Gerhard Schaden
Rapport de mission
Participation au colloque
Empirical, Theoretical and Computational Approaches to Countability in Natural Language

Ce colloque a été organisé par Tibor Kiss et son équipe du département de Sciences du langage de la Ruhruniversität Bochum en Allemagne, et a réuni une trentaine de participants du 22 au 24 septembre 2010.

Argumentaire général du colloque (en anglais)

The distinction between mass and count nouns has been addressed in a variety of linguistic (and also extra-linguistic) approaches. Initially, it has been suggested that the distinction is a property of lexemes, or that it can be derived from properties of the objects denoted by the respective nouns. This assumption has been severely challenged by a variety of approaches, leading to the assumption that countability is a property of constructions and phrases. Yet, a critical survey of the most advanced work on the count-mass distinction has shown that multiple, partially conflicting views on this phenomenon are still competing.

As an illustration for unsettled questions, consider the following:

- If the mass-count distinction is actually dependent on formal syntactic and/or semantic marking, how are nouns to be classified that lack such a marking, e.g., nouns in preposition-noun combinations (determinerless PPs)?
- If mass is taken to be a basic property of nouns to which syntactic marking must be added to transform the noun into a count noun, why do certain languages already require such marking for mass terms (e.g., Romance languages)?
- How can the apparent tension between theoretical constructional (i.e., token-based, and hence construction-specific) and computational (i.e. primarily type-based, and hence possibly lexical-class-based) classification be resolved?

The goal of this conference is to bring researchers from all areas of linguistics together to clarify the numerous existing theories concerning the count-mass distinction and also to offer a platform for new insights and constructive criticism.

Pour l'abstract de ma communication, cf. les pages suivantes.

Problematic Feature Mappings in Number

1 Introduction

Countability and number are central issues in the linguistic analysis of noun phrases. The problem has been treated from two different perspectives: semantics and morphology (or morphosyntax). Countability underlies the mass-count divide, and from a semantic point of view, it makes a difference if something can be counted or not, and if there are single instances of it, or several. Considerable work in formal syntax and semantics, starting from Link (1983/2000), has been dedicated to the problems associated with these issues.

On the other hand, number is also a pervasive issue in morphology, as testified by books like Corbett (2000). Yet, as noted by Harbour (2008), morphology- and semantics-based approaches to number do not interact that much. One does hardly ever find any explicit reference to morphology in semantics literature, or to semantics in morphology literature. One interpretation (the optimistic one) of this state of affairs is that the mapping between the semantic category of number (henceforth: S-number), and the morphological category of number (henceforth: M-number) is entirely transparent: a category of S-number is systematically encoded as the equivalent category in M-number, and vice-versa. If this were the case, of course, disregarding one or the other aspect of number would not be a problem.However, should the mapping turn out not to be that transparent, by not treating one aspect of number, one might miss important parts of the puzzle.

The general aim pursued in this paper is extremely modest: I will argue that the mapping from M-number to S-number — even in languages like German or English — is not obvious. More specifically, I will try to show that indeed i) there are cases of conflict between M- and S-number; ii) that M-number maybe does not encode what has been analysed as S-number; and iii) that for the moment being, our strategy should be to clearly keep these two notions apart. To the end, I will present data from Southern German indefinite articles, which exemplify both complicated alignment patterns and an interaction between S- and M-number.

2 Background

Much of linguistic theorizing has been trying to account in some way or another for a similarity in behaviour between mass and (count) plural nouns. However, it has not so often been looked at what separates these two categories: mass seems to be singular (M-singular, at least) after all, and as we will see, other diagnostics fail to unite these two as well.

2.1 Mass Nouns as Lexical Plurals

A dominant idea in the area of semantics is that mass nouns are lexical plurals (cf., e.g., Chierchia, 1998a; Chierchia, 1998b; Rothstein, 2010). Like plurals, mass have the properties of cumulativity and divisivity (cf. Kiparsky, 1998; Krifka, 1998), and as mass is M-singular, one might consider that we are faced with a possible source of conflict between M- and S-number. One of the motivations for the mass-is-plural hypothesis is combinatorial: it has often been observed (for instance with indefinite articles) that mass nouns — while being M-singular like count singulars — pattern with respect to determiners with M-plural count nouns. A suggested explanation is that they are both S-plural.

A prediction associated to the mass-is-plural hypothesis is that there are no determiner that combines exclusively with singular mass nouns and singular count nouns, while being inappropriate with plural count nouns. Stated differently: determiners do not care for morphology, they only care for semantics. I will present in section (3) a

counterexample to the prediction, but let us first see some more general inconvenients for this idea. First, as has been observed by Borer, 2005, one cannot say things like (1): M-number seems to be important and independent.

(1) *Water are refreshing.

The inacceptability of (1) might just be a fact about morphosyntactic agreement. However, the analogy of mass being plural might be problematic for semantic reasons as well. It has been observed by Farkas and Swart (2010, 6:11f.) that (personal) pronouns can be sensitive to semantic rather than morphosyntactic features.

(2) a. Három gyerek elment | *elment**ek**   b. Mari nem látta őket | *őt
    three child leave.PAST.SG | leave.PAST.PL   Mari not see.PAST III.PL.ACC | III.SG.ACC
    'Three children left.'   'Mary didn't see them.'

The point of (2) is the following: in Hungarian, cardinals require obligatorily singular M-number agreement on the noun and the verb. However, anaphora obligatorily take M-plural (cf. (2b)). This means that in Hungarian, discourse anaphora by pronouns agree with S-number rather than with M-number.

Now, should mass be semantically plural in any pertinent sense, one would expect pronouns that are anaphoric to mass nouns to allow for M-plural as well. Semantic agreement, however, does not seem to be even possible in these circumstances. I will illustrate the case with German, since there are known conflicts between morphological and "semantic" features which can be resolved either way, for instance with gender:

(3) Ich habe ein Mädchen$_i$ gesehen. Es$_i$ | Sie$_i$ stand auf dem Tisch.
    I have a girl.NEUT seen.   III.NEUT | III.FEM stood on the table
    'I saw a girl$_i$. She$_i$ was standing on the table.'

So, one would expect that — if there were a conflict between S- and M- features in number — it could be resolved in a similar way.[1] This, however, is not the case for German, nor for English.

(4) a. Der Kellner brachte uns Wasser$_i$. #Sie$_i$ waren kalt.
    the waiter brought us water.   they were cold.
  b. The waiter served water$_i$. #They$_i$ were cold.

So, the question remains whether this is a case where S-number would be for some reason inaccessible — in which case we could maintain the idea of 'mass as plural' —, or if mass and plural share some properties, but are not identical in all aspects.

2.2 An Alignment Puzzle

A second case in point to show that the mapping between S- and M-number is maybe not that simple is the fact that they have opposite poles of markedness: S-plural is unmarked, but M-plural marked, and M-singular is unmarked, but S-singular marked.

Generally, it is expected (by Horn's division of Pragmatic Labour, cf. Horn (1989), or other, similar principles) that marked forms receive marked interpretations, and that unmarked forms go with unmarked interpretations. Yet, as has often been noted (cf. Bale, Gagnon, and Khanjian, in press for an overview and a detailed examination), with number, the Horn pattern does not seem to obtain.

The problem is the following: in languages such as English, morphologically, it is quite obviously the plural that is the marked form, since it has an *-s* suffix on what looks like the singular, whereas the singular seems to be identical to the root. So, the plural form F_{Pl} is more marked as the singular form F_{Sing} (cf. (5a)). However, the semantic side

---

[1]This assumes that a given language resolves such feature conflicts in a coherent manner, which might not be a realistic assumption. In modern German, articles or relative pronouns have to agree with M-number (this might be a locality issue). Moreover, the pattern we observe in (3) might be a relatively recent one. For instance, in the Grimm brother's version of *Cinderella*, agreement systematically proceeds with (neuter) M-number, whereas a version by Bechstein (roughly one generation later) displays a rather modern mixture of agreement with M-number and S-number.

seems to be precisely the opposite: it has been convincingly argued by Sauerland, Andersen, and Yatsushiro (2005) that the plural meaning Σ_{Pl} is the unmarked one, and that the singular meaning Σ_{Sing} is marked (cf. (5b)).[2]

(5) a. $F_{\text{Pl}} > F_{\text{Sing}}$ b. $\Sigma_{\text{Sing}} > \Sigma_{\text{Pl}}$

Under a Horn alignment pattern, one should get the combination in (6a) — reflecting the pairing of the unmarked form with the unmarked meaning, and of the marked form with the marked meaning —, rather than the combination in (6b) — reflecting a combination of the marked form with the unmarked meaning, and vice versa (which is a so-called Anti-Horn configuration) —, which is what is effectively attested in languages like German or English:

(6) a. $\langle F_{\text{Sing}}, \Sigma_{\text{Pl}} \rangle$ and $\langle F_{\text{Pl}}, \Sigma_{\text{Sing}} \rangle$ b. $\langle F_{\text{Pl}}, \Sigma_{\text{Pl}} \rangle$ and $\langle F_{\text{Sing}}, \Sigma_{\text{Sing}} \rangle$

According to Bidirectional OT (cf., e.g., Blutner, 1999), the expected pattern of alignment would be (6a). The fact of obtaining an anti-Horn pattern may or may not be something desirable; but with Farkas and Swart (2010), I think it would be preferable that a theory of number displays a Horn-pattern here as well.

While on a general level, the assumed anti-Horn pattern with number might be a linguistic curiosity, but not condemning for an analysis (since it might be a fact about language, not about theory), from a derivational point of view, the anti-Horn pattern may be highly disturbing. For instance, (7) describes what Alonso-Ovalle and Menéndez-Benito (to appear) assume as their semantics for singular vs. plural. (7), of course, is not the only possibility, but this explicit statement illustrates the problem extremely well.

In order to keep notation short, I only note three individuals in the domain, namely a, b, and c. Alonso-Ovalle and Menéndez-Benito assume the root noun to be number-neutral, and to contain therefore a complete semi-join lattice; the singular contains only atoms, and the plural denotes again a full semi-join lattice.

(7) a. $[\![\sqrt{\text{student}}]\!] = \{ a, b, c, a \oplus b, a \oplus c, b \oplus c, a \oplus b \oplus c, \dots \}$
  b. $[\![\text{singular}]\!] ([\![\sqrt{\text{student}}]\!]) = \{ a, b, c, \dots \}$
  c. $[\![\text{students}]\!] = \{ a, b, c, a \oplus b, a \oplus c, b \oplus c, a \oplus b \oplus c, \dots \}$

All these assumptions are (semantically, and individually) well motivated. However, as soon as one adds morphology to the derivational elements in (7), the picture becomes rather puzzling: The root being identical (on surface, at least) to the singular form, one gets for free the reduction to atoms, as encoded by the singular. However — and it does not really matter how the plural is derived — an operation which in the end does not have any impact on the sense of the root has to be 'paid' for by a morpheme. Why cannot we simply take the root in order to express that meaning? Stated differently: shouldn't (7c) be blocked by the mere existence of (7a)? If one believes in economy playing some role in natural language, the pattern in (7) does not make any sense.

3 The Southern German Indefinite Article

In this section, I will introduce and discuss data from a Southern German dialect, which shows that an analysis taking into account only M- or only S-number has to fail, and that the interaction of these categories has to be taken into account.

3.1 Relevant Data

The Southern German indefinite article exemplifies the pattern that is predicted to be impossible by Chierchia, 1998a; Chierchia, 1998b, namely a determiner that is possible only with M-singular nouns, irrespective of mass- or count status, and ungrammatical with M-plural nouns, as illustrated in examples (8) – (10). (8) and (10) are identical to the English or standard German pattern; the crucial data is the one in (9).

(8) a. mir hond an hund.
        we have a   dog.

  b. *mir hond hund.
        we have dog.

(9) a. mir hond a meal.
        we have a flour.
        'We have (some) flour.'

  b. ??mir hond meal.
        we have flour.

(10) a. *mir hond a hünd.
        we have a dogs.

  b. mir hond hünd.
        we have dogs.

Crucially, (9a) is the normal version, and not (9b) (which sounds possible only as interference of Standard German). It needs to be stressed that (9a) is not the fruit of some kind of coercion-effect. The dialect in question has — like English — an opposition between a cardinal determiner 'one' and an indefinite article 'a', and not in all contexts where the indefinite is appropriate, it can be substituted by the cardinal determiner. Notably, the indefinite is acceptable with nouns like *money*, which are notoriously difficult to coerce into something countable:

(11) a. dr schorsch hot a gealt.
        the George   has a money.
        'George has (some) money.'

  b. *dr schorsch hot uj   gealt.
        the George   has one money.
        'George has one money.'

Similarly, the indefinite article does not influence the judgements with respect to homogeneity of a count or a mass noun, contrary to what the cardinality article does, which inforces countability:

(12) a. a wassr ⊕ a wassr = a wassr
        a water   a water   a water

  c. an hund ⊕ an hund ≠ an hund
        a dog   a dog   a dog

  b. uj   wassr ⊕ uj   wassr ≠ uj   wassr
        one water   one water   one water

  d. ujn hund ⊕ ujn hund ≠ ujn hund
        one dog   one dog   one dog

One might argue that this indefinite article is purely concerned with M-number, has no semantic impact, and no influence whatsoever on the count-mass distinction. However, this is not true. In generic judgements, there is a clear difference between mass and count M-singulars:

(13) a. *a īs   isch kolt.
        a ice is   cold.

  b. īs   isch kolt.
        ice is   cold.

(14) a. an hund isch a fich.
        a dog is   an animal.

  b. *hund isch a fich.
        dog is   an animal.

Intuitively, (13a) is impossible because it means something like *a quantity of ice is cold*. However, being cold is not a property of a quantity of ice, but of the substance 'ice' itself.

Whatever the reason of the divergence of acceptability beween (13) and (14) may be, three things seem quite clear: i) this indefinite article has semantic import; ii) it cannot be reduced to superficial (i.e., morphosyntactic) agreement; and iii) it does not coerce mass into count-like entities.

3.2 Discussion

A viable account of the Southern German indefinite article could procede in two directions: first of all, assume in the tradition of Link, 1983/2000 that there is a semantic difference between mass and plural, and some similarities between mass singular and count singular. Second, one might try to investigate more precisely the interaction and mapping from S-number to M-number involved here. Certainly, both ways are valuable enterprises, and probably, a successful account will combine elements of both, since I do not think that neither S- nor M-number can be evacuated from the discussion. Yet, here I would like to dedicate the remainder of this paper to the discussion of M- and S-plurality. This area seems critical to me, given the generally unquestioned hypothesis that M-plural denotes S-plural, an assumption that is rather dubious (cf. Borer, 2005 for discussion).

3.2.1 The Semantics of Plurality

The paper by Sauerland, Andersen, and Yatsushiro (2005) compares two theories of pluralities, an exclusive (intuitively: excluding the denotation of the singular) and an inclusive theory (where the denotation of the plural includes the denotation of the singular), and argues conclusively for the latter. The theoretical background of the paper is solidly anchored in the Link mereologic tradition, assuming that plurals are constructed by sum (or maybe set-) formation out of atomic elements, and denote lattices. This means that these theories operate under the assumption that plurals are about 'multitudes' of things. One should be aware that this is actually a theory of S-plurality, and not necessarily of the semantic import of M-plurality in natural languages.

For Borer (2005), however, plurals are actually classifiers, associated with a [DIV](isivity) feature. Borer constructs in her book a theory of the semantic impact of M-plurality. Part of the motivation comes from the following data from German or English:

(15) a. a point_   c. zero point**s**   e. 1.2 point**s**
      b. 2 point**s**   d. 0.3 point**s**   f. 1.0 point**s**

The picture that emerges is that the singular comes with the basic unit, and that all the rest takes a plural. However, only (15b) is strictly speaking an instance of S-plurality. The question is now what is the semantic content associated with the DIV-feature, since, unfortunately, Borer did never define explicitly what this feature is about.[3]

While in (15), M-plural marking does not seem to be about S-pluralities, precisely this is often the case. The typical effect of pluralising a mass noun is to impose some discrete structure on it. It seems therefore that there is some link between M-plurals and S-pluralities. But how can this be reconciled with the data in (15)? The basic idea I would like to put forward is that an M-plural presupposes that there exists some (possibly contextually given) unit in connection with a predicate such that that unit can serve as a basis to counting, whereas it does not specify what such a basic unit may be.

I assume such a content of M-plural to be presuppositional in nature and that it may give rise to presupposition failures. I also assume with Rothstein (2010) that a count noun comes with some indication as to what is to be counted in a given context, whereas mass nouns lack such an indication. Let us now look at what M-plural does under these assumptions. The semantics of a plural will not have that much impact on a count noun, since by definition, the basic unit will already be (at least partly) defined. However, with a mass noun, the existence of a unit will be coerced upon the underlying semantics — whenever possible — or else, the interpretation of the plural formation will fail.[4]

So, under these assumptions, plurals guarantee countability, not S-plurality *per se*. Apart from the semantic import of M-plurals, another point that deserves to be looked at is the question of the nature of these units — in case of count nouns as well as with pluralised mass nouns. For Rothstein, the indications what is to be counted are (largely) contextual, and independent of the predicate. While this is on the right track for some cases, it seems to be missing important points.

3.2.2 What Do We Count When We Count?

A first observation to be made is that, at some point, one can count nearly anything (with the notable, but prima facie confusing exception of 'money', cf. (11b)). The important point is that with count nouns, one counts individuals, but that with pluralised (or cardinalised) mass nouns, one counts consumption units, (sub-)kinds — anything but individuals.[5] This can be illustrated with the following thought experiment: assume somebody pours water from

---

[3]The content of [DIV] clearly cannot correspond to divisibility in the mereological tradition, as defined in Kiparsky (1998).

[4]Notice that therefore, M-plurals come without existential import, and are not about 'S-plurality' in any Link-compatible way.

[5]The notion of 'individual' used here is certainly problematic, but I do not know of a better term. Cf. also the discussion on the notion of 'integrated whole' in Moltmann, 1998.

a single bottle onto a table, where two, clearly distinct puddles form. The point that has often been made is that counting is not about reference or denotation. One could say the following in this circumstances:

(16) a.   There is water on the table.   c.   #There are two waters on the table.
     b.   There are two puddles on the table.

The entities as such are the same, and they are clearly bounded under our assumptions; yet, one simply cannot utter (16c) in these circumstances. The fact that there are two puddles does not allow for the assertion that there are two waters. There should be some kind-distinction, but this is not warranted here, since both units come from a single bottle, and a puddle does not qualify as consumption-unit. Where the two puddles formed of two different types of water — like sparkling vs. tap water — one could say (16c). The lesson I think to be learned from (16) is that we do not count entities, but entities *as instantiations of predicates* in a given context. The predicate and its possibilities of type-coercion (the type to be taken as in the tradition of Pustejovsky (1995)) determine the countability of an entity just as much as context, which is not something that is captured in Rothstein (2010).

4 Conclusions and Perspectives

In this paper, I have argued that — at least at the current state of affairs — it is important to play close attention to the fact that 'number' refers to quite different entities in natural language, according to the linguistic subfield one belongs to, and that the mapping between morphology and semantics is problematic. I have also argued that there M-plurals are quite probably not about S-plurals in a Link-Chierchia compatible sense.

My hope is that by raising conscience of the possibly conflicting issues of the two notions of number involved, progress in the analysis of the numerous puzzles I merely alluded to here will be made easier.

References

[1] Alonso-Ovalle, L. and P. Menéndez-Benito (to appear). "Plural Epistemic Indefinites". In: *Proceedings of NELS 40*. [2] Bale, A., M. Gagnon, and H. Khanjian (in press). "On the Relationship Between Morphological and Semantic Markedness: The Case of Plural Morphology". In: *Journal of Morphology*. [3] Blutner, R. (1999). "Some Aspects of Optimality in Natural Language Interpretation". In: *Papers on Optimality Theoretic Semantics*. Ed. by H. de Hoop and H. de Swart. Utrecht Institute of Linguistics OTS: pp. 1–21. URL: http://www.blutner.de/optimal.pdf. [4] Borer, H. (2005). *In Name Only*. Oxford: Oxford University Press. [5] Chierchia, G. (1998a). "Plurality of Mass Nouns and the Notion of "Semantic Parameter"". In: *Events and Grammar*. Ed. by S. Rothstein. Dordrecht: Kluwer: pp. 53–103. [6] Chierchia, G. (1998b). "Reference to Kinds Accross Languages". In: *Natural Language Semantics* 4.6: pp. 339–405. [7] Corbett, G. G. (2000). *Number*. Cambridge: Cambridge University Press. [8] Farkas, D. and H. de Swart (2010). "The Semantics and Pragmatics of Plurals". In: *Semantics and Pragmatics* 3: pp. 1–54. DOI: 10.3765/sp.3.6. [9] Harbour, D. (2008). *Morphosemantic number. From Kiowa Noun Classes to UG Number Features*. Dordrecht: Springer. [10] Horn, L. R. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press. [11] Kiparsky, P. (1998). "Partitive Case and Aspect". In: *The Projection of Arguments*. Ed. by M. Butt and W. Geuder. Stanford: CSLI: pp. 265–307. [12] Krifka, M. (1998). "The Origins of Telicity". In: *Events and Grammar*. Ed. by S. D. Rothstein. Dordrecht: Kluwer: pp. 197–235. [13] Link, G. (1983/2000). "The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretical Approach". In: *Formal Semantics. The Essential Readings*. Ed. by P. Portner and B. H. Partee. Oxford: Blackwell: pp. 127–146. [14] Moltmann, F. (1998). "Part Structures, Integrity, and the Mass-Count Distinction". In: *Synthese* 116.1: pp. 75–111. [15] Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge: MIT Press. [16] Rothstein, S. D., ed. (1998). *Events and Grammar*. Dordrecht: Kluwer. [17] Rothstein, S. D. (2010). "Counting and the Mass–Count Distinction". In: *Journal of Semantics*: pp. 1–55. DOI: 10.1093/jos/ffq007. [18] Sauerland, U., J. Andersen, and K. Yatsushiro (2005). "The Plural is Semantically Unmarked". In: *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*. Ed. by S. Kepser and M. Reis. Berlin: Mouton de Gruyter: pp. 413–434.