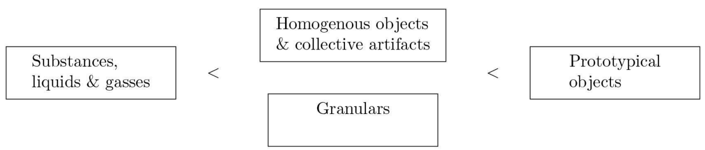


Figure 5: Partial Order between the four noun classes

It is interesting, given that Grimm derives his scale from morphological and syntactic factors that are independent to our own, that we

have arrived at approximately converging conclusions.

6.2. Mass and Count in the Two-Dimensional Semantics

As Table 2 shows, mass encoding can come either from Π-sensitivity or $\mathcal{C}$-sensitivity. However, when only one form of sensitivity is present, it is possible to find count encoding too. Our explanation for this is that there is a lexical choice in whether counting bases are indexed to the contexts of utterance (c_i, π_i), or the null contexts (c_o, π_0). We will show how, prototypical objects, and for substances liquids and gasses, this choice makes no difference. Crucially, for the classes of granulars, homogenous objects, and collective artifacts, we show how this choice translates into count nouns when applying c_i, π_i, and mass nouns when applying c_0, π_0.

We adopt Landman's (this volume) style of representing the lexical entries of nouns as a pair (he uses $\langle \mathbf{body}(X), \mathbf{c\_base}(X) \rangle$). It should, however, be greatly stressed that our definitions are not the same as Landman's. Our pairs are given as: $\langle \mathbf{qual}(X), \mathbf{c\_base}(X) \rangle$. For basic lexical noun entries, at least, $\mathbf{qual}(X)$ is the denotation of the number neutral predicate. This follows Krifka (1989) in representing the predicate 'stripped' of any quantitative criteria (leaving only the qualitative criteria, hence the notation **qual**). The counting base $\mathbf{c\_base}(X)$ is the notional set of entities for counting Xs.

It is important to emphasize that there is a key point of difference that separates our account from Landman's (2011; this volume). Unlike Landman (this volume), who requires that the **base** generates the **body**, we do *not* require that **c_base** generates **qual**. The generator condition is central to Landman's account, but not to ours. The reason for this is that, for granular nouns such as *lentil-s*, we claim that the counting base floats 'above' the bottom of the number-neutral predicate denotation. For example, parts of lentils (cooked in a stew), fall under the number neutral **qual** set *LENTIL*, but parts of lentils do not form the counting base (whole lentils do), therefore **c_base** sets cannot (always) generate **qual** sets.

We distinguish, semantically, between substances and objects. Prelinguistic infants can distinguish substances from objects (Soja et al. 1991). This does not mirror the mass/count distinction (mass nouns such as *furniture* denote objects, not substances). It does mirror the distinction between what we can and cannot intuitively individuate on perceptual and functional grounds. For example, with basic lexical nouns such as *cat*, *chair*, *kitchenware*, *fence*, *fencing*, we can, possibly relative to a context, determine, either on the basis of perceptual properties of their referents, or on the basis of more functional criteria, what counts as 'one' (item/piece for *kitchenware* and *fencing* like nouns). In contrast, the denotations of basic lexical nouns such as *mud* and *blood* cannot be so individuated. This conceptual difference is prelinguistic, and we propose to encode this distinction in the lexical entries of concrete nouns as shown in (17).

Presupposition: $\exists\pi_i \exists c_i \exists x[IND(P)(c_i)(\pi_i)(x)]$

$$[\![\text{n}]\!]^{c_i,\pi_i} = \begin{cases} \langle N, N\rangle & \text{if } N \text{ fails Presupposition} \\ \langle N, IND(N)\rangle & \text{otherwise} \end{cases} \tag{17}$$

The counting bases for substance denoting nouns are the same as the **qual** sets, i.e., number neutral predicates, because predicates in this class fail the presupposition. The counting bases of all other nouns will be the *IND* function applied to the **qual** set (*IND* applied to the number neutral predicate in the case of basic lexical nouns).

Following Krifka (1989) and Rothstein (2010), on our account, there is a typal distinction between count and mass nouns. However, the type difference will not itself draw the count/mass boundary. Instead, properties of counting bases will. Inspired by Landman's (2011) *overlap simultaneously in the same context*, on our account, mass nouns are saturated with the null context c_0 and the null precisification context π_0. Below is the schema for a noun lexical entry and the schemas which show how these entries differ for count and mass nouns.

$$[\![\text{n}]\!]^{c_i,\pi_i} = \begin{cases} \langle \textbf{qual}(N)(\pi_i), \textbf{c\_base}(N)(\pi_i)(c_i)\rangle & \text{if n is } [+C] \\ \langle \textbf{qual}(N)(\pi_i), \textbf{c\_base}(N)(\pi_0)(c_0)\rangle & \text{if n is } [-C] \end{cases} \tag{18}$$

With respect to **qual** sets, both count and mass nouns are indexed to the precisification context of utterance (π_i). This reflects that the standards of precision (precisifying context) may vary for count and mass nouns. For example, the interpretation of *lentils* is sensitive to precisifying context in the same way that the interpretation of *rice* is.

With respect to the counting base, not only is there an index to a precisifying context (π_i for count nouns and π_0 for mass nouns), but also

to a counting context (c_i for count nouns and c_0 for mass nouns). Recall that the inclusion of the counting context as an argument is introduced as part of the semantics of *IND*.

Our proposal is that when the counting base is non-empty and disjoint, we expect a count noun. An empty counting base provides no schema for counting and so leads to a mass encoding. An overlapping counting base makes counting go wrong. We expect a mass nouns otherwise. Variation in mass/count lexicalization will be expected when the disjointness and/or emptiness of the counting base turns on whether the indices are c_i and π_i or c_0 and π_0.

6.3. Examples of lexical entries and further predictions

We now go through some examples for each of the classes of nouns described in Table 1 and show how the *sensitivity* of counting bases to either precisification contexts or counting contexts makes correct predictions about the behaviors of the nouns in each case.

6.3.1. Prototypical objects: *cat*

The lexical entry for cat will be as in (19). The **qual** set is the number neutral predicate *CAT*. The counting base is a maximally disjoint subset of the set of individual cats, which is just the set of individual cats. This set is non-empty and disjoint, therefore *cat* is grammatically countable.

$$[\![\text{cat}]\!]^{c_i,\pi_i} = \lambda x.\langle CAT(\pi_i)(x), IND(CAT)(\pi_i)(c_i)(x)\rangle \tag{19}$$

However, suppose, just for fun, we were to try to give *CAT* a lexical entry in the mold of a mass noun as in (20). Critically, the result should still be a count noun. The counting base set is non-empty and disjoint, therefore even $\text{cat}_{\text{as if mass}}$ should be grammatically countable.

$$[\![\text{cat}_{\text{as if mass}}]\!]^{c_i,\pi_i} = \lambda x.\langle CAT(\pi_i)(x), IND(CAT)(\pi_0)(c_0)(x)\rangle \tag{20}$$

A question raised against our account relates to the stability of prototypical objects. Landman (p.c.) raises the case of potatoes. In the Netherlands (and in the United Kingdom), potatoes are frequently served halved or quartered after boiling (or roasting). In dining contexts, it is felicitous to refer to an individual half or quarter as *a potato* e.g.

Could you squeeze in two more potatoes? would mean .. *two more halves/quarters*. Krifka (p.c.) raised a similar case of when children are being fed apples that have been cut into quarters. In these contexts, in German, it is reportedly felicitous to refer, for example, to two quarters as *Zwei Äpfel* ('two apples'). Arguably, such cases could show, for some nouns in the prototypical objects class, that they are $\mathcal{C}$-sensitive in that what counts as 'one' is sometimes the parts and sometimes the whole (apples or potatoes). However, we suspect these cases are more likely cases of routinised coercion where constructions such as *three potatoes/apples* are coerced into something like *three pieces of potato/apple*. (Such coercion could be prompted by there being no whole apples/potatoes visible in the context.) If they are coercions of this kind, then they pose no problem for our account.

We do not, however, take *potato* to be a central case in the prototypical object category of nouns. Potatoes, especially in potato-eating cultures, are often numerously stocked in cupboards. We expect many everyday contexts in which one potato is insufficient to count as, for example, *having potatoes for dinner*. Arguably, therefore, *potato* is Π-sensitive. As such, we would expect to find cross-linguistic variation i.e. a mass-counterpart for *potato*, and indeed we do. For example, the Russian *kartoška* ('potato') is mass as shown by the infelicity of (21) (excluding packaging or taxonomic coercion).

(21) ?Ona kupila tri kartoški
she buy.PST.SG.F three potato.GEN
Int: 'She bought three potatoes'

6.3.2. Collective artifacts: furniture versus huonekalu

The lexical entry for *furniture* will be as in (22). The **qual** set is the number neutral predicate *FURN*. The counting base is the set of items of furniture that count as 'one'. The counting base set is non-empty, but it is not disjoint, since it is indexed to c_0. For the same reasons as given by Landman (2011), overlapping counting base sets make counting go wrong, hence *furniture* is mass.

$$[\![\text{furniture}]\!]^{c_i,\pi_i} = \lambda x.\langle \mathit{FURN}(\pi_i)(x), \mathit{IND}(\mathit{FURN})(\pi_0)(c_0)(x)\rangle \quad (22)$$

However, now if we try providing a lexical entry in line with the count schema, we see a different effect. Now the counting base set is indexed to the counting context of utterance/evaluation c_i, not the null context c_0. This means that at every context, the counting base set is disjoint and non-empty and so fit for use in counting. In other words, this lexical entry would be appropriate for a count noun with the same denotation as furniture such as the Finnish *huonekalut* ('(items of) furniture') as in (23).

$$[\![\text{huonekalu}]\!]^{c_i,\pi_i} = \lambda x.\langle \mathit{FURN}(\pi_i)(x), \mathit{IND}(\mathit{FURN})(\pi_i)(c_i)(x)\rangle \quad (23)$$

6.3.3. Homogenous objects: fence versus fencing

The same pattern as with collective artifacts emerges with homogenous objects. Indexing to a specific counting context yields a count noun such as *fence* in (24), because the counting base is non-empty and disjoint. Indexing to the null counting context yields a mass noun entry, such as one for *fencing* in (25), because the counting base is non-empty and overlapping.

$$[\![\text{fence}]\!]^{c_i,\pi_i} = \lambda x.\langle \mathit{FENCE}(\pi_i)(x), \mathit{IND}(\mathit{FENCE})(\pi_i)(c_i)(x)\rangle \quad (24)$$

$$[\![\text{fencing}]\!]^{c_i,\pi_i} = \lambda x.\langle \mathit{FENCE}(\pi_i)(x), \mathit{IND}(\mathit{FENCE})(\pi_0)(c_0)(x)\rangle \quad (25)$$

6.3.4. Granulars: rice versus lentils

The lexical entry for *rice* is given in (26). The **qual** set is the number neutral predicate *RICE* interpreted relative to a precisification context. In some precisification contexts, this will include single grains of rice, in others, it will not. The counting base set does *not* contain the single grains of rice, since it is interpreted relative to the (supervaluated) null precisifying context π_0. This set is, therefore, empty, so provides no counting schema. Therefore, even when a single grain is sufficient to count as *rice* in the context of utterance, the counting counting base does not pick single grains out for grammatical counting. *Rice* is mass.

$$[\![\text{rice}]\!]^{c_i,\pi_i} = \lambda x.\langle \mathit{RICE}(\pi_i)(x), \mathit{IND}(\mathit{RICE})(\pi_0)(c_0)(x)\rangle \quad (26)$$

Similarly to *rice*, the lexical entries for mass nouns such as *lešta* ('lentil', Bulgarian) and *čočka* ('lentil', Czech) are also indexed to the null counting and precisification contexts so are also mass as in (27)

$$[\![\text{lešta/čočka}]\!]^{c_i,\pi_i} = \lambda x.\langle LENTIL(\pi_i)(x), IND(LENTIL)(\pi_0)(c_0)(x)\rangle \quad (27)$$

However, now if we try providing a lexical entry unsing *LENTIL* in line with the count schema, we apply the counting context of utterance, and get a lexical entry for *lentil* as shown in (28). The **qual** set is the number neutral predicate *LENTIL* interpreted relative to a precisifying context.

$$[\![\text{lentil}]\!]^{c_i,\pi_i} = \lambda x.\langle LENTIL(\pi_i)(x), IND(LENTIL)(\pi_i)(c_i)(x)\rangle \quad (28)$$

There are now two possible cases. One is where the precisifying context includes single lentils. The other is where it does not. When π_i does admit single lentils, then the counting base is the set of single lentils. The counting base set is non-empty and disjoint since it is indexed to π_i. Interestingly, in the case where the precisifying context of utterance does *not* include single lentils in the denotation of *lentils*, this account predicts that access to the individual lentils should be obscured.

At first this may not seem like a validated prediction. Is not lentils a straightforward count noun? Yet, in contexts, say where lentils are being served for dinner, perhaps in a stew, it is decidedly strange to ask the question in (29):

(29) How many lentils would you like?[8]

The extent to which one can interpret this is actually more akin to a mass-count coerced packaging reading (e.g. How many SPOONFULS OF lentils would you like?). The lexical entry in (28) leads to a prediction that the question in (29) should not be answerable unless the precisifying context is shifted. Note that one could answer (29) with "Around 500", but this is almost some kind of joke. Perhaps part of the joke is answering as if from a different standard of precision.

6.3.5. Substances, liquids and gases: mud

The lexical entry for *mud* is given in (30). The **qual** set is the number neutral predicate interpreted relative to a precisification context, and in line with (17), due to the presupposition failure, the counting base is also the number neutral predicate interpreted relative to a precisification context. The counting base set is clearly overlapping, it also

provides no intuitive counting schema, so *mud* is mass.

$$[\![\text{mud}]\!]^{c_i,\pi_i} = \lambda x.\langle MUD(\pi_i)(x), MUD(\pi_0)(x)\rangle \quad (30)$$

Furthermore, were we to try to give an entry using *MUD* in line with the count schema, namely by applying the precisification context of utterance, the counting base is still overlapping (counting contexts are not relevant here since none have been introduced by the *IND* function):

$$[\![\text{mud}_{\text{as if count}}]\!]^{c_i,\pi_i} = \lambda x.\langle MUD(\pi_i)(x), MUD(\pi_i)(x)\rangle \quad (31)$$

As a slight aside, there may well be a graded scale between central cases of substance-denoting nouns like *mud* and central cases of granular nouns like *rice*. For example, *sand* denotes very small grains and *dust* arguably denotes specks. We have characterised granulars as nouns which have in their denotations single grains, granules etc. that intuitively count as 'one'. Substance-denoting nouns, we argued, fail to be so individuated. On the current formalism, there is not that much room for manoeuvre here. The partial function *IND* is either defined for a predicate or it is not (however see endnote 7). Furthermore, arguably for some of these borderline granulars, what counts as 'one' is context sensitive and/or vague. (A speck of dust can be made, sorites-fashion, arbitrarily bigger or smaller and still count as *a speck of dust*. The same does not hold for a grain of rice. Landman p.c.). There is some evidence, from languages which use diminutives as *unit extractors*, that for example, *dust*-like nouns pattern with nouns like *rice* and not with nouns like *mud*. (Examples (32) and (33) are from Landman (p.c.).)[9]

(32) rijst – rijstje
rice$_{[\text{-c}]}$ – rice.DIM$_{[+\text{c}]}$
'rice' – 'grain of rice'

(33) stof – stofje
dust$_{[\text{-c}]}$ – dust.DIM$_{[+\text{c}]}$
'dust' – 'speck of dust'

(34) modder – ??moddertje
mud$_{[\text{-c}]}$ – mud.DIM

To represent this gradience, and to include potential context-sensitivity into the *IND* function itself, arguably, we would need a formal language better suited to representing graded concepts. We have worked with such a formalism in modelling the count/mass distinction (see Sutton & Filip 2016c,d, for details). It would be interesting, in further research to apply this framework to these data.

6.4. Summary

In this section we have seen how sensitivity to precisifying contexts and sensitivity to counting contexts gains a great deal in developing a semantics for countability in nouns. We can predict the distributional patterns of mass and count we should expect to find cross- and intralinguistically based on semantic criteria. Furthermore, building on this, we have given representations for nouns across the classes that do and the classes that do not display mass/count variation. Based on these representations, we have been able to explain why the lexicalization of both a mass noun and a count noun is probable in some cases (granulars, collective artifacts, and homogenous objects), and much less probable in others (prototypical objects, and substances liquids and gasses).

7. LACUNA

Despite making progress in unraveling some of the complexities and puzzles in the semantics of the mass/count distinction, our account leaves a sizable lacuna. A question we have left unaddressed, is: What are the mechanisms that underpin *IND*, namely the mapping from number neutral predicates, to the entities that count as 'one' for that predicate? It should be stressed, however, that similar questions remain lacunae in most accounts of the mass/count distinction in that most accounts take *individuation* to be, at some level, pretheoretical. Furthermore, a similar issue imbues, to perhaps an even greater degree, research on aspect and the progressive in the verbal domain. Within these fields of research, issues surrounding when an event counts as 'one' with respect to a predicate has proven to be a highly slippery matter. (For example, see Landman (1992) for an overview of some classic approaches to the progressive and some discussion of this problem.)

We briefly address the lacuna in this section by assessing what has been said on this matter in the mass/count literature that we have discussed in this paper, and in relation to work we have done in a

frame-based framework, Type Theory with Records (TTR, Cooper 2012; Cooper et al. 2015).

To some extent, our *IND* function plays a similar role to Krifka's (1989) natural unit measure function (**NU**), except that we have enriched this notion to be counting context sensitive (following Rothstein 2010), and to be applied to a predicate at a precisification context. Krifka does briefly discuss some constraints on **NU** (which is the equivalent of **OU** in Krifka (1995):

> "we can assume that **NU** yields the same measure function for entities of a similar kind. For example, the unit for all living beings is constituted by the organism. Then, **NU**(**cattle**′) and **NU**(**game**′) should denote the same measure function." (Krifka 1989, p. 84)

> "The operator **OU** could reasonably be interpreted in such a way that it yields the same measure function for, say, bears and cats, that is, $\mathbf{OU}_1(\mathbf{Ursus}) = \mathbf{OU}_1(\mathbf{Felis})$ in both cases the unit is derived from the notion of a biological organism and may be identified with $\mathbf{OU}_1(\mathbf{animal})$." (Krifka 1995, p. 401)

As noted by Zucchi & White (1996, 2001) and later Rothstein (2010), Krifka's notions of *natural unit* and *quantization* do not apply well to nouns such as *string* and *fence* which do not come in natural units. However, it is also not clear how the above considerations affect the individuation of nouns which denote artifacts such as *chair*, let alone more complex cases such as *furniture* and *kitchenware*. Indeed, it pushes the use of 'natural' to describe artifacts as natural units at all. Perhaps a better nomenclature for artifacts would be *functional unit*.

The notion of counting as 'one' in Chierchia (2010), insofar as he has one at all, is reduced to the notion of *stable atom*. However, as we have seen, this wrongly predicts the availability of crosslinguistic count/mass pairs such as *lentil-s*$_{[+C,PL]}$ and *čočka*$_{[-C]}$ ('lentil', Czech). Or, at the very least, it wrongly predicts the availability of intralinguistic pairs such as *oat-s*$_{[+C,PL]}$ and *oatmeal*$_{[-C]}$ (Section 3.1).

Counting as 'one' also features in both Rothstein (2010) and Landman (2011). Rothstein (2010) enhances Krifka's *natural units* by making this function context sensitive (a device that we have adopted here).

However, part of this notion is still pretheoretical, since there are no constraints in Rothstein's account on what an acceptable or admissible counting context is. For example, why a portion of picket fencing in London and a portion of chain-link fencing in Berlin cannot, as a sum, count as 'one'.

Landman constrains counting as 'one' with his definition of *generator set*, namely a set that must generate the full denotation under complete mereological sum. However, on Landman's account, the *chosen* generator set for a predicate is only one of many possible such sets. In other words, for any predicate, the mechanism that selects a generator set as *the* set of entities that counts as 'one' (in a context) is assumed to be pretheoretical.

One response to the lacuna is to discharge oneself from the responsibility of filling it. After all, there are enough puzzles, complexities and data to account for even if one does take *counting as 'one'* as a pretheoretical notion. Indeed, this is not an incoherent position to adopt. One way of framing this position could be to draw a divide between compositional and lexical semantics. *Counting as 'one'* in relation to a predicate could be classed as a task for lexical semantics and so dispensed as a duty for projects in compositional semantics. However, this does not exclude the possibility of combining insights from lexical semantics with compositional semantics and outlining how they feed into compositional semantic representations.

Grimm's work (Grimm 2012), is one example of how this can be done. Grimm (2012) argues that mereology is insufficient to capture the notion of *individual* and argues that semantics should be enriched with mereotopological relations. For example, he gives the following lexical entry (Grimm 2012, p. 151):

$$[\![\text{dog}]\!] = \lambda x_0[R(x_0, \text{Dog}) \wedge \text{MSSC}(x_0)] \tag{35}$$

Where MSSC stands for the mereotopological predicate *maximally strongly self-connected*: "An [mereological]-individual is Maximally Strongly Self-Connected relative to a property if (i) every (interior) part of the individual is connected to (overlaps) the whole (Strongly Self-Connected) and (ii) anything else which has the same property and overlaps it is once again part of it (Maximality))" (Grimm 2012, p. 135). An alternative approach is argued for in Sutton & Filip (2016c,d).

Sutton and Filip develop a mereological version of probabilistic Type Theory with Records (Cooper et al. 2015). A major benefit of adopting a TTR framework is that it combines frame theoretic lexical semantic representations inspired by Fillmore (1975, 1976), with compositional semantics in the Frege-Montague tradition. Sutton & Filip (2016d) begin to investigate and formally model how perceptual information and conceptions of function interact with semantic learning and can lead to different conceptions of counting as 'one'. There, we argue (in line with Landman's early considerations regarding *perspective* (Landman 1992)) that counting as 'one' is a highly context dependent notion in which mereotopological factors are just one ingredient that does not determine the way that some stuff is (or is not) individuated.

The benefit of adopting probabilistic Bayesian frame semantics is that one gets, for free, the ability to represent underspecified information and defeasible reasoning on the basis of information drawn form a number of domains (semantic, perceptual, doxastic etc.). A major drawback of the two-dimensional semantics presented in this article, which is detailed in a relatively conservative model-theoretic framework, is that it is fundamentally unclear how one could include such a rich variety of information within the constraints of this framework.

To put things simply, suppose we call the investigation into the notion of what 'counts as one' the development of a *theory of individuation*, and the development of an account of the mass/count distinction a *theory of countability*. It is then likely that the development of a theory of individuation will inform the development of a theory of countability, and *vice versa*. However, if that is the case, then both should be explored within a formal framework capable of representing a rich enough range of information to encapsulate both. To this end, the probabilistic, mereological frame-based formalism developed by Sutton & Filip (2016c,d) is a good contender.

8. SUMMARY AND CONCLUSIONS

The account we proposed here can accommodate a much broader range of mass/count data than other leading proposals. For the classes of nouns we have identified:

Prototypical objects: Nouns in this class have a strong tendency to be count. This is because they are counting-context and precisification-context insensitive. They have non-empty, disjoint sets of entities that count as 'one' at all counting and precisification contexts. They even have non-empty, disjoint sets of entities that count as 'one' at the null counting and precisification contexts.

Collective artifacts and Homogenous objects: Nouns in this class display a lot of mass/count variation. This is because they are counting context sensitive, but not precisification context sensitive. They have non-empty sets of individuals that count as 'one', but this set is disjoint only when interpreted at specific counting contexts. It is not disjoint at the null counting context. This means that mass/count lexicalization turns on whether they are indexed to a specific counting context or the null-counting context. Hence, we expect mass/count variation.

Granulars: Nouns in this class display a lot of mass/count variation. This is because they are precisification context sensitive, but not counting context sensitive. When non-empty, the set of entities that count as 'one' are disjoint, but this set is not non-empty at some precisifications. This means that mass/count encoding turns on whether these nouns are interpreted at a specific precisification context or at the null precisification context. Hence we expect mass/count variation in this class.

Substances, liquids, and gasses: Nouns in this class have a strong tendency to be mass. They are undefined for both counting context and precisification context sensitivity, because they lack a defined set of entities that count as 'one'. Our considerations of the prelinguistic distinction between substances and objects (Soja et al. 1991) justified treating the lexical entries of these nouns differently. The **c_base** set, we suggested should be identical to the **qual** set, namely, the number neutral predicate (e.g. $\lambda\pi\lambda x MUD(\pi)(x)$). Such sets have no counting context argument and are not disjoint at any precisification context (or at the null precisification context). This is why we expect fairly consistent mass encoding for nouns in this class.

We have argued that at least two semantic sources must lie behind the semantics of mass and count nouns. This is because, with a single, categorical criterion for whether a noun will be mass or count, one cannot account for the full range of variation in mass/count lexicalization

patterns. However, there are further puzzles that our framework may be able to untangle. There are severe restrictions on the felicity conditions for cases of mass-to-count coercion. For example, granular and collective artifact mass nouns resist object 'packaging' readings: *Three furnitures* cannot be read as 'Three ITEMS OF furniture', and *three rices* cannot be read as 'Three GRAINS OF rice'. The fact that mass nouns, on our account, are indexed to the null counting and precisification contexts may hold the key to explaining this. For example, to pragmatically substitute ITEM OF into *three furnitures* would constitute removing the null counting context from the lexical entry for furniture, and replacing it with , the counting context of utterance. However, this is an operation above and beyond the usual 'packaging' coercion found in cases such as *three waters* meaning e.g., 'Three GLASSES OF water'. (For an account of such phenomena within the framework presented here see, Sutton & Filip 2016a)

Representing how different aspects of context interact and allow, or debar both mass and count conceptualizations of entities better accommodates mass/count lexicalization patterns, can lead to predictions regarding the precise behavior of nouns in different classes, and, in future research, may help to explain hitherto unaccounted for restrictions on mass/count coercion.

ACKNOWLEDGEMENTS

First, we would like to thank Fred Landman for his invaluable comments and suggestions for the revision of this manuscript. Second, many of the ideas of this paper were improved as a result of attending the *7th International School in Cognitive Sciences and Semantics* in Riga (Latvia) and the *11th International Symposium of Cognition, Logic and Communication*, dedicated to *Number: Cognitive, Semantic and Crosslinguistic Approaches*, both organized by Jurğis Šķilters and Susan Rothstein. We received very valuable feedback at these events from Scott Grimm, Jenny Doetjes, Daniel Hyde, Fred Landman, Suzi Lima, Susan Rothstein, and Yasutada Sudo. We also wish to thank the audiences at the *Sinn und Bedeutung 21* and *SALT 26* conferences and the participants of our *ESSLLI 2016* class and the *Formal Grammar 21* conference. Finally, we thank our colleagues and collaborators in the Department of

Linguistics and the affiliated *Collaborative Research Center 991* at Heinrich Heine University, and in particular the participants of the *Semantics and Pragmatics Exchange (SemPrE)* colloquium. This research was funded by the German Research Foundation (DFG) CRC 991, Project C09.

Notes

[1]The nouns *fence* and *cow* may be context sensitive in other respects.

[2]We use the standard notation, $^*X$ for the upward closure of a set X under $\sqcup$.

[3]Some native speakers seem to conceptualize nouns such as *fencing* like substance-denoting nouns such as *mud*. On this conception, *fencing* would denote *stuff from which fences are made*. We do not share this intuition. For us, there is a distinction between, for example, concrete posts and chain-link rolls that we consider as *fencing*, and, say, bags of concrete and lengths of wire (the stuff the posts and chain link rolls are made of). We do, however, recognize that our account does not fully capture the sense in which fencing can be more disassembled than a fence/some fences.

[4]As suggested to us by Landman p.c. and Grimm and Levin p.c., one could, potentially, think up an exotic context in which a single girl would not count as falling under the predicate *GIRL* (perhaps in some dystopian world in which a single human life was not held in high regard). Likewise, a single cow might not, in rare contexts, fall under the predicate *COW* (perhaps a farmer of an industrial scale farm would not consider a single cow to be sufficient to count as falling under *COW*). Such 'extreme' contexts could be argued to be defeaters for a vagueness-based account of the mass/count distinction. However, we mean to discuss only everyday, common contexts. These are the sorts of contexts that humans are exposed to when learning natural language. A better account of vagueness, such as one based on graded probabilistic judgements, would remedy such objections. We propose such an account in Sutton & Filip (2016b,c).

[5]An exception is arguably in Brazilian Portuguese (Pires de Oliveira & Rothstein 2011) in which all count nouns potentially have mass readings when in the bare singular.

[6]We are aware, however, that in some languages such as Yudja (Lima 2014), all nouns seemingly have straight-forward, non-coerced count uses (e.g. can be combined directly with numerals). Our current account does not accommodate these interesting data, however, we plan to address them in future work.

[7]An interesting topic for future work is to think of *IND* as a vague term. If vagueness is understood, in supervaluationist terms, as a form of underspecification, then *IND* could be underspecified (vague) for some nouns, namely substances, liquids and gasses. However, this would open up the possibility of resolving this vagueness in context. For languages in which substance denoting nouns are directly countable (such as Yudja (Lima 2014)), arguably, the facilitation of counting could arise as the resolution of vagueness in the *IND* predicate.

[8]It is also forced to ask *How much lentils would you like?*, this has to do with a clash with the input requirement of the determiner *much*. We also find it hard to interpret *How much lentil would you like?* except in cases where the answer would be, for example, half a lentil. In this sense, lentil behaves a stubbornly count noun. The fact that nouns like *lentil-s* resist both 'how much' and 'how many' questions is a puzzle, we intend to return to in more detail at a later time.

[9]More judgements should be collected before we put too much weight on these data. In our very small sample of three other native Dutch speakers, all three rejected the felicity of *rijstje* (rice.DIM) to mean 'grain of rice'.

References

Cheng, Lisa Lai-Shen. 2012. 'Counting and Classifiers'. In Diane Massam (ed.) 'Count and Mass Across Languages', 199–219. OUP.

Chierchia, Gennaro. 2010. 'Mass Nouns, Vagueness and Semantic Variation'. *Synthese* 174: 99–149.

Cooper, Robin. 2012. 'Type Theory and Semantics in Flux'. In R. Kempson, T. Fernando & N. Asher (eds.) 'Philosophy of Linguistics, Handbook of the Philosophy of Science', 271–323. Elsevier.

Cooper, Robin, Dobnik, Simon, Larsson, Staffan & Lappin, Shalom. 2015. 'Probabilistic Type Theory and Natural Language Semantics'. *LILT* 10, no. 4.

Fillmore, Charles J. 1975. 'An Alternative to Checklist Theories of Meaning'. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society* 1: 123–131.

——. 1976. 'Frame semantics and the nature of language'. *Annals of the New York Academy of Sciences* 280, no. 1: 20–32.

Grimm, Scott. 2012. *Number and Individuation*. Ph.D. thesis, Stanford University.

Krifka, Manfred. 1989. 'Nominal Reference, Temporal Constitution and Quantification in Event Semantics'. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.) 'Semantics and Contextual Expression', 75–115. Foris Publications.

——. 1995. 'Common Nouns: A Contrastive Analysis of English and Chinese'. In G.N. Carlson & F. J. Pelletier (eds.) 'The Generic Book', 398–411. Chicago University Press.

Landman, Fred. 1992. 'The Progressive'. *Natural Language Semantics* 1, no. 1: 1–32.

——. 2011. 'Count Nouns – Mass Nouns – Neat Nouns – Mess Nouns'. *Baltic International Yearbook of Cognition, Logic and Communication* 6: 1–67.

——. this volume. 'Iceberg Semantics for Count Nouns and Mass Nouns: The evidence from portions'. *The Baltic International Yearbook of Cognition Logic and Communication* 11.

Lima, Suzi. 2014. 'All notional mass nouns are count nouns in Yudja'. *Semantics and Linguistic Theory (SALT)* 24: 534–554.

Link, Godehard. 1983. 'The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretic Approach'. In Paul Portner & Barbera Partee (eds.) 'Formal Semantics - the Essential Readings', 127–147. Blackwell.

Pires de Oliveira, Roberta & Rothstein, Susan. 2011. 'Bare singular noun phrases are mass in Brazilian Portugese'. *Lingua* 121: 2153–2175.

Rothstein, Susan. 2010. 'Counting and the Mass/Count Distinction'. *Journal of Semantics* 27, no. 3: 343–397.

Soja, Nancy, Carey, Susan & Spelke, Elizabeth. 1991. 'Ontological categories guide young children's inductions of word meaning: Object terms and substance terms'. *Cognition* 38: 179–211.

Sutton, Peter R. & Filip, Hana. 2016a. 'Counting in context: Count/mass variation and restrictions on coercion in collective artifact nouns'. *Semantics and Linguistic Theory (SALT)* 26: 350–370.

——. 2016b. 'Vagueness, overlap, and countability'. *Proceedings of Sinn und Bedeutung 20* (To appear).

——. 2016c. 'A probabilistic, mereological account of the mass/count distinction'. *Proceedings of TbiLLC 2015* (To appear).

——. 2016d. 'Probabilistic Mereological Type Theory and the Mass/Count Distinction'. Under review for *JLM* (proceedings of Type Theory and Lexical Semantics (TYTLES) workshop at ESSLLI 2015).

Zucchi, Sandro & White, Michael. 1996. 'Twigs, Sequences and the Temporal Constitution of Predicates'. In Teresa Galloway & Justin Spence (eds.) 'Semantics and Linguistic Theory (SALT) 6', 223–270. Ithaca, NY: Linguistic Society of America.

——. 2001. 'Twigs, Sequences and the Temporal Constitution of Predicates'. *Linguistics and Philosophy* 24, no. 2: 223–270.