

Gennaro Chierchia
Harvard University
September 2019

Mass vs. Count: Where do we stand? Outline of a theory of semantic variation.

To appear in T. Kiss, F. J. Pelletier and H. Husiĉ (eds) *The Semantics of the Mass/Count Distinction: Recent Developments and Challenges*, Cambridge University Press, Cambridge

Abstract. DP structure, number marking, and the morphosyntax of the mass/count distinction appears to be subject to a great deal of variation. Language systems with clear evidence of two classes of nouns, those that allow direct combination with numerals and those that don't, are by now fairly well combed through. As are languages that *disallow* direct combinations of numerals with *any* noun, namely generalized classifier languages (Mandarin, Japanese, etc.). Finally, there are languages that do *allow* free combination of numerals with *any* N, whether conceptually mass or count, like Nez Perce, Yudja, Indonesian, ..., which have also been well documented at this point. This variation has given rise to theories of the mass/count contrast where the link between the pre-linguistic/cognitive basis of the distinction and its grammatical manifestation is weakened to the point of disappearance: basically any 'concept' can have a mass or a count grammatical representation (cf. e.g. Chierchia 1998a, Borer 2005, Rothstein 2010, Landman 2011, De Vries et al. 2018, a.o.). I am going to argue that this position is not supported by the available evidence: All of the languages mentioned above retain essentially the *same* notion of countability. I will, accordingly, propose an approach consistent with the thesis that the mass/count contrast rests on an underlyingly universal structure. To use one of Chomsky's favorite metaphors, if Martians were to be exposed to Italian, Mandarin and Yudja, they would think that they count things the same way, modulo minor phonological differences.

1. Plan.

A lot of the debate on the nature of the mass/count distinction centers on what the ultimate determinant of (un)countability is. Here are some of the current stances on this matter:

- (1) Mass nouns cannot combine directly with numerals because:
 - a. They are not atomic, under some formally supplied notion of atomicity (Bunt 1979, Link 1983)
 - b. Their generator sets have overlapping parts (Landman 2011, 2016)
 - c. Count nouns have contextually supplied minimal parts; mass nouns don't. (Rothstein 2010)
 - d. The minimal parts of mass nouns are not 'maximally connected.' (Grimm 2012b)
 - e. Count nouns denote states with singular/atomic participants. Mass nouns denote states with non-singular/atomic participants (Schwarzchild 2011)
 - f. The 'minimal components' of mass nouns are specified too vaguely to be used in counting. (Chierchia 2010, 2015, 2017)

A main problem that theories like those in (1) face is that of variation. I will illustrate and discuss the extent to which grammars may vary vis-à-vis the mass/count distinction with three main phenomena and use this as a springboard for outlining a theory of semantic variation, with a universal logical basis. The first form of variation concerns the most widespread empirical test associated with the mass/count distinction, namely how noun phrases (NPs) combine with numerals. The second is that of the so called 'fake' mass nouns (like *furniture*). The third concerns alternations between mass vs. count interpretations of nouns like *beer* or *chicken*.

From the point of view of *Numeral-Nouns* combination, three types of languages have been so far convincingly documented.

- (2) a. Type I languages.
 Numerals combine directly with some nouns, but not with others; for the latter a classifier or measure phrase needs to be interpolated.
 Example: *three chairs* vs. **three blood(s)* vs. *three ounces/drops of blood*
 Most of the Indo European languages belong to this type.
- b. Type II languages.
 Numerals cannot combine with *any* NP directly. A classifier/measure phrase is always needed, whether the noun is cognitively count or mass.
 Example: i. san *(ge) ren
 Three CL person 'three people'
 ii. san *(bang) yintao
 three pound cherry 'three pounds of cherries'
 Languages of this type are Chinese (e.g., Cheng and Sybesma 1998,1999), Japanese, Nuosu Yi (Jiang 2017), Bangla (Dayal 2012).
- c. Type III languages.
 Numerals freely combine with any type of nouns. In combination with cognitively count nouns (e.g., *three cats*) numerals have the meaning they do in English. In combination

with mass nouns they have a ‘container’ or a ‘quantity of’ reading (not necessarily ‘standard’ quantity of).

- Example: i. lepit cickan Nez Perce (Deal 2017)
 two blanket ‘two blankets’
 ii. lepit kiek’e’t
 two blood ‘two quantities (e.g. drops) of blood’

Languages of this sort include Indonesian (Darlymple and Morfu 2012), Yudja (Lima 2014), Nez Perce (Deal 2017).

This typology is meant to be descriptive: to arrive at the ‘right’ theoretical classification in terms of grammatical mechanisms is our main objective. In stating the typology above, I rely on the notion of ‘*cognitively* count vs. mass’ noun/concept. The definition I have in mind here is the one stemming from the research of cognitive psychologist like Carey and Spelke (1996).¹ This rich line of work shows that children, *before* language clearly distinguish concepts like ‘teddy bear’ or ‘car’ that have relatively well defined units (which retain their identity upon moving through space and entering in contact with each other), from concepts like ‘water’ or ‘sand’ do not have readily accessible minimal parts, and whose samples don’t retain an identity when moving or congregating. Establishing precisely the nature of this dichotomy is part of the problem, but that it exists is at this point well established. I use the label ‘cognitive’ in connection with the distinction just outlined to underscore the fact that it is present in the cognitive system of children before any manifestation of language, and, for that matter, it is also present in other non-human primates (cf., e.g. Hauser and Carey 2003).

A second form of variation concerns cases of ‘misalignment’ with respect to the cognitive contrast just characterized. In English, as in most IE languages, there is a class of nouns that is cognitively count but patterns with the mass ones with respect to the tests prevailing in the language: *furniture*, *kitchenware*, *jewelry*, *luggage*,... Half a table is no more a good instance of the table-concept, than it is of the furniture-concept. Furniture units are as well defined as table- and couch-units. And yet, *furniture* patterns with mass nouns with respect to pluralization, combination with numerals, use of the indefinite article, etc. I like to call nouns of this sort ‘Fake Mass’, not to downplay the phenomenon, but because Fake Mass nouns take on the grammatical behavior of mass nouns while being cognitively count, and patterning tendentially with count concepts in tasks that do not involve language.² Variation with respect to Fake Mass nouns manifests itself in three ways. First, there is variation that involves specific lexical entries across languages. For example, nouns like *luggage* or *jewelry* are Fake Mass in English but count in Italian (*bagagli* ‘luggages’, *gioielli* ‘jewels’); *servitù* ‘servants’ is Fake Mass in Italian, count in English. Second, the phenomenon of Fake Mass seems to be absent from Type II and Type III languages; this claim requires some factual justification that will be provided shortly. And third,

¹ See also, e.g., Carey (1985), Soja, Carey, and Spelke (1991), Feigenson et al. (2003).

² Or tasks which involve language only marginally; See, e.g., Barner and Snedeker (2005); for recent developments of this experimental paradigm, cf. Scontras et al. (2017). Barner and Snedeker call ‘Fake Mass’ nouns ‘object mass’. The terminology on them varies a lot. Doetjes (1997) calls them ‘count mass’; Rothstein (2010) calls them ‘superordinate mass’; Landman (2011) calls them ‘neat mass’; Grimm (2012b) calls them ‘functional aggregates’.

only a subset of Type I languages seems to have Fake Mass nouns. Greek is a Type I language that lacks them altogether (cf. Tsoulas 2009; also pc).

The claim that Type II and Type III languages lack Fake Mass nouns rests on the following observations, that have to do with the way in which the mass/count distinction manifests itself in these languages. In Type II languages, also known as Generalized Classifier-languages, the mass/count distinction manifests itself in the classifier system. Cheng and Sybesma (1998) a.o. have shown that languages like Mandarin or Cantonese have a grammatically identifiable system of count classifiers, which combine with cognitively count nouns, but not with mass ones; they include the ‘sortally unspecific’ classifiers, like *ge* in Mandarin, as well as some category specific ones like *ben* (the classifier for books, periodicals, journals, etc.). A Fake Mass noun in a language like Mandarin would, accordingly, be something that lexicalizes a cognitively count concept like *jiaju* ‘furniture’, but does *not* combine with count classifiers. As Cheng and Sybesma point out, such Ns do not exist: nouns like *furniture* naturally combine with unspecific count classifiers like *ge* and/or with category specific ones (like *jian jiaju*) with a count interpretation (‘whole piece’ of furniture). And in so far as I know this holds across all Type II languages. In Type III languages, the issue takes a different form. In these languages, *Numeral-Noun* combinations behave as indicated in (2c): if the noun is cognitively count, $n + N$ means the same as in English: n natural units of N ; if the noun is mass, $n + N$ means n (*natural*) quantities of N (e.g., *three bloods* = *three drops of blood* or *three puddles of blood*). So, a Fake Mass noun would be an N that is cognitively count, like, say *canoe* or *table*, but in combination with numerals freely allows for an interpretation parallel to that of mass nouns (i.e. *three canoes* = *three aggregates of canoes* or *three pieces of a canoe*), i.e. a cognitively count noun that gets ‘counted’ as a mass one. For the time being, I am not aware that this ever happens. If these observations withstand further empirical scrutiny, the conclusion that the phenomenon of fake mass nouns is attested only in (a subset of) Type I languages seems warranted, which cries out for an explanation.

The case of fake mass Ns is profoundly different, I think, from the case of *ambiguous* Ns like *rope* or *rock*, which give raise to a third form of variation. There are two equally natural senses of *rope*. One is that of a bounded, continuous string of some material suitable to tie things together. The other is any recognizable, not necessarily detached part of such a string. In this second sense, part of a rope still qualifies as *rope* (but not as *a rope*, unless it is cut off from the rest). Related to this is the ambiguity of Ns like *chicken*, that applies to whole animals in its count sense, and to animal parts (viewed mostly as food) in its mass sense. Or the ambiguity of words like *beer* that is equally felicitous as name of an alcoholic liquid, and as referring to some *standardized* serving of the latter (a bottle or glass of beer). With Ns of this sort, some kind of reconceptualization seems to be going on in switching back and forth between its mass and count uses, conceptualization that crucially affects what counts as a minimal unit of the relevant concept: the minimal unit of the count sense of *beer* is a glass or a bottle of beer (half a bottle of beer doesn’t qualify as a beer); the minimal units of the liquid are not really known, much less rigidly set by our lexicon, and any recognizable small amount of the liquid may be a candidate; but they differ from the minimal units of *beers*. With nouns like *furniture*, instead, what counts as a minimal unit does not seem to change significantly with respect to the categories (tables, chairs, couches) of which it is a superordinate.

Potentially ambiguous Ns are subject to variation both within and across languages. For example, arguably the word *hair* is predominantly mass in English: one says things like *my hair is white* not *my hairs are white*. The latter is, however, how a literal translation of the

functionally equivalent Italian phrase *i miei capelli sono bianchi* would sound. At the same time, English does admit things like *I found three blond hairs on your jacket*, showing that it tolerates count uses of *hair*, in specific contexts. This type of variation/elasticity is pervasive: I think we will find manifestations of it in any language. It being so widespread suggests that the variation across potentially ambiguous Ns is more socially and culturally determined, than grammatically determined. We all can have fun in imagining a society of vampires in which one could walk in a bar and order a blood on the rocks; or a society of wood worms where one does not digest *shelf*, but *table* is fine. Even this cartoon like imagery has its limits, however: geometrical shapes like circles or triangles, and other shape based concepts like that of *hole* are notoriously hard to massify (see Gathercole 1986, for an early discussion of these matters).

My goal in this paper is to sketch a theory of these three forms of variation, which are summarized in (3):

- (3) Three forms of variations in the mass/count phenomenology
 - a. Macrovariation in how numerals combine with mass vs. count nouns (i.e. the typology in (2)).
 - b. Variation affecting Fake Mass nouns, that appears to blur the cognitive basis of the mass/count distinction.
 - c. Variation in potentially ambiguous nouns and (re)conceptualization.

It should be underscored that one can easily conceive more options than those in (2). For example, a priori there might be a Type II language with clear cases of Fake Mass nouns. Or one could expect there to be mixed systems: e.g. a system which is basically like English, with two types of Ns, but where a subset of the cognitively mass nouns behaves like in Yudja, i.e. allow for direct combination with numerals, under the interpretation *quantity of* without standardization. So far, systems of this sort are unattested. In working out a proposal, it is worth keeping in mind that it has proven useful to work with frameworks that err in the direction of restrictiveness as opposed to frameworks that allow to describe attested and unattested patterns with equal ease.

One can talk of the forms of variations in (3) in radically different ways. At one extreme, one might be tempted to say that while Type I languages divide their lexica in a count and a mass region, Type II and Type III languages don't: in Type II languages every noun is mass, while in Type III languages every noun is count. At the opposite end, one can try to maintain that the mass/count distinction is universal, modulo fairly minor socially driven variations in the way potentially ambiguous concepts are lexicalized. And intermediate positions are also conceivable. I have already given some reasons (if tentatively, pending further empirical inquiry) for following the second perspective. I will explore the thesis that the only 'substantive' form of variation is the one that involves (re)conceptualization of potentially ambiguous nouns. If this is so where do the striking grammatical differences between Type I-III languages come from? The thesis to be developed is that they come from two sources: type theory and covert uses of certain general classifiers (which are grammatical formatives in their own right). Let me draw a couple of analogies to give some sense of where I am headed. Certain concepts can be intensionally coded in very different ways, even though they are extensionally equivalent. For example, relations can be viewed as a set of ordered pairs, or as a 'Curried' function-valued function. Sets of ordered pairs vs. Curried functions differ only intensionally. By the same token, a given concept, say *piece(s) of furniture*, can be coded in different ways, which in turn may affect the

way in which it undergoes operations like pluralization and counting, while ‘referring’ to one and the same thing. etc. I think that the difference between Type I and Type II languages and the phenomenon of Fake Mass nouns are to be thought along these very lines: languages exploit differently certain semantic types/categories that are extensionally equivalent from a mathematical stand point. Another familiar way in which languages vary is in whether they exploit grammatical resources like pronouns only overtly or also covertly, a prime example of the latter being pro-drop phenomena. I think that at the basis of Type III languages is the covert use of a classifier like function analogous to *(natural) quantity of*.

Whether or not one is able to articulate this still quite tentative thesis cogently, or even just coherently, the mass/count distinction keeps being a very useful testing ground for theories of meaning. Because we are constantly finding out more empirical aspects of the problem and they are game changing as they bring to light prima facie stunning forms of variation. And how can we have a universal theory of meaning, counting, etc. in light of that much variation? At the same time, as our grasp of semantics grows, our theoretical options may be getting a little clearer. In working out a framework for understanding the variations in (3), I will try to abstract away from what is the ultimate determinant of the mass/count distinction in the sense of the approaches in (1), to the extent that it is possible. But I won’t be able to hide my belief that some of the approaches in (1) appear to be better suited than others in understanding the nature of the generalizations in (3).

2. A Base-line framework.

In the present section, we review how approaches to the mass/count distinction grow out of theories of plurality. All modern theories of plurals employ a domain U of individuals ordered by a relation ‘ \leq ’ and closed under a plural sum operation ‘+’, a practice we will follow as well. The assumptions about plurals we will make for the purposes of this paper are meant to be uncontentious, to the extent that anything is.³ Theories of plurals based on structures $\langle U, \leq \rangle$ rely for counting on a notion of (‘structural’) atomicity. (Absolute) atoms are the members AT of U that are minimal with respect to the ordering \leq (i.e., $AT(x) =_{df} \forall y [y \leq x \rightarrow x = y]$).⁴ Properties are modelled in the usual way, as functions from worlds to (characteristic functions of) subsets of U . The introduction of properties affords us a useful notion of ‘relative’ atomicity: an individual x is an atom relative to P (a P -atom) in w iff no other individual of which P is true in w is a proper part of x . Generally, when we combine numerals with a property P , we count P -atoms (e.g. *three bears* means three individual/atomic things of which the property *bears* is true). A standard assumption is that the singular property *bear* is true of individual bears, i.e. bear-atoms. The plural property *bears*, instead, is true of any sum of bears, i.e. it is closed with respect to ‘+’, or, equivalently, it is a *cumulative* property. The singular property *bear* can be viewed as the generator set of its plural counterpart *bears*. We will also say that a singular property P is ‘quantized’, because all of the things of which it is true have the same size with respect to how P -

³ In particular, we adopt a lattice-theoretic terminology, a la Link 1983. But a set theoretic one (cf. Landman 1989, Schwarzschild 1996) would work equally well for our purposes.

⁴ We will assume here that U and \leq are pragmatically set; once pragmatically set, they remain constant across worlds (i.e. in modal reasoning). This is different from Chierchia (2010) where both U and \leq are taken to be relativized to worlds. The present approach is meant as a slight (?) simplification of the view developed there.

atoms are counted: each thing of which *bear* is true in any w counts as *one* bear. Its plural counterpart *bears* is not quantized, because it applies to sums of varying numbers of bear-atoms. These considerations take us to the following morphisms between properties, often used in counting:

- (4) a. $AT(P)$ ‘extracts’ from P the P -atoms:
 $AT(P) = \lambda w \lambda x P_w(x) \wedge \forall y [P_w(y) \wedge y \leq x \rightarrow x = y]$ Type: $\langle\langle s, et \rangle, \langle s, et \rangle\rangle$
- b. $*P$ closes P under sum:
 $*P = \lambda w \lambda x \exists Y [Y \subseteq P_w \wedge x = +Y]$ Type: $\langle\langle s, et \rangle, \langle s, et \rangle\rangle$
 where Y is of type et and $+Y$ is the sum of the extension of Y
- c. Numerals (first approximation)
 i. $3(\text{bears}) = \lambda w \lambda x \exists Y [Y \subseteq AT(\text{bears})_w \wedge |Y| = 3 \wedge x = +Y]$
 ii. $3 = \lambda P \lambda w \lambda x \exists Y [Y \subseteq AT(P)_w \wedge |Y| = 3 \wedge x = +Y]$
- d. Numerals (presuppositional version)
 i. $3 = \lambda P: *(AT(P)) = P. \lambda w \lambda x \exists Y [Y \subseteq P_w \wedge |Y| = 3 \wedge x = +Y]$
 (Plural agreement languages; e.g. IE languages)
 ii. $3 = \lambda P: AT(P) = P. \lambda w \lambda x \exists Y [Y \subseteq P_w \wedge |Y| = 3 \wedge x = +Y]$
 (Singular agreement languages, e.g. Ugro-Finnish, and Type II-III languages)

The world argument is notated as a subscript on a property P . Whenever the world argument is either understood or irrelevant, we will ignore it. $AT(P)$ pulls out P the P -atoms (which may or may not be atoms in the absolute sense). For singular properties like *bear*, it holds that $AT(\text{bear}) = \text{bear}$. The characterization of numerals in (4c) is a first approximation that needs comment. For one thing, we are assuming that *numeral* + N combinations have a basic predicative meaning of type $\langle s, et \rangle$ and a predictable Generalized Quantifier variant obtained via \exists -closure.⁵ Second, the number marking on NPs in *numeral* + N combination varies across language type: in IE languages, NPs (other than in combination with the first numeral) are plural marked; in Ugro-Finnic and Turkic languages NPs are always marked in the singular, in combination with any numeral. This variation can be modulated, e.g. by adding suitable domain restrictions in the definition of numerals, as illustrated in (4d). Third, the definition of numerals should be made compatible with how complex numerals like *thirty three* are compositionally built up. An important reference in this connection is Ionin and Matushansky (2018). The definition in (4c) is designed to be implementable within their approach, but to do so here would take us too far afield. And fourth, we are ignoring here the event-argument; taking it into account should not change dramatically the architecture of the present proposal.

The sketch just given reflects a widespread view on the nature of counting, which relies on a notion of (*relative*) atomicity. When mass nouns enter the scene, a notion of ‘non atomic’ property needs to be determined and this is where theories diverge widely, as sketchily indicated

⁵ Here is one way of doing so:

$$(a) \ 3(\text{bears})_{GQ} = \lambda P \lambda w \exists y [3(\text{bears})_w(y) \wedge P_w(y)] \\ = \lambda P \lambda w \exists y \exists Y [Y \subseteq AT(\text{men})_w \wedge |Y| = 3 \wedge y = +Y \wedge P_w(y)]$$

(It is also possible to assume that the GQ version of *numeral* + N combinations is the basic one and extract the predicative variant out of it).

in (1). The general idea is that non atomic properties are not good companions to numbers because they somehow don't allow access to units that can be readily counted. Non atomic properties like *water*, *blood*, etc. apply to samples that can be measured. But the measure has to be 'external' to the property, unlike what happens with atomic properties. In other words, water samples cannot be counted because they lack an atomic structure (in some sense to be specified), but their size can be measured using some general measure suitable for liquids. The problem is how to spell out this very rough intuition about (un)countability, and the theories in (1), or rather the slogans through which they are summarized, are different attempts to do so.

2.1. Atomic vs non atomic properties.

I will now outline, briefly and informally, the view on (non)atomicity I favor, just to give some intuitive content to it. But as I said, much of what is proposed on variation would go through on other approaches. On Chierchia's (2010, 2017) proposal, lack of atomicity has to do with vagueness: while the nature of the units associated with count properties are clear enough to be counted, those associated with mass properties aren't. To spell this out, worlds are taken to be ordered with respect to the *standards of precision* prevailing in them; $w \propto w'$ is to be understood as conveying that the standards of precisions in w' are at least as sharp (and possibly sharper) than those in w .⁶ Properties in general tend to be partial: any (non empty) P in any w has a positive extension (the things of which P is true in w) a negative extension (the things of which P is false in w) and typically also a vagueness band (things for which P is undefined in w). Precifications of w (i.e. $\{w' : w \propto w'\}$) are worlds in which the vagueness of each property P is resolved (partially or totally), by sharpening the criteria for having P : i.e., some things for which P is undefined in w , get to be assigned to the positive or negative extension of P in w' . Precifications affect vagueness in a monotonic fashion with respect to the ordering \propto : if u is in the positive (or res. negative) extension of P in w and $w \propto w'$, then u has to stay in the positive (resp. negative) extension P in w' . 'Base' worlds are those that are ' \propto -minimal'.⁷ Intuitively, they are the worlds where standards of precisions are set (typically through implicit conventions) by a community of speakers so as to ensure reliable and successful communication. Thus Base-worlds are historically and socio-pragmatically determined.

Here is how this frame is meant to help us understand vagueness in general and 'massiness' in particular. A property like *table* in any base world w will be true, say, of your dining table; it will be false of your chairs; it will also be false of the left half of your dining table. However, it can be undefined for the old table with three legs missing sitting in your garage. You need not take a stand as to whether that is (still) a table or not. This means that for certain purposes (e.g. if you intend to prop it somehow and use it for emergencies) it may still count as a table; while for other purposes (e.g., if you plan to use to as burning wood) it may not count as a table. Contrast this with the water you just spilled that now forms a puddle on the floor. It is definitely water, while, e.g., the floor it sits on is not water, but wood. Half of that puddle is also definitely water. But as you consider smaller samples, you become uncertain. There rapidly come parts of the puddle that are too small to be identified/recognized/perceived and one doesn't know whether

⁶ As usual, ' \propto ' is taken to be reflexive, transitive and antisymmetric. The basic idea for this formalization of vagueness is based on supervaluations (e.g. Fine 1975) and from 'data-semantics' (Veltman 1985; see also Landman 1991).

⁷ More explicitly, w is a Base-world iff for any w' , $w' \propto w \rightarrow w = w'$.

the property *water* applies to them in *w*. In any run of the mill Base-world, there are pretty small (e.g. droplet-sized) water amounts to which *water* applies; but we know that are a precisifications of the water-property in which parts of those water droplets may also turn out to be water or not as the case may be. This knowledge (or lack thereof) is what blocks combining numerals directly with water. *Three water(s)* doesn't work, in languages like English, because *three* looks for reliable/stable/lexically determined, water-atoms. But we know that all perceivable water samples may turn out to be aggregates, sets of smaller water samples, once sharper criteria are applied. What about water single H₂O molecules? Does that count as *water*? Honestly, I am not sure it does. Possibly, in some precisifications of the epistemic state (i.e. set of Base-worlds) common to my fellow speakers, it does. But in actual communication, the issue just doesn't arise. We live happily by leaving the matter unsettled: in Base-worlds we assume that *water* is undefined for water molecules.

So here is a formal candidate for a useful notion atomicity, relevant for counting. A property *P* is count iff there are (possible) Base-worlds *w*, in which $AT(P)_w$ is non empty and, for any precisification *w'* of any base world *w*, $AT(P)_w \subseteq AT(P)_{w'}$; a property is mass iff it is not count. Being not count entails that what may qualify as a smallest *P*-sample is some base world *w*, turns out to be an aggregate (i.e. a sum of smaller *P*-samples) in some precisification thereof. Any non empty property *P* has a non empty generator set $AT(P)$ in some base world *w*; only count properties have *stable* generator sets, for which I will use the boldface $AT(P)$, which grow monotonically across precisifications:

- (5) a. $AT(P) = \lambda w \lambda x P_w(x) \wedge \forall y [P_w(y) \wedge y \leq x \rightarrow x = y]$ P-atoms
 b. $AT(P) = \lambda w \lambda x P_w(x) \wedge \forall y \forall w' [w \propto w' \wedge P_w(y) \wedge y \leq x \rightarrow x = y]$ Stable P-atoms
 c. Modified definition of numerals
 i. $3 = \lambda P: *(AT(P)) = P. \lambda w \lambda x \exists Y [Y \subseteq P_w \wedge |Y| = 3 \wedge x = +Y]$
 (Plural agreeing languages; e.g. IE languages)
 ii. $3 = \lambda P: AT(P) = P. \lambda w \lambda x \exists Y [Y \subseteq P_w \wedge |Y| = 3 \wedge x = +Y]$
 (Singular agreeing languages, e.g. Ugro-Finnish, and Type II-III languages)

$AT(table)$ (or $AT(heap)$)⁸ by hypothesis/definition is non empty, in at least some possible Base-worlds; $AT(water)$ by hypothesis/definition is empty in any Base-world. Numerals need stable atomicity to work in construction with a property, as seems plausible enough and spelled out in (5c). Accordingly, *three(bears)* is a well defined, contingent property, but *three(water)* is going to be the logically empty property.⁹

Base-worlds for *water* are worlds in which *water* is true of vaguely specified (e.g. droplet-sized or larger) amounts of water. This entails that minimal water amounts are going to overlap, a feature the present approach shares with, e.g., Landman's (2011, 2016) take. However, natural Base-worlds for *rice* or *sand* include some in which we regard whole grains as minimal rice amounts, and we leave, e.g., half grains in the vagueness band. In such worlds, minimal rice amounts are *not* going to have material overlaps. I regard this as a plausible consequence of the

⁸ I refer to Chierchia (2010, 2017) for a discussion of the conceptual and formal differences between mass nouns, with vague atoms, and inherently vague concepts like *heap* or *cloud* with context dependent, but stable atoms).

⁹ See Gajewski (2002, 2009) and Chierchia (2013, Ch 1) a.o. for how logical falsehood may systematically give rise to ungrammaticality.

view just sketched, and as a (small) advantage over Landman's theory that links massiness to overlap.¹⁰ Be that as it may, readers are welcome to replace, for the purposes of the following discussion, the present notion of stable atomicity with their preferred one and explore the extent to which the goal of achieving a restrictive theory of cross-linguistic variation may be reached.

2.2. Ambiguous concepts and reconceptualization.

Armed with *some* formally explicit notion of count vs. mass property, we can tackle the issue of ambiguous nouns like *rope*, *beer* or *chicken*. The basic idea is that the relevant concepts come in related pairs. On the one hand, *rope* is associated with a sharp enough notion of minimal unit: any continuous, undetached string of material suitable for tying occurring in nature. On the other hand, it also comes with a more vaguely specified notion of amount of rope, possibly undetached from a larger rope portion, in principle sufficient to tie up something. By the same token, *chicken* on the one hand can be used to talk about whole animals, and on the other hand to talk about chicken parts, viewed primarily as food. Similarly for *beer*, etc. Much work on the mass/count distinction (e.g., Pelletier and Schubert 1989, Landman 1991, Krifka 1994) provides us with a natural way of thinking of this phenomenon in terms of (partial) functions that link a mass or count property P to its natural counterpart:

- (6) a. Packaging: For any property mass property P_M , $\Sigma(P_M)$, if defined, is the count property corresponding to P_M .
- b. Grinding: For any count property P_C , $G(P_C)$, if defined, is the mass property corresponding to P_C .
- c. 'Axioms' on Σ/G : i. $\Sigma(G(P_C)) = P_C$ ii. $G(\Sigma(P_M)) = P_M$
E.g.: $\Sigma(G(\text{beer}_C)) = \text{beer}_C$; $G(\Sigma(\text{beer}_M)) = \text{beer}_M$

An important condition on Packaging is that $\Sigma(P)$, when defined, is generally true of *conventional* or *standardized* units of P-amounts, not just of any natural quantity of P occurring in nature. For example, if beer_M is mass, $\Sigma(\text{beer}_M)$ is true of glasses or bottles of beer, not of drops or puddles thereof.¹¹ In the same spirit, if chicken_C is count, $G(\text{chicken}_C)$ is true of chicken parts that qualify as food (e.g., it wouldn't apply to, say chicken blood, or to chicken DNA).

The way in which one can imagine these morphisms to be used is that each community of speakers adopts slightly different variants $\Sigma, \Sigma', \Sigma'', \dots$ G, G', G'', \dots of packaging and grinding functions, defined for different properties, perhaps in slightly different ways. For example, standard Italian uses systematically a count version of the noun for *hair*, namely $\text{hair}_C = \Sigma(\text{hair}_M)$; it doesn't really use $\text{hair}_M = G(\text{hair}_C)$; so we could say $G_{\text{ITALIAN}}(\text{hair}_C)$ is undefined. And one can imagine a society of vampires where $\Sigma_{\text{vampire}}(\text{blood}_M)$ is defined. Intralinguistically, it looks like use of grinding and packaging may be tied in sometimes idiosyncratic ways to

¹⁰ Moreover, all group-level nouns like *group*, *bunch*, *set*, *quantity*,... are going to allow for overlaps in their parts while being perfectly count, which forces a theory like Landman's to distinguish 'bad for counting' overlaps from 'good for counting' ones.

¹¹ Another standard packager often used is bases on the notion of *kind* or *type of P*, exemplified in sentences like:

(a) There are only three wines I like, pinot, chardonnay, and malbec.
Adding it to the present framework presents no particular problem.

specific constructions. So, as we mentioned, English uses $hair_M$ predominantly, but in it allows *I found three blond hairs (= $*\Sigma_{ENGLISH}(hair_M)$) on your jacket*. And so on.

It seems reasonable to speculate that an approach based on packaging and grinding provides us with a good away of accounting for concept ambiguity and reconceptualization. Further formalization of these phenomena, which appear to occur to some extent in every language, may eventually be worked out, but it would be perhaps premature within the horizon of our current concerns. What is important to underscore is that functions like Σ and G , that link variants of closely related properties, bring about a change in the basic units to which properties apply: $\Sigma(beer_M)$ has an extension distinct from that of $beer_M$, in any world. To the extent that Fake Mass nouns like *furniture* apply to *whole* pieces of furniture and sums thereof and its extension coincides in any world with that of the Italian count concept $MOBILE_C$ ‘(whole) piece(s) of furniture’, treating them via extensions of packaging and grinding would appear to be unwarranted.

2.3. Quantized properties vs sum-closed ones; count kinds vs. mass kinds.

In sections 2.1-2.2 we discussed how count properties like $bear_C$ are taken to be true of individual members of U that retain an atomic status in any world. Quantized count properties like $bear_C$ generate via the star-operator their sum closed counterpart $*bear_C$. I will stick to the graphic convention using small characters for quantized properties like $bear_C$ and caps for their closure under sum $BEAR_C = *bear_C$. The distinction between generator sets and sum closed properties has a reflex in the domain of mass properties, at least for some theories. Usually, a mass property is taken to be sum-closed, e.g. $BLOOD_M$. On some approaches,¹² sum-closed mass properties will have a generator $blood_M$ true in any world only the minimal blood samples in that world, whatever they may be. It is controversial whether generators of mass properties are used in the languages of the world. I will argue below that they are - following Renans et al. (2018) - in languages like Modern Greek that allow for systematic pluralization of mass nouns, with an ‘abundance’ meaning;¹³ such languages are typologically not so infrequent.

NPs in English are property denoting when used predicatively (e.g., *that is blood*) or to restrict quantifiers (e.g., *I found some gold*), but in some argumental positions they have been argued to refer to kinds:

- (7) a. Gold is in short supply
- b. Bears are widespread

No single bear or single piece of gold can be in short supply or widespread per se: these properties have to do with the distribution of instances of kinds across space and time. Considerations of this sort have led to the conclusion that sum-closed properties stand in one-one correspondence to kinds. Kinds can be taken as primitives; but is also natural to regard them as functions from worlds into the total sum of the instances of that kind. Following this second option, Chierchia (1998) proposes the a way of linking sum-closed properties and kinds:

¹² In, e.g., all of those in (1) except Bunt (1979) and Link (1983).

¹³ Here is an example, from Renans et al (2018):

(a) Trehun ner-a apo to tavani
Drip-3PL water-PL from the ceiling ‘a lot of water is dripping from the ceiling’

(8) a. $\cap P = \lambda w. \iota x P_w(x)$ defined only if P is a sum-closed property

b. $\cup k = \lambda w \lambda x. \begin{cases} x \leq k_w, & \text{if } k_w \text{ is defined} \\ 0, & \text{otherwise} \end{cases}$

c. Examples:

i. $\cap \text{BEAR}_C = \lambda w. \iota x \text{BEAR}_{C,w}(x)$

ii. $\cup \cap \text{BEAR}_C = \cup \lambda w. \iota x \text{BEAR}_{C,w}(x) = \lambda w \lambda x. x \leq \iota x \text{BEAR}_{C,w}(x) = \text{BEAR}_C$

iii. in short $\text{supply}_w(\cap \text{GOLD}_M)$

iv. $\text{widespread}_w(\cap \text{BEAR}_C)$

Notice that the \cap -operator is undefined for quantized properties like *three bears* or the singular *bear_C*. Sentences like (7a-b) have the logical form in (8c.iii-iv).¹⁴ Kinds on this view are technically individual concepts of type $\langle s, e \rangle$, but I am often going to abbreviate the type of kinds as e_k . The functions defined in (8a-b) make explicit an isomorphism between the notion of sum-closed property and that of kind. While a sum-closed property is a function of type $\langle s, et \rangle$ that in any world maps any sample/instance of the property into truth, a kind is of type $\langle s, e \rangle$ and in any world directly picks out the maximal sum of what the corresponding property is true of. In a sense, sum-closed properties and kinds constitute different ways of coding the same information.¹⁵ Their types (and hence their logico-linguistic roles) are different; but there is a

¹⁴ The one in (7) constitutes the simplest (and least controversial) forms of kind predication, that involves kind level predicates. There are also cases where bare NPs occur as argument of object-level predicates:

(a) Bears hibernate in winter

(b) Bears were just sighted in this area

On one analysis, the bare nominals in (a)-(b) are still kind denoting, but they give raise to a quantificational interpretation, determined, essentially, by the aspectual properties of the relevant sentences. Generic sentences like (a), are understood as saying something about all bears; episodic sentences like (b) seem to involve some instances of the bear-kind. See Chierchia (1998), Dayal (2004) a.o. for analyses along these lines. For comparisons with other approaches, cf., e.g., Dayal (2011).

¹⁵ The only difference is that in a world w with, e.g., no bears, $\text{BEAR}_{C,w}$ is the empty characteristic function, while $\cap \text{BEAR}_C(w)$ is undefined. That's the reason for the disjunctive characterization of the \cup -operator in (8b).

It should also be born in mind that there are different notions of kind. In particular, Dayal (2004) argues that singular definite generics (like *the dinosaur* in *the dinosaur is extinct*) denote *taxonomic* kinds that have a group-level denotation, distinct from that of regular kinds. Moreover, Carlson (1977), who has spearheaded the relevance of the notion of kind for semantics, thought that NPs like *people in the next room* are not kind-denoting, on grounds that sentences like (a) are deviant:

(a) ?? People in the next room are rare

(b) People like those in the next room are rare

However, things are more complex than that, as (b) illustrates. In fact, Dayal (2013) suggests that there are reasons for considering NPs like those in (a)-(b) kind denoting as well, and dubs them 'indexical kinds'. I am going to follow here Dayal, and assume that the correlate of *any* sum-

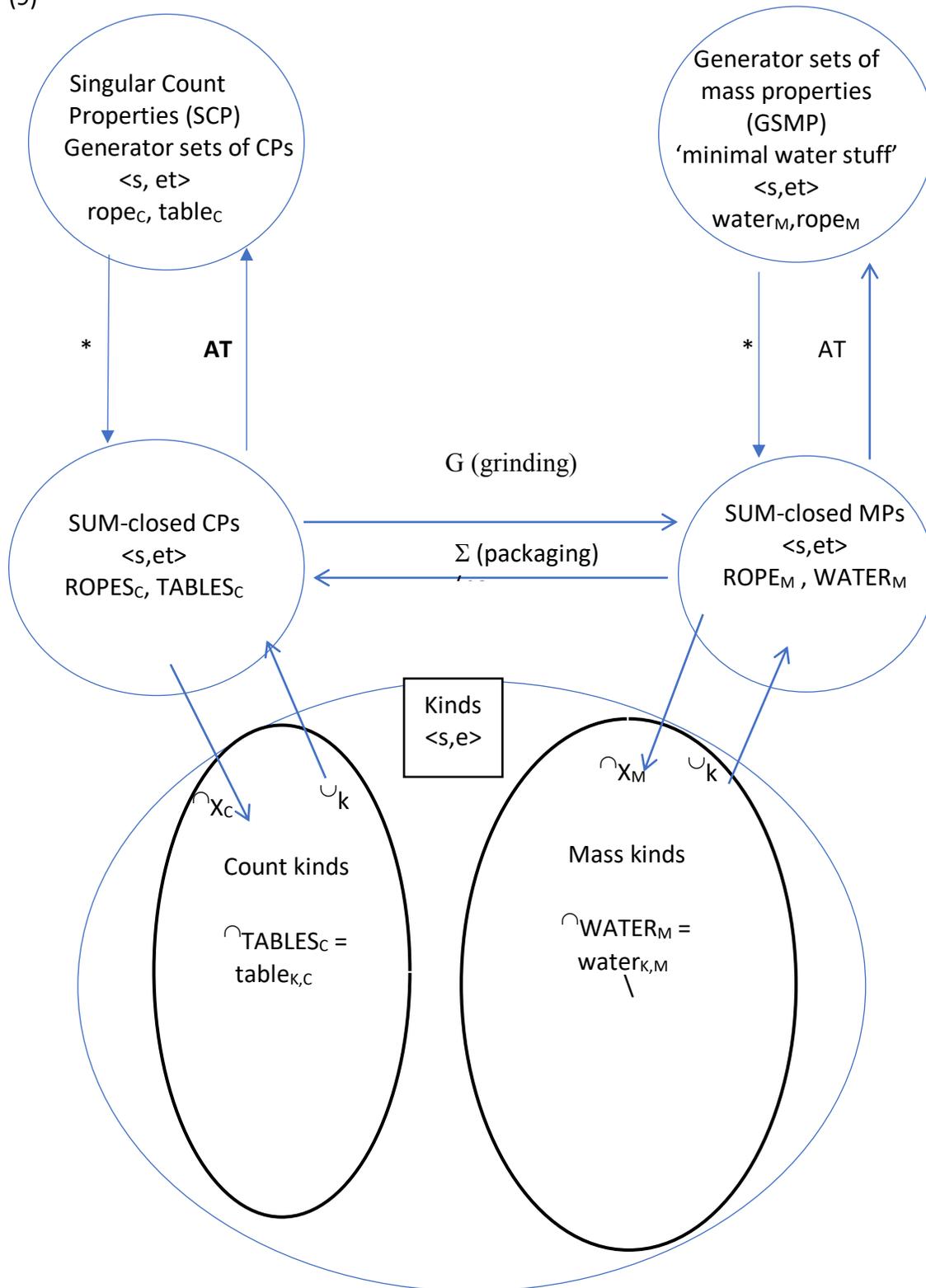
clear isomorphism between them, which the pair of operations in (8) seeks to capture. This will play an important role in theories of variation. And just like there are count and mass properties, so there will be count and mass kinds, a straightforward consequence of there being two kinds of properties, and an isomorphism between any sum-closed properties and a (certain notion) of kind. This set up can be visualized in the following diagram:

INSERT FIGURE 1

closed property is a kind. Predicates like *rare*, or *extinct* are sometimes marginal with indexical kinds, because they select for something like ‘well-established’ or ‘regular’ kinds, i.e. kinds for which some sort of regular behavior can be contextually identified. This hypothesis enables us to maintain a full-blown isomorphism between sum-closed properties and kinds.

Figure 1

(9)



Most of this apparatus is meant to be weakly theory bound. Every theory countenances count vs. mass properties and some ways of packaging and grinding them; and every theory should have room for some systematic correspondence between sum-closed properties and kinds. A point of controversy is the existence of generator sets for mass properties, which some theories (e.g. Link or Bunt) do not countenance. But the greater controversy concerns how to use an apparatus of this sort in understanding variation across languages (and, I think, also diachronic change). In what follows I am going to refer to structures like those in (9) as ‘Base-line (Mass/Count) structures/theories.’

3. What varies and what doesn’t.

A preliminary general issue should be addressed before getting into specifics. I think most authors would regard Base-line structures like (9) as a ‘logic’ in the same sense as, say, Link’s classical work: it constitutes a set of categories, with specific relations and operations on them that precipitate a notion of logical consequence, once a mapping onto language is provided. Where this logic comes from remains a matter of speculation for now (certainly within the limits of the present paper). My inclination is to regard Base-line theories as a ‘schematism’ that supervenes on our natural computational capacities (i.e., something like ‘merge’ and ‘copy-merge’ or the untyped λ -calculus) under the pressure of a preexisting conceptual space of ‘substances’ and ‘objects’. To put it somewhat coarsely, our computational endowment enables us to readily cast and generalize the pre-existing ‘natural’ divide between objects and substances. In particular, my notion of atomicity, which is linked to vagueness in a formally explicit way, maps directly onto the notions of ‘Spelke object’ vs. ‘Spelke substance’: a Spelke object is something whose identity conditions are clear enough to allow tracking across space, and to enable operations like summation without loss of individuality; a Spelke substance is made up of entities are not so finely identified (because of vaguer identity criteria). Accordingly, on my approach, a Spelke substance *has* to be coded as mass, a Spelke object *has* to be coded, at some level, as count. In this sense, my theory is arguably more restrictive than its competitors, and any apparent mismatch with the cognitive mass/count contrast is a *prima facie* problem for it. We will deal with some such apparent mismatches in Sections 3.1-2. On any other theory in (1), it seems to me that that the link between the logico-linguistic characterization of atomicity and the cognitive one is way more flexible, and ultimately arbitrary. But the latter is not a picture supported by the available linguistic evidence, in so far as I can tell, and what follows is an attempt to substantiate this claim.

3.1. Classifier languages.

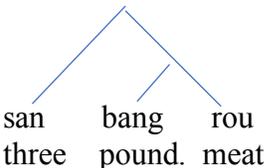
It is useful to start exploring variation through the lenses of Type II languages. Besides being by now fairly well studied, these languages will give us the opportunity to discuss some of the different notions of classifier and classifier phrase that play a role in the mass/count distinction. In section 2, we have given an analysis of numerals as property modifiers of type $\langle\langle s, et \rangle, \langle s, et \rangle\rangle$, in line with Ionin and Matushansky and many others. Suppose this is so universally. Suppose, furthermore, that Ns in classifier languages are kind denoting. It follows that Ns in these languages can never directly combine with numerals because of a type mismatch. One needs to interpolate some operation that affects the type of NPs, turning them into properties to allow for numerals to modify them. And one might expect such type changing

operations to be typically overtly morphologically realized. This arguably provides us with a natural function for classifiers. One very simple way of turning kinds into properties is by extracting their instances via the **AT**-function along the following lines:

- (10) a. $ge = \lambda x_k. \mathbf{AT}(\cup x_k)$ Type: $\langle e_k, \langle s, et \rangle \rangle$
 b. $ge(\text{ren}_{K,C}) = \mathbf{AT}(\cup \text{ren}_{K,C}) = \text{person}_C$
 where $\text{ren}_{K,C}$ is the (count) person-kind

Classifiers of this sort make Ns capable of combining with numerals, as *classifier-N* combinations like (10b) wind up (i) with the right semantic type to be arguments of numerals and (ii) with (stable) atoms as their extensions (and hence capable of satisfying the atomicity presupposition typical of numerals). Classifiers like *ge*, will be undefined for mass nouns, as the latter lack (stable) atoms. Hence the function in (10) constitutes a strong, wholly general candidate for what Cheng and Sybesma call count-classifiers (and are also sometimes called ‘individual’ classifiers).

There are various lines that can be taken on what Cheng and Sybesma call ‘massifiers’, by which they mean classifiers that are not inherently restricted to count nouns. With respect to the latter, I will pretty much follow here the line developed by L.J.Jiang (2011, forthcoming). In illustrating it, I will limit myself to a few basic observations and also stick to English as a metalanguage, for ease of illustration. I am going to consider just two main kinds of classifier-functions: measure phrases and container phrases. The former can be exemplified as in (11a), the latter as in (11b):

- (11) a. $\text{pound}: \lambda x_k \lambda n \lambda w \lambda y. \cup x_{k,w}(y) \wedge \mu_{PD,w}(y) = n$ Type: $\langle e_k, \langle n, \langle s \text{ et} \rangle \rangle \rangle$
 Where ‘n’ is the type of numbers and $\mu_{PD,w}$ is a measure function that maps individuals in a world into their weigh in pounds.
 b. $\lambda y. \cup \text{meat}_{k,w}(y) \wedge \mu_{PD,w}(y) = 3$

 c. i. $\text{yibei} = \lambda y \lambda w \lambda x. \cup \text{cup}_{k,w}(x) \wedge \text{fill}_w(x)(y)$ Type $\langle e_k, \langle s \text{ et} \rangle \rangle$
 There were three cups of coffee (on the table) \Rightarrow
 $\exists y \exists 3(\lambda w' \lambda x. \cup \text{cup}_{k,w'}(x) \wedge \text{fill}_{w'}(x)(\text{coffee}_{k,w'}))_w(y)$
 $= \exists y \exists X [X \subseteq \mathbf{AT}(\lambda x. \cup \text{cup}_{k,w}(x) \wedge \text{fill}_w(x)(\text{coffee}_{k,w})) \wedge y = +X]$
 ii. $\text{yibei} = \lambda x_k \lambda n \lambda w \lambda y. \cup x_{k,w}(y) \wedge \mu_{CUP,w}(y) = n$ Type: $\langle e_k, \langle n, \langle s \text{ et} \rangle \rangle \rangle$
 John drank three cups of tea \Rightarrow
 $\exists y [\cup \text{tea}_{k,w}(y) \wedge \mu_{CUP,w}(y) = 3 \wedge \text{drank}_w(y)(\text{john})]$

A measure phrase combines with a kind *k* and a numeral *n* and return a property true of individuals or sums of kind *k* that measure *n*, by some specific measure, in this case a weigh-based one. To say that an individual (or a sum) *x* weighs *n* pounds is to say that it can be

partitioned into n one-pound sized parts.¹⁶ As can be seen by the types in (11), the inputs and outputs of measure phrase classifiers are the same as those of count-classifiers, but the former have a number as an additional argument. The function in (11a.i) naturally goes with the constituency in (11a.ii), which appears to be well motivated for languages like Mandarin, but will require some local modification for other languages. Container phrases-classifiers are very varied and have several readings (see Rothstein (2016), Jiang (*forthcoming*), and references therein), which we will not consider in detail; typically they involve relational nouns (like, *slice, part, quantity, head, ...*), or nouns associated with a ‘content’ (like, *cup, basket, bucket, but also drop, puddle, ...*). Two recurrent construals of this type of classifier are illustrated in (11b). The first interpretation is in a sense analogous to that of count classifiers, the second to that of measure phrases. The important thing for our purposes is that classifiers employ atomizing relational nouns or measure functions (i.e. functions from individuals into numbers) whereby a numerical value may be attached to entities; consequently, classifier phrases denote quantized properties.¹⁷

The main points of these cursory remarks are three. The first is that the obligatory interpolation of quantizing functions in Type II languages upon combining numbers with nouns can be understood/explained in simple type theoretic terms: Ns in such languages are kind denoting, and something is needed to map them into quantizable properties. **AT** is one such functions; measure functions are others; container phrases can be used either way (as **AT** forming and as measures). The second point is that one expects, under this type driven view, that the mass/count distinction will be reflected in the classifier system; in particular, since not all kinds are atomic, the **AT** function will be partial, and classifiers based on **AT** will be ungrammatical with mass kinds. And third, one would not expect there to be Fake Mass nouns, for that would involve making **AT** undefined for some count kind (so that, e.g. *ge* would not apply to them). Thus, the main typological features of Generalized Classifier languages fall into place, while sticking to the horizon of a strictly universal Base-line theory of the mass/count distinction.

There may well be other ways of deriving the existence of Type II languages from Base line structures. One that often finds credit¹⁸ is to assume that in these languages, Ns are not differentiated between count vs. mass, and classifiers are needed to introduce differentiations that make them useful in counting. On this view, Ns must denote properties P that are ‘disjunctive’ in that they are true of both P -atoms (stable units) and P -amounts (unstable units). I have two main arguments against this line of approach, one more conceptual, the other empirical. The conceptual argument is that properties appear to be universally differentiated in mass vs. count, modulo a few potentially ambiguous cases. People, teddy bears, artifacts, come organized in natural units to the pre-linguistic child; sand, water, etc. do not. Does it make sense that upon learning her language the child would obliterate such a robust distinction in favor of a genuinely mass/count neutral notion of *person* or *cat*? The thesis that any property can be mass/count

¹⁶ I.e., given that $n = \langle\langle s \text{ et} \rangle\rangle$, $\langle s \text{ et} \rangle\rangle$, we define $\mu_{PD,w}(x) = n$ as follows

(a) $\mu_{PD,w}(x) = n \equiv_{df} n(\lambda w' \lambda y \exists \Pi_{PD} [\Pi_{PD,w'}(x)(y)])(w)(x)$

Where Π_{PD} is a function that maps an individual (or sum) x in a world w into a set of non overlapping and jointly exhaustive parts of x each weighing one pound.

¹⁷ Also like *numeral-N* construction, measure phrases will have generalized quantifier variants, to be obtained via existential closure.

¹⁸ See, e.g., de Vries et al. (*this volume*) and references therein.

‘neutral’ is strongly reminiscent of Quine’s ‘ontological relativism’ thesis, which Spelke and collaborators have proven wrong. On the present account, the presence of generalized classifiers is type driven and does not affect the cognitive dimension of nouns.

The more empirical argument against the ‘undifferentiated property’ view is that properties are not of an argumental type: their main logico-linguistic role is that of being predicated of something, or perhaps that of restricting a quantifier. Thus, if the category N is mapped onto properties, one should expect there to be languages in which bare Ns, or at least some subclasses thereof, display a non argumental behavior, e.g. by being disallowed in all or some argumental positions. We’ll see shortly that Type I languages are indeed languages where these distributional restrictions on bare nouns occur extensively. But in contrast with this, not a single Type II language has been found that disallows bare NPs, *any* bare NP, from occurring in any argumental position. This is true even of IndoEuropean languages like Bangla that have developed a generalized classifier system (Dayal 2012). It is even true of those typologically rare Type II, generalized classifier languages, that have developed something like a definite article, such as Nuosu Yi, as documented in Jiang (2018). On an approach where the presence of an obligatory classifier is systematically tied to the fact that Ns are kind denoting, the argumental behavior of Ns is wholly expected and predicted.

3.2. Basic IndoEuropean: English and other Type I languages.

IE languages typically have a clear distinction between mass and count NP, where the latter can directly combine with numerals (without classifier phrases), while the former cannot. IE languages also have very extensive singular-plural (and sometimes dual, etc.) morphologies. A natural and widely agreed upon way of interpreting IE-like number morphemes is as regulators of the denotation of the Ns of type $\langle\langle s, et \rangle\rangle$, $\langle s et \rangle$. In what follows I will sketch how a system of this sort fits the Base-line theory, focusing first on count nouns.

3.2.1. *Count NPs in IE*. The following is a fairly standard implementation of the idea that number morphology ‘regulates’ NP denotations (cf. e.g. Sauerland 2003, Sudo 2014):

- (12) a. i.
- ```

graph TD
 A[NumP, cat_C] --- B[SG]
 A --- C[NP]
 B --- D[∅]
 C --- E[cat]

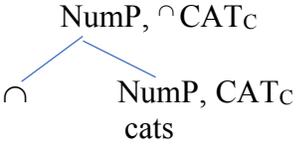
```
- ii.
- ```

graph TD
    A[NumP, CAT_C] --- B[PL]
    A --- C[NP]
    B --- D[s]
    C --- E[cat]
  
```
- b. i. $SG = \lambda P:AT(P) = P.P$
 ii. $PL = \lambda P:*AT(P) = P.P$

The trees in (12a) specify the syntax of number phrases (NumPs). The (functional) head of the phrase is the number morpheme, into which the (lexical) head of NP incorporates, via head-

raising. The SG morpheme is interpreted only once,¹⁹ but can be ‘exposed’ in multiple morphemes scattered throughout the DP (as is overtly visible in e.g. Romance languages), and triggers agreement phenomena on the VP. The semantics of number morphemes is that of restricted identity maps: they introduce presuppositions. The presupposition of singular NPs is that they are true of P-atoms (where P is the basic denotation of the N-head); the presupposition of plural NPs is that they are sum-closed.²⁰ The *-operator applies freely in the semantic composition, but the semantics of number features guarantees that the outcome is the correct one, so that, e.g., the singular NP *cat* winds up denoting the atomic property cat_C and the plural NP *cats* the sum closed property CAT_C , as annotated on the trees in (12a).

This hypothesis on the interpretation of number features is predicated on the assumption that, unlike what happens in Generalized Classifier languages, Ns in IE are property-denoting (and not kind denoting). An immediate consequence of this assumption is that (bare) NPs in IE cannot occur in argument position, without their denotation being further compositionally modified. This is a welcome result. Some IE languages, like French virtually disallow bare arguments in any position. Others, like English, disallow bare singular count NPs, but allow bare plurals. The latter circumstance is probably due to the covert use of argument forming operators as in the following example:

- (13) a. Cats are common
 b. 
 c. $common_w(\wedge CAT_C)$

The covert availability of the argument forming operator ‘ \wedge ’ is clearly a language particular option. It is immaterial for the present purposes whether this option is realized via a phonologically null determiner D, or by simply adjoining a null operator to NumP (as in (13b)). Notice that since ‘ \wedge ’ is restricted to sum-closed properties, this immediately predicts that bare singulars are disallowed in English.

The details of how argument formation takes place across various IE languages (including those that lack articles altogether and allow generalized bare arguments, like Hindi or Russian) are too rich (and controversial) to be dealt with here.²¹ What is important to note is that in (most) IE languages, NumPs are of the right type for combining with numerals; moreover, the presupposition embedded in numerals - that the Ns be (stably) atomic - is appropriately enforced by number morphology, as illustrated in (14):

- (14) a. $three\ cats \Rightarrow 3(\lambda P: *AT(P) = P.P(CAT_C)) = 3(CAT_C)$
 b. $three\ bloods \Rightarrow 3(\lambda P: *AT(P) = P.P(BLOOD_M)) = UNDEFINED$

¹⁹ We make this assumption for simplicity. See Sudo (2014) for the view that each occurrence of *number* is meaningful.

²⁰ The actual semantics of plurality and singularity in full blow sentences is much richer than what we can get in here. It requires among other things resorting to implicatures. See e.g. Spector (2007), and references therein.

²¹ See Dayal (2011) and references therein.

The *numeral-N* combination in (14a) yields a well-defined, quantized property, while the one in (14b) turns out to be deviant. This is a good start. But it rules out too much. Mass nouns have singular morphology. This is visible in English mostly through agreement, because in English the singular morpheme is null. But in many languages (e.g. in all the Romance languages) singular morphemes are overt, and clearly visible throughout within the DP. Number morphemes, on the present hypothesis, are linked to (stable) atomicity, a property that mass nouns lack and it is not obvious how to modify this assumption without creating havoc with count nouns.

A standard move in this connection is to assume that singular morphology is ambiguous. There is a ‘meaningful’ one, that combines with count Ns, and a ‘meaningless’ one that combines with mass nouns. Now positing ambiguities is certainly necessary sometimes (and we will see an instance of it in connection with Nez Perce), but it is hardly ever, per se, a source of insight. Moreover, it is very unclear under this view how one would account for the phenomenon of Fake Mass nouns, which, on the present approach, are as atomic as it gets. What could force them to take on the behavior of mass nouns?

3.2.2 *Mass nouns in IE*. Clearly something in the grammar of number marking needs to be modified for English and (most) other IE languages. One possibility that comes to mind is to weaken the meaning associated with number morphology, say, along the following lines:

- (15) a. i. $SG = \lambda P:AT(P) = P.P$ Original setting (stable atomicity gets checked)
 ii. $PL = \lambda P:*AT(P) = P.P$
 b. i. $SG = \lambda P:AT(P) = P.P$ New (proposed) setting (lack of +-closure gets checked)
 ii. $PL = \lambda P:*AT(P) = P.P$

On the new proposed setting, SG extracts the generator set from a property, without caring whether a property is (stably) atomic or not. (This will only work for approaches to atomicity in which mass properties do have generator sets). Similarly, pluralization checks whether a property is closed under sum, regardless of (stable) atomicity. In a language with this setting, we would expect any kind of noun, whether count or mass, to take on both plural and singular morphology. However, numbers have independently built into them a (stable) atomicity presupposition (cf. 5c above). Hence one would still expect mass nouns to be unable to combine with numerals. These seem to be features of Greek, where indeed mass nouns do *not* combine with numerals, but do pluralize freely, with an abundance interpretation. The latter should arise as an implicature, much like the *more than one* implicature comes about for plural count nouns (cf. Renans et al. 2018). So maybe the settings in (15b) do characterize some languages. But not English, nor the majority of IE languages.

For the latter, I am going to adopt an idea due to Giorgio Magri.²² The idea is to turn mass properties into properties that are (stably) atomic but in a sort of trivial manner, through a type theoretic trick which makes them retain their mass behavior with respect to pluralization and counting. That way, the singular morpheme can retain its original setting. Magri’s idea is that mass properties P can be coded as singleton properties P_{SGL} , i.e. functions of type $\langle s, et \rangle$ that at

²² Magri developed this idea in his undergraduate honors thesis in philosophy at the University of Milan.

each world are true of just of the maximal entity of which P is true (if there is one, else P is empty). One can define a ‘singulative’ function along the following lines:

- (16) a. $SGL(P) = \lambda P:P \in MASS \lambda w \lambda x. P_w \neq \emptyset \wedge x = +P_w$
 b. Suppose that the extension of P_w is the set $\{a, b, c, a+b, b+c, a+b+c\}$; then: the extension of $SGL(P)_w$ will be the singleton set $\{a+b+c\}$

Singleton properties are atomic, in that, when non empty, they are true of just one thing.²³ Note the similarity between Magri’s notion of singleton property and our notion of kind: they only differ from each other in a trivial way. In fact, this similarity predicts that our kind-forming operator ‘ \cap ’ should be able to apply to singleton properties generated via SGL, for they are necessarily sum-closed, and therefore different from the regular singular count properties, like *cat_C*, etc, which will be singletons in some worlds, but not in others. Hence, we expect mass nouns to undergo kind-formation as freely as their count plural cousins, which allow them to occur bare, without determiners, in argument position in spite of being morphologically singular. The general idea is that the SGL function applies freely (like the *-operator). Singular morphology makes sure that SGL applies to mass properties, so that they can pass through the number marking gate. At the same time, turning mass properties into singletons, while being just a type theoretic maneuver, will have consequences on semantic composition. I will now sketch the main ones, if only in informal terms. Since singleton properties are trivially closed under sum, it makes sense that the *-operator (i.e. pluralization) should not apply to them, on the ground that languages dislike trivial applications of morpho-semantic operators. Moreover, the cardinality of singleton properties P_{SGL} is clearly logically determined: whenever non empty, they necessarily contain one P_{SGL} -atom. This seems to constitute a plausible enough reason why they should be unable to combine with numerals. The numeral *one* is logically true of them, and any other numeral is logically false. These consequences of resorting to SGL (lack of pluralization, impossibility of combining with numerals) give us the basic behavior of mass nouns.²⁴

A further bonus of treating mass properties as singletons is that it creates a very natural niche for Fake Mass nouns: they too can be treated as singleton properties. This readily explains why they pattern with mass nouns with respect to pluralization, combination with numerals, quantification, etc., while retaining their atomic structure vis-à-vis less linguistically driven mental operations. In a sense, the existence of Fake Mass nouns can be viewed a copy-cat effect, a type-theoretic re-dressing of cognitively count properties: the SGL function, which is independently available for mass nouns, is idiosyncratically extended to some sum closed atomic properties.²⁵ The functional effect of this move is to de-emphasize their inherent atomicity: once

²³ Moreover, singleton properties are stably atomic in our sense: since precisifications grow monotonically, $AT(P_{SGL})_w$ has to be included in $AT(P_{SGL})_{w'}$, whenever w' is a precisification of w .

²⁴ Treating mass properties as singletons carries also requires some adjustments in how modification and quantification gets implemented. For example, the DP *some clean water* must have a denotation along the following lines:

(a) *Some clean water* $\Rightarrow \lambda P \exists y \exists x [water_w(x) \wedge y \leq x \wedge clean_w(y) \wedge P_w(y)]$

Arguably, these adjustments are independently needed for partitives (some of the water/apples).

²⁵ It is straightforward to extend the SGL so as to include some count, sum closed properties:

a sum-closed property is turned into a singleton one like *furniture*, the definite description *the furniture* will have to refer to the totality of furniture around; it won't be able to carry the uniqueness presupposition associated with true singulars like *the table*. I.e. we won't be able to use *furniture* to readily pick out singularities. This makes fake mass nouns particularly suitable for superordinate nouns, which collapse together more basic categories (like table, chair, etc.). An approach based on a type-theoretical re-dressing like the present one also makes sense of the fact that there doesn't appear to be any truth-conditional/referential difference whatsoever when a noun is coded in its canonical singular/plural count way (like Italian *mobile/i*) and when is coded as fake mass (*furniture*). Everything appears to kind of fall into place and we have a plausible account of a phenomenon of an apparent grammar/cognition mismatch.

The present take also has interesting cross-linguistic consequences. The phenomenon of Fake Mass nouns ought to be a side effect of the existence of a meaningful, strong singular morphology, and should therefore happen only in languages that do have the relevant morphosyntactic settings, which are not found in, e.g. Type II languages, where pluralization works very differently.²⁶ Moreover, they should not be found in languages where the singular/plural contrast, while pervasive, has the weak setting in (15b). Since singularity in these languages is not based on stable atomicity, but on lack of sum-closure, there is no pressure to resort to the SGL-function. Languages that allow pluralization mass nouns, therefore, ought to lack the Fake Mass noun phenomenon, and this is indeed the situation reported for Greek.

It may be useful a quick comparison with approaches to Fake Mass nouns which have been or could be developed by within some of the other theoretical lines in (1). Consider, for example, approaches like Link's or Bunt's. It is really hard to imagine how nouns like *furniture* or *footwear* could be viewed as being non atomic in the sense developed there: a piece of furniture, say a table ought to be made up of other pieces of furniture, all the way down. One could say that the sense in which something is non atomic is conceptual, and shouldn't be narrowly linked to our cognitive system. But if furniture can be conceptualized as non atomic in Link's sense, then surely any (other) Spelke object also could, equally well. And we basically lose the connection with between the grammatical mass/count contrast and the parallel one that seems to hold in pre-linguistic cognition. Not to mention the fact that one would expect the phenomenon of Fake Mass to occur randomly across any language. But it seems, instead, that it is tied to rather specific grammatical settings. Similar considerations apply, it would seem, to overlap based approaches, like Landman's. Furniture or kitchenware do not overlap more for English than for Italian speakers. But Italian has both a literal Fake Mass counterpart to *furniture* (namely, *mobilia*) next to a perfectly count near synonym (*mobile/i*), and lacks altogether a Fake Mass counterpart to *kitchenware*, *footwear*, etc. Now, in spite of this, one might try to develop a line where closely related concepts lexicalize different components of some natural class of objects: *mobile* goes for single, whole pieces of furniture, *mobilia* goes for sets of pieces of furniture that allow for overlap, somehow. Again, if this is so, Fake Mass nouns should be possible in any

$$(a) \text{ SGL}^+(P) = \lambda P:P \in \text{MASS} \cup D. \lambda w \lambda x. P_w \neq \emptyset \wedge x = +P_w$$

Where D is a subset of the set of sum-closed count properties such that $\text{FURNITURE}_C, \text{JEWEL}_C, \dots \in D$

Which count properties are coded as fake-mass has to be learned by the child on a case-by-case, language particular basis, which seems to be correct.

²⁶ See, e.g., Jiang (2017) for the syntax and semantics of plurals in Mandarin.

language, regardless of whether a language has a strong singular/plural morphology. The evidence we have so far does not support this expectation.

We have illustrated in broad strokes an approach to IE languages that at its core is fairly uncontroversial: Ns are property denoting, number features operate on properties and trigger presuppositions of singularity/atomicity vs. plurality. In these languages Ns clearly fall into two distinct categories, which we accounted for with the simple idea that numbers are property-modifiers with an atomicity presupposition built in. A problem that comes up in this connection is the role of singular morphology: it imposes atomicity requirements on count nouns (e.g., *the table* has a uniqueness presupposition that must be determined by its being singular), but no such requirement, it seems, on mass nouns. The standard move in this connection is to hypothesize that singular morphology is ambiguous between a meaningful version and a meaningless one, a move we have considered above. We have departed from this line, and considered what would it take to preserve the hypothesis that singular morphology is always meaningful. How could we modify the notion of atomicity in such a way that mass properties could count as atomic? We have come up with two hypotheses. The first is that for a property to be atomic is simply for it to be not closed under sum. The other is a type theoretic trick that rides on the partially ordered structure of the domain, namely to code a sum-closed property as a singleton property. We have seen that the first hypothesis matches up pretty well with Greek, where mass nouns pluralize but do not combine with numerals; the second hypothesis matches pretty well with English and the majority of the IE languages. On this second hypothesis, we have an explanation for the phenomenon of Fake Mass nouns, i.e. an apparent mismatch between grammar and cognition/perception. For *any* sum-closed property, count as well as mass, can be coded as a singleton property in an information preserving manner. So some count properties can be thus represented in the lexicon and thereby take on the behavior of mass nouns. Clearly, it is in the logic of the proposed system that the inherent differences between mass vs. count properties are never going to be obliterated: Mass properties *have to* be realized as singleton ones, to get around singularity checking; Count properties can readily pass singularity checking without the extra coding. So count nouns will choose the SGL option more rarely, perhaps as a device to de-emphasize individuals with respect to their aggregates.

In summary, while the departure from the traditional view that singular morphology can be meaningless leads to a representation of ‘massiness’ admittedly more complicated than usual, the overall outcome is arguably a more explanatory account, which provides a theory of variation fully in line with the Base-line take on the mass/count contrast.

3.3. Languages with generalized numerals: Type III.

The main characteristic of Type III languages is that they allow numerals to combine freely and directly with any kind of noun. The languages of this type documented thus far share some other features, besides the behavior of numerals: they lack determiners, and allow bare NPs in any argumental position. The patterns of pluralization across Type III languages seems to vary more. From the limited data reported in Lima (2014), it looks like pluralization in Yudja could bear similarities to that of Mandarin; as for Nez Pearce, there is a fairly detailed analysis of pluralization by Deal (2017), which reveals patterns more similar to those found in IE. All known Type III languages, however, are uniform in the way *numeral N* combinations are interpreted. If the N is count, no significant difference in interpretation emerges with respect to IE languages: numerals count the ‘natural units’ identified by N. In combination with mass

nouns, contextually relevant quantities or parts or units are counted, which may turn out to be bags, drops, piles, ... This naturally leads one to hypothesize that the generalized numeral option is somehow linked to the covert use of some very general classifier-like element similar to *quantity*. To see what is involved, it might be useful to start with a quick look at how context-dependent classifiers like *quantity* work in English and, on that basis, subsequently turn to Yudja and Nez Perce.

3.3.1 *Quantity*. In English, context-dependent classifiers like *quantity* combine with bare NPs (both count and mass – (17a)), in what are known as ‘pseudopartitive’ constructions. Use in partitives (17b) is also possible.²⁷ They also allow for an amount/measure construal, such as the one illustrated in (17c), which we will ignore here (see Scontras 2017 for an interesting take on them, consistent with the present approach):

- (17) a. There were two quantities of water/apples on the floor.
 b. One quantity of John’s blood was used for a transfusion for his brother.
 c. I consumed the same quantity of food you did.

In all their uses, quantity-phrases create count DPs, which justifies regarding them as context dependent classifiers. In terms of our approach to atomicity, this entails that the DPs created via context dependent classifiers of this sort identify stable atoms, suitable for counting. A possible approach, outlined in Chierchia (2010), is via context dependent functions $quantity_1$, $quantity_2, \dots$, $unit_1$, $unit_2, \dots$ etc., which partition their complements into countable parts. A function $quantity_{n,w}$ applies to an argument x , which can be a kind, in pseudopartitives, or an individual or a sum of individuals, in partitives; $quantity_{n,w}(x)$ is of type $\langle et \rangle$ and $quantity_{n,w}(x)(y)$ holds iff y is a part of x disjoint from any other part z in the same partition $quantity_{n,w}(x)$, and such that the total sum of the members of the partition $+quantity_{n,w}(x)$ is the same as x .²⁸ Here is an example of the denotation of the underlined portion of the DPs in in (17a,b):

- (18) a. In a world w with four apples a, b, c, d , where $a+b$ are in a basket, and $c+d$ just sitting on the floor:
 $quantity_{3,w}(\cap APPLE_c) = \{a+b, c+d\}$
 b. In a world w where a is the part of John’s blood used in a transfusion and b is the rest of John’s blood:
 $quantity_{7,w}(j’s\ blood) = \{a, b\}$
 c. One quantity of people in that room is very upset
 d. Two quantities of salt in that box were mixed with some other powder.

²⁷ Traditionally, the term *pseudopartitive* is reserved to structures of the form *classifier of NP*, where the NP is bare, while partitive is for structures of the form *classifier of DP*, with a full, definite DP.

²⁸ The formal definition, within the approach to atomicity we are pursuing, might go as follows:

(a) $quantity_n = \lambda x \lambda w \lambda z : \exists Y \forall w' (w \in w' \rightarrow \Pi_{n,w'}(x) = X). \Pi_{n,w}(x)(z)$ Type: $\langle e, \langle s, et \rangle \rangle$

The presupposition in (a) ensures rigidity across precisifications. For any x , $\Pi_{n,w}(x)$ is of type $\langle e, t \rangle$ and partitions x into a set of individuals that jointly make x up. If x is a kind, then the partition divides up its instances; if x sum, it divides it up into its components. If it is a plain individual (as in *a quantity of that apple*) it breaks it down into some of its mereological parts.

Going via the kind in (18a) in giving the semantics for *quantity* is a simple way to ensure that the property associated with the NPs *apples* or *water* in pseudopartitives are sum-closed, for ‘∩’ is undefined for quantized properties.²⁹ The cells in a partition $quantity_n$ are typically ‘maximally connected’ sub-units of the relevant totality (like baskets, puddles, piles,...), but sometimes can be more loosely identified (as in, e.g.,(18d)).

One feature of quantity-phrases in English is that when they apply to a *count* noun P_C it yields a partition that does *not* coincide with the one in P-atoms. Something like *there are three quantities of apples on the table* virtually never means *there are three apples on the table*. This is very natural: if we want to talk about apple-atoms, the basic noun suffices, and so quantity-phrases are used to get at a different quantization of apple aggregates. Bear this detail in mind as we consider Yudja or Nez Perce.

There are many imaginable variants of the analysis just sketched that might work for our purposes. All we have tried to provide here is an approach to context dependent classifiers which is compatible with the (English) data and explicit enough to test the claim that they have covert counterparts in generalized number languages, to which we now turn.

3.3.2 *Yudja*. In Chierchia (2015) I have begun to explore a rather simple minded approach to Yudja, which is however consistent with data presented in Lima (2014), in so far as I can tell. The basic idea is that Yudja is a kind oriented language like Mandarin (and therefore, ultimately, a Type II language), which has only one very general classifier that recruits a null counterpart to something like $quantity_n$. This classifier goes unexpressed probably, precisely because it is so general. Consider:

- (19) Txabiū ali eta awawa
 three child sand get
 ‘Children got three sand(s)’
 (The) children got three containers with sand

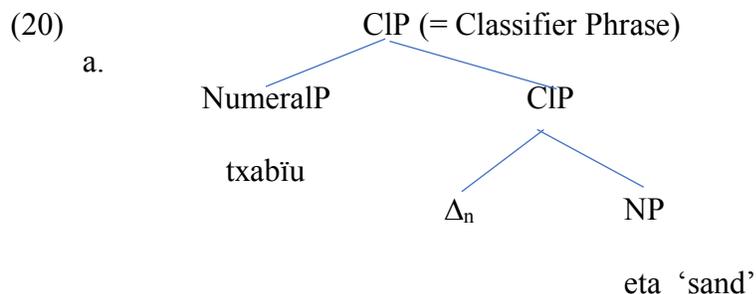
For the present purposes, and in line with Lima (2014), we assume that the numeral *txabiū* ‘three’ and the N *eta* ‘sand’ in (19) form underlyingly a constituent as in (20), constituent which is then split up by movement (presumably, leftward float of the numeral).

²⁹ In plain terms, I am assuming that the bare NP underlined in a pseudopartitive like (a) is kind denoting:

(a) Two kilos/baskets of rice/apples

Nothing of any substance changes if turns out that it is better to regard it as a sum-closed property, instead.

Recall also that we are using here a notion of kind broader than the one usually found in the literature that includes also indexical kinds. See fn. 14.



b.

$$\Delta_n(k) = \begin{cases} AT(\cup k), & \text{if defined} \\ \text{quantity}_n(k), & \text{otherwise} \end{cases}$$

The null classifier notated as Δ_n in (20a) is interpreted as in (20b). If the null classifier Δ_n combines with a count noun as in, say, *txabiü ali* ‘three + child’, it respects its natural atomicity, so that *three + child* never means ‘three aggregates of children’ or ‘three parts of a child’; it just means ‘three children’. If, on the other hand, Δ_n combines with a mass noun, like *eta* ‘sand’ it will acquire the meaning of some contextually salient, natural partition of the (relevant) sand. The null quantifier differs from its overt counterparts, like *quantity of* in English in that the latter typically overrides the atomicity of its argument, and re-calibrates it into something else, as we saw in Section 3.3.1. while the null classifier Δ_n doesn’t seem to allow that. The hypothesis of a null atomizing function constitutes a way of understanding why all nouns in Yudja freely combine with numerals, and also why they are interpreted the way they are. It is also readily consistent with the possibility of having generalized bare arguments (because every N is kind denoting, and hence argumental) and the extreme scantiness of plural marking. Further work is of course needed to verify/falsify this hypothesis.

3.3.3. *Nez Perce*. One of the ways in which *Nez Perce* differs from *Yudja* is in having apparently a richer system of number marking, which we will now summarily describe (referring for details to Deal’s work, from which all examples are taken).

(21) The number marking system in *Nez Perce*.

a. Plural is overtly marked on:

i. Animate Ns:

'aayat	ha-'aayat
woman-SG	woman-PL

ii. Some adjectives:

ki-kuckuc taam'am
small-PL egg

b. The Quantifier-system of *Nez Perce* is constituted by the following Qs:

'oykala	la'am	'ileni	miil'ac	tato's
all ₁	all ₂	a lot	a few/little	some of (partitive)

c. With count nouns, Qs in *Nez Perce* requires syntactically plural forms:

i. 'oykal-o	ha-'aayat/	*'aayat
-------------	------------	---------

- all-HUM woman-PL/ *woman-SG
 ii. 'oykala *kuu'pnin/ k'i-uupnon' tiim'en'es
 all *broken-SG/ broken-PL pencil 'all broken pencils'

With mass nouns, we get two forms with different meanings:

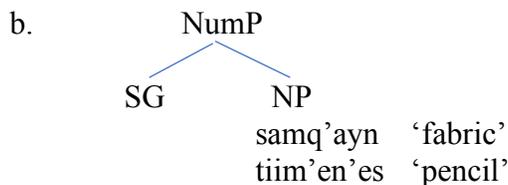
- iii. a'ilexeni cimuuXCimux samq'ayn
 a lot black-SG fabric 'a lot of black fabric'
 iv. a'ilexeni cimuuxcimux samq'ayn
 a lot black-PL fabric 'a lot of pieces of black fabric'

There are three main ways in which the singular/plural contrast manifests itself. First, it is overtly marked on some human Ns; second it is marked on some adjectives; note that adjectives in the plural form, in combination with mass nouns, induce an atomic interpretation (just like numerals do) – cf. (20c.iv). And third, the Nez Perce Q-system selects for sum-closed NPs: this requirement is evident from the fact that with count nouns quantifier select for Ns in plural forms while with mass nouns, quantifiers go either with the (atomized) plural or with the singular form interpreted cumulatively, similarly to what happens in English with partitives and pseudopartitives. The atomization of mass nouns (when they pluralize) does not involve a count reconceptualization of the latter (i.e. 'standardized' packaging), but some contextually salient partitioning.

In what follows, I'll sketch a slight (?) modification of Deal's analysis. First the presence of a fairly articulated number marking system suggests that Nez Perce is a property-oriented language, i.e., in essence, a Type I language, with generalized bare argument formation (like Russian or Hindi). This is so because, as we saw, number marking is generally associated with property oriented, presupposition inducing operators like SG and PL. Second, and this is the substantive parameter, number marking in Nez Perce recruits the same atomizing function Nez Perce uses as a generalized classifier. We hypothesize that the interpretation of SG and PL in Nez Perce is as in (22a), with a simple illustration provided in (22b):

$$(22) \quad \text{a. i. SG} = \begin{cases} \lambda P. \Delta_n(P), & \text{if defined}^{30} \\ \lambda P. P, & \text{otherwise} \end{cases}$$

$$\text{ii. PL} = \lambda P: *SG(P) = P. P$$



³⁰ I assume that Δ_n is polymorphic:

- (a) if α is of type e_k , then $\Delta_n(\alpha) = AT(\cup\alpha)$, if defined, and $\text{quantity}_n(\alpha)$, otherwise
 (b) if α is of type $\langle s, et \rangle$, then $\Delta_n(\alpha) = AT(\alpha)$, if defined, and $\text{quantity}_n(\cap\alpha)$, otherwise

SG morphology employs Δ_n ; in combination with a count noun like *pencil*, $\Delta_n(\text{PENCIL}_C)$ is defined and picks the PENCIL_C -atoms (i.e., the atomic property *pencil_C*) with individual pencils in its extension. Hence, counts nouns are expected to behave pretty much like in English. The quantifiers of Nez Perce require sum-closed properties, so *pencil_C* will have to be pluralized to combine with any Q; pluralization will be visible only on animate NPs and on adjectives – cf. (21c.i-ii). In combination with a mass noun like *fabric*, i.e. $\Delta_n(\text{FABRIC}_M)$ there are two options: either some partition (in terms of, say, pieces of fabric) is contextually salient, and then $\Delta_n(\text{FABRIC}_M) = \text{piece} (\cap \text{FABRIC}_M)$; or no such partition is contextually salient, and then SG lets through just FABRIC_M (i.e. the sum-closed, non (stably) atomic property itself). To satisfy the presupposition of a numeral or of plural marking, however, some salient partition *must* be available. This analysis of SG yields the two options observed with quantifiers when they combine with mass nouns: either the SG marked mass noun has to be the basic sum-closed property, e.g. FABRIC_M ; or it has to be the sum-closed, atomized (i.e. plural) $*\Delta_n(\text{FABRIC}_M)$. This seems to deliver the observed paradigm.

(23) a. Count Ns with quantifiers

i. *'oykala kuu'pnin tiim'en'es

all broken-SG pencil

$\Rightarrow \text{all}(\text{SG}(\text{broken}(\text{pencil}))) = \text{all}((\text{broken} \cap \text{pencil}_C)) = \text{undefined}$

ii. 'oykala k'i-uupnon' tiim'en'es

all broken-PL pencil

$\Rightarrow \text{all}(\text{PL}(\text{broken}(\text{pencil}))) = \text{all}(*(\text{broken} \cap \text{pencil}_C)) = \lambda P. *(\text{broken} \cap \text{pencil}_C) \subseteq P$

b. Mass Ns with quantifiers

iii. 'oykala cimouxcimux samq'ayn

all black-SG fabric

$\Rightarrow \text{all}(\text{SG}(\text{black}(\text{fabric}))) = \text{all}(\text{black} \cap \text{FABRIC}_M) = \lambda P. (\text{black} \cap \text{FABRIC}_M) \subseteq P$

iv. 'oykala cimouxcimux samq'ayn

all black-PL fabric

$\Rightarrow \text{all}(\text{PL}(\text{black}(\text{fabric}))) = \text{all}(*\Delta_n(\text{black} \cap \text{FABRIC}_M)) = \lambda P. *quantity_n(\text{black} \cap \text{FABRIC}_M) \subseteq P$

Deal convincingly points out that the limited character of the overt evidence available to the learner in Nez Perce provides a strong poverty of stimulus argument in favor of the universality of the mass count distinction.

Summing up, what we have called Type III languages are characterized by the use of a context dependent atomizing function, that resembles closely classifiers like *quantity_n*. Such atomizing function is essentially identical to *quantity_n* for mass nouns, while for count nouns it retains the inherent atomicity of the noun. The atomizing function Δ_n is employed in two ways. If the language is kind oriented, it is employed as a null classifier; if the language is property oriented, it is employed in the definition of number marking. What determines the kind vs. property orientation of the language are the same factors that do so for languages that lack Δ_n , namely the pattern of pluralization (whether it is more similar to IE languages, or to generalized classifier languages). The unavailability in languages like English of Δ_n may be a blocking effect in the spirit of Chierchia (1998), due to the overt presence in English of a rich inventory of overt

context dependent classifiers like *quantity, unit, amount, aggregate*, etc. Whether this last speculation has a real basis is for further research to determine.

4. Summary and concluding remarks.

The approach to variation in Section 3 is meant to provide us more with an existence proof than with a conclusive theory, if there ever was such a thing. The starting point is a Base-line theory in which nouns are interpreted ‘naturalistically’, i.e. each noun has a natural atomicity or inherently lacks it (or else, it is inherently ambiguous between a mass vs. count variant vis G/Σ), with rather minimal, culturally driven variations (like, e.g., *bear* is predominantly count in English, but might be ambiguous in a bear-eating society). Grammatical variation comes from two sources: type theory and covert use of grammatical resources. Their respective role can be summarized as in the following scheme:

	Property oriented languages	Kind oriented languages	Availability of Δ_n
(24)	Most IE languages (Type I)	Classifier Languages (Type II)	NO
	Nez-Perce	Yudja	YES, Type III

In this analysis, Type III languages, cut across Type I and Type II languages.

Property-oriented languages are subject to further variation in the way number is interpreted. Here is what we found:

(25) Variation number marking

a. English (and most IE languages)

- i. $SG = \lambda P:AT(P) = P.P$
- ii. $PL = \lambda P:*SG(P) = P.P$
 - No numerals with mass nouns
 - No pluralization of mass nouns
 - Presence of Fake Mass Ns (SGL function is active)

b. Greek

- i. $SG = \lambda P:AT(P) = P.P$
- ii. $PL = \lambda P:*SG(P) = P.P$
 - No numeral with mass nouns
 - Pluralization (of “abundance”) of mass nouns
 - No Fake Mass Ns (SGL inactive)

c. Nez Perce

- i. $SG = \lambda P. \Delta_n(P)$, if defined; $\lambda P. P$, otherwise
- ii. $PL = \lambda P:*SG(P) = P.P$
 - Numerals with any kind of Ns
 - Pluralization of mass noun (with a quantity of interpretation)
 - No Fake Mass Ns (SGL inactive)

Note that each characterization of singularity has its plausibility. English has the strongest one: a property P is true of some individual u only if that individual can be regarded as *one* P -atom. The SGL-function is activated by the stricture of this characterization. Greek opts for the notion of not being sum-closed (anti-cumulativity). This allows mass nouns to pluralize, while still banning combination with numerals. Nez Perce uses atomization in the form of the covert classifier Δ_n , which retains the atomic character of count nouns but introduces context dependent atomizations of mass nouns; when no atomization is pragmatically available, mass nouns are let through in their more natural forms, as sum-closed, non atomic properties – this is the ‘ambiguous singular morphology’ option. And there are surely other plausible microvariations in what makes a noun singular/atomic. The notion of *plural*, on the other hand, remains constant, as is the sum-closure of the singular, and hence a function of the latter.

There are various ways of implementing the kind- vs. property-orientation of languages. The one we will suggest here is to assume that the interpretation of the noun root, e.g. $\sqrt{\text{woman}}$ in English or its counterpart $\sqrt{\text{nürén}}$ in Mandarin is unspecific among its natural semantic interpretations, which include the set of woman-atoms, the sum-closure of the latter and the corresponding kind (cf. (26a)). A nominal root is then merged with little n , merger that drives the rest of the syntactic derivation as in (26b), as proposed within the ‘Distributed Morphology’ approach – see Lima (2014), Deal (2017). The semantic role of little n is similar to the role of functional heads: it introduces a (language specific) presupposition. This presupposition determines the type of the noun denotation that part takes into further semantic compositions higher up in the syntactic derivation: a property or a kind.

(26) Noun interpretation and Distributed Morphology

- a. $\|\sqrt{\text{woman}}\| = \|\sqrt{\text{nürén}}\| = \{\text{woman}_C, *\text{woman}_C (= \text{WOMAN}_C), \text{woman}_{k,C} (= \bigcap \text{WOMAN}_C)\}$
 b. kind orientation

$$\begin{array}{c} n, \lambda_{x_k.x_k}(\|\sqrt{\text{nürén}}\|) = \bigcap \text{WOMAN}_C^{31} \\ \swarrow \quad \searrow \\ n, \lambda_{x_k.x_k} \quad \sqrt{\text{nürén}}, \|\sqrt{\text{nürén}}\| \end{array}$$

- c. $n, \lambda_{P_{\langle s,et \rangle}.P_{\langle s,et \rangle}}(\|\sqrt{\text{woman}}\|) = \{\text{woman}_C, \text{WOMAN}_C\}$

$$\begin{array}{c} n, \lambda_{P_{\langle s,et \rangle}.P_{\langle s,et \rangle}} \sqrt{\text{woman}}, \|\sqrt{\text{woman}}\| \end{array}$$

- d. i. Kind oriented languages: $n \Rightarrow \lambda_{x_k.x_k}$
 ii. Property-oriented languages: $n \Rightarrow \lambda_{P_{\langle s,et \rangle}.P_{\langle s,et \rangle}}$

In Mandarin, little n triggers a presupposition that the interpretation of its sister is kind level. Hence a classifier will be needed to combine with numerals. In English, little n triggers a presupposition that N s are properties. Which property (atomic or plural) will be further determined by number marking. The setting of this parameter is learned in the way any parameter is learned: through its morphosyntactic manifestations. For example, upon finding out that her language allows direct *numeral* – N combinations and/or that plural is overtly marked

³¹ The formula $\lambda_{x_k.x_k}(\|\sqrt{\text{nürén}}\|)$ constitutes a slight abuse of notation. It is to be understood in terms of ‘pointwise’ function application:

(a) $\lambda_{x_k.x_k}(\|\sqrt{\text{nürén}}\|) = \{a: \lambda_{x_k.x_k}(a) \text{ is defined and } a \in \|\sqrt{\text{nürén}}\|\}$

and multiply exposed (e.g., through verb agreement) the child will determine that the interpretation of little *n* is as in (26c).

A fairly simple base-line theory, which determines a formal notion of atomicity, strongly linked to the pre-linguistic distinction, but more general than it,³² appears to account for a rather dazzling set of macro- and micro-differences across languages. This happens with no variation at all in the underlying logic of mass vs. count properties, and some variation in the *types* a language chooses to exploit (which is a matter of labelling/categorization) together with some variation in which functors may be used covertly. No deep differences in ‘mode of thinking’ emerge. But some ‘zooming in/out’ effect in how natural units are put in focus does emerge. One might hope that all forms of semantic variation could be understood along similar lines.

ACKNOWLEDGEMENTS: I am indebted to A. R. Deal, Jeff Pelletier, Roberta Pires, and Susan Rothstein for criticisms of versions of this paper. I miss terribly discussing and ‘crossing sabers’ with Susan, a wonderful scholar, long-time friend, and twin soul in many of our intellectual curiosities.

REFERENCES

- Barner, D. and J. Snedeker (2005) “Quantity judgments and individuation: Evidence that mass nouns count,” *Cognition*, 97: 41-66.
- Borer, H. (2005) *In Name Only. Structuring Sense*, Volume 1, Oxford University Press, Oxford.
- Bunt, H. (1979) “Ensembles and the Formal Semantic Properties of Mass Terms” in F.J. Pelletier (ed) *Mass Terms: Some Philosophical Problems*, Kluwer, Dordrecht
- Carey, S. (1985) *Conceptual Change in Childhood*, Bradford Books, MIT Press, Cambridge, Ma.
- Carey, S. and E. Spelke (1996) “Science and core knowledge,” *Philosophy of Science*, 63(4): 515-533.
- Carlson, G. (1977) *Reference to Kinds in English*, Ph. D. Dissertation, University of Massachusetts at Amherst. Published in 1980 by Garland, N.Y.
- Cheng, L. and R. Sybesma (1998) “Yi-wang Tang, yi-ge Tang: Classifiers and Massifiers.” *Tsing Hua Journal of Chinese Studies*, XXVIII (3), pp. 385-412.
- Cheng, L. and R. Sybesma (1999) “Bare and not so bare Nouns and the Structure of NP”, *Linguistic Inquiry*, 30.4: 509-542.
- Chierchia, G. (1998a) “Plurality of Mass Nouns and the Notion of Semantic Parameter’ in S. Rothstein (ed) *Events in Grammar*, Kluwer, Dordrecht.
- Chierchia, G. (1998b) “Reference to Kinds across Languages”, *Natural Language Semantics*, 6: 339-405.
- Chierchia, G. (2010) “Mass Nouns, Vagueness, and Semantic Variation,” *Synthese*, 174: 99-149.
- Chierchia, G. (2013) *Logic in Grammar*, Oxford University Press.
- Chierchia, G. (2015) “How universal is the mass/count distinction? Three grammars of

³² For example, the mass/count contrast extends to eventuality-denoting nouns: *knowledge* is mass but *belief* is count, and to abstract nouns: *beauty* is mass, *quality* is count. See Husic (2019b) for a very intriguing attempt of extending a vagueness-based approach to the mass/count contrast in the realm of abstract nouns.

- counting,” in Y. A. Li, A. Simpson, and W.D. Tsai (eds) *Chinese Syntax in a Cross-linguistic Perspective*, Oxford University Press.
- Chierchia, G. (2017) “Clouds and Blood. More on the Mass/Count Distinction”, *Synthese*, 194.7: 2523-2538.
- Darlymple, M. and S. Morfu (2012) “Plural Semantics, Reduplication, and Numeral Modification in Indonesian,” *Journal of Semantics*, 29.2: 229-260.
- Dayal, V. (2004) “Number Marking and (In)definiteness in Kind terms” in *Linguistics and Philosophy* 27.4: 393-450.
- Dayal, V. (2011) “Bare Noun Phrases” in C. Maienborn, K. von Stechow and P. Portner (eds), *Semantics: An International Handbook of Natural Language Meaning*, De Gruyter Mouton, Berlin.
- Dayal, V. (2012) “Bangla Classifiers: Mediating between Kinds and Objects”, *Italian Journal of Linguistics* 24.2: 195-226.
- Dayal, V. (2013) “On the Existential Force of Bare Plurals Across Languages”, in I. Caponigro and C. Cecchetto (eds) *From Grammar to Meaning: The Spontaneous Logicality of Language*, Cambridge University Press, Cambridge, UK.
- De Vries, H., G. Tsoulas, R. Folli, A. Renans, and J. Romoli (this volume) ‘Exoskeletal Iceberg Semantics: a Best-of-both-worlds Approach to the Mass/Count Distinction.
- Deal, A. R. (2017) “Countability Distinctions and Semantic Variation”, *Natural Language Semantics* 25.2: 125-171.
- Doetjes, J. (1997) *Quantifiers and Selection*, The Hague: Holland Academic Graphics.
- Feigenson, L., S. Carey and E. Spelke (2002) “Infants' Discrimination of Number vs. Continuous Extent”, *Cognitive Psychology*, 44, 33-66.
- Fine, K. (1975) “Vagueness, Truth, and Logic,” *Synthese*, 30: 275-300.
- Gajewski, J. (2002) “L-Analyticity and Natural Language,” manuscript, University of Connecticut at Storrs, jon-gajewski.uconn.edu/wp-content/uploads/sites/1784/2016/08.
- Gajewski, J. (2009) “L-Triviality and Grammar,” manuscript, University of Connecticut at Storrs, <https://jon-gajewski.uconn.edu/wp-content/uploads/sites/1784/2016/08/Logic.pdf>
- Gathercole, V. (1986) “Evaluating Competing Linguistic theories with Child Language Data: the Case of the Mass-Count Distinction”, *Linguistics and Philosophy*, 9.2: 151-190.
- Grimm, S. (2012) *Number and Individuation*, Ph. D. Dissertation, Stanford.
- Hauser, M. and S. Carey (2003) “Spontaneous Representation of Small Numbers of Objects by Rhesus Macaque: Examination of Content and Format”, *Cognitive Psychology*, 47: 367-401.
- Husic, H. (2019) “A Vagueness Based Analysis of Abstract Mass Nouns” Paper presented at Sinn und Bedeutung 19, Osnabrück.
- Ionin, T. and O. Matushanski (2018) *Cardinals. The Syntax and Semantics of Cardinal Containing Expressions*, MIT Press, Cambridge, Ma.
- Jiang, J. L. (2017) Mandarin associative plural -men and NPs with -Men” *International Journal of Chinese Linguistics* 4:2: 191-256.
- Jiang, J. L. (2018) “Definiteness in Nuosu Yi and the Theory of Argument Formation”, *Linguistics and Philosophy*, 41.1: 1-39.
- Jiang, J. L. (forthcoming) *Nominal Arguments and Language Variation*, Oxford Studies in Comparative Syntax, Oxford University Press.
- Krifka, M. (1994) “Mass Expressions,” in R. E. Asher & J.M.Y. Simpson (eds.), *The Encyclopedia of Language and Linguistics*, Pergamon Press, Oxford.

- Landman, F. (1989) "Groups I", in *Linguistics and Philosophy* 12: 559-605.
- Landman, F. (1991) *Structures for Semantics*, Kluwer, Dordrecht
- Landman, F. (2011) 'Count nouns, mass nouns, neat nouns, mess nouns,' in: Skilters, Jurgis (ed.) *Papers from: Formal Semantics and Pragmatics: Discourse Context and Models, The Baltic International Yearbook of Cognition, Logic and Communication*, Riga.
- Landman, F. (2016) 'Iceberg semantics for count nouns and mass nouns: the evidence from portions.' In: Susan Rothstein and Jurgis Skilter (eds.), *Number: Cognitive, Semantic and Cross-linguistic Approaches, The Baltic International Yearbook of Cognition, Logic and Communication*, Vol 11, Manhattan, KS: New Prairy Press
- Lima, S. (2014) *The Grammar of individuation and counting*, Ph. D. Dissertation, University of Massachusetts, Amherst.
- Link, G. (1983) "The logical analysis of plurals and mass terms. A lattice-theoretic Approach," in R. Bauerle, C. Schwartz, and A. von Stechow (eds) *Meaning, Use and Interpretation of Language*, de Gruyter.
- Pelletier, J. and F. Schubert (1989) "Mass expressions," in D. Gabbay and F. Guentner (eds) *Handobook of Philosophical Logic* (Vol. IV), Kluwer, Dordrecht.
- Renans, A., J. Romoli, M. Makri, L. Thieu, H. de Vries, R. Folli, and G. Tsoulas (2018) "The Abundance Inference of Pluralized Mass Nouns is an Implicature: Evidence from Greek", *Glossa: A Journal of General Linguistics*, 3.1:103; DOI: <http://doi.org/10.5334/gjgl.531>
- Rothstein, S. (2010) "Counting and the Mass/Count Distinction", *Journal of Semantics*, 27.3: 343-397.
- Rothstein, S. (2016) "Counting, Measuring, and the Semantics of Classifiers" in S. Rothstein and J. Skilters (eds) *Baltic International Yearbook of Cognition, Logic and Communication*, Vol 11, Riga, <https://newprairiepress.org/biyclc/vol11/iss1/>.
- Sauerland, U. (2003) "A new semantics for number," in *Proceedings of SALT 13*, CLC Publications, Cornell University, Ithaca, N.Y.
- Schwarzschild, R. (1996) *Pluralities*, Kluwer, Dordrecht.
- Schwarzschild, R. (2011) "Stubborn Distributivity, Multiparticipant Nouns and the Count/Mass Distinction", In S. Lima, K. Mullin and B. Smith (eds), *Proceedings of the 39th Meeting of the North East Linguistic Society*, 39, Amherst, Ma.
- Scontras, G. (2017) "A New Kind of Degree", *Linguistics and Philosophy*, 40.2: 165–205.
- Scontras, G., K. Davidson, A. R. Deal, and S. E. Murray (2017) "Who has More? The Influence of Linguistic Form on Quantity Judgments", *Proceedings of the Linguistic Society of America*, 2, 41:1–15.
- Spector, B. (2007) "Aspects of the pragmatics of plural morphology: On higher order implicatures," in U. Sauerland and P. Stateva (eds) *Presuppositions and Implicatures in Compositional Semantics*, Palgrave.
- Soja, N., S. Carey and E. Spelke (1991) "Ontological Categories Guide Young Children's Inductions of Word Meanings: Object Terms and substance Terms," in *Cognition*, 38: 179-211.
- Sudo, Y. (2014) "Dependent Plural Pronouns with Skolemized Choice Functions", *Natural Language Semantics*, 22: 265-297.
- Tsoulas, G. (2009) "On the Grammar of Number and Mass Terms in Greek", in C. Halpert, J. Hartman, and D. Hill (eds), *MIT Working Papers in Linguistics*, Vol. 57, Cambridge, Ma.
- Veltman, F. (1985) *Logics for Conditionals*, Ph. D. Dissertation, University of Amsterdam.