

MASS EXPRESSIONS

Ms 1097

Author: Harry Bunt

Author Address: Tilburg University

P.O. Box 90153

5000 LE Tilburg

The Netherlands

Email: harry.bunt@uvt.nl

Abstract

This article deals with the count/mass distinction, which is found in many languages and is chiefly a distinction among nouns. Mass nouns, like *water*, *furniture* and *measles*, do not have both a plural and a singular form, and combine with different determiners and adjectives than count nouns. Semantically, mass nouns differ from count nouns in that they do not 'individuate': they have no built-in way of dividing their reference into individual objects. This latter characteristic can be captured formally by using mereological or ensemble-theoretical part-whole structures as mass noun denotations, rather than set-theoretical structures.

By 'mass expressions' one usually means expressions formed with so-called 'mass nouns', words like *water*, *rice*, *poetry* and *garbage*, which differ morphologically and syntactically from count nouns, like *book*, *apple*, *chair* and *house*, in that they do not have both a singular and a plural form, and differ in the possible combinations with numerals, determiners, and adjectives. In particular, mass nouns do not admit combination with numerals, with the indefinite singular article, and with a range of quantifiers: **a water*, **both rice*, **five poetry*, **many garbage*, **several music*. Count nouns, on the other hand, do not combine well with certain quantifying adjectives, such as English *much* and *little*; Spanish *mucho* and *poco*; or Danish *meget* and *lidt*, which combine only with mass nouns. In contrast with count nouns, mass nouns allow the formation of determinerless singular noun phrases: *There's furniture in this room* versus **There's chair in this room*. Such bare NPs are often called 'mass terms' and include phrases like *imported furniture*, *eau de Cologne*, *mousse au chocolat*, *orange juice from Brazil*, and *refined pure Cuban cane sugar*.

The count/mass distinction is found in many languages, but is not universal and has different manifestations in different languages. The Hopi language has been mentioned by Whorf (1939) as a language that has no mass nouns, and several Asian languages such as Chinese and Japanese do not mark the count/mass distinction and have been claimed to have only mass nouns (Sharvy, 1978). Also, what is described by a count noun in one language may be described by a mass noun in another; e.g., English *fruit* is a count noun, whereas Dutch *fruit* is a mass noun (*a fruit* can be translated either as *een stuk fruit* (*a piece of fruit*) or as *een vrucht*. In English, by far the majority of mass nouns is morphologically and syntactically singular; a minority of mass nouns is syntactically plural. A plural example is *measles*: it would be funny to ask **How many measles have you got* (except maybe in a conversation between two doctors, as a way of referring to measles patients). In other languages, plural mass nouns are quite frequent, e.g. in Swahili the word for *water* is the plural mass noun *maji*; in Italian the many pasta varieties are syntactically plural, as they are when imported in other languages. One should, for example, say *How much spaghetti do you want* rather than **How many spaghetti do you want*, just as in Italian (*Quanto spaghetti...; *Quanti spaghetti...*). A morphologically remarkable phenomenon in Dutch is the formation of diminutive forms of mass nouns to refer to portions with certain properties of the stuff that the mass noun refers to. For example, the mass nouns *snoep* (candy), *drop* (liquorice), *chocola* (chocolate), *plastic*, *papier* (paper), *brood* (bread) have the diminutive forms *snoepje*, *dropje*, *chocolaatje*, *plasticje*, *papiertje*, *broodje* which are count nouns referring to particular types of physically well-defined pieces of candy, liquorice, chocolate, plastic, paper, and bread, respectively. The interesting point is that diminutive forms of mass nouns differ systematically from those of count nouns in that diminutive count nouns, as opposed to mass noun diminutives, invariably refer to small exemplars in the set of objects denoted by the count noun. For example, *appeltje* and *boekje* refer to small apples and little books, but the diminutive form of a mass noun, like *broodje* or *plasticje*, does not refer to a small piece of bread or plastic, but to specific pieces, like a roll and a folder for storing papers.

While the count-mass distinction as just outlined may seem intuitively clear, it turns out to be difficult to make it precise, especially to make it sufficiently precise for being incorporated in a formal grammar. For example, while *apple* may seem a clear example of a count noun, it is possible to say things like *Don't put too much apple in the salad*, using *apple* as a mass noun. David Lewis has invented a hypothetical device to show that every count noun can be used as as a mass noun. This device, the Universal Grinder, can take as input any objects, denoted by a count noun, like apples, books, or crocodiles; it grinds these and spits out the stuff that the objects were made of: apple-stuff, book-stuff, crocodile-stuff. This machine could be said to turn *apples* into *apple*, *books* into *book*, and *crocodiles* into *crocodile*. One can also imagine a device that works in the other direction. This device, that we might call the Universal Packer, takes as input a continuous stream of any stuff that a mass term *M* may refer to, and outputs packages containing amounts of *M* that are appropriate in a given context. This device illustrates that one can in general construct a count use of a mass noun by finding a context in which the stuff, that

the mass noun normally refers to, comes in certain standard portions, like cups of coffee in a restaurant, where it is quite common to speak of *two coffees*.

These considerations show that virtually every count noun can be used as a mass noun and vice versa. We should therefore not classify nouns as count or mass, but instead view the count/mass distinction to be one of different ways of using nouns, or perhaps not even as a syntactic or morphological distinction, but as a semantic one. Intuitively, the fundamental difference between a mass noun like *apple-sauce* and a count noun like *apple* is that there is a clear notion of what is one apple, but there's no clear notion of what is one apple-sauce. In the words of Jespersen (1924): "There are many words which do not call up the idea of some definite thing with a certain shape or precise limits. I call these *mass words*". Or as Quine (1960) put it: "Inherent in the meaning of a count noun, like 'apple' is what counts as one apple and what as another... such terms possess built-in modes of dividing their reference, ... while mass nouns do not divide their reference." A common way of expressing this is that count nouns 'individuate' their reference, while mass nouns do not.

The non-individuating way of referring that is characteristic of mass nouns closely relates to the phenomenon that mass nouns can be used to refer to each of several objects as well as to the whole formed by these objects. For example, in a situation where there are several puddles of water on the floor, the term *water* in the sentence *Please mop up the water on the floor* may refer to the individual puddles as well as to the totality of all the water formed by the puddles. This phenomenon is known as 'cumulative reference': "Any sum of parts that are water is water" (Quine, 1960). Similarly, suppose one is served a bowl of rice; when one has eaten half of *the rice in the bowl* what remains would also be called *rice in the bowl*. In general, for a mass noun *M*, any part of something that is *M* is again *M*. This phenomenon is called 'distributive reference'. The term 'homogeneous reference' has been used both as synonymous with distributive reference and as denoting the combination of cumulative and distributive reference.

The property of distributive reference has been a matter of discussion among linguists and philosophers. Quine (1960) has rejected the idea that every part of something to which a mass noun may refer may also be referred to by the noun. He notes that "There are parts of water, sugar, furniture too small to count as water, sugar furniture" since the parts of an H₂O molecule are not water, the legs of a chair are not furniture, and the parts of a grain of sugar would perhaps not be called sugar. Instead, he posits the 'Minimal Parts Hypothesis', which says that for each mass noun *M* there is a specific minimal size that parts of its referent may have in order to count as *M*. There is not much support for this position however; semanticists generally agree that mass terms should be treated as referring homogeneously, in spite of the fact that their referents in the physical world may have minimal parts. In the standard formalization of count noun meanings, the intuition that count nouns individuate their reference is captured by construing the extension of a count noun as the set of all individuals that correspond to the built-in individuation of its reference. So *apple* refers to the set of all apples. Since a mass noun does not individuate its reference, this leads to the question what kind of things mass nouns denote. Many authors on mass terms believe that the answer to this question requires something else than sets. Quine (1960), Burge (1972), Moravcsik (1973), Ojeda (1993) and several others propose to make use of mereology, a theory of non-atomic part-whole structures that has been developed as an alternative to set theory (Lesniewski, 1929; Leonard and Goodman, 1940); Bunt (1979; 1985) proposes to use part-whole structures called 'ensembles', defined in an extension of standard set theory called 'ensemble theory' (see also Lewis, 1991); Parsons (1970) has proposed an altogether different notion of 'substances'.

The idea that mass terms do not individuate their reference explains why they cannot be combined with numerals: one wouldn't know what to count, so the numerical information would make no sense, and neither would quantifiers like *several*, *many*, *both*, that presuppose countability. As mass terms do not refer to well-delineated objects, it would also be strange to apply adjectives describing properties like shape, size or weight to mass nouns. This is

confirmed by the fact that it is strange to speak of **small wine*, **square water*, or **heavy sugar*. (And *heavy water* cannot be used to refer to water which is heavy, but only to refer to the substance deuterium oxide (D2O), formed by oxygen and the hydrogen isotope deuterium, of atomic weight 2.) It has therefore been suggested to make a distinction between count and mass adjectives, depending on whether the adjectives share with mass nouns the property of referring homogeneously (Bunt, 1980; cf. Moravcsik, 1973). For instance, *square* refers neither cumulatively nor distributively, since the whole formed by two square objects is in general not square, nor are the parts of a square object; *heavy* refers cumulatively but not distributively, and *small* refers distributively but not cumulatively.

Ter Meulen (1980) has suggested a count/mass distinction among verbs, depending on whether a verb refers distributively, that is, whether it denotes events that have sub-events which could be described using the same verb. Performance verbs, like *write* and *reach* would be count verbs, whereas activity verbs such as *travel* and *think* would be mass verbs. Syntactic phenomena that support this distinction are that it is strange to say that **Harry was reaching the airport for an hour* while it is fine to say that *Harry was travelling to the airport for an hour*, and, relating to the use of mass terms as direct objects, it is strange to say that **Alice was writing a poem for an hour* but there's nothing wrong with *Alice was writing poetry for an hour*.

It may be noted that some of the above observations on mass nouns do not really apply to *all* mass nouns. There is a subclass of mass nouns which do in fact individuate their reference; examples in English are *furniture*, *cattle*, *clothing*, *footwear*, *luggage*. Using Quine's terminology, inherent in the meaning of *furniture* is what counts as *one piece of furniture*. This may explain why such nouns can be modified by adjectives that do not refer homogeneously, as in *small furniture*, *heavy furniture*. Quantification is also different for these nouns than for other mass nouns. Whereas *All the water in this area is clean* says that every water part that you can take in this area is clean, *All the cattle in this area have been vaccinated* clearly applies only to animals, not to arbitrary cattle-parts. A related morphological phenomenon is that mass nouns of this kind in Dutch do not admit a diminutive form (see **meubilairtje*, **veetje*, **kledinkje*, **schoeiseltje*, **bagagetje*) – which makes sense, since in these cases the mass noun itself already denotes individuals of the kind that the diminutive form would otherwise denote. Whereas mass nouns in general have syntactically much in common with plural count nouns, the ones in this particular subclass moreover are semantically no different from count nouns. It has therefore been suggested they be assigned to a separate category, called 'collective' mass nouns (Bunt, 1985).

Mass terms are a challenge to the formal linguist, not only because it is difficult to pin down the count/mass distinction on morphological and syntactic grounds, but also because they call for a logical representation of the intuitions about non-individuating reference in such a way that sentences involving mass expressions are systematically assigned correct semantic interpretations through the application of a set of rules in a formal grammar.

Concerning the first part of this challenge, Pelletier and Schubert (1995) have argued that the count/mass distinction can be formalized in two ways: in terms of *occurrences* of nouns and as different *senses* of nouns. An occurrence approach characterizes the use of a noun syntactically as count or mass. An implementation in a formal grammar would typically assign features 'count' and 'mass' to occurrences of nouns and certain other expressions. By contrast, a sense approach considers all nouns to be just nouns, avoiding any 'count' or 'mass' labelling, and interprets a noun occurring in a certain syntactic ('mass' or 'count') context in its 'mass' or in its 'count' sense. Since virtually every noun can be used either way, the main virtue of an occurrence approach is not to assess the syntactic well-formedness of expressions, but rather to characterize the syntactic environments that force one interpretation of the noun or the other. For instance, characterizing the quantifier *much* as 'mass', we can force the occurrence of *apple* in *not too much apple* to be interpreted as apple-stuff, rather than individual apples. Some authors have suggested to make finer distinctions among nouns than count/mass (or count/mass/collective), depending on the noun's syntactic preferences for occurring in certain

syntactic environments. Based on Allan (1980)'s 'noun countability preferences', Bond et al., (1994) distinguish besides pure count and mass nouns also plural-only nouns (like *scissors*, *pants*), easily convertible count nouns (like *cake*, *stone*) and easily convertible mass nouns (like *beer*, *coffee*). This may be of practical use e.g. in machine translation (Baldwin and Bond, 2003) or in language learning (Nagata et al., 2005).

Concerning the second, semantic part of the challenge, Link (1983) and Landman (1991) have suggested that the models for a model-theoretic semantics of natural language should include a non-atomic Boolean algebra (or more specifically, a join semi-lattice) supporting part-whole structures without minimal parts as semantic interpretations of mass terms. Such a structured model can be used to assign logically adequate interpretations to sentences with mass terms, i.e. interpretations that have the desired logical properties (such as rendering *Water is water* necessarily true, and supporting inferences like *This puddle is water*, *Water is transparent*, therefore *This puddle is transparent*, but not supporting the inference *Water is scarce here*, *This is a puddle here*, therefore *This puddle is scarce here*).

The fundamental question what kind of thing a mass noun denotes is not answered in these formal approaches, other than that mass terms have denotations figuring in a non-atomic part-whole structure (which would be formally correct only for noncollective mass terms). The use of mereological wholes for these denotations, which many language philosophers have embraced, does provide non-atomic part-whole structures, but has the drawback that mereology is an alternative to set theory, and that mereological concepts as such do not fit in set-theoretical frameworks. In this respect the use of ensemble theory, which formalizes both atomic, non-atomic, and partly atomic part-whole structures within an extension of standard set theory, offers better possibilities for an elegant, integrated treatment of the semantics of mass expressions (cf. Lewis, 1995).

Bibliography

Allan, K. (1980) Nouns and countability. *Language* 56(3), 541--567.

Allan, K. (2001) *Natural language semantics*. Oxford, UK: Blackwell.

Baldwin, T. and F. Bond (2003) Learning the countability of English nouns from corpus data. In *Proc. of the 41st Annual Meeting of the Association for Computational Linguistics*, Sapporo, Japan. 463--470.

Bond, F., K. Ogura and S. Ikehara (1994) Countability and number in Japanese-to-English machine translation. In *Proc. of the 15th International Conference on Computational Linguistics (COLING'94)*, Kyoto, Japan. 32--28.

Bunt, H. (1976) The formal semantics of mass terms. In F. Karlsson (ed.) *Papers from the 3rd Scandinavian Conference of Linguistics*, Helsinki, Finland. 71--82.

Bunt, H. (1979) Ensembles and the formal semantic properties of mass terms. In F.J. Pelletier (1979), 249--277.

Bunt, H. (1980) On the why, the how, and the whether of a count-mass distinction among adjectives. In J. Groenendijk and M. Stokhof (eds.) *Formal methods in the study of language*. Amsterdam: Mathematical Centre. 51--77.

- Bunt, H. (1985) *Mass terms and model-theoretic semantics*. Cambridge, UK: Cambridge University Press.
- Burge, T. (1972) Truth and mass terms. *Journal of Philosophy* 69 (10), 263—282.
- Chierchia, G. (1983) On plural and mass nominals and the structure of the world. In T. Borowski and D. Finer (eds.) *University of Massachusetts Occasional Papers VIII*. Amherst: GLSA.
- Jespersen, O. (1924) *The philosophy of grammar*. London: Allen and Unwin.
- Landman, F. (1991) *Semantic structures*. Dordrecht: Kluwer Academic Publishers.
- Leonard, H. and Goodman, N. (1940) The calculus of individuals and its uses. *Journal of Symbolic Logic* 5, 45--55.
- Lesniewski, S. (1929) Grundzüge eines neuen Systems der Grundlagen der Mathematik. *Fundamenta Mathematicae* 14, 1--81.
- Lewis, D. (1991) *Parts of classes*. Oxford: Blackwell.
- Link, G. (1983) The logical analysis of plurals and mass terms: a lattice theoretic approach. In R. Bäuerle, C. Schwarze and A. von Stechow (eds.) *Meaning, use and interpretation of language*. Berlin: De Gruyter. 303-323.
- Lønning, J.-T. (1987) Mass terms and quantification. *Language and Philosophy* 10(1): 1--52.
- McCawley, J. (1981) *Everything the linguist always wanted to know about logic*. Chicago: University of Chicago Press.
- Moravcsik, J. (1973) Mass terms in English. In J. Hintikka, J.M.E.Moravcsik and P. Suppes (eds.) *Approaches to natural language*. Dordrecht: Reidel, 301—288.
- Nagata, R., F. Masui, A.Kawai and N. Isu (2004) An unsupervised method for distinguishing mass and count nouns in context. Forthcoming in *Proc. of the Sixth International Workshop on Computational Semantics (IWCS-6)*, Tilburg, The Netherlands.
- Ojeda, N. (1993) *Linguistic individuals*. Menlo Park: Center for the Study of Language and Information.
- Parsons, T. (1970) An analysis of mass terms and amount terms. *Foundations of Language* 6: 363--388.
- Pelletier, F.J. (ed.) (1979) *Mass terms: some philosophical problems*. Dordrecht: Reidel.
- Pelletier, F.J. and Schubert, L.K. (1995) Mass expressions. In: D. Gabbay and F.Guenther (eds.) *Handbook of Philosophical Logic*. Dordrecht: Reidel. 327--407.
- Quine, W.V.O. (1960) *Word and object*. Cambridge, MA: MIT Press.
- Sharvy, R. (1978) Maybe English has no count nouns: Notes on Chinese semantics. *Studies in Language* 2, 345--365.
- Ter Meulen, A. (1980) *Substances, quantities and individuals*. Ph.D. dissertation, Stanford University.

Whorf, B. (1939) The relation of habitual thought and behaviour to language. In J. Carroll (ed.) *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press. 134--159.

List of keywords

collective noun, count-mass distinction, count adjective, count noun, countability, cumulative reference, diminutive, distributive reference, ensemble theory, homogeneous reference, individuation, mass adjective, mass expression, mass noun, mereology, minimal parts hypothesis, quantification

Cross references:

article, determiner, diminutive, noun, numeral, quantification, quantifier, reference

Author Bio:

Harry Bunt (1944) is professor of linguistics and computer science at Tilburg University (Tilburg, the Netherlands), since 1983. He graduated from the University of Utrecht in physics and mathematics and obtained a Ph.D. in linguistics at the University of Amsterdam (1981) with a dissertation on the semantics of mass terms. From 1970 to 1983 he performed research in Artificial Intelligence and language understanding systems at Philips Research Laboratories in Eindhoven. His recent research interests and publications are focused on computational semantics and pragmatics, dialogue theory, knowledge representation, and multimodal human-computer interaction.