Mass nouns, vagueness and semantic variation

Gennaro Chierchia

Abstract The mass/count distinction attracts a lot of attention among cognitive scientists, possibly because it involves in fundamental ways the relation between language (i.e. grammar), thought (i.e. extralinguistic conceptual systems) and reality (i.e. the physical world). In the present paper, I explore the view that the mass/count distinction is a matter of vagueness. While every noun/concept may in a sense be vague, mass nouns/concepts are vague in a way that systematically impairs their use in counting. This idea has never been systematically pursued, to the best of my knowledge. I make it precise relying on supervaluations (more specifically, 'data semantics') to model it. I identify a number of universals pertaining to how the mass/count contrast is encoded in the languages of the world, along with some of the major dimensions along which languages may vary on this score. I argue that the vagueness based model developed here provides a useful perspective on both. The outcome (besides shedding light on semantic variation) seems to suggest that vagueness is not just an interface phenomenon that arises in the interaction of Universal Grammar (UG) with the Conceptual/Intentional System (to adopt Chomsky's terminology), but it is actually part of the architecture of UG.

Keywords Atomicity · Classifiers · Mass/count · Parameters · Singular/plural · Supervaluations

1 Introduction

The mass/count distinction keeps being at the center of much attention among cognitive scientists, as it involves in fundamental ways the relation between language

G. Chierchia (✉)
Department of Linguistics, Harvard University, Cambridge, MA 02138, USA
e-mail: Chierch@fas.harvard.edu

(i.e. grammar), thought (i.e. conceptual systems not necessarily rooted in grammar) and reality (i.e. the physical world). In the present paper, I wish to explore the view that the mass/count distinction is a matter of vagueness. While every noun/concept may in a sense be vague, mass nouns/concepts are vague in a way that systematically impairs their use in counting. This idea has never been systematically pursued, to the best of my knowledge. I believe that doing so will lead us to a better understanding of the nature of the distinction and, more generally, of the relation between grammar and extralinguistic cognitive systems. The idea that the mass/count distinction is based on vagueness sets the following tasks for us. We must make such an idea sufficiently precise for empirical testing; and we must show that the model that comes out of it affords us a better understanding of the relevant phenomenology than other current alternatives.

What makes this task difficult is that the relevant set of morphosyntactic and semantic facts, while presenting a stable core, is also subject to considerable variation across languages both at a micro-level (among closely related languages/dialects) and at a macro-level (among genetically unrelated language families). Most classic proposals on the mass/count distinction up to, say, Link (1983) while showing awareness of this fact, provide models that directly apply only to number marking languages (i.e. languages which obligatorily mark their nouns as singular/plural) of the familiar Indo European type. Recently, there has been an exciting growth in crosslinguistic semantic investigations and a discussion from such a point of view of where we stand in our understanding of the mass/count distinction across languages is perhaps timely. Accordingly, our goal is to come up with a model that on the one hand can accommodate the universal properties of the mass/count distinction and at the same time paves the way for understanding how languages may vary along such a dimension.

Our discussion should help bring into sharper focus the important role of vagueness in grammar. In particular, if it turns out to be right that the main determinant of the mass/count distinction is vagueness, then the interface of grammar with the conceptual system has to be somewhat reassessed. Under the view that Universal Grammar (UG) is, essentially, a lexicon together with a recursive computational apparatus, vagueness wouldn't just play a role at lexical interfaces with extralinguistic systems, but it would shape some aspects of the computational apparatus itself. More specifically it would have a rather direct and perhaps surprising impact on the morphosyntax of number. Another way to put it is that the non lexical component of UG (unlike other conceivable computational devices) is pre-wired to handle vagueness; if one views UG as the structure of a 'language organ', one can readily speculate on why a computational system that is capable of handling efficiently vague information would be a better fit with our environment that a system that isn't.[1]

I will be taking the notion of vagueness pretty much as given and adopt a 'supervaluation' approach to it,[2] in the particular version developed by Veltman (1985),

---

[1] Similar conclusions can be reached on the basis of, e.g., work on comparatives like Klein (1980). Kennedy (2007) (and much work related to it) takes a different stand on the relation between vagueness and gradability. Even on Kennedy's position, however, many extremely interesting issues pertaining to how vagueness permeates grammar remain open and will have to be reassed if the present work is on the right track.

[2] Standard references include Van Fraassen (1971), Fine (1975), Kamp (1975).

also known as 'data semantics'. In adapting to my goals a supervaluationist approach, I will try to remain neutral between 'indeterministic' versus 'epistemic' stands on vagueness, even though my own bias towards the former may sometimes seep through.[3] I also won't have anything to say on issues pertaining to higher order vagueness.

Here is the plan for this work. In the remainder of this introduction, I will say something in general on the relation between the cognitive basis of the mass-count distinction and its grammatical manifestations. In Sect. 2 I will review well known facts that constitute the main empirical manifestations in language of the mass/count distinction trying to tease apart some of its universal properties from some of its language particular ones. This will constitute the empirical basis against which the theory will be tested. In Sect. 3, I will sketch (without being able to defend) the necessary background for the present proposal. Such background includes some formal aspects (e.g. which logic/representation language will be assumed) and some substantive, linguistic aspects (pertaining to the theory of plurals and to kind reference). In Sect. 4 I will develop the specifics of my vagueness based approach to singular/plural structures and illustrate how a theory of masses and countables flows from it. In Sect. 5, I will discuss how such a theory accounts for number marking languages. In Sect. 6, how it accounts for (aspects of) cross linguistic variation. Finally, in Sect. 7 we will conclude, after having discussed some of the available alternatives.

Let us begin with some terminological choices. As is well known, fluids (water, air), pastes (dough, clay), minerals (gold), assorted materials (wood, bronze, sand) are typically associated with mass nouns. A common way to put this is that mass nouns denote/refer to/are semantically associated with 'substances,' where the latter are to be thought of as material aggregates that may be joined or split without changing their nature. I will refer to this somewhat loosely defined category of mass nouns as 'canonical', 'prototypical' or 'core'. There are many mass nouns in English that are not canonical in this sense. They include concrete 'superordinate' nouns (like *furniture* or *footwear* as opposed to *table* or *shoe*), abstract nouns (like *beauty* or *knowledge* as opposed to *virtue* or *belief*) and eventuality denoting nominals (like *jumping* as opposed to *jump*). According to some, the mass/count contrast also invests the aspectual/Aktionsarten system as pervasively as the nominal one. In the present work, I'll consider the mass/count distinction only limitedly to concrete nouns.

An old and difficult issue is whether the mass/count distinction is linguistic (stemming, that is, from grammar) or extralinguistic (i.e. rooted in language independent cognitive systems or, if you, wish, in reality as such). Much of the relevant debate has in W. v. O. Quine a main crossroad; he developed an influential relativistic stand in a number of writings (e.g. 1960, 1969) in which he argued that a frame of

---

[3] What is meant here by 'indeterminism' versus 'epistemic stand', can be very roughly illustrated by an example as follows. Suppose that uncle David falls right between bald and non bald, by any reasonable definition of baldness we care to give; you hold an indeterminist view, if you believe that that there is no fact of the matter that would settle the question whether 'uncle David is bald' is true or false. The criterion for being bald is just not sharp enough. On the other hand, you hold an epistemic stand if you believe that there is a fact of the matter as to whether uncle David is bald or not, but we do not know what it is. Please bear in mind that the epistemic stand, as developed by Williamson (1994), is much more compelling than what my off hand remarks may let through. For a balanced discussion, see, e.g., McGee and McLaughlin (1998).

reference arises holistically through language. One way in which he illustrated his thesis is precisely through the mass-count distinction. He argued, in particular, that the pre-linguistic child really has no notion of 'object' as opposed to 'substance' and works with a radically undifferentiated frame of reference. Each noun is first understood as being mass like and referring to qualitatively similar portions of experience (e.g. 'mama' refers to any portion of mama-experience). What qualifies as an 'object' only comes about as a result of language development and of the mastery of the quantificational system of language. Quine's ontological relativism meshed well with the then popular claim (often associated with the Sapir-Whorf hypothesis) that languages embody indefinitely varying ontological/conceptual systems that in turn reflect the forms of life of particular groups of speakers.

The basis of Quine's views, namely the claim that the pre-linguistic child lacks any notion of 'individuated object' has been challenged in 80s and 90s by the work of cognitive psychologists like S. Carey and E. Spelke. Using diverse and sophisticated experimental techniques, they have shown that at a few months of age the child has fairly structured expectations as to how objects differ from substances. Closely paraphrasing their words, the child knows that

(1)   Objects have boundaries, move as wholes along continuous paths, retain their identity upon colliding with each other, while substances don't.

This is not the place to review the type of experiments that have lead to these conclusions. Nor is it the place to decide whether the relevant notion of object develops from getting attuned to aspects of the physical world, or it is an innate schematism that we superimpose on the world. What is relevant for our purposes is that such notion of object is (i) clearly language independent but (ii) aligns fairly well with the distinction between core mass nouns versus count nouns. In fact, Soja et al. (1991) argue that 2 year old children in inducing the meaning of new words crucially exploit the categories in (1). Confronted with something that can be parsed as an object, they generalize a new name to other tokens of the same object (regardless of the material that makes it up), while confronted with a substance, they generalize according to the material (regardless of, e.g., shape). Moreover, Marc Hauser and his associates have shown that also non human primates, e.g. rhesus monkeys, seem to be endowed with similar forms of knowledge (cf., e.g., Hauser and Spaulding 2006). Like the preverbal child, rhesus monkeys seem to expect objects to have different properties from, say, fluids.

I take it that this line of research has established that the object versus substance contrast is robustly present in language independent cognitive systems. The question that arises, then, is how this language independent contrast comes to be encoded in grammar. The existence of non canonical mass nouns, like *furniture* (that clearly applies to objects in Spelke's et al. sense)[4] shows that the mapping is non trivial.

There is an even more radical question to be asked. Namely, *why* does language bother at all with the mass/count distinction? To see what is involved in this question, consider locutions like *quantity/amount of water* (known in linguistics as psuedopartitive structures). Certainly, anything which is water is also a quantity of water; and vice versa. But *quantity of water* seems to be a *count* noun phrase. It takes the indefinite

---

4   See Barner and Snedeker (2005) for very relevant discussion.

article, it pluralizes, it can even be used with numerals as in 'I am now going to spill three quantities of water on your floor'. It seems somehow difficult to maintain that the differences between 'that water', 'that quantity of water' or 'that water amount' etc. are somehow truly semantic in nature. Yet, we can say 'three quantities of water' but not as naturally 'three waters' (and, for that matter, 'three pieces of furniture' but not 'three furnitures'). Is it just syntax? Is the awkwardness of combining *three* with *water* just a syntactic constraint?

We may have here the making of a conceptual paradox. First, the object versus substance contrast is quite clearly pre- (or extra-)linguistic. Second, every language encodes it in a number of conspicuous morphosyntactic ways. Third, every language could in principle do without encoding it, by resorting to whatever it is that nouns like *quantity* or *amount* do. It is not just the existence of such lexical items that matters. It is the availability of the corresponding operation of 'apportioning'. If such an operation is available (as it is lexicalized in words like *quantity*, and perhaps used covertly when we say things like 'bring three waters to that table') why does grammar insists on treating mass nouns differently from count ones? Why each time that we want to count using a mass noun don't we simply, automatically 'apportion' it as needed? What is to prevent us from *interpreting* 'water' as meaning something like 'water amount' or 'water quantity'?

What I take these considerations to show is that the existence of the mass/count distinction in grammar is neither a logical nor, perhaps, a communicative necessity. The universal insistence of natural language on a separate combinatory for mass versus count nouns is not easy to figure out. I see only one way if not to dispel, at least to begin making sense of this puzzle. And it involves separating off language and natural language semantics from 'general cognition' and the like. Language, viewed as specifically human aggregate of cognitive capacities, must have developed an autonomous apparatus responsible for the mass/count system. Such an apparatus is perhaps filogenetically related to the pre linguistic categorization in objects and substances, and certainly interpretively related to it. What I mean by 'interpretively related' is that there is a natural mapping between the grammatical basis of the mass/count contrast and the pre-linguistic categories discovered by Carey and Spelke, along lines I will try to make clear. More specifically, I am going to propose that the impossibility of directly counting with mass nouns has two sources. Both are, in some sense, formal (i.e. logical/computational properties of grammar). One is rooted in how vagueness is coded in grammar and that, by its very nature, cannot be subject to much variation across languages; the other is a clash of logical categories, a type theoretic mismatch in the compositional constructions of noun phrases, which is where languages may vary.

2 Variable and constant aspects of the mass/count distinction

In this section I will first address what seem to be universals of the mass versus count contrast. Then I will turn to how such a contrast shows up in three different language families. I conclude with a more detailed discussion of number marking languages like English.

2.1 Universal properties of mass nouns

There are relatively few properties of mass nouns that are truly tendentially constant across languages. I think the main candidates are mainly three, which I will dub the 'signature property', the 'mapping property' and the 'elasticity property' respectively.

2.1.1 The signature property

Perhaps the most steady grammatical property associated with mass nouns in general is the marked status of their direct combination with a numeral expression: regardless of word order, constituents of the form [$_{NumP}$ Num N$_{MASS}$]^5 are either outright ungrammatical or are felt as requiring a reinterpretation of sorts ('coercion' or 'type-shifting,' on which more below, under the rubric 'elasticity').

(2)   a. Thirty three tables/stars/pieces of that pizza   [$_{NumP}$ Num N$_{COUNT}$]
      b. *Thirty three bloods/waters/golds   [$_{NumP}$ Num N$_{MASS}$]

It should be noted that this generalization is independent of the fact that numeral-noun combinations in languages English require plural marking on the noun (for numerals other than *one*). The impossibility of combining a mass noun with a numeral holds also of subject-predicate structures, where number marking does not 'interfere':

(3)   a. Those boys are at least thirty   [$_{PredP}$ NP$_{MASS}$ Pred Num]
      b. *That gold is at least thirty
      c. That gold is at least thirty pounds

Moreover, there are languages like Finnish, in which numeral-noun constituents require *singular* agreement on the noun, but still the combination of a numeral with a mass noun is deviant and/or induces coercion:

(4)   a. Yhdeksän omena-a   b. *Yhdeksän vesi-a
         Nine-nom apple-part-SG   Nine-nom water-part-SG

In general, to felicitously combine a N$_{MASS}$ with a numerical expression, we need to interpolate a suitable measure phrase (like *pound*, *liter*, etc., as is *three pounds of sugar*) or a classifier like phrase (e.g., container words like *cup*, *spoon*, *truckload*, etc. as in *three cups of sugar*).

    The impossibility (in the sense specified above) of directly combining a numeral with a mass noun is universal, in so far as I know. Here are examples from languages as diverse as Mandarin or Dëne Sųline', an Athapaskan language:

(5)   a. Mandarin
            i. *san rou   ii. san bang rou
               three meat   three CL meat 'three pounds of meat'
      b. Dëne
            i. *solaghe bër   ii. solaghe nedadhi bër
               five meat   five pound meat

---

^{5} Where Num = Numeral (number expression), NumP = Numeral Phrase. A 'Numeral Phrase' is something like *three apples*. I use 'Measure Phrase' for expressions like *three pounds of apples*. Finally, 'number marking' is used to refer to singular/plural contrasts.

2.1.2 The mapping property

The second universal pertains to how the grammatical manifestations of mass/counts relate to the conceptual, pre-linguistic contrast object/substance discussed in the introduction. I think such a relation can be compactly characterized as follows:

(6)   In any language L, substances are coded as mass by the tests prevailing in L

The idea is that each language will have specific morphosyntactic generalizations that distinguish mass from count (just like, say, every language has criteria to tease subjects from objects). By those tests, whatever they turn out to be, in no language the basic words for say *blood* or *air* will come out as count. This is an extremely strong and substantive universal. Notice that the symmetric mapping principle does not hold:

(7)   *In any language L, (Spelke-like) objects are coded as count in L

English clearly counterexemplifies the mapping principle in (7). Nouns like *furniture*, *footwear*, *jewellery*, etc. are clearly mass by the tests operative in English. Yet they apply to things that the pre-linguistic child would classify as objects.

2.1.3 Elasticity

The mass/count distinction appears to be 'elastic' in a sense that might be crucial to our understanding of what is going on. Many nouns appear to admit count or mass uses, if not with equal frequency, with apparent equal ease:

(8)   a. I need three ropes   rock, tape, …
       a′. I need a lot of rope
       b. I drank three beers   tea, coffee,…
       b′. I drank a lot of beer

In these cases, it seems straightforward to describe what is going on by starting from the mass use. Rock, rope and beer are naturally conceived as substances or materials (which yield prototypical mass nouns). However it is also the case that such substances naturally occur in either lumps or standard servings/amounts. For example, rocks (count) are lumps of rock (i.e. continuous, bounded amounts of rock); beers (count) are standard servings of beer (bottles, glasses); ropes might be either (i.e. continuous, bounded pieces of rope or standard amounts like, say, coils). In cases like these, we might want to talk of a conceptually prior mass nature of the relevant noun, with a concomitant, equally natural, but conceptually 'derived' count use.

Here are some further examples of predominantly mass nouns that allow, with various degrees of ease, also count uses:

(9)   a. I drank three waters
       b. I drank three bloods
       c. I ate three breads
       d. I bought three golds

Something like (9a) is still fairly natural, but might require more context than say (8b); on the other hand, (9b) would make sense only in a quite exotic 'vampire bar'

situation. In cases like (9c,d), it might be appropriate to talk of 'coercion' into count moulds.

These phenomena have been addressed in terms of the notion of a 'universal packager', to borrow D. Lewis's term, i.e. a function that turns masses into countables. Such a packager may be involved in the shifts illustrated in (8) and (9); but its application isn't equally smooth on every mass noun.

While on the topic of M(ass) → C(ount) shifts, there is a second major way of doing so, illustrated in (10).

(10)   a. I like only three wines: chardonnay, pinot, and chianti.
        b. I like only three dogs: Irish setters, golden retrievers, and collies.

Three wines in (10) seems to have the interpretation 'three kinds/sorts/types of wine.' This type of shift seems to be possible with nearly every noun, with varying degrees of naturalness. It is not limited to mass nouns; as (10b) illustrates it seems to be possible in pretty much the same way for count nouns.[6]

So, there are two main mechanisms that turn a conceptually mass noun into a count one. One involves appealing to standardized or otherwise naturally occurring bounded amounts of a substance/material. A second one turns a noun associated with a kind of objects or substances (dogs, wine) into a noun of subkinds of those objects/substances. The latter, while not being restricted to mass nouns, has as a side effect that of making a mass noun countable. These two conceptual devices seem to pretty much exhaust M → C shifts.

Let us now turn to C → M shift. Here too we find cases in which a noun that is conceptually count has a nearly as natural mass alternant.

(11)   a. I ate three apples   chicken, rabbit, …
        b. There is apple in the soup

Here we seem to be able to massify a count N, using an operation that apparently involves the notion of 'material part.' (11b) describes a situation in which there are apple parts in the soup. Again, this type of shift can be generalized well beyond the cases in (11), but the results are decidedly more marked:

(12)   There was table/bicycle all over the floor.

The notion used in the literature in this connection is that of 'universal grinder'.

These phenomena are not restricted to English. Consider the case of Mandarin. A noun like *ji* 'chicken' goes equally well with a classifier like *zhi*, usually restricted to whole objects, and with one like *pan* 'dish' that is not so restricted; moreover, at a colloquial register, *ji* can also go with the general classifier *ge* (that typically only combines with count nouns) with *both* interpretations:

(13)   a. san pan ji   'three portions of chicken'
        b. san zhi ji   'three chickens'
        c. (?) san ge ji   'three chickens' or 'three portions of chicken'

---

[6] Dayal (2004) makes crucial use of the fact that count nouns can be shifted from object level to kind level in her proposal on the singular generic uses of the article *the* (as is *the dodo is extinct*).

In contrast with this, use of the generic (count) classifier *ge* with a mass noun that has infrequent count uses like *xue* 'blood' appears to be more marked (and/or to require highly special contexts):

(14)   ??san ge xue 'three portions of blood'

This seems to confirm that also in Mandarin while *chicken* has equally natural count or mass uses, *blood* does not.

In conclusions there are a couple of ways of shifting mass to count and one way to shift count to mass. These shifts seem to be always available across languages; but they do not appear to be fully general: they are heavily context dependent and give rise to graded judgements. At the same time the pattern of gradability appears to be fairly constant across languages (with, e.g., beer more prone to being count than blood, etc.).

While there appear to be some interesting universal tendencies in the mass/count distinction, there is also systematic cross linguistic variation, to which we now turn.

2.2 Three major ways of encoding mass/count

As is well known, in languages like Mandarin Chinese *no* noun can directly combine with numerals. A classifier (i.e. a word that denotes something like a measure, a container, or shape based words that express something like 'unit') is always needed:

(15)   a. san *(ge) nanhai   b. yi *(ben) shu
            three CL boy 'three boys'   one CL book   'one book'

Languages of this sort are often referred to as classifier languages. Classifier languages do not have obligatory number marking on nouns and, in fact, it is controversial whether the singular/plural contrast is at all attested.[7] While it can be maintained that in such languages every noun has a macro syntax similar to that of mass nouns in English, it is, however, important to bear in mind that the mass/count distinction is also active in their grammar. In particular, shape based classifiers, or very unspecific ones like *ge* either do not combine with prototypical mass nouns or force a count interpretation on them, as we saw in example (13c) of the previous section. Such 'count' classifiers have also a syntactic distribution that differs along various dimensions from that of classifiers that are not restricted to count nouns (see, e.g. Cheng and Sybesma 1998). For example, insertion of the particle *de* (which indicates, roughly, modification) is possible after classifiers that go with mass nouns, but not after count classifiers:

(16)   a. *san ben de shu   b. san bang de shu
            three CL-de book   three pound-de book 'three pounds of books' or
                                                        'three pound book'

So the way in which the mass/count distinction is coded in the grammar of languages like Chinese is through the syntax and semantics of classifiers.

---

[7] In Mandarin, we find forms like *men* that in combination with nouns, indicates 'group' and is associated with the idea of plurality. Li (1999) claims that *men* actually is a real plural morpheme. However, its syntax and semantics appears to be very different from that of number marking morphemes. E.g. it is restricted to humans, it can appear on proper names, it only has a definite interpretation, etc. For an alternative to Li's analysis, cf., e.g., Kurafuji (2003).

This contrasts with what happens in number marking languages. I refer by this label on languages that have overt number features which obligatorily appear on nouns (and may appear to other components of the nouns phrase). IndoEuropean languages, UgroFinnish, etc. are languages of this type. As we will review in the next section, the mass/count contrast in number marking languages is typically centred on the distribution of singular versus plural morphology.

There is a third important type of languages, that seem to lack both obligatory number marking and obligatory classifier systems. Dëne Sųłine, for example, is explicitly described in this fashion by Wilhelm (2008). Let us briefly review what the pattern is, according to Wilhelm.[8] Sentences such as (17) show that the basic form of the noun has both singular and plural uses.

(17)   a. tth'ay   thiłtsį   sį
          dish   perf-1SG-make SG O perf   assert 'I made one dish'
        b. tth'ay   ghighą   sį
          dish   perf-1SG-make PL O perf   assert 'I made several dishes'

Moreover, inspection of the syntax of [Num N] constituents immediately reveals that Dëne also lacks obligatory classifiers.

(18)   a. sǫlaghe dzoł   b. sǫlaghe k'asba   c. *sǫlaghe bër
          five   ball   five   chicken   five   meat

At the same time, as shown in (5b.i) above, repeated here as (18c), numerals cannot combine with mass nouns. So the mass/count distinction in languages of this sort seems to show up in the syntax of numerals. Assuming that obligatory classifiers and number morphology are two different modules of the nominal agreement system (broadly construed), languages like Dëne do not show clear signs of a system of this sort. Let us call them, accordingly, '(nominal) number neutral' languages.

This gives us three quite different ways in which the mass count distinction is coded. In a schematic way, in classifier languages we detect a different behaviour of mass versus count in the classifier system; in number marking languages the distinction affects the distributions of plural versus singular morphemes, while in nominal number neutral languages it shows up in the distribution of numerals. I am not ready to speculate on whether this comes close to exhausting the main dimensions of variation in this domain. The very existence of at least these three types immediately raises the question of why the mass/count distinction should take these specific forms. In a parametric perspective, one would like to arrive at the identification of a uniform base, part of Universal Grammar (UG), and parameters of variation that concur in determining the grammars of specific languages. Readers that expect to find this program fully worked out here will be disappointed. Such a parametric approach has to deal with a host of issues in the syntax and semantics of Noun Phrases (NPs) that are as of yet not fully understood. However, I will offer some relatively precise conjectures as to where to look for some understanding of the differences between classifier and number marking languages (which are relatively better studied), as an example of how

---

[8] There is a wide variety of languages that may fit this mold. E.g. Austronesian languages like Tongan (Polinsky, p.c.).

issues of variation in semantics might be addressed.[9] The incorporation of languages like Dëne in a full blown theory of the mass/count distinction requires further work (but see, again, Wilhelm (2008) for an interesting start).

2.3 Mass nouns in number marking languages

Let us now turn to a more detailed characterization of number marking languages. First of all, let us point out that also within these languages we will find variation (albeit of a more local nature). Consider for example pluralization. Prototypical mass nouns are infelicitous in the plural in English.

(19)   a. That blood is RH Positive
        b. ?? Those bloods are RH Positive
        c. That gold weighs two ounces
        d. ?? Those golds weigh two ounces

While this is a fairly steady characteristic, variation is attested. According for example to Tsoulas (2006), Modern Greek allows fairly freely for mass noun pluralization:

(20)   Epesan nera   sto   kefali   mu
        Fell-3pl water-pl-neut-nom on-the head-neut-sg my 'water fell on my head'

As discussed in Tsoulas's work, plural marking on mass nouns seems to function similarly to an intensifier like 'a lot of,' so that (20) winds up conveying that the amount of water falling on my head is substantial.[10] In what follows, we will have to put variations such as these aside and focus on English like languages to get a more detailed sample of how the mass/count contrast may manifest itself in number marking languages.

Besides pluralization, the Determiner (D-) system (which includes articles and quantifiers) appears to be highly sensitive to the mass/count distinction. For example, while the definite article goes with any kind of noun (21a), the indefinite one only goes with count Ns (21b). The same holds for several other quantifiers like *every*. In contrast with this, there is a class of quantifiers that go with plural or mass nouns to the exclusion of singular count (21c).

(21)   Combinatorial restrictions in the D system
        a. the/some boy   a′. the/some boys   a″. the/some water
        b. a/every boy   b′. *a/every boys   b″. *a/every water
        c. *most/all boy   c′. most/all boys   c″. most/all water

Here too there are crosslinguistically stable tendencies (e.g. the existence of determiners that go with plural and mass, but not with singular count), with some variation (e.g. *no* goes with any type of N in English, while *nessuno* in Italian only goes with singular count).

---

[9] In this connection, I will be building on, and updating, Chierchia (1998a,b).

[10] Exactly the same pattern seems to hold in several native North American languages. For example, Mathieu (2009) argues that Ojibwe, an Algonquian language has a formal mass/count distinction but allows pluralization of mass nouns with the same interpretation as Greek. Gillon (2009) makes a similar claim for Innu-aimun, a dialect of Eastern Cree.

It is worth emphasizing that patterns such as (21) show the limitations of a purely lexical coding of the mass/count distinction. For suppose we simply mark some nouns as -COUNT and we use this abstract marking to drive combinatorial restrictions such as those in (21). Would that make us understand why, say, *the* goes one way and *every* the other? Or why one doesn't find Ds restricted to singular nouns (count or mass) to the exclusion of plurals? Hardly so, it would seem. While an explanatory theory of the pattern in (21) still might not be available, theories based on lexical features and selectional stipulations are unlikely to offer much insight on this score.

We have already remarked on the fact that languages like English displays the phenomenon of nouns that are cognitively count but have the distribution of mass ones: *furniture, mail, cutlery, footwear*, etc.

(22)   a. *I bought three furnitures   a′. I bought three pieces of furniture
       b. *I don't have many furnitures   b′. I don't have much furniture

I will refer to nouns of this class as 'fake mass' noun.[11] Fake mass nouns have an interesting property that we will want to try to understand. While their syntactic behaviour is substantially parallel to that of prototypical mass nouns, Schwarzschild (2007) has detected a class of predicates that actually differentiates between the fake and the prototypical. He dubs such predicates 'Stubbornly Distributive Predicates' (STUBS). They are exemplified in (20), along with the relevant paradigm:

(23)   a. STUBS: large, small, cubical, big,…
       b. Those violets are small   only distributive
       c. Those violets occupy little space   distributive/collective
       d. That furniture is small   only distributive
       e. ?? That snow is small

With plurals, stubs only have a distributive reading. Contrast (23b) with (23c); in principle (23b) might mean something like (23c). Yet (23b) only has one of the readings associated with (23c). The same goes for fake mass nouns: (23d) is well formed and has to be interpreted distributively. Instead, with prototypical mass nouns like *snow* in (23e), stubs are deviant. So with respect to this class of predicates, fake mass nouns pattern with plurals and not with prototypical mass nouns. While fake mass nouns behave just like the core ones in many respects, there are corners of grammar where they resemble more plural count nouns.

A further characteristic of fake mass nouns is that they are subject to a fair amount of variation (within closely related number marking languages). For example, French and Italian have perfectly count counterpart of *furniture*. They have, in fact, minimal pairs of the following form:

(24)   mobile/mobili   mobilia
       piece of furniture/pieces of furniture (count)   furniture (mass)

---

[11]   It is useful to underscore that fake mass nouns are not collective like, say, *group*. While a single person is not a group, a single lamp or chair would qualify as furniture:

(a) That chair is cheap furniture.

Some researchers call fake mass nouns 'count mass nouns' or 'individual mass nouns'.

The Italian translation of *footwear*, namely *calzatura/calzature* is also count, and so on. If Tsoulas (2006) is right, Greek might not have fake mass nouns at all.

What makes fake mass nouns interesting is that they constitute a fairly recurrent type of non canonical mass nouns, and yet they are subject to micro-variation among closely related languages. For all we know, the phenomenon of fake mass appears to be restricted to number marking languages. It is unclear that classifier languages like Mandarin and number neutral languages like Dëne display a class of cognitively count nouns with the morphosyntax of mass nouns.[12] In view of this intricate behaviour, fake mass nouns arguably constitute a good testing ground for theories of the mass/count distinction.

Summing up, I have laid out some properties that appear to be universally associated with the mass-count distinction (schematically summarized in (25a)). I have then looked at some major typological variations (cf. 25b). Finally, I have considered in more detail the properties associated with such a distinction in number marking languages (cf.25c).

(25)　　a.　Universals
　　　　　　　i.　The signature property: *three bloods
　　　　　　　ii.　The mapping property: *blood* is mass in every language
　　　　　　　iii.　Elasticity: some nouns are ambiguous, most can be coerced
　　　　　b.　Three nominal systems:
　　　　　　　i.　classifier languages
　　　　　　　ii.　number marking languages
　　　　　　　iii.　number neutral languages
　　　　　c.　Characteristics of mass versus count in number marking languages
　　　　　　　i.　mass nouns do not pluralize (but cf. Greek)
　　　　　　　ii.　the determiner system is sensitive to the mass/count distinction
　　　　　　　iii.　existence of fake mass nouns

The properties in (25) constitute an ideal testing ground for theories of the mass/count distinction, as they force us to face the issue of the universality of the distinction vis-à-vis the language particular character of its manifestations. In what follows, I will offer a new account of these properties.

3 Background

The present section is devoted to a quick review of the assumptions I will be making. I do this to be precise enough to allow my claims to be duly checked. But I hope that most of the present work will be understandable without necessarily following all of the formal details. I adopt a minimalistic set up for syntax, where Logical Forms constitute the level at which entailment, presupposition, etc. are defined. Such a definition will take the form of a compositional mapping of LFs into representations in a formal language, that are stand ins for model theoretic objects. In the first subsection below, I briefly review matters related to the relevant representation language and the notational conventions I am assuming. In the second subsection, I review recent work

---

[12] Or, to put it differently, in Mandarin every count noun sort of behaves like a fake mass one.

on plurals and kind reference that forms the theoretical backbone of the approach to
be developed.

3.1 The representation language

I adopt as a representation language the typed system TY2 Gallin (1975), with basic
types e ('entities'), t ('truth values') and s (worlds). Complex types are recursively
built out the basic ones through the function space construction; accordingly they have
the form $\langle a,b \rangle$, the type of all functions (partial or total) from objects of type a into
objects of type b. As a logical basis, I will adopt a strong Kleene logic for connectives
and quantifiers. An intensional object will have the form $\langle s,a \rangle$ of functions from worlds
(or situations) to objects of type a. So for example, the property of running will be
represented as in (26a):

(26)   a. $\lambda w \lambda x$ [run(w)(x)] $\langle s,\langle e,t \rangle \rangle$
       b. $\lambda w \lambda x$ [run$_w$(x)]
       c. λx [run$_w$(x)] $\langle e,t \rangle$
       d. {x: run$_w$(x)}
       e. A∩B

I will usually put the world coordinate as a subscript as in (26b). In fact, often I will
omit making the world variable explicit, pretending to work in an extensional setting
(and that will largely suffice for most of our purposes). When I do so, however, I will
still be talking of expressions of type $\langle e,t \rangle$ as properties (as I really am assuming a
intensional entity with the world variable filled).

I will also assume that TY2 is enriched with set theoretic notations such as those in
(26d,e). In particular, (26d) is understood as an expression of type $\langle e,t \rangle$ that denotes a
total function (intuitively, the completion of the function in (26c)); '∩' as an expression
of type $\langle \langle a, t \rangle, \langle \langle a,t \rangle, \langle a,t \rangle \rangle \rangle$, for any type a, etc.

I assume that TY2 may contain indexical elements (e.g. constants like 'I', 'you',
etc.). Accordingly, the interpretation for TY2 expressions will be relativized to a con-
text c (and, of course, a value assignment to variables g). A model for (this version of)
TY2, will, therefore look as in (27a) and the semantics will take the form of a standard
recursive definition of the interpretive function in (27b), which I won't spell out here.

(27)   a. A model M is a tuple of the form $\langle U, W, C, F \rangle$ where U is the set of indi-
          viduals, W the set of worlds, C the set of contexts and F the interpretation
          function; for any a, D_a is the domain of entities of type a; moreover:
          i. Basic types: $D_e = U$ ii. $D_s = W$ iii. $D_t = \{0,1\}$
          iv. Complex types: $D_{\langle a,b \rangle} = [D_a \rightarrow D_b]$
          v. For any constant α of type a and any context c ∈ C, $F(\alpha)(c) \in D_a$
       b. Recursive definition of $||\beta||_{M,c,g}$.
          i. if β is a constant of type a, $||\beta||_{M,c,g} = F(\beta)(c)$; if it is a variable,
             $||\beta||_{M,c,g} = g(\beta)$
          ii. $||\beta(\alpha)||_{M,c,g} = ||\beta||_{M,c,g}(||\alpha||_{M,c,g})$, etc.
       c. Truth. An expression of type t is true in a model M, relative to a context c
          and an assignment g, iff $||\varphi||_{M,c,g} = 1$

3.2 Singularities, pluralities and kinds

The outcome of much recent work on the singular/plural[13] distinction has led to the hypothesis that in each situation (state of affairs/world) w, the domain of discourse (of entities of type e) U_w has a structure that can be schematically represented as follows:

(28)
$$a \cup b \cup c \cup d \ \dots..$$
$$a \cup b \cup c \qquad\qquad a \cup b \cup d \ \dots$$
$$a \cup b \quad b \cup c \quad a \cup c \quad a \cup d \ \dots.$$
$$a \qquad b \qquad c \qquad d \qquad e \ \dots.$$

Such a domain constitutes an atomic semilattice closed under a join operation '$\cup$' ('group formation')[14] and partially ordered by a 'part of' relation '$\leq$' (so that, e.g., $a \leq a \cup b$, $a \cup b \leq a \cup b \cup c$, etc.). The atoms of this domain are the elements at the bottom of the structure in (28). They are formally defined as entities that have no proper parts in the relevant sense[15] and they are typically used to represent the denotation of definite singular expressions (like *John* or *that cat*). Generally speaking, the notion of 'atomicity' is a relative one; it is helpful to think of it in terms of a function AT that extracts from properties their smallest elements and from individuals their smallest components as follows:

(29) a. Examples:
      i. $AT(\{a, b, a \cup b\}) = \{a,b\}$
      ii. $AT(\{a \cup b, b \cup c, a \cup b \cup c\}) = \{a \cup b, b \cup c\}$
      iii. $AT(a \cup b \cup c) = \{a, b, c\}$
   b. Definition of AT:
      i. If P is of type $\langle e,t \rangle$, $AT(P) = \{x \in P : \forall y \in P \ [y \leq x \rightarrow x = y]\}$
      ii. If x is of type e, $AT(x) = AT(\lambda y[y \leq x])$
   c. Remark: for any x, $AT(x) \subseteq AT(U)$ (i.e. AT(x) always yields absolute atoms)

So, AT applied to a property P returns its 'relative atoms'; at the same time AT applied to the universal property U or to individuals embodies an absolute definition of atomicity. Non atoms (like $a \cup b$) are used in first approximation to represent the denotation of plural definite expressions, like *John and Bill* or *those cats* (but this will have to be partially qualified later on).

Nouns divide the domain of discourse in sortally homogenous subspaces. For example, if a, b and c are all the cats there are, the extension of the singular noun *cat* and that of the plural noun *cats* might be represented as follows:

---

[13] Stemming from Link (1983), Landman (1989), Schwarzschild (1996).

[14] From now on, I will use '$\cup$' ambiguously for set theoretic union and for the join operation over entities of type e. In particular, if A is of type $\langle e,t \rangle$, $\cup$ A is of type e and is the member of U that constitutes the sum of all the things A is true of.

[15] Being an atom in the technical sense just introduced does not entail lacking structure. There might be complex 'part of' relations among atoms (in the sense in which, say, my arm is part of me).

I leave open here the status of singular group level expressions like *committee* or *group*. For discussion of the latter, see, e.g. Schwarzschild (1996), Barker (1992), among others; but cf. also our discussion of nouns like quantity in Sect. 5.2.

(30) a.
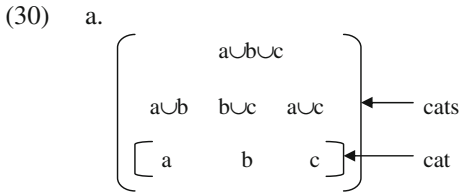
b. Pluralization: for any P, $*P = \lambda x \, \exists Q[Q \subseteq P \wedge x = \cup Q]$

The diagram in (30) embodies the view that the singular noun *cat* is true just of cat-atoms. From the set of cat atoms we can obtain the corresponding plural property via the closure operation '*' defined in (30b). Following Sauerland (2003) and much previous research (and contra e.g. Chierchia 1998a), I am assuming that the plural noun *cats* includes singularities in its extension (and hence it is, in some sense, number neutral). I will use the caps CAT to refer to such a number neutral property. The main argument in favour of the view that plurals are number neutral has to do with the semantics of plural negative quantified Determiner Phrases (DPs) like *no cats*. In a nutshell, if *cats* excluded singularities a sentence like *there are no cats on the mat*, would be saying that there is no group of cats on the mat. Such a sentence would be expected to be true in a situation where there is just one cat on the mat, which seems wrong.[16] To get sentences of this sort right, singularities have to be included. This raises the issue as to why positive sentences like *there are cats on the mat* cannot be used to describe a situation in which there is only one cat on the mat. The idea that is being pursued in this connection is that this effect might be due to a scalar implicature.[17]

There is one further aspect of the structure in (28) that turns out to be useful to discuss. Consider the plurality $a \cup b \cup c$ in (28), under the hypothesis that the cat-property is as in (30). The plurality $a \cup b \cup c$ is special: it constitutes the totality of the cats there are. This is different from dishomogeneous pluralities (like, say, the join of my cat and Brooklyn bridge). A homogeneous plurality like $a \cup b \cup c$ might be identified with the cat-kind. So in general, maximal homogeneous pluralities can be thought of as 'kinds'.

The notion of kind has proven very useful in analyzing certain types of nominal arguments, particularly bare plurals and, on some views, definite singular generics. This is often illustrated through examples like (31):

(31) a. Cats are rare/come in different sizes/are everywhere
      b. The cat is rare/evolved from the Saber tooth tiger

---

[16] Chierchia (1998a) tries to get out of this by proposing that *no* quantifies over parts of the supremum. However, a properly plural property would denote the empty set in a world with just one cat. In such a world there would be no supremum on whose parts to quantify. This leads to various technical complications.

[17] In a situation with only one cat on the mat, the sentence *there is a cat on the mat* would be a better (logically stronger) characterization of the facts than *there are cats on the mat*, and this would trigger an implicature. As is well known scalar implicatures are sensitive to the polarity of their local context and thus canonical implicatures tend to be absent from downward entailing contexts, which explains the asymmetry under discussion in the interpretation of plurals. On these issues, see Chierchia (2004), Chierchia et al. (2008), Spector (2007) and references there in.

No single cat or group of cats is rare or comes in different sizes or can be everywhere. These seem to be properties that only the cat-kind can have. It is on the basis of examples of this sort that it has been proposed that sometimes bare plurals in English are kind denoting. In fact, according to certain analyses, English bare plurals are *always* kind denoting (cf. Carlson 1997; Chierchia 1998b). Be that as it may, what is important, from the present stand point is that (i) some NPs are kind denoting, (ii) the domain of quantification extends to kinds, (iii) kinds are related to pluralities.

The hedge 'are related to' is due to the fact that at a closer look, totalities like $a \cup b \cup c$ might be too extensional to represent kinds (just like characteristic functions are too extensional to represent properties). For example, two extinct kinds, like dinosaurs and dodos, don't have any manifestation/instances, yet they are clearly distinct (just like the corresponding properties are distinct). So, we may actually want to represent kinds as individual concepts of type $\langle s,e \rangle$ (i.e. functions that at any world yield the totality of the manifestations of that kind in that world). Kinds versus properties are different and at the same time related. Let us focus on their differences first. Minimally, there is a difference in semantic type, i.e. in logico-semantic role. In the present setting, kinds are a special sort of individual concepts. Properties apply to individuals (and, in case of properties like *rare*, *extinct*, etc., to individual concepts) to yield truth or falsity. From an ontological point of view, we might regard kinds as relatively 'concrete' inhabitants of the world, and properties as something more 'abstract' (perhaps mental capacities). [18] Semantically speaking, I will be assuming (simplifying somewhat) that plural NPs denote kinds when they occur in argument position as in (32a); they denote properties when they occur in predicate position or as restrictions of determiners as in (32b).

(32)   a. i. Cats are common   ii. Dogs have evolved from wolves
          b. i. Those are cats   ii. Most cats like their whiskers

We have dealt so far with how kinds differ from properties. Now let us consider their relatedness. Look at the cat-kind, as construed here, and compare it with the number neutral property CAT (which we take to represent the extension of the plural N *cats*). Clearly, if you know what the cat kind is, you also know what the CAT property is and vice versa. In some sense, kinds and properties constitute two distinct ways of coding the same information. They 'classify' the domain of individuals in parallel ways. In the simplified model we are building up, we can define very simple morphisms that take us from kinds to properties and vice versa. Such morphisms 'commute' as shown in (33c):
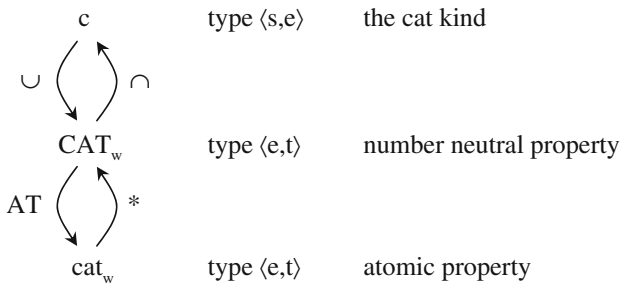
(33)   a. for any kind k, $^{\cup}k = \lambda w \lambda y \, [y \leq k_w]$[19]
          b. $^{\cap}P = \lambda w [\iota x P_w(x)]$
          c. i. $^{\cup \cap}CATS = CATS$ ii. $^{\cap \cup}c = c$ (where c is the cat kind)

Summing up, we get the following picture:

---

[18] I leave it open here whether kinds are a subtype of the type e of individuals. The requirement that, in my terms, $\langle s,e \rangle$ be subtype of e creates a potential cardinality problem (since the function space $D_{\langle s,e \rangle}$ must be essentially bigger than D_e), which calls for some modification of the background theory. Property theoretic frameworks (e.g. Chierchia and Turner 1988; Fox 1998) or some other non standard set theory are quite useful in this connection.

[19] In worlds w where k_w is undefined, $^{\cup}k_w = \lambda y \, [y \neq y]$

(34)   The 'semantic triad':

$$
\begin{array}{ll}
c & \text{type } \langle s,e \rangle \qquad \text{the cat kind} \\
\end{array}
$$

$$
\begin{array}{ll}
\text{CAT}_w & \text{type } \langle e,t \rangle \qquad \text{number neutral property} \\
\end{array}
$$

$$
\begin{array}{ll}
\text{cat}_w & \text{type } \langle e,t \rangle \qquad \text{atomic property} \\
\end{array}
$$

The diagram in (34) represents the denotation space of a noun in its various syntactic roles; nouns take their values from this semantic triad depending on the syntax and morphology of various languages. Some aspects of NP semantics might be universal, others may reflect language particular choices. Be that as it may, it is important to bear this picture in mind, as it will enable us to relate my proposal on mass/count directly to current work in semantics.

4 A vagueness framework for singular/plural structures

Virtually all natural concepts are vague and various scholars have argued that vagueness plays an important role in grammar.[20] Our claim is that if we integrate singular/plural structures with vagueness, we can't really miss where the mass/count distinction comes from and the mass/count contrast gets to be explained in terms of independently needed notions. This will constitute the main argument, I submit, for the present proposal. In what follows, I first present the idea in informal terms, then I spell it out with supervaluations/data semantics.

4.1 Properties with vague minimal arguments

The present proposal comes in two parts: (i) numbers are cardinality predicates of sets (/property extensions) and (ii) the sets we can count with must have *minimal* elements that are sufficiently well defined (= not too vague). Let me elaborate on these two ideas.

That counting always involves individuating a level at which to count, can be illustrated by paraphrasing a famous example by Kratzer (1989). Suppose I ask you how many things/objects there are in this room. First, you may have problems understanding my question; but more importantly an answer like the following will never do: "Here is a table. That's one object. Then, there is the table's right leg, and that is a second object. Then there is the upper part of the right leg…" The role of sortal nouns in combination with numerals is precisely that of making explicit the level at which counting is to take place. In particular, if we use a noun true of aggregates of varying size (such as a plural), we count the smallest things to which the noun applies. In the case of mass nouns such 'smallest things' are too vaguely specified to be counted.

---

20 Cf. Kennedy (2007), Lasersohn (1999), Pinkal (1989), among many others.

Let us flesh this out a bit (still informally). Consider the number neutral property CAT that we suggested might be the denotation of the plural noun *cats*. As we observed, virtually every property, including CAT has vague boundaries. Take an individual cat, modify it along some of its dimensions and it will come a point where you will become uncertain as to whether it still is a cat. Consider age, for instance. A newborn kitten is a cat; as it probably is a cat also just before coming out of its mother's womb. But at what point a cat embryo becomes an actual cat? While there may be a fact of the matter, it is hard to determine what it is even in principle. Or take a cat and imagine surgically removing various parts of it. At some point (when?) you won't have a cat anymore. Or, finally, and (only slightly) less gruesomely, imagine cat breeders that perform genetic engineering in search of new breeds. When exactly a new genetically altered descendant of the original cat breed ceases to be a cat? Let us dub this kind of vagueness 'inherent vagueness'. Inherent vagueness is such that even an expert will typically be unable to provide uncontroversial criteria according to which a property applies or not to a certain individual or a group thereof. It is a matter of how, for various purposes, we may wish to cope with unclear, 'marginal' instances of a certain property.

It may be useful to compare, in a preliminary way, the inherent vagueness of CAT with the vagueness of scalar adjectives like 'tall'. According to, e.g. Kennedy (2007) and related work, the analysis of "that is tall" is roughly: what we are pointing at stands out in height with respect to some contextually determined degree d. Many factors enter into the determination of this contextual standard: what kind of objects we are pointing at (houses? people?), what the average height of the relevant comparison class is, how much higher than the average the object in question is, etc. Even after all these questions have been settled, we still will be left with a range of values that might fit the bill, and that constitutes the inherent vagueness of *tall*. Inherent vagueness affects how we count: the inherent vagueness of *cat/s* determines how many cats there are; the vagueness of *tall* determines how many tall things (buildings, people, or what have you) there are. The inherent vagueness of mass nouns affects the way we count in a more radical way: it prevents us from counting with mass predicates.

Here is why. Consider again the number neutral property CAT that applies to individuals as well as to groups of cats. When we use it to count as in *three cats*, we count individuals, i.e. the cat-atoms (which constitute the smallest things to which CAT applies). Now, there are plenty of cat-atoms that are *not* vaguely specified. There are plenty of things that we (or the relevant experts) are sure fall under the cat concept, have the cat-property or however you want to put it. In other terms, the boundary of the property 'cat/s' is such that that there definitely are x's that fall under it, such that no proper part of x does. We have a reasonably clear idea of what qualifies as a (more or less 'whole') cat atom.

Now contrast this with what happens with mass nouns. Consider rice. A spoonful of rice is rice. What about a single grain of rice? In many contexts we would not consider a single grain of rice enough to reach the threshold of significance. To a child saying she has finished her rice, no parents in their right mind would reply 'no you have not' upon detecting a single grain. Yet for some other purposes, we might consider a grain of rice, rice. But, then that applies to

half grains as well. And to quarters of grains. In certain cases, we may regard rice flour as rice (as when we say 'there is rice in this cake'). The point is that there is no systematic basis for deciding which rice amounts qualify as rice atoms.

Kennedy compares what he calls relative gradable adjectives like *tall* to absolute gradable ones like *dry* and observes that while for the latter there is a natural cut off point that separates dry things from non dry things (even though its exact characterization might still be vague), for adjectives like *tall* such a natural cut off point separating, say, tall people from non tall ones is much more elusive. The same, I suggest, applies to the mass/count contrast. In considering smaller and smaller instances of the property CAT, there is a cut off point such that if you go smaller, you won't have a cat anymore (even though where such a cut off point lies may be somewhat vague); on the other hand in considering ever smaller water samples the cut off point that separates water from non water remains way more elusive.

You see where this is heading. If in counting directly with a property P we count P-atoms, and such atoms happen to be all vaguely specified, as all fall outside of the 'safe' boundaries of the relevant property, we are stuck. We don't know what to count, not even in principle (although we will of course be able to measure). It is also clear how this links up with the prelinguistic contrast between objects and substances as defined by Spelke. Objects in Spelke's sense are clearly bounded and hence constitute good candidates for being atoms of the relevant property. Substances manifest themselves as aggregates, pluralities whose components elude our perceptual system and/or our pragmatic need for clear identification.

Finally, it should also be evident how this immediately explains why 'core' mass nouns are going to be so stable (with respect to our capacity to count with them) across languages. The perceptual properties of our species are uniform and are not going to afford us access to the minimal components of fluids or minerals. And the pragmatic/communicative interests of a community in dealing with things like rice are unlikely to be enhanced by fixating on, say, whole rice grains (as opposed to other ways of apportioning that cereal in slightly bigger or slightly smaller units).

My goal in the following section is to build a model against which to test this basic intuition.

4.2 Supervaluations and the mass/count distinction

In this subsection, I first introduce the concept of supervaluation (in a way directly inspired by Veltman (1985)), then I show how it can be integrated with singular/plural structures. Vague predicates P are interpreted by partial functions from individuals into truth values (relative to a context c that fixes any context dependent parameter P may have). In virtue of such partiality, a property P_w of type $\langle e,t \rangle$, is going to be associated a positive extension $P_w{}^+$ (the set of all things for which it yields 1) and a negative one $P_w{}^-$ (the set of all things for which it takes value 0). Things for which P_w is undefined are said to fall into P_w's truth value gap, which we take here to represent

P's vagueness band.[21] The following is a graphic representation of a vague (atomic) predicate:

(35) P_w^+ P_w^-

          a   b c   d   e f   g   h   i   j   k

          positive extension   gap   negative extension
                        P's vagueness band
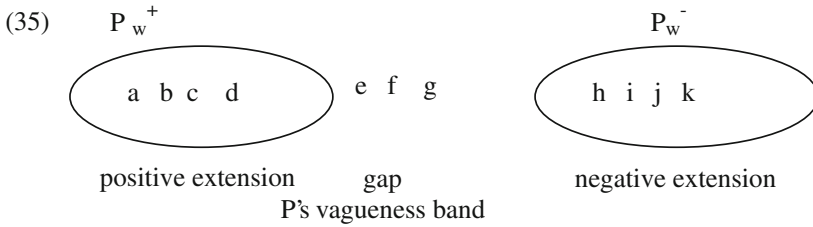
Now we can imagine sharpening our criteria for being P_w by shifting from c to a context c′ in which fewer things fall within the vagueness band. Through successive sharpenings (i.e. through further context changes) we may reach a point in which P_w becomes total. Clearly, there is no single way of making P_w total. But we can be sure that the things that are in P's positive extension relative to the base context c will remain stable no matter how we make c more precise. We can represent this by ordering contexts as follows: we say that c ∝ c′ (to be read as "c′ is a precisification of c") iff for every P and every world w, P_w's vagueness band relative to c′ is smaller or equal to P_w's vagueness band relative to c. We can then say that a formula ϕ of type t is definitely true (in symbols Dϕ) relative to a context c iff ϕ is true relative to every total precisification c′ of c. This has the usual effect of supervaluations: tautologies like $\neg[P(a) \wedge \neg P(a)]$ come out as definitely true and contradictions as definitely false even if P(a) does not have a stable value relative to every context.

Let us spell out the relevant notions.[22]

(36)   a. A vagueness Model for TY2 is of the form ⟨U, W, C, ∝, F⟩, with ∝ a
            partial order over C; contexts that are minimal with respect to ∝ are called
            'ground' contexts; contexts that are maximal relative to ∝ are called total
            precisifications.
        b. F is defined as in (27a.iv); furthermore, for any constant P of type ⟨s,⟨e,t⟩⟩,
            any context c, c′ and any world w,
            i. if c ∝ c′, then $F(P)(c)(w)^+ \subseteq F(P)(c')(w)^+$ and $F(P)(c)(w)^- \subseteq F(P)$
               $(c')(w)^-$
            ii. If c is ∝-maximal, then F(P)(c)(w) is total
        c. For any expression β, the interpretation function $||\beta||_{M,c,g}$ is as before, with
            the following addition:
            $||D\phi||_{M,c,g} = 1$ iff for every total precisification c′ of c, $||\phi||_{M,c',g} = 1$

---

[21] This picture is a simplification. In reality, properties have range (a set of things to which they can meaningfully apply) typically determined by their presuppositions. For example, one might argue that for something to qualify as cat or not it has to be (at least) animate. So a predicate P may come with a range R of things to which P may in principle apply. The vagueness band of P is intended to fall within such range. If being a cat is restricted to animate things (under normal circumstances), then its vagueness band should also be thus restricted. These issues are complex and cannot be properly addressed here. I think, however, that the complications that arise are orthogonal to our present concerns.

[22] The definition to be given will remain somewhat imprecise, in, I hope, an innocent way. For example, I limit myself to a consideration of partiality only relative to one place properties. Its generalization to other types is routine.

We now integrate singular/plural structures with the approach to vagueness just sketched. In essence, what has to happen is that being atomic (i.e. AT) becomes a vague and context dependent notion. A vague model M for plurals will contain a set of individuals U, and a *partial* join operation ∪ over U. The individuals, relative to a context c, will be divided between things that we are sure are atoms $AT_c(U)$, things we are sure are sums $\Sigma_c(U)$ and things we don't know whether they are atoms or sums. The latter constitute the 'unstable' individuals, as they may fall on either side of the atom/sum divide in different precisifications. So, some sums will be the sums of stable individuals (i.e. are grounded on definite atoms relative to c). Others are not; they are the sums of unstable individuals. This type of 'partial sums' ($P\Sigma_c(U)$) lend themselves naturally to the modelling of mass nouns (reflecting the idea that the latter are aggregates whose components are unknown). In what follows, I first lay out the axioms that govern vague plural structures; then I enrich the language of TY2 so as to reflect these changes.

(37)   Vague plural structures.
        For any context c, AT_c, Σ_c are functions such that:

   a.   i. $AT_c(U) \subseteq U$, $\Sigma_c(U) \subseteq U$; $AT_c(U) \cap \Sigma_c(U) = \varnothing$
      The atoms and the sums are disjoint subsets of U.

   b. $STB_c(U) = AT_c(U) \cup \Sigma_c(U)$
      The stable individuals are the definite atoms and the definite sums

   c. $STB_c^-(U) = U - STB_c(U)$
      The unstable individuals are the complement of the stable ones.

   d. $P\Sigma_c(U) = \{x \in \Sigma_c(U): \exists A\ (A \subseteq STB_c^-(U) \wedge x = \cup_c A)\}$
      The partial sums are the sums generated by unstable individuals.

   e. $\cup_c$ is a partial join operation on U; for any $A \subseteq U$, $\cup_c A \in \Sigma_c(U)$ (if defined).
      If $\cup_c A$ is defined, $\cup_c A$ is a sum.

   f.   For any $A \subseteq AT_c(U)$, $\cup_c A \in \Sigma_c(U)$;
      For any subset of the atoms, their sum exists and is in $\Sigma_c(U)$.

   g.   For any $x \in \Sigma_c(U)$, there is an $A \subseteq U$ such that $x = \cup_c A$
      All sums are sums of something (and hence have parts)

   h. $x \leq_c y =_{df} y = x \cup_c y$
      Standard definition of the partial order '≤'; but now $x \leq_c y$ is defined iff $x \cup_c y$ is.

   i.   If P is of type ⟨e,t⟩, $AT_c(P)(u) = \{x \in P^+: \forall y \in P^+\ [y \leq x \rightarrow x = y]\}$
      For any P, $AT_c(P)$ selects the smallest P-members like before.

   j.   If x is of type e, $AT_c(x) = AT(\{y: y \leq_c x\})$
      If x is an individual, AT_c selects its components like before. But '$\leq_c$' is now a partial, context dependent relation.

(38)   Additions to the language of TY2.
   a.   ∪, ≤ and AT are constants of type ⟨e,⟨e,e⟩⟩, ⟨e,⟨e,t⟩⟩ and ⟨a,⟨e,t⟩⟩ (where a = e or a = ⟨e,t⟩), respectively.
   b.   For any context c,

    i. $F(\cup)(c) = \cup_c$
   ii. $F(\leq)(c) = \leq_c$
    iii. $F(AT)(c) = AT_c$

We can now define a notion of stable atom using the 'definitely' operator:

(39)   **AT**$(P) = \lambda x\, D[AT(P)(x)]$

The stable atoms of a property are the atoms of that property in every precisification of the original context. By construction, everything which is a P-atom relative to a context c, will remain such in every precisification. Stable atomicity will differentiate count from mass nouns, as we will see.

4.3 Sample lexical entries

What we are doing can be informally summarized through the following picture. Here is how in any context c, the domain of individuals is going to look like:

(40)
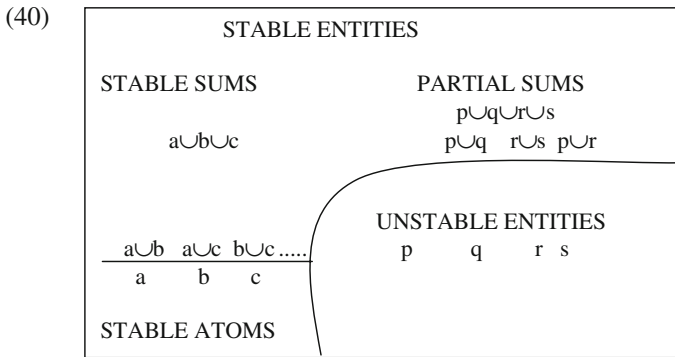
The structure of the domain U (relative to a context c)

These formal structures can be put to use by letting mass nouns be generated out of unstable entities and count nouns out of stable ones. Considering the 'semantic triad' in (34), I am going to assume, for explicitness' sake, that nouns map onto number neutral properties on which number morphology operates, but nothing hinges on that (as we will see below, the very same results can be obtained by mapping nouns into kinds).

(41)   For any model M, ground context c and assignment M:
        a. If N is count, $||N||_{M,c,g} = {}^*A$, where $^*A$ is the closure under $\cup$ of some partial function $A \subseteq AT_c$ (U)
        c. If N is mass $||N||_{M,c,g} = {}^*A$, where $^*A$ is the closure under $\cup$ of some partial function $A \subseteq STB_c^-$.

The lexical entries in (41a,b) are self explanatory. For count nouns, the ground context chooses the level of vagueness for each lexical entry (deciding, e.g. whether mature cat embryos count as cats, as non cats or whether they fall in the truth value gap) and supervaluations do the rest. As for mass nouns, they are treated as $\cup$-closed sets

of partial sums. Counting is subject to two laws: (i) we count the minimal elements to which a property applies and (ii) the property used for counting must have stable minimal entities (i.e., the same ones in every precisification). In case of count nouns like cat/s we are guaranteed by axiom (41a) that there are stable individuals to be counted; in the case of, say, water, instead, we know that they are generated by unstable individuals and hence we do not really know what to count. The vagueness band of a count noun can be viewed as 'horizontal': it takes place at the level of the stable atoms; the vagueness band of a mass noun is 'vertical': the generating elements may be split into components more than one way.

It might be worth considering various types of vagueness affected nouns, for comparison. Let us start with the paradigmatic vague noun *heap*. The problem is that heaps form a continuum with respect to their height and it seems impossible to give firm criteria for when something along such a continuum is a heap.

(42)

$\quad$ a. $\qquad\qquad\qquad$ b. $\qquad\qquad$ c.

Heap being so vague, in different contexts all, none or some of the things in (42) may count as heaps. But in every context c where F(heap)(c) is non empty it has to be a subset of $AT_c(U)$ in virtue of (41a); and once something is in $AT_c(U)$ it will remain so for every precisification of c. If for example you decide to count (42c) as a heap-atom in a context c, there can't be any precisification of c in which (42c) is going to turn into more than one heap. This is what I mean when I use the metaphor that the vagueness of a count noun is 'horizontal'. Let us dramatize this further, by building on an example I owe to Alan Bale. Consider the word *mountain* and imagine something like the following:

(43)

$\quad$ a

$\quad$ b $\quad$ b'
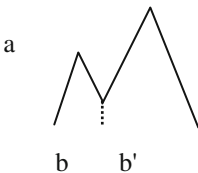
Do we have in (43) one or two mountains? Take your pick/peak. Puns aside, many factors may enter into this decision. If such factors weigh in favour of treating (43a) as a single mountain-atom, then you will be setting yourself in a (ground) context c which is committed to that choice and in no precisification of c (43a) is going to become two mountains; if on the other hand you opt for treating (43b) and (43b′) as two mountains, then you are going for a context c′ in which (43b) $\cup_{c'}$ (43b′) (which coincides spatiotemporally with (43a)) is a plurality of mountains. What is crucial is that c′ *cannot* qualify a precisification of c. Let me show why.

(44) $\quad$ a. (42a) $\in AT_c(U)$, (42b) $\in AT_{c'}(U)$, (42b′) $\in AT_{c'}(U)$; by hypothesis
$\qquad$ b. if c $\propto$ c′, then $AT_c(U) \subseteq AT_{c'}(U)$; by definition of '$\propto$'
$\qquad$ c. suppose that (43b) $\cup_{c'}$ (43b′) $=$ (43a) then:
$\qquad\qquad$ i. (43a) $\notin AT_{c'}(U)$; by definition of AT

ii. $AT_c(U) \nsubseteq AT_{c'}(U)$; by definition of '$\subseteq$'
iii. $\neg c \propto c'$; from (b) and (c.i), by contraposition.

Notice that in principle, the system leaves us free to regard (43a), (43b) and (43c) as three distinct mountain atoms, something we would never want to do. But the same is of course true of any current approach to the plural/singular contrast (regardless of vagueness related issues). We must independently require (on anyone's theory) that for any concrete sortal noun N, its atoms are chosen so as not to overlap spatiotemporally.

To put it differently, a disagreement over whether what you see in (43) is one or two mountains is, in the first place, a disagreement on how to resolve contextual parameters. The key difference between count nouns like *heap* or *mountain* and mass nouns like *rice* is that minimal rice amounts, once contextually set, can still be viewed units or aggregates without re-negotiating the ground rules. Again, the contrast in question is best illustrated by means of an example. Here is a shot at defining the rice-character:

(45)   For any ground context c, world w, individual u:
   i. $F(rice)(c)(w)(u) = 1$ iff $u \in P\Sigma_c(U)$ is a rice aggregate of size s(c);
   ii. $F(rice)(c)(w)(u) = 0$ iff u is not rice;
   iii. $F(rice)(c)(w)(u)$ is undefined otherwise.

This definition leaves rice amounts smaller than the measure s(c) set for ground contexts in the vagueness band. Which guarantees that there will be precisifications of ground contexts in which smaller amounts of the relevant protein will count as rice. So *rice* has contextually supplied smallest parts, but lacks stable atoms (because the vagueness of the smallest rice parts will be resolved differently in different precisifications). The following are consequences of the axioms in (41):

(46)   a. For any M, c, and g,
   if N is count $||\mathbf{AT}(N)||_{M,g,c} \subseteq AT_c(U)$ if N is mass $||\mathbf{AT}(N)||_{M,g,c} = \varnothing$
   = any count N has stable atoms   = no mass nouns has stable atoms

Summing up, heaps and mountains are vague and context dependent in very different ways than rice or sand.[23] The present theory offers a formal take on such differences. One that matches well with the plain observation that the basic components of core mass nouns are not accessible to our perception and not worth our while dwelling too much over. This won't impair an effective use of the noun, but it will impair using it for counting.

5 The mass/count distinction in number marking languages

In the present section I put to further test the theory developed in Sect. 4, by applying it in more detail to number marking languages. Focussing on English, I will in turn discuss (i) numbers, (ii) relational nouns like *quantity of/part of*, (iii) elasticity, and (iv) number marking proper. This should give a fair idea of the scope of the present approach.

---

[23] The same holds of nouns like *wall* or *fence*, discussed by Rothstein (2008).

5.1 Numbers

Many linguists have argued that numbers are adjectival in nature. This fairly traditional claim, whose philosophical roots can, in fact, be traced back to Frege's and Russell's conception of numbers, has been recently pursued within frameworks close to the one adopted here in, e.g. Landman (2004), or Ionin and Matushansky (2006). I am going to specifically follow the latter and assume that numbers are of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$, i.e. the combine with properties to return properties. *Three* combines with *cats* and returns a property true of groups of exactly three cats. When embedded in a sentence, as in *three cats are purring*, the NP *three cats* is subject to a default operation of existential quantification (Existential Closure), so that the truth conditional import of the sentence in question can be spelled out as "there is a group of three cats that is purring". This is summarised in (47).

(47)   a. Three cats are purring
       b. $\exists x$ [three(cats)(x) $\wedge$ purring (x)] $= \exists x[\ \mu_{AT,\text{cats}}(x) = 3 \wedge$ purring(x)] [24]
       c. $[_{\text{NumP}}$ n[cats]] $= \lambda x[\mu_{AT,\text{cats}}(x) = n]$ NumP = Number Phrase
       d. For any x, and any P, $\mu_{AT,P}(x)$ is defined only if i. P(x) holds and ii. P has stable atoms;
         if defined, $\mu_{AT,P}(x) =$ the number of *stable* P-atoms that are part of x

The approach in (47) builds into numbers a presupposition that makes them defined only relative to properties that have stable atoms. If a property is mass, it will lack stable atoms; and hence any measure function based on the counting of stable atoms will fail to be defined. In principle, mass nouns can be coerced into count construals; so one might imagine that in combining a numeral with a property P, if P is not stably atomic, P might be automatically coerced into its count counterpart. But as we have already observed and as we shall see in more detail shortly, such a coercion is subject to fairly severe limitations, and it doesn't work in all cases. Hence mass to count coercion can't be expected to rescue failures of numeral-noun combinations across the board.

    This is one of several ways of formally capturing the observation that numbers cannot, in general, combine with properties that are not stably atomic.

5.2 Quantities and parts

We now turn to an analysis of what in English appear to be relational nouns like *quantity of*/*part of*. This is useful for two reasons. First, it enables us to understand how it

---

[24] I am finessing on many things here. In particular, I am ignoring the fact that *purr* is a distributive predicate. For a discussion of the relevant issues, see, e.g. Schwarzschild (1996) and references therein. I am also not spelling out exactly how existential closure takes place (see, e.g. Heim 1982; an alternative way to go is Chierchia (2005)). Note, moreover, that while the lexical semantics of *three* is 'exact' (i.e. *three cats* is true of groups of exactly three cats), at the sentence level we get an 'at least' semantics (i.e. (40b) is true if at least three cats purr). I assume that the exact semantics at the sentential level arises as an implicature (see Horn (1989), Levinson (2000) for discussion; recent relevant developments of the theory of implicatures can be found in Chierchia et al. (2008)). Finally, the present analysis is compatible with a compositional analysis of complex numbers; but I won't pursue this here.

can be that expressions like *that water* versus *those three quantities of water* can refer to the very same entity while having a different status with respect to the mass/count distinction. Second, such an analysis paves the way for investigating the phenomenon of elasticity, to be dealt with next.

Compare the two sets of expressions in (48) and (49).

(48) a. i. a quantity of apples   ii. two quantities of gold
      b. i. a quantity of those apples[25]   ii. two quantities of that gold
      c. i. (?) a quantity of that person   ii. two quantities of that pizza

(49) a. i. *a part of apples   ii. *a part of gold
      b. i. a part of those cakes   ii. some parts of those cakes
      c. i. a part of that person   ii. two parts of that rice

Quantity combines with bare nouns (cf. (48a)), in what is known as the pseudopartitive construction, while *part* does not (cf. (49a)). Moreover *quantity* (in the singular) is slightly degraded in combination with singular count nouns. This may suggest that *quantity* combines felicitously with expressions denoting something over which the ≤-relation is (non trivially) defined, i.e. plural count and mass properties (like *apples* or *gold*) and plural count or mass definite NPs. *Part* is instead defined over individuals (expressions of type e) regardless of whether they are sums or singularities.

Building on this observation, we may analyze *quantity* as a function of type $\langle a, \langle e, t \rangle \rangle$, where $a = e_\Sigma$ (i.e. the subtype of e ranging over sums) or $a = \langle e, t \rangle$. To spell out its meaning more, we need to borrow from the work on plurals the notion of partition. A partition Π is any function of type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$, such that for any property P, $\Pi(P)$ satisfies the following requirements:

(50) a. $\Pi(P) \subseteq P^+$
          A partition of P is a total subproperty of P[26]
      b. $AT(\Pi(P)) = \Pi(P)$
          If x is a member of a partition of P, no proper part of x is (relative atomicity)[27]
      c. $\forall x [\Pi(P)(x) \rightarrow \forall y [\Pi(P)(y) \rightarrow \neg \exists z [z \leq x \wedge z \leq y]]]$[28]
          No two members of a partition overlap

My proposal is that words like *quantity of*, *amount of*, etc. are simply variables over partitions. Such variables have to be thought of as indexical elements, proforms whose value is supplied by the context. Expressions are interpreted relative to a model M, a

---

[25] For unclear reasons, these type of nouns are unacceptable with coordinated structures:
(a) *A quantity of John, Paul and Bill.
This restriction is operative also in partitives:
(b) i. Two of those boys   ii. *Two of John, Paul and Bill

[26] Notice that we don't require that $\Pi(P) = P$; this distinguishes our characterization of partition from its standard set-theoretic counterpart.

[27] Condition (44a) could equivalently be rewritten as follows:
(a) $\forall x [\Pi(P)(x) \longrightarrow \forall y [y < x \longrightarrow \neg \Pi(P)(y)]]$

[28] In the work on plurals, it has been argued that this condition should be weakened and than rather than with partitions, one should work with 'covers'. The same might be true in connection with our analysis of *quantity of*. It is only for expository purposes that I don't pursue this here.

context c and an assignment g. Free proforms are interpreted via the assignment g (as they are free variables) and an interpretation is felicitous if an appropriate value for the free proforms can be found.

(51)   i. $||\text{quantity}_n||_{M,c,g} = g(\text{quantity}_n)$, where $g(\text{quantity}_n)$ satisfies the conditions in (50).[29]
         ii. *quantity$_n$* translates into TY2 as Π_n.

To illustrate with a concrete example, imagine a context in which there are five apples (a, b, c, d, e) two of which (a and b) are disposed in a bowl and the remaining three (c, d, e) on a tray. It would be natural to refer to them as *two quantities of apples*. The denotation of *quantity/quantities of apples* in such a context would be as shown in (52), where the plural is simply the closure under join of the denotation of the singular (as usual).

(52)   a. $[\text{quantity}_n \text{ of apples}]^{30} \rightarrow \Pi_n(\text{apples}) = [a \cup b, c \cup d \cup e]$
       b. $[\text{quantities}_n \text{ of apples}] \rightarrow {}^*\Pi_n (\text{apples}) = \begin{bmatrix} a \cup b \cup c \cup d \cup e \\ a \cup b \quad c \cup d \cup e \end{bmatrix}$

This analysis extends to mass nouns without changes; *quantity of water/amount of rice/piece of gold* are partitions over the denotation of mass nouns.

   The main point of this proposal is that *quantity of* is radically context dependent, but not vague. Accordingly, the value of *quantity* is a total function that remains constant across precisifications.

(53)   For any M, c, c′, and g, $||\text{quantity}_n||_{M,c,g} = ||\text{quantity}_n||_{M,c',g} = g(\text{quantity}_n)$

This guarantees that *quantity of apples* has stable minimal parts (if defined). Consider again the example in (52). In any precisification of that context, there may be further elements that qualify as apples; however the only things that will be in *every* such precisification will be what we have in the ground context (by the monotonic character of precisifications); hence if the original partition is defined, it will remain defined in the very same way, no matter how we precisify. What is interesting is that this applies in the same way to *quantity of water*, for nothing in our reasoning depends on whether the original property has stable atoms or not. It follows that numbers, as defined in the previous section, can apply to *quantity of water* (as they can, of course, to *quantity of apples*) without any problem:

(54)   a. There are two quantities of water on the table: the one in the cup and the one in the bottle.
       b. $\exists x [\text{two}(\text{quantities}(\text{water}))(x) \wedge \text{on the table } (x)]$
          $= \exists x [\mu_{\textbf{AT}}, \Pi_{(\text{water})}(x) = 2 \wedge \text{on the table } (x)]$

---

[29] The index n on the proform *quantity* is due to the fact that such proforms may occur more than once in the same discourse with different interpretations. Consider for example:
(a) This quantity$_{\{a,b,c\}}$ is clearly more than this quantity$_{\{a,b\}}$
Imagine uttering (a) while pointing at, say, a scoop of rice (represented by {a,b,c}) and a proper subpart of it (represented by {a,b}). Such an utterance would be felicitous. One can conjecture that the indexical act that accompanies each occurrence of the word *quantity* brings to salience two different partitions: Π_1 (between {a,b,c} and the rest of the rice in the world) and Π_2 (between {a,b} and the rest).

[30] I am assuming that the preposition *of* is a semantically vacuous element preset for case theoretic reasons.

In essence, *quantity of water* fixes its minimal parts of water by taking a contextually salient partition thereof. The distinction between vagueness versus context dependency is absolutely crucial in this way of analyzing things.

Among other things, this analysis predicts that sentences of the following sort are going to be synonymous.

(55)   a. We bought a lot of gold
         a′. We bought a large amount of gold
         b. The rice we have is insufficient
         b′. The quantity of rice we have is insufficient
         c. That gold is old
         c′. That piece of gold is old

For example, (55a,a′) have essentially the same semantics, viz.:

(56)   a. $\exists x[\text{ gold}(x) \wedge \text{large}(x) \wedge \text{bought}(we, x)]$ $= (55a)$
         b. $\exists x [\Pi(\text{gold})(x) \wedge \text{large}(x) \wedge \text{bought}(we, x)] = (55a')$

Formula (56b) entails (56a); and modulo plausible assumptions about the contexts (viz. that uttering (55a′) brings to salience a partition between the gold we bought and the one we didn't buy), (56a) entails (56b). At the same time, since the denotation of *water* and *quantity of water* are not identical (with the latter a 'stable subset' of the former), we will be able in principle not only to explain their different behaviour vis a vis numbers but also to differentiate more generally grammatical combinations (like *every quantity of water on that table*) from ungrammatical ones (like **every water on that table*), by simply making the relevant functions (like *every*) defined only over stable properties. These are welcome results.

So words like *quantity* denote variables over partitions of properties of sums (or, directly partitions of sums)[31], henceforth S-partitions. It is useful to contrast such an analysis, with that of words like *part*, which can actually be quite similar. While *quantity* is primarily defined over properties of sums (bare plurals or bare singulars), *part* appears to be primarily defined over individuals, which suggests that the basic type of *part* is $\langle e, \langle e,t \rangle \rangle$, i.e. that of functions from entities into sets of entities. As with *quantity*, we can interpret *part* as variable over partitions of individuals (i-partitions). The notion of i-partition can be made clear through an example first:

(57)   a. part_n of John $= \pi_n(j) = \{$j's right arm, j's left arm,...$\}$
         b. Suppose that (i) holds:
              i. *that rice* $= (a \cup b \cup c)$
                 where a, b, are, say, two bowls of rice and c a pile of rice next to them, and they jointly constitute the rice we are pointing at. Then, we might interpret *part of that rice* in the ways shown in (c).
         c. i. [part_1 of that rice] $\to \pi_1(a \cup b \cup c) = \{a, b, c\}$
             ii. [part_2 of that rice] $\to \pi_2(a \cup b \cup c) = \{a \cup b, c \}$

---

[31] If x is a sum, like say, the denotation of *those apples*, then a *quantity$_n$ of those apples* is interpreted as
(a) $\Pi_n(\lambda y[y = x])$
In other words, a partition of sums $\Pi_{\langle e, \langle e,t \rangle \rangle, n}$ is defined in terms of partitions of properties as follows:
(b) $\Pi_{\langle e, \langle e,t \rangle \rangle, n}(x) = \Pi_{\langle \langle e,t \rangle, \langle e,t \rangle \rangle, n}(\lambda y[y = x])$

On the interpretation associated with (c.i) there are three parts of rice; on the one associated with (c.ii) the are two parts (the rice in the bowls and the pile). In the general case, for any n, we want part$_n$ to be a variable π_n of type $\langle e, \langle e,t \rangle \rangle$ subject to the following constraints:

(58)   For any model M, assignment g, and individual u, $g(\pi_n)(u)$, if defined, is an
         $A \subseteq U$ such that:
         a. for any a, b $\in$ A, neither $a \leq b$ nor $b \leq a$ (relative atomicity)
         b. for any $B \subseteq A$, $\cup B \in \Sigma$ (closure under $\cup$)
         c. $\cup A$ is spatiotemporally included in u   (the 'material' condition)

Conditions (a) guarantees that *part of x*, whenever defined, is atomic with respect to '$\leq$'. Condition (b) guarantees that the extension of *part of x*, is closed under '$\leq$', and hence the plural *parts of x* will be defined (cf. Sauerland (2003) for a detailed analysis of relational plurals compatible with the line adopted here).[32] Finally, (58c) establishes (somewhat loosely) some kind of 'material' relation between the original individual and its parts.

Summing up, *quantity* and *part* have a slightly different type, but are otherwise quite similar. Both are context dependent and, in some sense, subdivide a property or an entity into (non overlapping) parts, creating relatively atomic, countable properties.[33]

5.3 Analyzing 'elasticity'

The analysis of elasticity is central to a full understanding of the mass/count distinction. I will first consider 'ambiguous' mass-count nouns; then I will look at cases of 'coercion'.

Rope or *rock* constitute nouns that, from a conceptual stand point, can be regarded as primarily mass, but have well established, readily accessible count alternants. *Rabbit* or *apple* represent primarily count nouns with common mass uses. Our take on this phenomenon is based on the analysis of *quantity* and *part* from the previous section. Let us begin by analyzing nouns like *rope*; like any other mass nouns, it will denote a $\cup$-closed set generated by unstable individuals. But of course, rope also occurs in standardized bounded units. Such naturally occurring bounded units constitute a particular partition of rope, call it $\Pi_{ST}(rope)$, which constitutes a good candidate for the count sense of *rope*. This idea seems to be fully general. Nouns like *beer*, *stone* etc. come out as mass, as the size of their minimal units is unspecified. At the same time bounded instances (often with a uniform, recognisable function) occur frequently enough to constitute a particular partition salient enough for lexicalization. So, in a sense, count

---

[32] The *part of* relation is like the *having x as a father* relation; a single individual u cannot have as (biological) fathers two individuals a and b; similarly there cannot be one (material) part of two things, unless that part is in turn constituted by two subparts. For example, some c can be a material part of two pizzas A and B, only if $c = a \cup b$ where $a \in A$ and $b \in B$ (so that $a \cup b$ are parts of $A \cup B$).

[33] Interestingly, words like *piece* appear to be a mix of *quantity* and *part*. *Piece* acts like *part* with definite NPs and bare plurals and like *quantity* with bare mass nouns:

(1)   a. i. a piece of John   b. a piece of that pizza
       (b) i. ? a piece of those cakes   ii. Some pieces of those cakes
       (c) i. *a piece of apples   ii. a piece of gold

uses of predominantly mass nouns can be thought of as covert applications of a word like *quantity*, which anchors the noun to a standardized partition. One way of implementing this is by hypothesizing a operator (i.e. a logical *constant*) Π_{ST} that denotes the S-partition most salient in the context:

(59)   For any model M, any $c \in C$ and any $P \in D_{\langle e,t \rangle}$, $F(\Pi_{ST})(c)(P)$ is the partition for P most salient in c (the standard S-partition).[34]

Given that for any property P and any world w, $\cup\Pi_{ST}(P_w) \subseteq \cup P_w$, we predict that (60a) entails (60a′); we also predict that rope and ropes qua kinds will be different; that this is so can be appreciated from minimal pairs like (60b) versus (60b′):

(60)   a. I bought those ropes at CVS
        a′. I bought that rope at CVS
        b. Rope is in short supply
        b′. Ropes are in short supply

Clearly, (60b) and (60b′) have different truth conditions. Moreover, it is worth observing that while (60a′) does not logically entail (60a) (for example, a rope sample might be too small to qualify as a rope) in most contexts in which (60a′) is true, (60a) will also be (because what you can buy in stores are standardized, bounded amounts of rope). All this is evidence, I take it, that the present approach is on the right track.

Turning now to nouns like *rabbit* or *apple*, we expect that they in their mass uses denote ∪-closed sets generated by unstable individuals. It is natural to exploit i-partitions in this connection, as has been suggested many times. I-partitions divide objects into parts. Some such partitions will split the relevant individual into partial sums (i.e. sums whose components are unstable) and these are the ones we want to exploit for massification. Here is how we may proceed in defining the notion of *standard* i-partition π_{ST}. The operator π_{ST} is a logical constant (of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$) defined as follows.

(61)   For any model M, any $c \in C$, any $P \in D_{\langle e,t \rangle}$, and any u:
        i. $F(\pi_{ST})(c)(P)(u) = 1$, if $u \leq_c \pi(\cup P) = 1$ and u is of size s(c),
        ii. $F(\pi_{ST})(c)(P)(u) = 0$, if it is not the case that $u \leq_c \pi(\cup P)$
        iii. $F(\pi_{ST})(c)(P)(u)$, undefined otherwise
        where π most salient i-partition in c such that $\pi(\cup P) \subseteq P\Sigma$, if there is such
        a π.

The condition that $\pi(\cup P) \subseteq P\Sigma$ requires that a standard i-partition, if it exists, must be rooted into unstable individuals (reflecting the intuition that we don't know how they may be subdivided further). Notice π_{ST} that creates partial properties with unstable minimal entities that can be precisified in different ways. It follows from this definition that neither $\cup$rabbits $\leq \cup\pi_{ST}$(rabbit) holds nor vice versa. Hence we predict pairs like those in (62a,b) to have non equivalent truth conditions:

(62)   a. I bought rabbits at Trader Joe's
        a′. I bought rabbit at Trader Joe's

---

[34] So expressions of the form Π_n are variables and their value is set by the assignment function g; Π_{ST}, on the other hand, is an (indexical) constant and its value is set by the interpretation function.

  b.   *Rabbit evolved long ago
  b′. Rabbits evolved long ago

The functors Π_{ST} and π_{ST} are both partial and context dependent. They are our way of spelling out the 'universal packaging' function and the 'grinding' one, but, as we shall see in more detail below, our proposal differs from the latter in several ways. It makes sense to maintain that if a property P is count, then Π_{ST} applied to P coincides with P's atomicization. On the other hand, if P is mass, applying π_{ST} to it won't give us anything new. In other words, Π_{ST}/π_{ST} act as identity maps on count/mass nouns respectively.

(63)   a. For any model M, context c, and property P, if $AT(P) \subseteq STB$, $\Pi_{ST}(P) = AT(P)$.
       b. For any model M, context c and property P, if $AT(P) \subseteq STB^-$, then $\pi_{ST}(P) = P$

This guarantees that we can never shift a count noun into a different count noun (e.g. use *boy* to mean *gang of boys*), nor can we turn a mass noun into a different one (say, *rice* into *rice flour*).

Overall, the main feature of this approach to elasticity is that it uses (i) the lexical meaning of *quantity* and *part*, which any theory must provide, and (ii) a notion of (context dependent) standardization, which again any theory must provide. It is interesting to underscore the nouns created by the overt nouns *quantity* and *part* are always count. In particular, *part* splits an individual into a set of non overlapping stable individuals, and hence it gives rise to perfectly good (context dependent) count properties. Nonetheless, the massification operator, defined in terms of *part* is set up so as to always get mass nouns. To my knowledge, it is difficult to find in the literature an approach to the problem of elasticity that uses fewer stipulations.

Let us now turn to the issue of coercion. While nouns like *beer* appear to have equally natural mass versus count uses; words like *blood* do not.

(64)   a. I drank many beers   a′. I drank a lot of beer
       b. I need a beer   b′. I need beer/a lot of beer
       c. *I lost many bloods   c′. I lost a lot of blood
       d. *I need a blood   d′. I need blood/a lot of blood

Context, including the choice of predicate, helps a bit, as shown by examples like (49), even if such sentences remain, for many speakers, substandard:

(65)   a. ?? In the emergency room we always have at least ten bloods.
   b. ? I bought two rices: basmati and arborio.

The shifts from count to mass appear to be even more restricted. Even nouns that admit quite readily mass uses are not felicitous in most mass contexts:

(66)   a. I bought chicken   a′. ? I bought apple
   b. I couldn't find enough chicken   b′. ? I couldn't find much apple

And most count nouns are plainly out in mass contexts:

(67)   a. I bought rice/furniture   a′. *I bought table
       b. Rice is easy to draw   b′. *Triangle is easy to draw
       c. There is less dust around here c′. *There is less cousin around here

How can we explain these graded judgments, some of which seem to be pretty steady crosslinguistically? Let me address such question after having taken stock of where we stand. First we have argued that the independently established 'semantic triad' comes in two varieties, depending on whether kinds/properties have vague atoms or not. We have discussed how this is so with respect to properties. To see that the same is true of kinds, consider, for instance, the totality w_{MAX} of water that there is versus the totality c_{MAX} of cats. If you take the respective smallest parts of w_{MAX} and c_{MAX} under '$\leq$', they will differ. The smallest cat parts (in the relevant sense, i.e. under '$\leq$') will be individual cats, i.e. stable atoms. While the smallest water parts will be unstable (i.e. their nature as atoms versus sums being up for grabs). Based on this (formal) feature, we can talk of mass versus count *kinds* just in the same way as we've been talking of mass versus count properties. This can be schematized as follows:

(68)   a.

   b. i. $^{\cup\cap}$CAT = CAT ii. $^{\cap\cup}$c = c iii. *$\mathbf{AT}$(CAT) = CAT iv. $\mathbf{AT}$(*cat) = cat
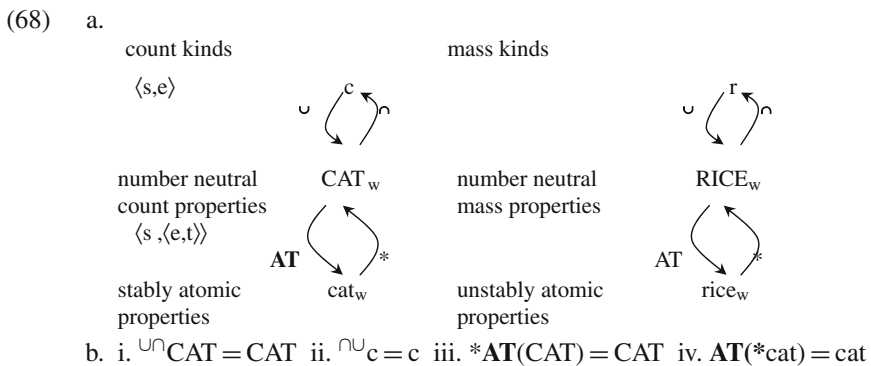
In principle, one can freely move among the components of a triad through the functors/type shifters that relate them; such functors generally 'commute' as in (68b). as we already mentioned, languages may vary in how their nouns map into the semantic categories in (68), and such variation will depend to a large extent on specific properties of the morphology of the language. Below, we will go over a hypothesis as to how the English system may work. The two columns in (68) are also related to each other by a small set of functors/type shifters. I summarize these at the level of number neutral properties (but I believe that analogous functors could be defined at the kind level).[35]

(69)   a.

---
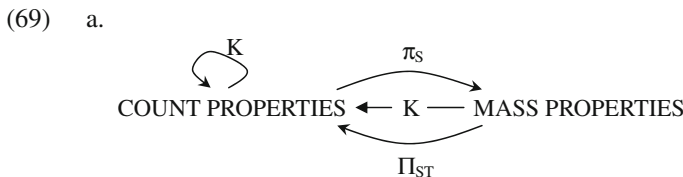
[35] Perhaps this is what happens in classifier languages where some of these functors are expressed as classifiers.

b.   RABBIT π_{ST}(RABBIT)

                TABLE π_{ST}(TABLE) = undefined

                Π_{ST}(BEER)   BEER

                Π_{ST}(BLOOD) = undefined   BLOOD

                K(DOG)

                K(WINE)   WINE

c.   i. $K(P) = \lambda x$ [subkind $(x, ^{\cap}P)$]

     ii. $\Pi_{ST}(\pi_{ST}$(RABBIT)) = undefined

     iii. $\pi_{ST}(\pi_{ST}$(RABBIT)) = π_{ST}(RABBIT)

Unlike the functors in (68), those in (69) do not, in general, commute (cf. (69c.ii)): there are no standard servings of rabbit-meat) or iterate (69c.iii); they are heavily partial and context dependent. It is this partiality and context dependence that makes them unsuitable to rescue, in general, bad numeral-noun combinations. Something like *three bloods* is bad because *three* cannot combine with a non atomic property like BLOOD and Π_{ST}(BLOOD) is either undefined, or defined only in very special contexts.

Which brings us back to the issue of how to understand the variability and graded character of our intuitions about mass-to-count and count-to-mass shifts. How do we make sense of the observation that while, e.g., mass uses of *table* are highly marked, they are, nonetheless, sometimes possible? Here is where a more substantive theory of the lexical interface becomes crucial. Let me sketch a couple of ways one might explore. A minimal approach could rely totally on context. 'Standardization' may well vary across contexts. This is reflected in the fact that the covert shifters are treated as indexicals. In most or all contexts standard servings of beer occur; contexts in which the same happens for blood are rarer. Similarly mass counterparts of *apple* are way more common than mass counterparts of *table*. One might imagine developing an approach in which contexts are partially ordered by standardization norms (as with modalities). But the basic point is that the (un) grammaticality of *three beers/bloods* depends on the choice of context.
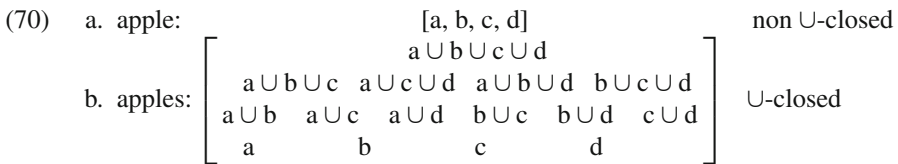
A different take would rely on 'stratified,' cyclic models of morphology (and phonology), as developed by e.g. Kiparsky (1982), and its more current Optimality Theoretic incarnations (e.g., Steriade 2008). According to these approaches, the lexicon is built in strata. First the core lexicon (of level 0) undergoes inflection and derivation with a first layer of (innermost) morphemes. The output of this process is a derivative lexicon, (of level 1) which then undergoes a second block of rules, and so on in a cyclic fashion. In such model, one would produce the mass version of *chicken* and the count version of *beer* in early strata, while the mass (respectively count) counterpart of *table* (respectively *blood*) could be created in more peripheral word formation strata.

Ultimately, I believe, a combination of both ideas will prove useful. Be that as it may, we have enough evidence to be hopeful that the phenomenon of elasticity is on its way towards being better understood. While our hypothesis needs to be tested further against a broader range of cross linguistic data, it looks like a prima facie

reasonable account of why mass-to-count and count-to-mass type shifting appears to be constrained and graded.[36]

5.4 On singular/plural marking

In this section, I would like to examine the specific contribution of singular/plural morphology in English. While plural morphologies have, obviously, a common core, it is plain that they are also subject to a fair amount of variation. The following proposal targets English like number marking and it is meant more as an illustration of some interesting consequences of the current analysis of mass nouns, than as a complete story. Let us begin with a puzzle linked to our analysis of relational nouns like *quantity*. Something like *quantity of apples* denotes a set of sums (i.e. pluralities). So how come the noun *quantity* shows up in the singular? In introducing singular/plural structures, I have led readers to believe that singular definite DPs (e.g. *that table*) denote atoms, while plural definites (e.g. *those tables*) denote sums[37] However, a phrase like *that quantity of apples* is singular even if it denotes a sum, accounting to our proposal. Can this be maintained while having a coherent view of the contribution of singularity? I think this is possible under the assumption that singular/plural marking operates at a 'low' level of DP-structure (say, at the NP level). In light of the analysis in Sect. 5.2, here is what we might do. At the level of nouns, singularity corresponds not so much to 'absolute' atomicity, but to 'relative' atomicity. A noun is singular if whenever it is true of an entity x, it isn't true of any proper part of x.[38] This generalizes in a natural way our previous view of singularity, for if a noun applies only to atoms (and hence P is 'absolutely' atomic), it also applies only to 'relative' atoms (and hence P is 'relatively' atomic; in other words, $P \subseteq AT(U)$ entails $AT(P) = P$. The definition of plurality is likewise to be generalized from closure of a set of atoms with respect to sum formation, to closure of a set of individuals (atoms or sums) with respect to sum formation. Here is a way to visualize the main idea, through a toy model with four apples. As usual, I give the extension of the relevant nouns:

(70) a. apple: [a, b, c, d]   non ∪-closed

$$\text{b. apples:} \begin{bmatrix} & a \cup b \cup c \cup d & \\ a \cup b \cup c & a \cup c \cup d & a \cup b \cup d & b \cup c \cup d \\ a \cup b & a \cup c & a \cup d & b \cup c & b \cup d & c \cup d \\ a & b & c & d \end{bmatrix} \quad \text{∪-closed}$$

---

[36] I believe that the methods developed here for words like *quantity* and for the phenomenon of elasticity applies with simple changes to words like *wall* or *fence*, discussed in Rothstein (2008). Such words are strongly context dependent just like *quantity*. Rothstein develops a theory of mass nouns based on her analysis of words like *fence*; however her proposal suffers from drawbacks discussed in Krifka (2008).

I also believe the present strategy has fruitful applications to the analysis of group level nouns like *group* or *bunch*. But to pursue this would take us too far afield.

[37] That this view is simplistic is evident from the fact that sometimes it seems possible to refer to the same entity with a plural as well as with a singular group-level noun:

(a) I saw those kids at school yesterday

(b) I saw that bunch of kids at school yesterday

[38] I am sure this idea has antecedents in the literature, but I am unable to locate them. For me, it came up in conversations with G. Magri.

    c. quantity of apples $[a \cup c \quad b \cup d]$ non $\cup$-closed

    d. quantities of apples $\begin{bmatrix} a \cup b \cup c \cup d \\ a \cup c \quad b \cup d \end{bmatrix}$ $\cup$-closed

A mass noun looks like a plural, except that its generator set (i.e. the bottom line in (70d)) falls in the unstable part of the domain. This idea is married with the notion of stability, which is also to be thought of as a characteristic of properties. A property is stably atomic iff it has a non empty set of stable atoms (i.e. atoms that remain such in every precisification) relative to some ground context c. As we saw, while *water* is not stably atomic in this sense, *quantity of water* is. These ideas seem to constitute a reasonable hypothesis as to what is the semantic basis of the singular/plural contrast in languages like English.

In the following subsection (which can be skipped without loss for the overall understanding of the basics of the present proposal), I will sketch an execution of these ideas within a syntactic framework based on minimalism. Following much recent work on NP structure, one can characterize a full expansion of the English noun phrase along the following lines:

(71)   a. I fed all those twenty three cats
         b. I fed cats
         c. [Q [Det [Num [PL [N]]]]]
            all those twenty three -s cat

Sentence (71a) provides an example of a 'large' nominal argument in object position, as opposed to sentence (71b) which provides a 'small' nominal argument (i.e. a bare argument). N in (71c) is a lexical category; the other elements along the 'spine' of this structure constitute the elements that form the 'extended projection', or functional layer, of the noun (Grimshaw 2000). I will assume that when these elements are absent, the corresponding positions in the structure are not, in general, projected. Q is the position of quantifiers, Det of articles and demonstratives, Num of number phrases, SG/PL is the locus of number marking (which gets spelled out on the noun). I will refer to the constituent [SG/PL N] as 'Atom Phrase' (AtP), as that is where Ns are checked for atomicity. In English (but not in, e.g., Romance) AtP can occur bare as arguments of the verb, as in (71b). Full fledged nominal arguments (such as the one in (71a)) are often loosely referred to as DPs (even when a quantifier is present), a practise I will follow. Number features (SG/PL) are semantically significant on the noun, and have no meaning on higher functional categories or on the verb. Accordingly, functional elements above AtP inherit number features from AtP (in a way that makes them visible for purposes of verb agreement). At the same time, determiners or quantifiers may semantically select for arguments of a particular semantic (sub)type; for example, *every* semantically selects a singular count argument, and syntactically inherits from it SG, which is then used for verb agreement purposes; *the*, instead, combines freely with singular, plural or mass properties and inherits whatever number feature its argument carries. Even though each of these assumptions may be modified, they form a widely shared theoretical apparatus.

Perhaps the single most important part of all this, for our purposes, is the role of AtP (i.e. number marking). A natural way of understanding it is as an atomicity check point. Number morphology lets a property go through if it is atomic. There are two

ways of qualifying as atomic, both based on the function **AT**; the first consist in being composed solely of stable individuals not closed under ∪ (and is coded in the singular morpheme); the second consists of being generated by a set of stable individuals via ∪-closure (and is coded in the plural morpheme). This can be spelled out as follows:

(72)   a.

    i.   AtP   ii.   AtP

        SG   N   PL   N
        ∅   apple   s   apple
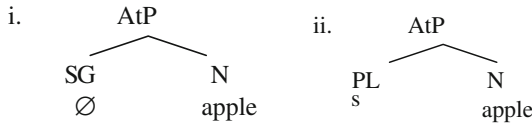
    b.   i. SG(P) = P, if for any u   ii. PL(P) = P, if for any u such that P(u) = 1
             if P(u) = 1, $\mu_{\mathbf{AT}(P)}(u) = 1$ there is some n such that $\mu_{\mathbf{AT}(P)}(u) = n$

(72a) gives the syntax of AtP and (72b) the corresponding semantics. Number features act as gates to higher functions in the semantic composition. They let an N through only if it is relatively and stably atomic or the ∪-closure of a relatively atomic property.[39] Technically, features are treated as identity maps, restricted to properties of the right kind. Thus, as discussed in Sect. 5.3, properties are freely created in the lexicon via the restricted set of operations UG makes available, and inserted in frames such as those in (72). The SG features requires its argument to be semantically singular (i.e. relatively and stably atomic) and hence, it won't let, say, a property like CAT through, but it will let through AT(CAT); on the other hand, PL will let CAT go through.

    Perhaps it is useful to see how DPs may be treated on this approach. Consider, for example, the articles. Adopting the (somewhat simplistic) view that *a* is an existential quantifier restricted to singular properties, while *the* is unrestricted and denotes the ι-operator, we get the following results.

(73)   a. a(P) = λQ∃x [P(a) ∧ Q(x)], if **AT**(P) = P (else undefined)
        b. the(P) = ιP, where ιP = ∪P, if ∪P ∈ P (else undefined)
            Examples:
        c. i. [$_{DP[PL]}$the [SG boy]] → ι**AT**(BOY) defined only when *boy* is a singleton
           ii. [$_{DP[PL]}$ the [PL boy]] → ι*****AT**(boy) defined only when *boys* is non empty
          iii. [$_{DP[SG]}$ the [SG quantity of apples on the table]] → ιx[Π(apples)(x) ∧ on
              the table(x)]
              (defined only when *quantity of apple on the table* is a singleton)
        d. i. [$_{DP[PL]}$ a [PL boys]] → undefined,
             because [PL boys] → BOY and **AT**(BOY) ≠ BOY
          ii. [$_{DP[SG]}$ a [SG boy]] → λQ∃x [**AT**(BOY)(a) ∧ Q(x)]

A similar analysis can be devised for other quantifiers and for the numerals. The point of this exercise is to show that the thesis that singularity corresponds to relative atomicity

---

[39] These operations are in an elsewhere relation: the stronger of the two must be selected, if they are both possible. Thus for example, a singleton property {a} satisfies both SG({a}) and PL({a}) are defined. In this case the first option (i.e. SG) must be selected. A way to implement this idea is via Heim's 'Maximize presupposition'. See Sauerland (2003) for one possible implementation.

has a fair amount of initial plausibility. It adequately unifies the semantic contribution of singularity versus plurality and it provides us with a semantic distinction that can be used to constrain the domain of higher functions.

A further consequence of this approach is that mass nouns in number marking languages, as it were, 'do not fit': they do not pass the atomicity check: SG(AT(RICE)) and PL(RICE) are undefined because these properties have no stable atoms (in any context). This means that mass nouns have to be treated somehow as 'exceptions' to canonical number marking, a welcome result. We are deriving, in an arguably principled manner, the fact that in number marking languages, mass nouns cannot have the same morphosyntax as count nouns (while in, e.g., classifier languages, they do). Minimally, one could say that in English mass nouns receive a semantically void, 'default' singular morphological marking. But, as we will see in the next section, by pushing our line of inquiry a bit further, we can gain further insight as to the nature of the mass/count distinction. Be that as it may, even if we stopped here, we would have sketched a seemingly promising analysis of the mass/count distinction in a number marking language like English.

6 Cross linguistic variation

In this section I will investigate how the framework outlined above deals with the aspects of crosslinguistic variation discussed in Sect. 2. I will, more specifically, consider two sets of data pertaining to fake mass nouns and to classifier languages, respectively. What I will be able to present here is largely informal, and a far cry from a complete theory of these phenomena. My aim, again, is just to illustrate how the approach developed here may help sort out the variable from the tendentially invariant in the domain of mass versus count nouns.

6.1 Fake mass nouns

Suppose we take seriously the idea that number morphology constitutes a check point for atomicity but we want to avoid resorting to a default morphological number feature, void of meaning for mass nouns. Then, we would have to somehow assign them an atomic denotation; such a denotation, however, should still prevent mass nouns from being pluralized and counted. Is this possible? Does it make sense?

Here what one might explore.[40] Imagine coding all mass nouns as singleton properties true of the totality of the manifestations of a substance. For example, the denotation of *water* would look as follows:

(74)   water $= \lambda x \; [x = w_{MAX,w}] = \{w_{MAX,w}\}$, where $w_{MAX,w}$ is the sum of all the water there is in w.

This denotation qualifies as stably atomic: its smallest element is a single individual (a sum) that remains constant in every precisification. What are the consequences of this

---

seemingly strange move? To start, it is useful to point out that such a proposal makes mass nouns look a bit like proper names in languages where the latter require articles (like some dialects of German, Northern Italian, Modern Greek, Portuguese, etc.). It is very plausible to analyze proper names in such languages as singleton properties:

(75) a. La Francesca 'the Francesca'
      b. francesca$_{\langle e,t \rangle}$ = λx [x = f] = {f}[41]

The analysis in (75) seems capable of accounting for the properties of proper names in the relevant languages. First, it would explain why they require a definite article. They do because they are properties and properties cannot be taken as arguments by verbs (without a type adjustment). Moreover, being *singleton* properties, they go well with the definite article which demands uniqueness of its N-argument (in the singular). Second, perhaps the singleton property hypothesis could also explain why proper names do not pluralize (without a shift in meaning, that is): there is no point in so doing, as proper names cannot be true of more than one individual.[42] In a similar fashion, we could explain why proper names do not combine with numerals: there would be no point of speaking of one Francesca, two Francesca's, etc. as if Francesca exists, it has to be unique (in the relevant sense).[43]

So in much the same way by assuming that mass nouns denote singletons, and hence they are in a sense proper name like, we would be able to immediately derive some of their key properties (like the lack of pluralisation and the impossibility of combining them with numbers). At the same time, as is well known, there are quantifiers that do go well with mass nouns (like *some, all, most,…*), and this is where mass nouns and proper names part ways. This too might have a reasonable explanation under the current hypothesis. Mass nouns are singleton properties of *sums*; proper names of absolute *atoms*. From the first we can extract properties that are good quantifier restrictions, by looking at their parts, from the latter we cannot (as atoms have no parts in the relevant sense). Here is one way of doing so, along with an example:

---

[41] This hypothesis is reminiscent of Quine (1960) analysis of 'Pegasus flies' as 'there is an x which pegasizes and flies'. Note that it is compatible with the view that a name is rigid (in the sense that its extension is constant). For a different view on the semantics of this type of proper names, see Longobardi (1994).

[42] The intuition here is that grammatical operations are subject in general to a non vacuouty constraint. In the case at hand, a way to spell out this constraint might be using 'maximize presupposition' as sketched in fn. 33.

Proper names can of course be turned into common nouns with the rough meaning 'person named x' (as in "I know many Johns"); in such a case they behave as regular common nouns. For striking evidence (based on the morphophonology of Rumanian) that the process converting a proper name into a common noun is an actual word formation rule, see Steriade (2008).

[43] The principle we need here is that something like *one (P)* or *two(P)* is well formed if it doesn't follow from the rules of grammar that P is a singleton. This reflects the intuition that (a) and (b) have a different status:

  (a) I met one queen of England
  (b) I met one John

Sentence (a) owes its status to a widely known fact (that queens of England, at any given time, are unique). Sentence (b) is deviant because it is a grammatical principle that proper names either refer uniquely or fail to refer. Formalizing the use of semantic principles as proof devices is not trivial. See Gajewski (2002) for relevant discussion (and Chierchia 1984, Chap. 3 for an antecedent to Gajewsky's discussion).

(76)   a. For any property P, *P = λx [x ≤ ∪P]
        b. *(λx [x = w_MAX]) = λx[x ≤ w_MAX] = {w1, w2, . . .w1∪w2, . . ., w_MAX}
        c. *(λx[x = f]) = λx[x ≤ f] = {f}
        d. most(P)(Q) = μ(∪(*P∩Q)) > μ (∪(*P − Q))   μ a suitable measure function[44]

Clearly, *most*, as defined in (76d), would yield something sensible when applied to a
mass or a count noun; applied to a proper name, it would yield something logically
false.

Here is a way of restating the above strategy. There are two planks to it. First, single-
ton properties are never good quantifier restrictions; however, from singleton proper-
ties that apply to sums we can extract good (i.e. non singleton) quantifier restrictions.
Second, there are mainly two types of quantificational elements; some (like num-
bers) count atoms (call them atomic quantifiers); others (like *most*) measure parts (in
the sense relevant to the singular/plural distinction; call them part-quantifiers). Mass
properties (whether extracted from singletons or not) are never good as restrictions
for atomic quantifiers; but they maybe good as restrictions for part-quantifiers. In
conclusion, by coding mass nouns as singleton properties, we understand why they
bear singular morphology (passing through the strictures of atomicity checking); at
the same time we retain a reasonable account of their distributional properties.[45]

One may wonder whether stipulating that sometimes number morphology is just
vacuous or defective might not be simpler that the singleton property stipulation. Here
is a potential argument against going this way. We are suggesting that mass nouns in
number marking languages take the form of singleton properties, in order to remedy
the fact that they lack a stable atomic structure. It is conceivable, in this situation, that
some nouns take on the same formal feature (that of denoting singleton properties)
even if conceptually their atoms are stable. All it would take is for such nouns to be
listed in the lexicon as singleton properties. We know from much work in morphology
(and phonology) that general processes of word formation stand in an 'elsewhere'

---

[44] The fact that quantifiers like *most* involve an implicit reference to a measure function can be seen from examples like:

(a) most numbers are not prime

We tend to regard sentences like (a) as true, even if the set of numbers and the set of prime numbers have the same cardinality. An adequate measure function for (a) would measure the frequency of occurrence of primes relative to the natural ordering of numbers.

[45] I believe that the singleton property hypothesis can also explain the difference between (a) and (b), discussed in Gillon (1992) and Chierchia (1998a):

(a) The tables in room a and the tables in room b resemble each other
(b) The furniture in room a and the furniture in room b resemble each other

Sentence (a) has two prominent readings (reading 1: the tables in room a resemble each other and the tables in room b resemble each other; reading 2: the tables in room a resemble those in room b and vice versa). Sentence (b) has only one reading (corresponding to reading 2). This can be understood under the assumption that the 'contrast property' of the reciprocal *each other* is sensitive to the antecedent NP, rather than to the DP (NP which, in the case of mass nouns would be a singleton). To spell this out any further would mean to spell out the grammar of reciprocals, which we cannot do within the limits of the present work. I believe that this solution (also due to an idea by Giorgio Magri) constitutes a step forward over the one explored in Chierchia (1998a).

relation to (i.e. are overridden by) more specific ones.[46] Any version of the elsewhere condition will allow us to derive something like following:

(77)   The core meaning of a noun stem W in language L is a number neutral, $\cup$-closed property P, unless $\langle$W, P$'\rangle$ is listed in the lexicon of L.

On this basis, one might assume that English chooses to list some nouns as follows:

(78)   a. $\langle$furniture, λx [x $= \cup$FURNITURE]$\rangle$ b. $\langle$footwear, λx [x $= \cup$FOOT-WEAR]$\rangle$,....

The presence of these pairings in the English Lexicon, will make the (candidate) core meanings FURNITURE, FOOTWEAR, etc. unavailable in English. It follows, then, that such nouns will pattern exactly as mass nouns, with respect to the phenomena considered above.[47]

The idea, in other words, is that fake mass nouns arise as a 'copy cat' effect from the way in which number marking languages react to unstably atomic nouns. Since listing a potentially count noun as a singleton property is essentially a matter of lexical choice, we expect there to be variation, even across closely related languages or language families on this score, which has, in fact, been often observed in connection with fake mass nouns.[48] Moreover, and more interestingly, this approach links the existence of fake mass nouns to the presence of (obligatory) number marking. The logic of this link is the following. If a language lacks obligatory number marking, there is no need to turn its mass nouns into singleton properties. And hence no copy cat effect can take place. As we will see shortly, classifier languages might indeed be a case in point.

Finally, even if listed as singleton properties, nouns like *furniture* do retain their atomic structure. Such structure can be extracted from their denotation (in fact, it has to be extracted by quantifiers like *some*, *all*, *most*,…). And such structure will be undistinguishable from that of plurals. This would explain why in some cases, fake mass nouns do pattern with count plural nouns and unlike core mass nouns, as with the 'Stubbornly Distributive Predicates', discussed in Schwarzschild (2007).

It is hard to see how these results can be derived if we simply stipulate that the singular morphology of mass nouns rather than being, in some sense, 'real', as we maintain, is defective. It should also be noted that we are really forced to this stand by

---

[46] See, e.g. Kiparsky (1982) for a classic point of reference; see Steriade (2008) for a more recent embodiment of similar ideas.

[47] In the framework we are assuming, singleton properties can be obtained by a simple modification the definition of **AT**:

(a) $\mathbf{AT}(P) = \begin{cases} \text{AT(P), if } \text{AT(P)} \subseteq \text{STB} \\ \lambda x[x \leq \cup P], \text{ otherwise} \end{cases}$

[48] Saying that the treatment of a potentially count noun as mass is a lexical option does not mean that it is totally arbitrary. For example, it seems natural to maintain that such choice is driven by 'lack of interest' in the atoms. This might account for the collective flavour of nouns like furniture, as for the fact they tend to be 'superordinate' category. However, the choice of treating count nouns as mass would be open only in number marking languages (or other languages where singleton properties are, for whatever reason, used).

Moreover, there may strategies different from the singleton property one to get around the requirement of number semantics. This has to be so, for example, if Tsoulas is right on modern Greek. But a discussion of the relevant facts (and of Tsoulas' own proposal) must await for another occasion.

the basic line we are taking, namely that mass is a matter of vagueness. The point is that tables and chairs are furniture, boots and shoes are footwear, and so on. Hence, it seems highly implausible that nouns like *furniture* are any vaguer than *table* or *chair*. It follows, that the mass like behaviour of nouns like *furniture* must come from a totally different source. Even if many details of the above account may well turn out to be wrong, this part of the present proposal follows from its basic tenets and seems to explain why variation of precisely this type should come about.

It may be useful to contrast the above take on *fake* mass nouns with the kind of variation that affects *core* ones. Some such variation is attested. For example, *bean/s* is count in English or Italian (*fagiolo/i*), but *fasole* is mass in Rumanian. In some varieties of Mandarin, the count classifier *ge* is acceptable with *mi* 'rice' (e.g. *san ge mi* 'three CL rice') with the meaning of 'grain of rice'.[49] What this suggests is that standardized partitions for the relevant substances are more readily available in such languages/dialects. This type of variation is a consequence of the fact that vagueness comes in degrees: some nouns may well be less vague than others, in the sense that a usable notion of 'smallest sample' can more readily be devised. Clearly, for example, defining smallest samples for (non commercially packaged) liquids is harder than for granular substances. On the present approach we expect that this type of variation, unlike that pertaining to fake mass nouns, not to be restricted to languages with oblig- atory number marking (as it stems from the inherent nature of substances, rather that from pressure by grammar). From the examples I've come across, reported above, this seems to be so.

In conclusion, I've sketched here an approach to the phenomenon of fake mass nouns. This approach is not a necessary consequence the thesis that the mass/count distinction is a matter of vagueness. The latter position may be adopted without buying into the thesis that fake mass nouns are singleton properties. What however follows from a vagueness based approach is that fake mass nouns *must* have a different source than the core ones.

6.2 The case of classifier languages: a sketch

So far we have been assuming that nouns start their life in the compositional construc- tion of complex meanings as properties. But in principle there is no reason to think that properties are in any sense more fundamental than kinds or vice versa. Hence, there is no reason to think that nouns should always be encoded in the lexicon as properties instead of kinds. I have argued (in Chierchia 1998b) that this question is to be thought of in terms of an option languages have to encode their nouns one way or another (the Nominal Mapping Parameter) and that the core properties of the nominal system of (pure) classifier languages may be derived by choosing the nouns as kinds option (henceforth N-as-K). It seems appropriate to consider briefly such a proposal in light of the view on the mass-count distinction explored here.

Suppose now that a language 'registers' its nouns as kind denoting, so that the noun *water* will denote the water kind, the noun *cat* the cat kind and so on. Suppose, in other

---

words, that in such languages kinds can (or, indeed, must) be chosen as noun meanings and that the syntax and compositional semantics operates, therefore, on kinds. What properties would such languages be expected to have? How would they differ from the Nouns-as-Properties (N-as-P) languages, like English? The answer to this question depends on many things. However, some general properties of such a language might be inferred by very general assumptions. In particular, we have argued that numbers are adjectival, of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$. Imagine that this is universally so. Assume, moreover, that when a number combines with its nominal argument, automatic type adjustments are disallowed (or marked). It then seems to follow that if in a language all nouns are kind denoting, by the time they combine with a number something must intervene to turn a kind into the corresponding property. That seems to constitute a natural slot for classifiers. The steps of this 'deduction' are quite simple:

(79) a. Nouns are uniformly mapped onto kinds
      b. Numbers are adjectival
      c. No automatic type adjustments are possible to turn kinds into properties in number-noun constituents.
         Consequence: overt morphemes (of type $\langle k,\langle e,t\rangle\rangle$) must intervene between numbers and their nominal arguments.

This is a way of understanding *why* there are generalized quantifier languages, i.e. languages where numerals can virtually never directly combine with nouns. Other explanations are of course conceivable. This one grounds the problem very directly in independently established semantic structures.

The above line of explanation for the existence of generalized classifier languages has a further consequence. Kinds, however construed, are of an argumental type. I.e. they are semantic entities that, as we know from English, can be directly taken as arguments of predicates. Suppose now that the mechanism of syntactic combination (i.e. 'merge') applies freely, modulo semantic coherence. Suppose, that is, you can in principle merge any syntactic category A with any syntactic category B forming [A B] if the result is semantically coherent. It follows that if a language has argumental, kind denoting nouns, one would expect them to combine freely and directly with verbs (the prototypical argument taking category). In other words we seem to derive the following generalization:
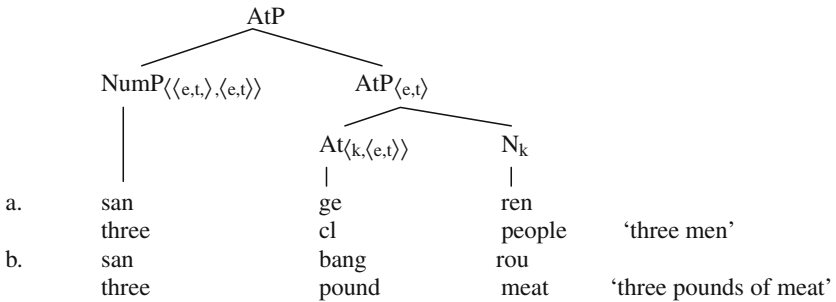
(80) If in a language N maps into kind, [V N] mergers are freely possible

Such languages, in other words, will freely allow bare arguments. We thus derive that generalized classifier languages also are bare argument languages. The idea is that what drives the existence of generalized classifier systems automatically makes the language a bare argument one. Caveat: this is a one way generalization: generalized classifiers → free bare arguments. The converse does not hold: there are plenty of bare argument languages (lacking articles altogether) that are not generalized classifier languages (e.g., Russian, Hindi, etc.; see on this Dayal (2004)).

Thus, by comparing classifier languages with English, we see that there are two types of morphology devoted to mediating between numbers and nouns: (obligatory) number marking (of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$); and (obligatory) classifiers ($\langle k, \langle e,t\rangle\rangle$). To borrow, Chris Kennedy's words, the former is a kind of a "filter", the latter are

sorts of "type shifters." This strongly suggests a complementarity between the two: if a noun has obligatory classifiers it lacks obligatory number marking and vice versa. Classifier morphology and number marking may be viewed as two different instantiations of AtP.[50]

(81)

$$
\begin{array}{c}
\text{AtP} \\
\diagdown \\
\text{NumP}_{\langle\langle e,t\rangle,\langle e,t\rangle\rangle} \qquad \text{AtP}_{\langle e,t\rangle} \\
| \qquad\qquad \diagdown \\
\qquad\qquad \text{At}_{\langle k,\langle e,t\rangle\rangle} \qquad \text{N}_k
\end{array}
$$

		At$_{\langle k,\langle e,t\rangle\rangle}$	N$_k$	
a.	san	ge	ren	
	three	cl	people	'three men'
b.	san	bang	rou	
	three	pound	meat	'three pounds of meat'

One would expect that certain classifiers appeal to the atomic texture of their kind argument (i.e. they mean something like *atom of*, sometimes with shape based restrictions tacked on); such classifiers will naturally be restricted to atomic kinds. Other classifiers, like measure phrases, superimpose their own atomizing criterion and hence will be equally happy with count as with mass kinds. These two types of classifiers might well have distinct distributional properties (as Cheng and Sybesma have indeed shown) which will provide formal criteria for the mass/count distinction in classifier languages. For example, one would expect combinations of the form [ge N] to be deviant (or to force coercion) when N is mass.

While all this constitutes, I submit, an interesting take on crosslinguistic variation, it is important to underscore that this typology, besides being extremely sketchy, is not meant to be exhaustive (a misunderstanding engendered by Chierchia's (1998b) original formulation of this parameter). There clearly are languages that have neither obligatory number marking, nor obligatory, generalized classifier systems. It remains to be seen how such languages would fit in a theoretically interesting typology of nominal systems, and there is little point in venturing further speculations on this score here. [51]

The reader may have noticed that treating all nouns as kind denoting is very close to the way we have treated *furniture* nouns in number marking languages.

(82)   a. furniture → λwλx[x = f$_{MAX,w}$]   type: $\langle s, \langle e,t\rangle\rangle$
        b. jiaju 'furniture' → λw f$_{MAX,w}$ type: $\langle s,e\rangle$

The main difference is that (82a) is non argumental, while (82b) is. Thus one might expect there to be number marking languages similar to English, in which bare arguments are disallowed. The Romance Languages (French, in particular) appear to be of such type.[52] Moreover, in some sense, the N-to-K mapping parameter might be viewed

---

[50] For partly converging ideas, cf. Borer (2005).

[51] We have already pointed to Wilhelm (2008) for an interesting attempt.

[52] According to this line of thought, bare arguments in N-as-P languages come about via the covert use of argumentizing operators. Such covert uses must be severely restricted; how they operate is the object of an intense debate. See again Chierchia (1998b), Dayal (2004) and references therein for discussion.

as embodying a suggestion that all Chinese *count* nouns are analogous to *furniture*, and hence, if you wish, fake mass. But just as is the case for *furniture* nouns in English, this does not obliterate the distinction between core mass nouns and the rest. That might also explain why the grammar of the mass/count distinction in Chinese seems to line up neatly along its conceptual/extralinguistic roots (i.e. there are no nouns that are conceptually count but take on the morphosyntactic behaviour of mass nouns). The main road to such an effect is blocked, since all nouns have a macro syntax similar to that of mass nouns in number marking languages.[53]

The purpose of this brief excursion in the domain of semantic variation is twofold. The first is to put some precise, if sketchy, hypotheses on the ground on where exactly languages vary in the mass/count distinction. The second is to underscore where the main divide should be according to the present proposal. If a noun is associated with either a property or a kind whose minimal elements are unstable (in the sense made precise in this work), there is no way to directly use it to restrict a number word (not without a change in the interpretation of the noun). English, Mandarin or languages where atomicity checking works very differently (if it at all) like Dëne are all alike on this score. The rest is potentially subject to variation; in particular, conceptually count nouns may take on in varying degrees the form of mass nouns. This means that probably the type of nouns (i.e. their logico-semantic role) is subject to a certain degree of systematic variation across languages. How type theoretic distinctions play out depend on how the rest of a specific language system work, which takes some figuring out. All of this section above is no more than a way of illustrating and giving concrete, falsifiable bite to what remains a still largely programmatic, if fascinating hypothesis: that the interface of syntax with a theory of entailment may vary across languages.

7 Comparisons and conclusions

In this section I will first briefly compare the vagueness based approach developed here with some of the proposals on the ground, without any pretence of completeness. Then it will be time for some brief concluding remarks. For comparison purposes, I single out, in particular, two approaches. The first is Link (1983) classic approach. The reasons for this first choice are that Link (1983) is among the most influential theories and, furthermore, it is representative of a vast class of approaches. The second approach I will briefly discuss is Chierchia (1998a); the reasons for this second choice are more personal.[54]

---

[53] Jim Huang in much recent work (e.g. Huang 2006) has been developing the idea that many differences between Chinese and English, including the fact that Chinese is a generalized classifier language, can be fruitfully analyzed in terms of a macroparameter involving *analyticity*. His proposal is consistent with the one made here. I think, in fact, that these two lines of inquiry at some level need each other. But to explore how they might be integrated in a principle based grammar must await for further research.

[54] For a broader overview in an historical perspective see, e.g., Pelletier and Schubert (1989). A more recent, extremely useful survey can be found in Krifka (2008).

7.1 Some alternatives

Link (1983) proposes that nouns take their denotation from different quantificational domains (and constitutes, therefore, a 'Double Domain' approach). Count nouns draw their denotation from an atomic semilattice (similar to the one adopted here). Mass nouns draw their denotation from a semilattice which is not atomic (or not known to be atomic). There are morphisms that project members of the mass domains into members of the count one and vice versa. In particular, the grinding function gr: $C \rightarrow M$ is a homomorphism from the count domain into the mass one; while the 'packaging function' p: $M \rightarrow C$ is an isomorphism from the mass into (a subset of) the count domain (terminology and notation from Landman (1991)). The reason for this set up can be summarized as follows. Any count object can in principle be grinded into a mass one, in a way that is structure but not identity preserving (whence the homomorphism requirement); at the same time any mass object can be injected in the count domain by prefixing to it words like *quantity* in a way that seems to be 'undoable' (in the sense that whenever we have, say, a quantity of water we can get back to the corresponding water).

In what follows, I discuss four main reasons to think that the vagueness based approach takes us further in our understanding of the mass/count distinction. The first two are conceptual, the second two empirical.

First, on a vagueness based approach there one domain. Masses are of the same sort as countables. Everything else being equal, this constitutes a simpler and more economic way of understanding what the distinction is about.

Second, the idea of a 'non atomic or not known to be atomic' domain is obscure. If we construe such a notion literally as 'atomless', the result is counterintuitive. The idea that rice can be infinitely subdivided while preserving its quality as rice makes little sense (at least since Democritus, we know that subdivisions of material objects come to an end). If, on the other hand, we construe the mass domain as something we do not know whether it is atomic or not, why should this lack of knowledge impair our capacity to count? Usually, ignorance is not sufficient ground for non-countability. Domains we know very little about but that are clearly countable abound (elementary particles, solutions to problems,…). If, finally, by 'non atomic' we mean 'with vague atoms', we should spell it out in terms of a theory of vagueness.
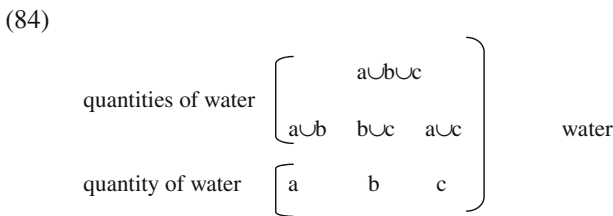
Third, consider the issue of fake mass nouns. The natural move to make for fake mass nouns on a Double Domain approach is to say that they draw their denotation from the *mass* domain. But this has several undesirable consequences. One of it is that the English noun *furniture* has to translate into Italian or French in a non reference preserving way. It is counterintuitive to think that your furniture changes its ontological nature by referring to it in English versus French. Second, and more importantly, treating core mass nouns and fake ones alike seems factually wrong even for English alone, as we now know from Schwarzschild's work. Of course, one might try to marry the Double Domain approach with a different treatment of fake mass nouns. But such an approach per se gives no clue as to why and how fake mass nouns should be different from core ones. On the present approach, instead, since *furniture* is not inherently vaguer than *table*, we know right off the bat that *furniture* cannot be treated on a par with *water*.

Fourth, consider the C → M shift. On the Double Domain approach it has to be some kind of isomorphism. On the present approach, we can rely on the notion of context dependent partition. The 'isomorphism' approach to which the Double Domain take is forced entails that *that gold* and *that quantity gold*, said by pointing at the very same object, must be coded as two different model theoretic entities. Thus to obtain the equivalence of sentences like (66a,b) something must be done:

(83) a. That gold comes from South Africa
      b. That quantity of gold comes from South Africa

One must resort to either meaning postulated or type shifting (see, e.g. Landman (1991) for suggestions). On the present approach the equivalence of sentences such as those in (66) follows as a matter of logic. All these seem to constitute reasons for favouring a vagueness based account.[55]

Chierchia's (1998a) theory is a single domain approach. The main idea explored there is that while in count nouns singularities are segregated from pluralities, mass nouns 'don't care' about the singular/plural distinction. Accordingly, a singular noun (like *table* or *quantity of water*) is true of singularities, a plural noun is true just of (the corresponding) pluralities (to the exclusion of singularities) and a mass noun is true in an undifferentiated manner of both. This idea can be schematically visualized as follows:

(84)

quantities of water

$$
\begin{bmatrix}
 & a \cup b \cup c & \\
a \cup b & b \cup c & a \cup c \\
\end{bmatrix}
$$

quantity of water $\begin{bmatrix} a & b & c \end{bmatrix}$

water

I tried to make a case in Chierchia (1998a) that all the distinctions that can be made on the Double Domain approach can also be made, as economically, on the basis of this rather minimal idea, and that the resulting approach was, therefore, to be preferred on economy grounds. (Also, the third argument discussed above goes through on my 1998a.)[56]

The main drawbacks of Chierchia's (1998a) proposal are two. The first is that it has become increasingly difficult to maintain for English the view of plurals adopted there. As pointed out in Sect. 2, there are good reasons for thinking that the denotation of plurals in number marking languages has to include also the singularities. This makes the denotation of plurals identical to that of mass nouns as understood there and makes it hard to account for their different behaviour. The second drawback has to do, again, with fake mass nouns (which provided much of the push behind that proposal). Clearly, fake and core mass nouns come out as having the same structure

---

[55] In so far as I can tell exactly the same drawbacks can be levelled towards theories like Bunt (1979) (that models mass nouns in terms of 'ensembles' with infinitely descending ∈-chains), and towards theories that model the mass/count contrast in terms of distinct Boolean structures, like Lonning (1987) or Higginbotham (1994).

[56] I discuss it in Chierchia (1998a) under the label 'the supremum argument'.

on Chierchia's (1998a) approach. As we now know, this is wrong on several counts. Why, for example, wouldn't there be a language in which every core mass noun is registered as count and every count noun as mass?

I was aware of this conceptual issue and pointed to vagueness as a way to address it, but didn't develop this idea any further. As it turns out, in view of the evidence discussed above, it is necessary to do so. The present proposal, in a way, marries the basic insight of my 1998a proposal, with the idea that core mass nouns have vague atoms. This addresses both the problem of the treatment of plurals and the inadequacies in the treatment core versus fake mass at once.

There are of course many other venues that can (and have) been taken to try to understand what is involved in mass versus count. But I have certainly abused enough of the reader's patience to engage in further comparisons.

7.2 Concluding remarks

The work on plurals consistently points in the direction of a formal semantic distinction between singularities and pluralities, often cashed out in terms of atomic/non atomic. We furthermore know that most properties/kinds we are able to conceptualize are conceptualized as having vague boundaries. The idea pursued here is that vagueness must invest the atom/non atom distinction as well. There are things that we don't know whether they are atoms or not ('unstable individuals'). How could it be otherwise? This much seems to belong to the realm of 'virtual conceptual necessity'. The more specific claim we have explored is that the mass/count distinction rides on it. Simplifying somewhat, a noun is count if there are at least some things it applies to that are clearly atomic (sometimes in an absolute sense, e.g. with *cat*; some other times in a relative sense as with *quantity of apples*). A noun is mass if *all* of its *minimal* instances/manifestations fall in the vagueness band (and can be construed as atomic or not depending on how we choose to make things more precise). This idea uses vagueness to identify a formal/grammatical distinction on how nouns are mapped out. Such a formal distinction naturally converges with the pre-linguistic notion of 'bounded individual', as defined by Carey and Spelke, that we (as well as other species) seem to be endowed with.

The present hypothesis provides a robust account of the signature property of mass nouns: the impossibility of using them in direct counting. To count we need a sortal property capable of identifying a suitable level at which to count. If a property can be applied to aggregates/pluralities of varying size, we go for the minimal samples, the atoms; or else we go for partitions (with relational nouns like *quantity*); if the status of such minimal samples is unclear (vis-avis atomicity), we are stuck, in so far as counting goes. This sets the boundaries of Babel (to steal Moro (2008) captivating phrase), i.e. it grounds the universal facets of the phenomenon and links it to its extra linguistic roots.

The variable aspects of how mass/count functions depend on choices languages may make in dealing with counting and atomicity. In summarizing my main proposals on this score, it should be borne in mind that they are in part independent from the main thesis and have, therefore a more open ended ('negotiable') status. In particular,

we have focussed on classifier versus number marking languages and conjectured that generalized classifiers and obligatory singular/plural marking are two different ways of making sure that a property is (relatively) atomic and hence useful for counting. The choice between the two depends on whether nouns are coded as kinds or as properties. If they are coded as kinds, given that numbers want properties, something must operate the shift (and it better create atomic properties). On the other hand, if nouns are properties, something must 'check' that they *are* atomic. As mass properties are not, number marking languages must treat them in a way that differs from other nouns. We have explored the possibility that one such way is treating them as singleton properties (thereby making them formally atomic). This creates a window (i.e. a formal niche) through which fake mass nouns slide in.

Many, perhaps all the ideas I have first sketchily presented and now even more sketchily summarized are likely to turn out to be wrong. However, the general architecture of the present proposal may well survive such (welcome) falsifications. It is a fundamental fact that certain properties of the distinction are tendentially universal (and clearly linked to a pre linguistic cognitive capacity) while certain, specific others vary in systematic ways. It is my hope that the present approach contributes to a better understanding of this fundamental divide.

References

Barker, C. (1992). Group terms in English: Representing groups as atoms. *Journal of Semantics, 9*, 69–93.

Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition, 97*, 41–66.

Borer, H. (2005). *Structuring sense*. Oxford: Oxford University Press.

Bunt, H. (1979). *Ensembles* and the formal semantic properties of mass terms. In F. J. Pelletier (Ed.), *Mass terms: Some philosophical problems*. Dordrecht: Kluwer.

Carlson, G. (1997). Reference to kinds in English. Ph. D. Dissertation, University of Massachusetts at Amherst. Published in 1980 by Garland, NY.

Cheng, L., & Sybesma, R. (1998). Yi-wang Tang, yi-ge Tang: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies, XXVIII*(3), 385–412.

Chierchia, G. (1984). Topics in the syntax and semantic of infinitives and gerunds. Ph. D. Dissertation, University of Massachusetts at Amherst. Published in 1989 by Garland, NY.

Chierchia, G. (1998a). Plurality of mass nouns and the notion of 'semantic parameter'. In S. Rothstein (Ed.), *Events and grammar*. Dordrecht: Kluwer.

Chierchia, G. (1998b). Reference to kinds across languages. *Natural Language Semantics, 6*, 339–405.

Chierchia, G. (2004). Scalar implicature polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond*. Oxford: Oxford University Press.

Chierchia, G. (2005). Definites, locality, and intentional identity. In G. Carlson & J. Pelletier (Eds.), *Reference and quantification*. Stanford: CSLI.

Chierchia, G., & Turner, R. (1988). Semantics and property theory. *Linguistics and Philosophy, 11*, 261–302.

Chierchia, G., Fox, D., & Spector, B. (2008). The grammatical theory of scalar implicatures. In C. Maierborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning*. Berlin: Mouton de Gruyter (forthcoming).

Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy, 27*(4), 393–450.

Fine, K. (1975). Truth, vagueness and logic. *Synthese, 30*, 265–300.

Fox, C. (1998). Plurals and mass terms in property theory. In F. Hamm & E. Hinrichs (Eds.), *Plurality and quantification*. Dordrecht: Kluwer.

Gajewski, J. (2002). L-analyticity and natural language. Manuscript, University of Connecticut at Storrs.

Gallin, D. (1975). *Intensional and higher order logic*. Amsterdam: North Holland.

Gillon, B. (1992). Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy, 15*, 597–640.

Gillon, C. (2009). *Mass, count and plurality in Innu-aimun*. Paper presented at SULA V, Harvard/MIT, May 2009.

Grimshaw, J. (2000). Extended projection and locality. In P. Coopmans, M. Everaert, & J. Grimshaw (Eds.), *Lexical specification and insertion*. Amsterdam: J. Benjamins.

Hauser, M., & Spaulding, B. (2006). Wild rhesus monkeys generate causal inferences about possible and impossible physical transformations in the absence of experience. *Proceedings of the National Academy of Science, 103*, 18.

Heim, I. (1982). The semantics of definite and indefinite noun phrases. Ph. D. Dissertation, University of Massachusetts at Amherst. Published in 1989 by Garland, NY.

Higginbotham, J. (1994). Mass and count quantifiers. *Linguistics and Philosophy, 17*, 447–480.

Horn, L. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Huang, J. (2006). Analyticity, variation and change. In *International conference of East Asian Linguistics*, University of Toronto.

Ionin, T., & Matushansky, O. (2006). The composition of complex cardinals. *Journal of Semantics, 23*, 315–360.

Kamp, H. (1975). Two theories about adjectives. In E. Keenan (Ed.), *Formal semantics of natural language*. Cambridge: Cambridge University Press.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy, 30*, 1–45.

Kiparsky, P. (1982). From cyclic phonolophy to lexical phonology. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representation*. Dordrecht: Foris.

Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy, 4*, 1–45.

Kratzer, A. (1989). An investigation into the lumps of thought. *Linguistics and Philosophy, 12*, 607–653.

Krifka, M. (2008). *Masses and countables*. Paper presented and the workshop "The Syntax and Semantics of Measurement", University of Tromsø.

Kurafuji, T. (2003). Plural morphemes, definiteness and the notion of semantic parameter. In *The proceedings of the third GLOW in Asia*.

Landman, F. (1989). Groups I. *Linguistics and Philosophy, 12*, 559–605.

Landman, F. (1991). *Structures for semantics*. Dordrecht: Kluwer.

Landman, F. (2004). *Indefinites and the type of sets*. Oxford: Blackwell.

Lasersohn, P. (1999). Pragmatic halos. *Language, 75*, 522–551.

Levinson, S. (2000). *Presumptive meaning*. Cambridge: MIT Press.

Li, Y.-H. A. (1999). Plurality in a classifier language. *Journal of East Asian Language, 8*, 75–99.

Link, G. (1983). The logical analysis of plural and mass terms: A lattice theoretic approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language* (pp. 302–323). Berlin: de Gruyer.

Lonning, J.-T. (1987). Mass terms and quantifiers. *Linguistics and Philosophy, 10*, 1–52.

Longobardi, G. (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry, 25*, 609–665.

Mathieu, E. (2009). *On the count/mass distinction in Ojibwe*. Paper presented at the workshop on parts and quantities, University of British Columbia, November 2007.

McGee, V., & McLaughlin, B. (1998). Vagueness by Timothy Williamson. *Linguistics and Philosophy, 21*, 221–235.

Moro, A. (2008). *The boundaries of Babel*. Cambridge, MA: MIT Press.

Pelletier, J., & Schubert, L. (1989). Mass expressions. In D. Gabbay & F. Guenthner (Eds.), *Handbook of philsophical logic* (Vol. IV). Dordrecht: Kluwer.

Pinkal, M. (1989). Imprecise concepts and quantification. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expressions*. Dordrecht: Foris.

Quine, W. (1960). *Word and object*. Cambridge, MA: MIT Press.

Quine, W. (1969). *Ontological relativity and other essays*. NY: Columbia University Press.

Rothstein, S. (2008). Counting and the mass/count distinction, ms. Bar Ilan University, Tel Aviv.

Sauerland, U. (2003). A new semantics for number. In R. Young & Y. Zhou (Eds.), *Proceedings of SALT 13*. Ithaca, NY: CLC Publications, Cornell University

Schwarzschild, R. (1996). *Pluralities*. Dordrecht: Kluwer.

Schwarzschild, R. (2007). *Mass nouns and stubbornly distributive predicates*. Paper presented at the Harvard Colloquium Series, September 2007, STUBs/ Pluralities

Soja, N., Carey, S., & Spelke, E. (1991). Ontological categories guide young children's inductions of word meanings: Object terms and substance terms. *Cognition, 38*, 179–211.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland & P. Stateva (Eds.), *Presuppositions and implicatures in compositional semantics*. Hampshire, UK: Palgrave-Macmillan.

Steriade, D. (2008). A pseudo-cyclic effect in romanian morphophonology. In A. Bachrach & A. Nevins (Eds.), *Inflectional identity*. Oxford: Oxford University Press.

Tsoulas, G. (2006). *Plurality of mass nouns and the grammar of number*. Paper presented at the 29th glow meeting, Barcelona.

Van Fraassen, B. (1971). *Formal semantics and logic*. NY: McMillan.

Veltman, F. (1985). Logics for conditional. Ph.D. Dissertation, University of Amsterdam.

Wilhelm, A. (2008). Bare Nouns in Dëne Sułine. *Natural Language Semantics, 16*, 39–68.

Williamson, T. (1994). *Vagueness*. London: Routledge.

Zamparelli, R. (2008). Bare predicate nominals in romance languages. In H. Høeg Müller & A. Klinge (Eds.), *Essays on nominal determination*. John Benjamins.