Degrees of Countability: A Mereotopological Approach to the Mass/Count Distinction

Theories of nominal semantics traditionally focus on two grammatical categories of countability, mass and count; however, many language families (e.g. Nilo-Saharan, Celtic, Slavic) morphologically recognize three countability categories. Welsh, for instance, not only has substance nouns which are "mass" (*dwr*, 'water'), and nouns with a singular-plural contrast (*llyfr*/*llyfr-au*, 'book/books'), but also has nouns of an intermediate category with a collective-singulative contrast (*tywod*/*tywod-yn*, 'sand/grain.of.sand'; *picwn*/*picwn-en*, 'wasps/a wasp'). I refer to these noun types as SUBSTANCE, COUNT and AGGREGATE, respectively. A clear semantic generalization underpins aggregates across a variety of languages: the referents of these nouns canonically "come together" in some fashion, in contrast to count noun referents, which canonically appear as individual entities. Welsh and the other systems examined here imply that semantic countability is, contrary to standard assumptions, not a binary contrast between mass and count, but rather scalar. Welsh and others divide the scale into three classes, whereby aggregates are morphologically recognized in a distinct fashion from typical countable or uncountable nouns. English divides nouns into two classes, aggregates being split between them. Standard theories, based purely on a part relation over elements in a model, face difficulties in accounting for the broader cross-linguistic data. This paper demonstrates that the data instead calls for enriching part structures with connection relations. This approach delivers a semantic basis for a scalar approach to countability while avoiding several recalcitrant problems in standard theories.

The distinction between count and mass terms is commonly related to the properties of being *atomic* (1) or *divisive* (2), respectively, defined over a part structure. Assimilating aggregates to one predicate type or the other makes a variety of wrong predictions. If aggregates were simply atomic, pluralization would be expected to be possible, as with regular count predicates; however, direct pluralization of aggregate terms only results in a "different kinds of" reading (cf. *wines*). Instead, pluralization of an aggregate requires the singulative form: *grawn*/*gron-yn*/*gron-ynn-au* 'grain'/'a single grain'/'grains'. Analyzing aggregates instead as divisive would not only be false for, e.g., *wasp*, but also predict the singulative could then apply to other divisive predicates, namely substances (e.g. *water*), contrary to fact. Further, divisiveness is itself problematic for many nouns that are uncountable in English, such as *sand* or *furniture*. Such nouns have clear minimal parts and, while grammatically uncountable, are logically countable: *Mary counted the sand/furniture/*water.* It is striking that all equivalents of such nouns in Welsh fall into the aggregate class.

(1) Atomic(x) relative to P = $P(x) \rightarrow \neg\exists y[y < x \wedge P(y)]$

(2) Divisive(P) = $[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$

Recent work in philosophy has shown that standard mereology can be profitably extended with connectedness relations, resulting in "mereotopology" (Smith 1996, Casati and Varzi 1999 *inter alia*). The basic connectedness relation C holds when two individuals (in the mereological sense) touch at least on their boundaries. This relation interacts with the pure mereological relations *overlap*, O, and *part*, ≤: if two individuals overlap and/or one is part of the other, it implies that they are connected. One fundamental motivation for the mereotopological approach is to distinguish between individuals forming integrated wholes and those forming only arbitrary collections. I relate count nouns to integrated wholes in a spirit similar to Moltmann (1997), but adopt the stricter notion of Maximally Strongly Self-Connected (MSSC) relative to a property (Casati and Varzi 1999). An individual satisfies MSSC relative to a property if every (interior) part of the individual is connected to the whole and anything else which has the same property and overlaps it is once again part of it. This guarantees that an integral whole will both be unique and not overlap with any other individual with the same property, although it may of course touch distinct individuals of the same type. Turning to connection relations, they may come in a variety of strengths. The two primary

types, stated in (3), are Strongly Connected, two individuals are connected via overlapping, and Externally Connected, two individuals are not connected by overlapping but by touching.

(3) (a.) $\text{StrongC}(x, y) = \text{O}(x, y)$ (b.) $\text{ExtC}(x, y) = \text{C}(x, y) \wedge \neg\text{O}(x, y)$

Different semantic classes of nominal predicates can be distinguished through which connectedness relations may or must hold among the individuals in their denotation, which I formalize as conditions on allowable covers over the domain of reference of a predicate. I will illustrate with the three core types. Let R be a realization relation holding between individuals and a kind/concept (Krifka 1995) and let $\{C \in C\}$ be a set of covers over the domain, with covers composed only of individuals with a property P abbreviated as C_p. (4) asserts that if an individual realizes a count predicate (*dog*), then there exists a cover containing that individual which is composed of MSSC individuals. This prevents strongly connected (overlapping) individuals from being in the extension of a count predicate. In contrast, substance predicates require their extension to be comprised of individuals which are strongly connected to other individuals of the same substance, as given in (5). This is satisfied, for instance, by a section of a pool of water—it overlaps other sections of the pool, which are again water. Also, since connectedness is implied by parthood, the whole pool is strongly connected to its parts which are water. Aggregates are a hybrid of the first two categories. Granular aggregates, defined in (6), have extensions which include both MSSC individuals and individuals which are externally connected to (viz. touching) other individuals of the same type. *Sand*, for example, is true of single grains (MSSC individuals), clumps of sand, which can be divided into multiple externally connected individuals, or combinations of the two. This analysis brings out the similarity between substances and granular aggregates, namely their referential domains include clusters of tightly connected individuals, but also shows how they differ, both in the type of connection and in granular aggregates' inclusion of natural minimal parts.

(4) $R(x, P_{\text{Count}}) \rightarrow \exists C_p[x \leq C_p \wedge \forall y[y \in C_p \rightarrow \text{MSSC}(y)]]$

(5) $R(x, P_{\text{Substance}}) \rightarrow \exists C_p[x \leq C_p \wedge \forall y[y \in C_p \rightarrow \exists y'[y' \in C_p \wedge y' \neq y \wedge \text{StrongC}(y, y')]]]$

(6) $R(x, P_{\text{Gran.Agg}}) \rightarrow \exists C_p[x \leq C_p \wedge \forall y[y \in C_p \rightarrow \exists y'[y' \in C_p \wedge y' \neq y \wedge \text{ExtC}(y, y')] \vee \text{MSSC}(y)]]$

Other types of aggregates parallel (6) but with different connection relations, such as Proximately Connected, which holds when two entities are co-located and near one another, appropriate for collective aggregates such as insects or berries. Ordering predicate types by the strength of the connection relation then generates a scale of individuation, which languages divide up into grammatical categories of countability: substance < granular aggregate < collective aggregate < count.

This proposal side-steps recalcitrant problems for traditional accounts. First, many substance nouns do not lend themselves to be infinitely divisible into the same type of stuff. *Soup*, for instance, may contain meatballs that, while part of the soup, are not in themselves soup, contrary to (2). The characterization of substance terms given in (5), however, accords with this scenario: any individual portion of a soup which in itself qualifies as soup will be strongly connected to another such individual. Second, defining count nouns through atomicity (1) stumbles with nouns such as *fence*, since a part of a fence can again be thought of a fence (Rothstein 2010). Defining count nouns in terms of MSSC entities does not face this problem, but rather guarantees that an utterance of *a fence* makes reference to the largest connected entity which satisfies being a fence, intuitively the correct result.

References: Casati, R. and Varzi, A. (1999) *Parts and Places*. MIT Press. Krifka, M. (1995) "Common Nouns: A Contrastive Analysis of English and Chinese", in *The Generic Book*, Chicago University Press. Moltmann, F. (1997) *Parts and Wholes in Semantics*. OUP. Rothstein, S. (2010) "Counting and the Mass/Count Distinction", *JoS* 27. Smith, B. (1996) "Mereotopology: A Theory of Parts and Boundaries," *Data and Knowledge Engineering* 20.