NUMBER AND INDIVIDUATION

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Scott Grimm
August 2012

Abstract

This dissertation investigates the semantic foundations of nominal countability. Standard accounts are typically concerned with a binary distinction between countable words (*dog*/ *dogs*) and non-countable words (*water*). This dissertation examines this issue from the perspective of languages with richer grammatical number systems. I develop a typological generalization that countability is a scalar phenomenon and propose new techniques to formally model these facets of nominal semantics by augmenting standard mereological accounts with topological relations.

Languages such as Welsh or Maltese grammatically recognize what I call aggregate nouns—nouns which designate entities that habitually come together, such as insects (*ants*) or granular substances (*sand*). These nouns are grammatically distinct from both non-countable nouns and countable nouns with a singular/plural contrast, instead they display a collective/singulative contrast. These grammatical number systems vividly demonstrate how a binary countable/non-countable distinction oversimplifies the typological space. I argue from the data from Welsh and Maltese, and even more complex fieldwork data from the Gur language Dagaare, that countability is a scalar phenomenon.

I propose that the morphosyntactic organization of grammatical number systems reflects the semantic organization of noun types according to the degree of individuation of their referents. Nouns of different types are individuated to different degrees and can accordingly be ordered along a scale of individuation: substances < granular aggregates < collectives < individual entities. Noun types which are less individuated are on the lower end of the scale and are cross-linguistically less likely to signal grammatical number, while the converse holds for highly individuated noun types. Understanding morphosyntactic number categories in light of a scale of individuation avoids the difficulties binary accounts

face, since languages may divide up the scale of individuation into any number of classes and at different points. For instance, languages with a collective/singulative recognize a grammatical number category corresponding to the middle region of the scale. At the same time, the proposal provides a predictive framework for how grammatical number systems are organized: the contrasts being made are common across languages, and, as a corollary, the endpoints of the scale (substances and individual entities) are predicted to be stable across languages. I show that this view of countability also answers many of the standard criticisms of accounts where a noun's meaning determines its grammatical behavior with respect to number marking.

I explore the implications of this broader typological view for formal semantic treatments of countability. Standard mereological accounts turn out to be not sufficiently expressive to model the aggregate nouns nor the grammatical number systems which distinguish them. I enrich the standard mereology framework with topological connection relations, resulting in the more expressive "mereotopology". Through using different connection relations, this framework is able to represent aggregate nouns and the ways in which entities may come together. Consequently, this framework is able to deliver analyses of particular grammatical number systems, such as Welsh. In addition, this more expressive framework resolves several recalcitrant problems noted for many treatments of countability, such as the "minimal parts" problem discussed in relation to nouns such as *sand* or *furniture* which, while non-countable, still have minimal pieces.

Acknowledgements

This dissertation has benefitted from a large number of people who have advised me, provided insights and criticisms, and who have generally helped me to develop as a linguist. First and foremost, I would like to thank my dissertation advisor Beth Levin. She has given me guidance in matters large and small throughout my time at Stanford—practically every aspect of my graduate career has benefitted from her advice. She has also been an ideal advisor for me. She has a rare ability to make the right suggestion at the right time, simultaneously leaving much freedom in how to follow up on the suggestion. I have learned an enormous amount working with her. For this, and much more, I am grateful.

The other members of my committee have had a similarly large influence on my development. Paul Kiparsky's breadth and depth of knowledge has been a source of awe for me throughout my graduate career. Many of the most thrilling intellectual moments of my graduate studies occurred in his classes and in meetings with him. In the context of this dissertation, he has been particularly helpful in both drawing my attention to many portions of my thinking which needed clarity as well as to encouraging me to always relate my proposals to the big questions. Chris Potts has provided me with extremely useful counsel during the course of this dissertation, and also advised me on projects concerning plurality that led up to the dissertation. I have consistently enjoyed being exposed to his infectious enthusiasm for innovative approaches in linguistics and have learned much about semantics, among many other things, through our interactions. I met Donka Farkas through various Bay Area semantics events and it has been a true privilege to have her as part of my committee. I am grateful to her for all the insights she has shared with me about plurals, markedness and semantic theory during our meetings on the idyllic UC Santa Cruz campus. I would also like to thank Tom Wasow for his participation in my dissertation

defense and for much valuable advice throughout my time at Stanford.

This dissertation grew out of my fascination with the Dagaare inverse number marking system. Arto Anttila and Adams Bodomo introduced me to the problem during my first year at Stanford and have provided encouragement for my work on it ever since. Thanks to Adams Bodomo, I got into contact with Mark Ali at the University of Education in Winneba, Ghana. Mark Ali devoted an enormous amount of time to working with me on Dagaare during two trips to Ghana—simply put, this dissertation would not have come into being without him. I can never thank him enough for his generosity, insight, and friendship.

As my dissertation research progressed, I gained much from conversations with Dave Barner, Jade Comfort, Cleo Condoravdi, Hana Filip, Jeff Good, Peggy Li, John Lucy, Frank Seifart, and Henriëtte de Swart. I would also like to thank the members of the "Mass-Count Collective": David Clausen, Alex Djalali, Sven Lauer, and Tania Rojas-Esponda. Our weekly meetings during Spring Quarter of 2010 shaped many aspects of my thinking which became foundational for this dissertation. During the final phase, Nadine Borchardt, Dan Lassiter, Tania Rojas-Esponda, and Middy Tice provided proof-reading and many valuable comments.

Part of the research for this dissertation was carried out during stays in Berlin, Germany. My first stay was as a guest researcher at Zentrum für Allgemeine Sprachwissenschaft (ZAS). I would like to thank Manfred Krifka and Barbara Stiebels for arranging for that stay, as well as providing me with office space on a subsequent visit, as well as to thank my ZAS colleagues for such a stimulating environment. I also would like to thank Tom Güldemann for supporting me during a second stay in Berlin as a Deutscher Akademischer Austauschdienst (DAAD) scholar at Humboldt University. During these stays, I had the pleasure of engaging with the vibrant linguistics community in Berlin. In particular, I would like to thank Laura Downing, Ines Fielder, Gudrun Miehe, Anne Schwarz, Kerstin Schwabe, Stephanie Solt, Uli Sauerland, Frank Seifart, Tonjes Veenstra, and Malte Zimmermann.

A large part of the richness of my graduate experience has been of course due to my fellow students. I would like to thank my cohort with whom I have spent a most enjoyable six years: Matthew Adams, Uriel Cohen Priva, Marie-Catherine de Marneffe, Jason Grafmiller, and Tyler Schnoebelen. Marie-Catherine deserves special mention: office-mate, frequent

co-author, and friend, she has also single-handedly prevented my French from falling into total ruin during the last six years. Merci! During my time at Stanford, I have overlapped with many other students who have provided me with stimulating conversations about linguistics as well as good times. Inbal Arnon, Lucas Champollion, Liz Coppock, Alex Djalali, Itamar Francez, Roey Gafter, Lauren Hall-Lew, Philip Hofmeister, Andrew Koontz-Garboden, Tatiana Nikitina, Lis Norcliffe, Nola Stephens, and Hal Tily—thank you all.

I gratefully acknowledge the financial support I have received for the research undertaken in this dissertation: a National Science Foundation Dissertation Improvement Grant (Grant No. 1023289, P.I. Beth Levin), a Stanford Center for African Studies Summer Fellowship, a Graduate Research Opportunity Grant from Stanford University and a Deutscher Akademischer Austauschdienst (DAAD) Visiting Scholar Grant.

Finally, I would like to thank my family who has provided me with a constant source of love and all manners of support during the various stages of my life. I am especially grateful to my parents who have always encouraged me in my interests and choices in life. This encouragement has been unwavering, even when it has meant that they have had to see me less, visit me in far-away places, or store an unfortunately large number of boxes of books. Last, but not least, I am thankful to Nadine Borchardt, my partner, who has helped me in uncountable ways while writing this dissertation. Above all, I am grateful for her love and the opportunity for us to travel through life together.

Glossing Conventions

3	3rd person
3DIM	3-dimensional
COP	copula
DEF	definite article
DEM	demonstrative
DETPL	determinate plural
DISTPL	distributive plural
HUM	human prefix
M	masculine
NHUM	non-human prefix
PART	particle
PL	plural
PRES	present
PROX	proximate
QUANT	quantifier
REDUPL	reduplication
SG	singular
SING	singulative
SCM	specific class marker

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This work is about the countability of nouns. Some nouns in English regularly use grammatical means to enumerate their referent—the noun *cat* regularly combines with the plural morpheme *-s* to designate multiple cats. Other nouns do not allow this—the noun *lightning* resists pluralization, and pluralized forms of this noun are typically found to be awkward (**lightnings*). Pluralization is one grammatical operation, among many, which bears witness to divergences in different nouns' allowances for counting. Determining the actual source of these differences among nouns has been an active area of research for decades and is the primary task this dissertation engages in.

There are several elements any theory of countability must confront: nouns, their interpretations, their referents, and the morphosyntactic means of encoding number distinctions associated with them. In the chapters to come, I examine these elements both on their own terms and in terms of the interrelations among them. The central problem is to determine the locus of countability. Does it solely concern language or does it concern the referents designated by the nouns? Perhaps it concerns neither directly, but is related to a third phenomenon. Each of these positions has been asserted by previous researchers.

Modern work in countability finds its starting point in the discussion of Jespersen (1913). He establishes a distinction between COUNT nouns, a class of nouns which allow pluralization (and other grammatical means implicating number), and a class of nouns he names MASS. Jespersen defines this class of words in terms of how speakers conceive of the noun's referent:

> "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': they may be either material, in which case they denote some substance in itself independent of form, such as 'silver', 'quicksilver', 'water', 'butter', 'gas', 'air', etc., or else immaterial, such as 'leisure', 'music', 'traffic', 'success', 'tact', 'commonsense', and ,'satisfaction', 'admiration', 'refinement', from verbs, or ... 'restlessness', 'justice', 'safety', 'constancy', from adjectives." (Jespersen, 1924, p. 198)

As will become clear, the proposal elaborated in the following chapters takes much of Jespersen's original intuition to be insightful. At the same time, much progress has been made in our understanding of countability, from the reporting of different grammatical systems to psycholinguistic experiments which bear upon these issues. Other proposals have been put forth as well: some have hypothesized that this linguistic distinction is grounded in foundational cognitive capacity, others have claimed that the distinction is epiphenomenal, and there have been a range of proposals across the spectrum spanning these.

The rest of this chapter sets the stage for an investigation of countability. I first present the various distributional properties used to differentiate nouns in terms of countability. I subsequently provide an overview of the primary positions that have been argued for to account for countability. With these foundations in place, I then outline what is to come in the remainder of this work.

1.1 Distributional Properties of Countable and Non-Countable Nouns

A wide range of distributional properties of nouns have been taken to reflect differences between countable and non-countable nouns. The permissible combinations of nouns and determiners or quantifiers have been taken as central, and I will refer to them as the CORE DISTRIBUTIONAL PROPERTIES. These core properties concern the distribution of elements in the noun phrase and appear to track whether the head noun permits direct counting, e.g. through cardinal numbers. Other distributional facts have been pointed out that may involve not

only grammatical acceptability, but may also indicate whether individuals, a term used informally here, are part of the noun's meaning. Thus, even at the initial stages of outlining what falls under the study of countability, one must reckon with grammatical and semantic facets of the problem.

As pointed out by Behrens (1995), the mode of presentation of different countability phenomena often reflects the adoption of a particular perspective on countability. Some researchers characterize types of lexical elements in terms of their distribution, viz. 'a mass noun may be modified by the quantifier *much*'. Others take the grammatical context as basic, deriving the noun's countability status from it, viz. 'nouns following *much* are mass nouns'. This seemingly innocent shift in perspective underlies important differences among the approaches: according to the first approach, it is the semantic nature of the lexical unit which constrains what it may co-occur with, while according to the second approach, the grammatical context constrains what the noun's interpretation is. Although often seen as pitted against one another, these are not actually incompatible views, and the approach developed in the later chapters attempts a reconciliation of the two approaches. In presenting the different distributional properties, I will try to remain neutral for the moment and simply present them as distributional facts.

At this point, it is useful to fix some terminology which will be used throughout this work. The distinctions that have just been discussed are normally described as falling under the "count-mass" distinction. Unfortunately, the various ways these terms have been used has rendered them confusing. In particular, "mass" has acquired a range of uses different from Jespersen's original usage. For some "mass" is been taken to designate a grammatical category (Bloomfield, 1933), but for others, it has a narrower range and aligns only with certain types of entities in the world, namely substances. I will instead describe nouns as COUNTABLE or NON-COUNTABLE. Countable nouns are able to combine with affixes, determiners, or quantifiers which indicate number, whereas non-countable nouns are those that can not. These two terms, then, are restricted to designating nominal behavior in terms of morphosyntactic characteristics, and make no reference to whatever semantic characteristics nouns may possess.

Morphosyntactic Traits	Singulars	Plurals	Non-countable
pluralization	✓	N/A	*
quantifiers implicating plurality (*many*, *several*)	*	✓	*
cardinal modification (*two*)	*	✓	*
much, *little*	*	*	✓
combine with measure terms (*two kilos/meters of*)	*	✓	✓

Table 1.1: Core Distributional Properties of Nominal Countability

1.1.1 Core Distributional Properties

The primary distributional characteristics in English taken as reflecting a countable/non-countable distinction are that countable nouns combine readily with numerals, determiners implicating cardinality, and the indefinite article *a*, and they accept pluralization. Non-countable nouns do not permit these combinations, nor do they have plural forms. Rather, there are several quantifiers which combine only with non-countable nouns, including *much* and *a little*. Two further distributional properties distinguish the singular form of countable nouns from non-countable nouns: non-countable nouns may appear bare and may combine with measure terms such as *three kilos/meters of*. These distributional criteria are summarized in table 1.1.

The picture of countability presented in this table leads to the false hope that one could sort the lexicon into nouns that are countable and nouns that are non-countable thanks to these distributional properties. The situation is, however, much more complicated. As will be discussed shortly in section 1.1.3, not all nouns which could be classified as countable or non-countable by *some* distributional properties behave identically with respect to *all* of the distributional properties. Another source of complication is that nouns often have several uses, some countable and some not. For example, as shown in (1), *wine* may designate a particular liquid, in which case it combines with quantifiers such as *much* which align with non-countable nouns, or it may designate servings of the liquid, in which case it combines with cardinals, *many*, and other determiners and quantifiers which align with countable nouns. Similarly, nouns which according to table 1.1 would appear to be non-countable, when designating a particular kind (i.e. taxonomic sub-kind) of that entity, may still be found with the indefinite article or in the plural, as shown in (2).

(1) a. much wine
b. three wines

(2) A wine from the Russian River is usually a good value.

I will discuss in detail in section 3.6.3 the issue of "nominal flexibility", where nouns may adopt different countability statuses depending on the surrounding context. For the moment, I refer to nouns such as *water*, which align with the distributional properties for non-countable nouns in their typical use, as non-countable nouns, even though they may have other uses.

1.1.2 Additional Properties

While the behavior of nouns with determiners and quantifiers is the basis for the core distributional properties for countability, many additional properties have been proposed, such as the behavior of nouns in relation to adjectival modification or comparative constructions. These additional properties, in turn, give rise to a classification which differs from that gained through the core distributional properties. This difference already foreshadows a point that will emerge several times in this study: considering the countable/non-countable distinction to be binary is an over-simplification of countability phenomena.

An additional reason to outline these different properties is their importance for examining languages beyond English: other languages may not have determiners, quantifiers, or pluralization of exactly the same kind as found in many European languages, but different countability classes may still be recovered through these alternate distributional properties. I now briefly detail these properties.

Bases for Comparison The comparative construction has been observed to differentiate countable and non-countable nouns in that countable and non-countable nouns in comparative constructions show different dimensions of comparison, as demonstrated by (3) from Bale & Barner (2009, p. 226). In (3-a), the basis for comparison is the number of individual objects, while in (3-b), the basis for comparison is volume.

(3) a. Esme has more cups/plates/candles than Seymour.

b. Esme has more butter/water/toothpaste than Seymour.

The comparison context also aligns nouns such as *furniture* with countable nouns in that individuals provide a basis for comparison, as shown in (4).

(4) Esme has more furniture/cutlery than Seymour.

Here, as with all the distributional properties to be discussed in this section excepting reciprocal resolution, this difference among nouns relates to the accessibility of individuals in the meaning of a given noun. In other words, comparison in terms of number of individual objects is available for nouns such as *furniture* because, presumably, *furniture* describes (a set of) individual objects, i.e. tables, chairs, etc., in contrast to, e.g. *oil*. This differs from the categorization results from section 1.1.1—*furniture* may make individuals in some way accessible, but that is not sufficient to permit it to combine with *each* or *two* (**each furniture/*two furniture(s)*).

Since two different types of nouns will figure prominently in the next sections, it is useful to fix some terminology for them. I will refer to nouns which (speaking informally) designate collections of particles or small pieces of matter, such as *rice* or *sand*, simply as AGGREGATES. I will refer to nouns such as *furniture* and *cutlery* as ARTIFACTUAL AGGREGATES.

Modification Adjectival modification patterns provide yet another perspective on different countability statuses. Quine (1960, p. 104) provides the first discussion of this, noting the oddity of adjectives which designate shape as modifiers of non-countable nouns, such as in (5-a). Building upon Quine's discussion, Bunt (1979, 1985) provides a detailed discussion of the acceptability patterns of adjectives of shape (*square, spherical, round*) and size (*large, small, tiny, huge*) with non-countable nouns.

(5) a. ??square water
b. ??huge wine

Yet, not all nouns classified as non-countable by the distributional properties concerning determiners and quantifiers from section 1.1.1 exclude modification by these adjectives. McCawley (1975, p. 319), in his discussion of *furniture*-nouns, notes that size adjectives

are however felicitous with nouns such as *furniture* or *crockery*. The interaction between non-countable nouns and size and shape adjectives is discussed at length in Schwarzschild (to appear). There several examples of artifactual aggregates that are given, some of which are repeated in (6). These nouns do not accept pluralization, but are able to be modified by shape and size adjectives, as given in (6).

(6) a. The furniture in that nightclub is round. (=individual pieces)
b. The mail in that bin is square and small. (=individual pieces)
c. The luggage she brought was big. (=individual pieces)

These observations can be extended to other non-countable nouns, namely aggregates.

(7) a. The rice was spherical/small.
b. The grass was long/large.

One might suppose that any non-countable noun permitting an interpretation in terms of discrete parts would also permit such modifiers. Yet, Bunt (1985) observes, giving the examples in (8), that non-countable nouns that are often individuated via conventional portioning, viz. links of sausage, still do not readily combine with size adjectives.

(8) Bunt (1985, p. 207) ??You have heavy sausage on your plate.

In the other direction, some non-countable nouns permit modification by shape adjectives if the context is such that discrete elements are made available. Bunt (1985, p. 208) discusses the example in (9), which relies on interpreting *sugar* as *lumps of sugar*. (This interpretation is no doubt facilitated by the fact that servings of sugar are standardly portioned out in cubes, although it seems to me that with sufficient contextual support *sausage* could be used in the same way as *sugar* in (9).)

(9) The blue boxes are filled with cubic sugar, the red boxes with rectangular sugar, and the white boxes with a mixture of cubic and rectangular sugar.

The upshot of these observations is that size and shape adjectives provide another test to distinguish countable from non-countable nouns, where some clear contrasts present themselves (??*square water* vs. *square book*). At the same time, the class of nouns which accept such adjectives overlaps, but does not coincide with the nouns classified as countable by the distributional properties of section 1.1.1. Some nouns which robustly resist pluralization (*furniture*, *rice*) accept modification by size and shape adjectives equally robustly. There is some further variation for nouns designating substances—they may be modified by shape adjectives given sufficient contextual support as is the case for *cubic sugar*.

Distributivity Countable nouns can also be distinguished by whether they support distributive elements and/or distributive interpretations. Gil (1996), for instance, clearly relates distributivity to the countable/non-countable distinction.[1] Gil (1996, p. 57) makes this point with examples such as those given in (10), where the quantifier *each* does not distribute over the non-countable and singular nouns, but does over the plural form.

(10) a. *Rice is fifty cents each.
b. *An apple is fifty cents each.
c. Apples are fifty cents each.

This distribution is similar to those of section 1.1.1 in that it revolves around whether the noun in question licenses a quantifier. Yet, some nouns may serve as a distributive share for *each* though they do not directly combine with *each* to form a well-formed NP.

(11) a. The cutlery is fifty cents each.
b. *Each cutlery is fifty cents.

A related property is whether a noun supports distributive interpretations for predicates with which it combines. Predicates such as *be heavy* are known to be systematically ambiguous between collective and distributive interpretations, as shown in (12).

[1] See also Gil (1987) and references therein for discussion of distributive numerals, such as *three each*.

(12) These spoons are heavy.
→ The spoons (taken together) are heavy. (✓ collective)
→ The spoons (taken individually) are heavy. (✓ distributive)

Nouns which designate substances, for instance, do not show this ambiguity, but only license the collective interpretation, as shown in (13). This is taken as evidence that there are no individuals in the denotation of these nouns which may license the distributive interpretation.

(13) The concrete is heavy.
→ The concrete (taken together) is heavy. (✓ collective)
↛ The concrete (taken individually) is heavy. (# distributive)

As discussed in Bale & Barner (2009), some nouns which are non-countable by the distributional properties in section 1.1.1 pattern like countable nouns with respect to this property. Again this is the set of artifactual aggregate nouns such as *furniture* or *cutlery*, as shown in (14).[2]

(14) This cutlery is heavy.
→ The cutlery (taken together) is heavy. (✓ collective)
→ The cutlery (taken piece by piece) is heavy. (✓ distributive)

This property, which distinguishes whether individuals can be detected in the denotation of a noun, differs from the core distributional properties in 1.1.1. As opposed to examining the permitted combinatorics of nouns and determiners, the distribution is related to whether a noun makes individuals accessible.

[2]Bale & Barner (2009) actually provide a larger discussion of this property, as well as nouns as the argument of *to count* (which will be discussed shortly), arguing that these properties are unreliable due to possible contextual variation. Instead, they argue, only examining nouns in the comparative construction (discussed previously in this section) is reliable. That one must carefully control for contextual effects is a point well taken, although, in work subsequent to Bale & Barner (2009), Grimm & Levin (2011) have shown that behavior in comparative constructions is also dependent on context. Thus, regardless of the particular construction being examined, one must carefully control for context.

Resolution of Reciprocals A property similar to the interpretation of predicates as collective or distributive is the resolution of reciprocals. Gillon (1992, p. 629) provides examples, given in (15), in which reciprocal resolution is also sensitive to the countable/non-countable distinction. In (15-a), two readings are possible: (i) the drapes resemble each other and also the carpets resemble each other or (ii) the drapes resemble the carpets. In (15-b), there is only one reading: the drapery resembles the carpeting.

(15) a. The drapes and the carpets resemble each other.
b. The drapery and the carpeting resemble each other.

In contrast to the distributions related to distributivity, the resolution of reciprocals does not appear to be successful with nouns such as *furniture*, as shown in (16).

(16) a. *The cutlery/furniture/rice resemble each other.
b. The pieces of cutlery/pieces of furniture/grains of rice resemble each other.

Argument of the Verb *to count* A very intuitive test for whether a noun or NP is countable is whether a noun is acceptable as the object of the verb *to count*. This test yields different results than the distributional properties in the last section involving determiners and quantifiers. Whether a noun is acceptable in this frame rests on whether the noun is what I will term LOGICALLY COUNTABLE. Whether what a noun designates is logically countable is determined by whether there is some way to match up the elements of the referent with the set of natural numbers. Nouns which are countable by the core distributional properties, e.g. *dog*, have this property. Yet, the range of nouns which are acceptable with *to count* includes many nouns which are not countable by the core distributional properties, such as *furniture*. Non-countable nouns such as liquids, however, are not acceptable. These contrasts are shown in (17).[3]

(17) a. John counted dogs.
b. (?) John counted furniture/traffic/rice/sand.

[3]I have marked the example in (17-b) with a question mark in parentheses since some speakers find these examples awkward; however, these same speakers still recognize a clear contrast between (17-b) and (17-c), which are impossible rather than just awkward.

c. John counted *water/*oil.

1.1.3 Degrees of Countability and Countability Preferences

The last two sections have presented a number of distributional properties of nouns which have been taken as diagnostics of different types of "countability". The work of Allan (1980) has in fact already shown that some of the standard distributional properties deliver mixed results which are dependent upon the noun and the determiner/quantifier at issue. The data shown in table 1.2, adapted from Allan (1980, p. 549), displays a discrepancy between nouns which combine with unit terms, such as the indefinite article, and nouns which combine with quantifiers of indefinite quantities, such as *several*.

Based on such observations, Allan (1980) proposes that countability is a matter of degree and that different nouns have different countability "preferences". For instance, in table 1.2, the noun *car* would have the highest degree of countability since it may appear with both unit quantifiers (*one*) and fuzzy quantifiers (*several*). Nouns such as *admiration* and *cattle* would both be of lesser degrees since they fail to appear in one of those two environments. Other nouns discussed by Allan, such as *equipment*, which do not combine with either unit or fuzzy quantifiers, stand at the lowest degree of countability.

Allan (1980) brings out several important points. Given the data he discusses, it seems unlikely that countability can be treated simply as a binary phenomenon. Further, the problem of understanding countability cannot be reduced to understanding only nouns or only other elements such as quantifiers, but is dependent on understanding the contribution of both types of elements. Finally, the observation that different nouns have different countability preferences directly leads to asking what, if anything, lies behind the nouns' different countability preferences. This in turn is one of the primary questions that I will consider in the following chapters. I now turn to examining the various approaches that researchers have taken in accounting for the countable/non-countable distinction.

Type of quantifier	*Head Noun in the Domain of the Quantifier*		
	car	admiration	cattle
unit, *a(n)*	✓	✓	*
'fuzzy' quantifiers, e.g. *several, about fifty*	✓	*	✓

Table 1.2: Discrepancies in Countability Statuses (adapted from Allan 1980, p. 549)

1.2 Approaches to the Countable/Non-countable Distinction

In this section, I will present a brief summary of the main positions that have been taken in the literature on the countable/non-countable distinction. There are already several excellent, detailed overviews of the countability literature (Pelletier & Schubert, 2004; Joosten, 2003; Payne & Huddleston, 2002; Lasersohn, 2011), so I will limit the exposition here to describing at a very general level the three main approaches to countability: the notional account, the morphosyntactic account, and the contextual account.

Much of the literature on countability has argued for one alternative at the expense of the others, i.e. countability is either a grammatical phenomenon or a semantic phenomenon, but not both. There is no a priori reason for such exclusivity, and it has unfortunately led to much effort spent on pointing out the failings of other proposals. I will be striving towards a more inclusive account, which accommodates the strong points of these different approaches.

1.2.1 The Notional Account

Probably the most widely held view on countability is that it is related in some manner to objects in the world, whether due to these objects' ontological status or due to speakers' conceptualization of objects.

An early proposal along these lines is due to Quine (1960), who suggested that count nouns "divide their reference", meaning that they permit individuation of entities and whereby speakers may trace their identity through space and time. Non-countable nouns, in contrast, do not individuate entities. According to Quine, this capacity for individuation

is provided by language, but there is a tight link to the real world. Several properties help to distinguish countable and non-countable nouns. For instance, non-countable nouns refer to real-world entities that have the property of "cumulative reference": viz. "any sum of parts which are water is water" (p. 91). This is in contrast to count nouns which are "bounded" entities—adding an apple and another apple results in apples, not an apple.

A related view is what is known as the "cognitive individuation hypothesis". This view holds that the use of countable or non-countable nouns corresponds to whether speakers interpret the referent as an "individual", or an "individuated" entity, or as a "non-individuated" entity (Bloom, 1990, 1994, 1996; Imai, 1999; Langacker, 1987; Mufwene, 1984; Wierzbicka, 1988; Wisniewski et al. 1996). This distinction between individuated and non-individuated is one step removed from the ontological level: this line of explanation relies on whether language users "conceptualize" of the referents as individuated or not, which may be independent of the real world. Proponents of this line of explanation claim that the individuated/non-individuated contrast connects with the distribution of countable and non-countable nouns.

> "For example, physical objects are prototypical individuals in being discrete, bounded entities that are separate from other aspects of the world. Substances are prototypical non-individuated entities in being continuous, unbounded, and arbitrarily divisible (e.g. mud divided into any-sized portion is still mud). Not surprisingly, physical objects are almost always labelled with count nouns (e.g., *a cat*, *a computer*, *a coffee cup*) and substances with mass nouns (e.g., *clay*, *honey*, *jelly*)." (Wisniewski et al. 2003, p. 586)

The individuation hypothesis is the view which, as far as I can tell, comes closest to matching naïve intuitions about what the distinction between countable and non-countable nouns signifies. Indeed, this hypothesis is akin to Jespersen's original definition of count and mass nouns quoted in the opening paragraphs of this chapter. Yet, there are serious challenges—while it covers many cases concerned with physical entities, extending this view to the full range of countable and non-countable nouns is not always obvious. Critics tend to discuss the countability properties of nouns which are abstract or differ from typical physical objects or subtances. For instance, in the following quote, Barner & Bale (2002)

criticize the putative mappings between countable nouns and objects and between non-countable nouns and substances.

> "... the problem again is that neither adults nor children show evidence of respecting such mappings. As pointed out by Bloom (1999), many words used as count nouns are not marked for +object: *dream*, *puddle*, *sound*, *thought*, *cause*, etc. Furthermore, ... many words used as mass nouns have nothing to do with substances. Consider, for example, the following mass nouns: *furniture*, *rice*, *pasta*, *infantry*, *traffic*, *footwear*, *toast*, *cutlery*, *drapery*, *fruit*, and *clothing*. Unlike substances such as water or glue, one could quite conceivably count the footwear or furniture in a room, or sit on the curb counting traffic." (Barner & Bale, 2002, p. 785)

The primary challenge then for maintaining this view is to address how the notion of individuation applies beyond the core cases of physical objects (*dog*) and substances (*water*).

1.2.2 The Morphosyntactic Account

A very different account argues that the countable/non-countable distinction is only a morphosyntactic or grammatical classification and is not reducible to a systematic distinction in meaning. This view has a distinguished pedigree with Bloomfield (1933) as an early proponent:

> "To describe the grammar of a language, we have to state the form-classes of each lexical form, and to determine what characteristics make the speakers assign it to these form-classes. The traditional answer to this question appears in our school grammars, which try to define the form-classes by the CLASS MEANING—by the feature of meaning that is common to all the lexical forms in the form-class. ... School grammar defines the class of plural nouns by its meaning "more than one" (person, place, or thing), but who could gather from this that *oats* is a plural while *wheat* is a singular? Class-meanings, like all other meanings, elude the linguist's power of definition, and in general do not

> coincide with the meanings of strictly defined technical terms." (Bloomfield, 1933, p. 266).

The quote from Bloomfield illustrates the primary argument: if the countable/non-countable distinction were based on meaning, there should not be words that designate similar entities, but have divergent grammatical behavior in terms of countability. A similar argument is made by Palmer (1971), where a second dimension is added to the argument—not only are there similar entities in a given language which have divergent countability preferences, but there are also cross-linguistic mismatches.

> "It is easy enough to show that grammatical distinctions are not semantic ones by indicating the many cases where there is not a one-to-one correspondence. ... examples are to be found in *foliage* [mass] vs. *leaves* [count], in English *hair*, which is singular, vs. French *cheveux*, plural. These distinctions are grammatical and do not directly correspond to any categories of meaning." (Palmer 1971, p. 34–35)

These arguments have also made their way into the formal semantic literature as well. For instance, Chierchia (1998a, 2010a) has been a proponent of the grammatical account.

> "What we find is that even closely related languages have somewhat diverging sets of mass vs. count nouns. For example, in English one says "I cut my hair". In Italian, one has to say "mi sono tagliato i capelli" (*I cut my hairs*, plural). *Hair*, used to refer to what grows on our head, seems to be mass in English, and count in Italian. Yet clearly we are referring to the same stuff. Your hair doesn't change, as we change language." (Chierchia, 2010a, p. 151)

This approach, by rejecting any correspondence between countability and the meaning of nouns, avoids having to explain any connection between grammatical countability facts and the referents of nouns—the distributional properties of countability are simply part of the grammar. While recognizing as legitimate the challenges to grounding countability in meaning, many researchers have found this conclusion too strong:

> "It is hard to believe that the uncountability of nouns such as *water*, *gold*, or *smoke*, and the countability of nouns such as *car*, *flower*, and *dog* would be purely coincidental. It seems unlikely that this general tendency for substances to be referred to by mass nouns, and for objects and animate beings to be referred to by count nouns, would be unrelated to any meaning distinction." (Joosten 2003, p. 219)

One of the aims of chapters 2 and 3 is to examine closely the challenges to grounding countability in meaning, in particular in connection to the variation of countability classification across languages. I will argue that a broader view of the data points to a way which can do justice to countability as a grammatical phenomenon, but which is grounded in nominal meaning.

1.2.3 The Contextual Account

The final major view on countability denies that there is a substantial classification of nouns in terms of countability in the first place. Instead, these researchers hold that the countability of a noun is only determined due to the overall grammatical context the noun is found in, as opposed to being determined by a noun's intrinsic meaning (Allan 1980, Borer 2005, Pelletier 1979). This view would claim that, for instance, *a dog* is countable as a result of the constraints imposed by the determiner.

The primary evidence for maintaining such a view is that many nouns are interpretable in contexts which enforce a countable interpretation as well as in contexts which enforce a non-countable interpretation. If nouns are interpreted so flexibly, it is reasonable to suppose that nouns in general are underspecified for countability, and that this information is supplied by the context in which the noun occurs.

The most famous context which enforces a non-countable interpretation is the 'Universal Grinder', credited to David Lewis, but first discussed in Pelletier (1979). A succinct description is provided in Pelletier (1991), p. 497:

> "Consider the 'Universal Grinder', a device that takes in an object corresponding to the count term and spews out the finely ground matter of which it is

> made. A hat, for instance is fed into it and afterwards there is hat all over the floor. This is so despite the fact that there is another word we might have used (for example, *felt* or *straw*). So for any word one would wish to call a count term, there is a related mass term designating, roughly, the stuff of which it is made."

A similar observation has been made that some contexts demand a countable interpretation even from nouns which are usually non-countable. A Universal Packager is often thought to be at work for context such as in (18).

(18) Two beers, please.

Although the Universal Grinder and Packager have been quite successful at convincing researchers about the flexibility of nominal interpretation, at several points in the literature it has been pointed out that its success is also limited. For instance, Ware (1975), considering the Universal Grinder, states:

> "... I do not think that all [words] with count occurrences have mass occurrences and vice versa. Words for orifices seem to have count but not mass occurrences, for example: *opening*, *hole*, and *mouth*. We would also be hard pressed to find mass occurrences for *peculiarity*, *trick*, *act*, and *occurrence*." (Ware 1975, p. 383)

Similarly, Galmiche (1989, p. 68) notes the unacceptability in French of **du kilo*, **de la catégorie*, and **du chapitre*, which would appear also to be the case for their English counterparts, **much kilo*, **much category*, or **much chapter*. I will examine the issue of nominal flexibility and its limits in 3.6.3, where other examples will be added, such as nouns designating abstract shapes such as *triangle*, which has only a countable use. Aside from these counter-examples, it would seem that the criticism by Joosten quoted above would hold here as well: there should be a general account for the cross-linguistic regularity of objects being realized by countable nouns and substance being realized by non-countable nouns.

1.3 Overview of the Dissertation

The rest of the dissertation is organized as follows. In chapter 2, I lay out data from several languages with grammatical number systems that are richer than those usually considered. These languages distinguish countable nouns from non-countable ones and morphologically realize a singular/plural distinction, as in English; however, they also have a distinct class for nouns where a "collective" value is the default, and the unit interpretation is brought about by a special morpheme termed a SINGULATIVE. I discuss the morphological realization of the collective/singulative class across the different languages and also the lexical semantic generalization that accompanies it: that types of entities falling in the collective/singulative class tend to be those which habitually appear together, such insects with swarming behavior or granular substances such as *sand*.

Chapter 3 investigates the implications of the data in chapter 2. After reviewing the relevant evidence from the psycholinguistic literature, I argue that nouns' countability status is related to the level of individuation of the entity described by the noun, i.e. entities that are construed as having properties related to being an individual, are more likely to be countable. I then argue that countability, while often taken to be a binary distinction between countable and non-countable nouns, should be viewed as a scalar phenomenon. In particular, the morphosyntactic organization of grammatical number systems reflects the semantic organization of noun types according to the degree of individuation of their referents. Nouns of different types are individuated to different degrees and can accordingly be ordered along a scale of individuation: substances < granular aggregates < collective aggregates < individuals. Noun types which are less individuated are on the lower end of the scale and are cross-linguistically less likely to signal grammatical number, while the converse holds for highly individuated noun types. I show how this scale works for the grammatical number systems of several other languages which have been discussed in the literature, and then discuss how this scale of individuation relates to animacy which influence countability and consider its relation to frequency. Finally, I discuss how understanding morphosyntactic number categories in light of a scale of individuation avoids the difficulties binary accounts face, since languages may divide up the scale of individuation into any number of classes and at different points. At the same time, the proposal provides

a predictive framework for how this component of the grammar is organized: the contrasts being made are common across languages, and, as a corollary, the endpoints of the scale (substances and individuals) are predicted to be stable across languages. Grammatical number categories, e.g. "mass" or "count", then are viewed not as designating mutually exclusive classes of entities across languages, but as referring to sets of entity types which cohere on particular segments of a scale of individuation. I discuss how this view answers some of the common criticisms against a semantic account of countability.

Chapter 4 then incorporates this view into a formal semantic framework. I provide an overview of previous formal accounts, all of which in one way or another make use of part structures, or mereologies. I enrich standard mereology with topological relations, which permits formal recognition of whole objects as well as of different types of connections that may hold between entities. The addition of connectedness relations allows for a richer set of denotation types for nouns. This approach resolves some long-standing problems associated with mereological accounts of countability. The remainder of the chapter shows how this richer framework accords with the generalizations from chapter 3 and provides an account of the grammatical number systems of English, as well as Welsh and Dagaare, which are discussed in chapter 2.

Chapter 5 provides a conclusion and discusses directions for future work.

Chapter 2

Cross-Linguistic Manifestations of Countability

The countability literature has primarily focused on the distinction between nouns which are countable and those which are non-countable. Yet, many other distinctions arise in grammatical number systems—the countable/non-countable distinction is simply the most frequently attested and the most frequently discussed. This chapter examines a range of grammatical number systems which express a greater number of countability distinctions in addition to the simple binary countable/non-countable contrast expressed in English. A full-scale typological survey of grammatical number systems is beyond the scope of this work, but a sampling of systems will be examined here to demonstrate that considering countability as a binary phenomenon is far too simplistic. The first two sections of this chapter examine languages from distinct language families and geographic regions, namely Welsh (Indo-European), Turkana (Nilo-Saharan), Maltese (Afroasiatic), and Dagaare (Niger-Congo). This will yield a small database of interesting grammatical number systems, which will be drawn upon in future chapters.

In examining each of these languages, the aim will be to determine which categories of number are grammatically recognized, and for each of these categories, which types of entities fall under it. As the purpose of examining these number systems is to look for distinctions beyond the countable/non-countable distinction, the focus of this investigation shifts from the typical nouns considered in the literature, such as *dog* and *water*, to nouns

such as *sand* or *ants* which, although logically countable, often behave differently from typical countable nouns in many grammatical number systems. As far as possible, I will explore the core distributional properties used as diagnostics for distinguishing countable from non-countable nouns, such as cardinal modification, within these languages and determine how these distributional properties relate to the other categories of number.

I have restricted this investigation in two primary ways. First, the focus will be on languages which code number distinctions through affixes which are unambiguously devoted to coding those distinctions. I will not discuss in detail languages with systems of classifiers, such as Chinese, or elaborate gender systems, such as many Niger-Congo languages. The reason for this limitation is that classifier and gender systems tend to code number along with other categories of nominal meaning; for instance, Niger-Congo gender systems typically code singular/plural contrasts alongside other semantic distinctions such as shape or animacy, although such distinctions may be obscure synchronically (see Denny, 1976; Katamba, 2003; Maho, 1999). (Dagaare, a language that will be examined in section 2.2, is a Niger-Congo language, but its gender system has for all effective purposes been lost.) As these further semantic distinctions may provide confounding factors in the investigation, I start with the clearest cases first.

Second, I restrict my focus to what I term NATURAL CONCRETE ENTITIES (*dog*, *water*), to the exclusion of ARTIFACTUAL ENTITIES (*hammer*, *furniture*) or ABSTRACT ENTITIES (*arrival*, *happiness*). The reason for this restriction is that natural concrete entities provide the firmest foundation for comparative studies. It is a reasonable assumption that words designating *apple* in different languages all refer to the same entity. If there are differences, say different species of apples, these difference are relatively easy to assess once they are recognized. It is much more difficult to discern whether words glossed as 'sadness' across different languages do in fact refer to the same (abstract) entity. Cultures differ widely as to how they describe and lexicalize abstract notions, such as emotions, and this wide variation obstructs comparison across languages. Further, natural concrete entities tend to be basic nouns, i.e. not derived. In contrast, artifacts and abstract nouns, as the examples given indicate, tend to be derived nouns, and this adds another layer of complexity which would hinder the investigation at this point.

Section 2.1 describes the grammatical number systems of Welsh, Turkana and Maltese,

all of which make three-way distinctions in their number systems, and section 2.2 describes the number system of Dagaare, which makes an even greater number of distinctions. For each language, I will discuss how the countable/non-countable distinction is manifested and proceed to discuss the other number distinctions that the language makes. As discussed in section 1.2, cross-linguistic variation is often taken as an obstacle for any meaning-based theory of the countable/non-countable distinction. I show here that while languages vary as to how they encode number for certain types of entities, the variation is not free, but constrained with respect to a restricted set of entity types.

2.1 The Collective/Singulative Contrast

The next simplest type of grammatical number system, after a binary grammatical number system such as in English which only makes countable/non-countable distinction, is one which makes a tripartite distinction. Although tripartite systems are familiar to those working on particular language families where they are common, e.g. Nilo-Saharan, they have yet to be fully integrated into the countability literature. I present in this section three instances of tripartite number systems, which, while functionally similar in making a three-way countability distinction, differ in the morphological means they employ, as well as the lexical semantic domains included in their three number categories.

2.1.1 Countability in Welsh

Several distributional properties, laid out in section 1.1, have been generally taken in the literature to support a division, at least in English, between countable and non-countable nouns. Contrastive morphological coding, viz. plural coding, acceptability of modification by cardinal modifiers and other determiners implicating plurality (*few*, *many*), all point to a broad division into two grammatical categories. Yet, as the discussion of Allan (1980) demonstrates, there are some more nuanced distinctions than a strictly binary division can represent. Unsurprisingly, different grammatical number systems also make distinctions that do not completely align with a simple division into countable and non-countable nouns.

Welsh provides an interesting case for comparison with the better-known system of

English: according to some of the canonical morphosyntactic diagnostics of countability, Welsh is comparable to English, yet in several other respects, such as the coding of singular and plural number values, Welsh parts ways with English. In Welsh, in addition to nouns where the singular value is zero-coded and the plural is overtly coded, there are some nouns where the plural value is zero-coded (collective) while the singular value is overtly coded (unit). In total, Welsh grammatically recognizes several classes of nouns: singular/plural, collective/unit, group, pluralia tantum and non-countable. I now discuss the different classes in turn.

Countable and non-countable nouns Grammars of Welsh typically distinguish countable and non-countable nouns based on the presence of contrastive morphological coding for singular and plural values. Countable nouns regularly show an alternation between singular and plural values which correspond to different codings. Non-countable nouns, however, are described as being "not usually found in the plural" (King, 2003, p. 35) or as those nouns that "do not normally have number contrast" (Jones & Thomas, 1977, p. 161). Examples of countable and non-countable nouns are given in (1) and (2), respectively. According then to the diagnostic of morphological number coding, Welsh distinguishes countable and non-countable nouns.

(1) Countable (plural-coded) nouns:
afal/*afalau* 'apple/apples'; *cadair*/*cadairiau* 'chair/chairs'; *dyn*/*dynion* 'man/men'; *merch*/*merched* 'girl/girls'; *ton*/*tonnau* 'wave/waves'

(2) Non-countable nouns:
glo 'coal'; *'menyn* 'butter'; *mêl* 'honey'; *llefrith* 'milk'; *llwch* 'dust'

This contrast between countable and non-countable nouns is also supported by co-occurrence behavior with different quantifiers: some quantifiers only occur with countable (plural) nouns, while others occur with both countable (plural) and non-countable (singular) nouns. Table 2.1 shows a set of quantifiers in Welsh and their acceptability patterns with countable and non-countable nouns. The acceptability patterns show that just as *many snow* is unacceptable in English, so in Welsh several quantifiers indicating cardinality are

not acceptable with nouns designating substances, such as the entities given in (2).

Quantifier	Gloss	With countable plural N	With non-countable N
lot	'lot'	lot o lyfrau	lot o eira
		'lot of books'	'lot of snow'
llawer	'many' ; 'much'	llawer o lyfrau	llawer o eira
		'many of books'	'much of snow'
digon	'enough'	digon o lyfrau	digon o eira
		'enough of books'	'enough of snow'
mwy	'more'	mwy o lyfrau	mwy o eira
		'more of books'	'more of snow'
mwya'	'most'	mwya' o 'r llyfrau	mwya' o 'r eira
		'most of the books'	'most of the snow'
rhan fwya'	'greatest part'	rhan fwya' o 'r llyfrau	rhan fwya' o 'r eira
		'part most of the books'	'part most of the snow'
rhai	'some'	rhai o 'r llyfrau	*rhai o 'r eira
		'some of the books'	'some of the snow'
nifer	'number'	nifer o 'r llyfrau	*nifer o 'r eira
		'number of the books'	'number of the snow'
mwyafrif	'majority'	mwyafrif o 'r llyfrau	*mwyafrif o 'r eira
		'majority of the books'	'majority of the snow'
amryw	'several'	amryw o 'r llyfrau	*amryw o 'r eira
		'several of the books'	'several of the snow'

Table 2.1: Acceptability of Quantifier-Noun Combinations (adapted from Jones & Thomas 1977, p. 175–176)

In the Northern dialects, a further distinction is made. Non-countable nouns, at least those which are concrete, readily co-occur with the quantifier *peth* ('some'), while this is disallowed for countable nouns, echoing the distribution of *much* in English.

(3) a. peth o 'r siwgr
some of the sugar
(Jones & Thomas, 1977, p. 161)

b. *peth o 'r llyfrau
some of the books
(Jones & Thomas, 1977, p. 176)

The evidence so far implicates that the distinction in English between countable and non-countable nouns finds a parallel in Welsh. Additionally, these two categories of nouns are comparable in terms of notional distinctions, at least restricting ourselves to natural concrete entities: countable nouns correspond to different entities which are in some sense individuals, while non-countable nouns correspond to material stuff, namely substances and liquids.

Countability Category	Singular	Plural	Gloss
Singular/Plural	cadair	cadair-iau	'chair'
Collective/Unit	cacyn-en	cacwn	'hornet'
Non-Count	llefrith		'milk'

Table 2.2: Grammatical Number Categories in Welsh

Collective Nouns Countable nouns can be further divided according to the mode of their morphological coding: some nouns possess morphological coding of the plural while for others the singular value is overtly coded. Grammars of Welsh have not settled on a term for this class, but I will follow the terminology of King (2003) and designate it as the COLLECTIVE/UNIT distinction and call the morpheme which codes the singular value the SINGULATIVE.

The examples in (4)–(5) (Jones & Thomas, 1977, p. 157–158) demonstrate these distinctions with the words *afal* ('apple') and *adar* ('birds') which differ as to their default form: for *afal*, the singular is the default (uncoded) form and the plural is coded, while for *adar*, the plural is the default form and the singular is coded. The number values of these forms can be also observed in the agreement of pronominal elements. This is shown in the responses to the questions in (4)–(5) where the pronoun in the response agrees in number with the noun in the question, both of which are in bold. I gloss the singulative morpheme as SING as opposed to SG, which glosses the singular.

(4) a. lle mae 'r **afal**? — Mae **o** ar y bwrdd
where is the apple — is he on the table
Where's the apple.SG? — It's on the table.

b. lle mae 'r **afal-au**? — Maen **nhw** ar y bwrdd
where is the apple-PL — are they on the table

Where are the apples? — They're on the table.

(5) a. lle mae 'r **adar**? — Mae **nhw** ar y wal
where is the bird — is they on the wall
Where are the bird.PL? — They are on the wall.

b. lle mae 'r **ader-yn**? — Mae **o** ar y wal
where is the bird-SING — is it on the wall
Where is the bird? — It's on the wall.

The grammatical number system of Welsh provides a clear instance of a tripartite number system. As displayed in table 2.2, three different categories of opposition in terms of countability are manifested: singular/plural, collective/unit, non-countable.

Descriptions of the number system in Welsh inevitably point to a notional distinction which accompanies this formal distinction, namely the distinction between those types of entities which habitually co-occur in the world and those which do not. According to this view, the fact that Welsh has two types of countable nouns is not some morphological oddity, but a systematic distinction. Stolz (2001) provides a detailed analysis of the lexical items which comprise the collective class in Welsh, stating that they are characterized as those nouns which designate entities which are "saliently perceived as collectivities rather than 'individuals'" (p. 65). Stolz's findings indicate that the collective class is cohesive in terms of its lexical semantic categorization, and his categorization is partially reproduced in table 2.3

A related, but different, notional contrast is pointed out by King (2003), who advises students of Welsh grammar not to confuse the Welsh collective/unit contrast with the singular/plural contrast in English:

> "While, for example, *coed* can be translated as *trees* (because English has only the plural to fall back on in any case), it has a strong sense of a homogeneous group about it that *trees* on its own does not convey. The alternative translation *wood* (sing. in English) conveys the idea of a single item or group, but cannot include any idea of the units that make up that group (the trees). Both English translations are perfectly adequate as far as they go, but the relationship between the group and its individual components is neatly expressed only in the [collective/unit] system." (King, 2003, p. 48)

small animal and insects: *llygod*/*llygod-en* ‘mice’/‘mouse’; *berdys*/*berdys-en* ‘shrimp’/‘shrimp’; *cacwn*/*cacyn-en* ‘hornets’/‘hornet’; *picwn*/*picwn-en* ‘bees’/‘bee’, *chwain*/*chwann-en* ‘flea’/‘fleas’; *morgrug*/*morgrug-yn* ‘ants’/‘ant’

middle-sized animals: piod/piod-en ‘magpies’/‘magpie’; *moch*/*moch-yn* ‘pigs’/‘pig’; *cwning*/*cwning-en* ‘rabbits’/‘rabbit’; *hwyaid*/*hwyad-en* ‘ducks’/‘duck’

vegatation/cereals/fruits/vegetables: *dincod*/*dincod-yn* ‘seeds’/‘a seed’; *ceirch*/*ceirch-en* ‘oats’/‘an oat’; *chwynn*/*chwynn-yn* ‘weeds’/‘weed’; *cnau*/*cneu-en* ‘nuts’/‘nut’; *danadl*/*danhadl-en* ‘nettles’/‘nettle’; *maip*/*meip-en* ‘turnips’/‘a turnip’; *cennin*/*cenhin-en* ‘leeks’/‘leek’; *afan*/*afan-en* ‘raspberries’/‘raspberry’; *gellyg*/*gellyg-en* ‘pears’/‘pear’; *bresych*/*bresych-en* ‘cabbages’/‘a cabbage’

granular mass: *tywod*/*tywod-yn* ‘sand’/‘grain of sand’; *marwor*/*marwor-yn* ‘embers’/‘an ember’; *llwch*/*llych-yn* ‘dust’/‘speck of dust’

Table 2.3: Sub-types of Collective Nouns in Welsh (adapted from Stolz 2001)

From this data, it seems reasonable to suppose that the morphological pattern of collective/singulative reflects a categorization of entities in the world.

Stolz (2001) argues that this grammatical category and the notional category which underlies it, have been both historically stable, and furthermore, have gone through expansion via borrowing. As noted both by Stolz (2001) and Jones & Thomas (1977), borrowing in Welsh appears to be affected by whether the referent of the borrowed term is perceived as a collectivity. Some borrowings from English are given in table 2.4.

English Term	Collective	Unit
bricks	*brics*	*bricks-en*
figs	*ffigys*	*ffigys-en*
gooseberries	*gwsberys*	*gwsberys-en*
peas	*pys*	*pys-en*
garlic	*garlleg*	*garlleg-en* ('clove of garlic')

Table 2.4: Borrowings from English into the Collective Class in Welsh

The typical pattern of borrowing is that the morphologically uncoded form is borrowed (see Tiersma, 1982); however, the borrowings in table 2.4 are remarkable in that the plural, and morphologically complex, English form is what is borrowed as the base form in Welsh. Welsh speakers have apparently taken the occurrence of multiple entities as the basic situation for these nouns, a situation designated through plural morphology in English, and thus borrowed the plural form, to which the singulative can then apply.

Two final noun types can be established through contrastive singular/plural marking and verbal agreement. First, a small set of nouns designating social bodies or organizations, characteristic "group" nouns (see Landman (1989), Joosten et al. (2007) and references therein for discussion of these nouns). These nouns manifest a singular/plural distinction; however, the singular form requires the plural form of the verb, as shown in (6) from Jones & Thomas (1977, p. 161) .

(6) beth mae 'r **llywodraeth** yn mynd i wneud? — **Maen nhw** 'n mynd i godi
What is the government in go to do — are they in go to raise
trethi
taxes
What is the government going to do? They are going to raise taxes.

Pluralia tantum, i.e. nouns which only appear in plural form, are another type of noun distinguished in Welsh, which include nouns such as *nefoedd* 'heaven' or *trigolion* 'inhabitants'. Some pluralia tantum are not overtly marked as plurals, such as *gwartheg* 'cattle', but can be distinguished from the other noun types through agreement with pronouns, which must be in the plural. Unlike group nouns or collective/unit nouns, they do not manifest a singular/plural contrast.

Altogether, the data from Welsh discussed here demonstrates that countability distinctions that are grammatically recognized in languages may be richer that a binary countable/non-countable contrast. The morphosyntactic patterns of the language isolate several classes of nouns: singular/plural, collective/unit, group, pluralia tantum and non-countable. In considering the implications of these different categories, I will for the moment just focus on the additional recognition of the collective category, putting the group and pluralia categories aside.

Implications for Countability The first question to be addressed is whether this additional category of collectives is actually relevant for understanding the distinction between countable and non-countable nouns. This is not obvious, since from the morphosyntactic evidence adduced so far, it could be that the collective is just a flavor of the plural. One line of argumentation would be that these nouns are simply countable nouns, as their agreement patterns would indicate, and therefore this distinction, whatever it ultimately may be, is one that does not pertain to non-countable nouns, or aid us in understanding the nature of the countable/non-countable divide. Coming from the perspective of a binary count/non-count contrast, one could question whether these three categories could not be wedged into a binary mold. Although that form of reductionism may seem appealing at first as it keeps to a simple two-way contrast, it runs into far too many difficulties to be maintained.

First, consider the point of view of the notional distinction between countable and non-countable nouns, where countable nouns correlate with individuals and non-countable nouns correlate with non-individuals. The core classes of entity types realized as non-countable nouns and countable nouns in Welsh and English are essentially identical, namely liquids/substances and individuated objects, respectively. Yet, the collective class does not fit comfortably in either. Many members of the collective class would make for extremely

dubious individuals, for instance in the case of the cognates of *dust* and *sand*, a point reinforced by the fact that these nouns are non-countable in English. This evidence is, however, only suggestive.

A more serious objection to conflating the collective class with simple plural nouns is that it is common for collective nouns to make a three-way distinction between collective, singular and plural values. An example from Welsh is given in (7). This potential to make a three-way distinction is not limited to Welsh but also occurs in the related Celtic language Breton, whose number system is similar to that of Welsh. A parallel example in Breton to (7) is shown in (8), taken from Press (2009, p. 445) who glosses the pluralized form as "individualized" to signal its meaning. As both examples make clear, the collective form and the plural form designate distinct meanings: the pluralized singulative form designates a set of individual pieces in contrast to, e.g. a heap of grain. Treating collectives as plurals faces a serious challenge in accounting for these three-way contrasts.

(7) Welsh (Stolz, 2001, p. 70)

a. grawn
grain
'grain'

b. gron-yn
grain-SING
'a single grain'

c. gron-ynn-au
grain-SG-PL
'grains'

(8) Breton (Press, 2009, p. 445)

a. deil
leaf
'leaves' (collective)

b. deli-enn
leaf-SING
'leaf'

c. deli-enn-où
leaf-SING-PL
'leaves' (individualized)

In sum, the distributional differences reviewed above show three distinct major classes. The implication for theories of countability, then, is that a binary categorization of nouns into countable and non-countable may be necessary, but is not sufficient. Both a simple grammatical distinction between nouns compatible with number coding and those not, as well as a notional distinction between individuals and non-individuals, oversimplifies the typological space. The data from Welsh's collective class also indicates that the nouns which are not typical examples of countable or non-countable nouns, such as *leaves* and *ants*, provide much interest for understanding countability contrasts.[1] These types of nouns will reappear again and again in the other languages that are examined in this section.

2.1.2 The Collective/Singulative in Nilo-Saharan Languages

A reflex similar to that seen in Welsh's grammatical number system appears across other, unrelated, languages in disparate parts of the world. Languages of the Nilo-Saharan family, spoken across a large area in Central and Eastern Africa, show a widespread use of a collective/unit distinction in addition to a countable/non-countable distinction. I will primarily exemplify this pattern with examples from Turkana, spoken in Kenya, as it has been well-described in terms of grammar and vocabulary (Dimmendaal (1983) and Ohta (1989), respectively).

Turkana displays a countable/non-countable distinction: countable nouns have both a singular and plural form, while the form of non-countable nouns is either exclusively singular or exclusively plural. Number is coded in Turkana by virtue of both number suffixes and through agreement prefixes, which code gender as well. The example in (9) displays the singular and plural form of the word *-kòrì* 'giraffe', where the singular form has the singular, masculine agreement prefix *e-* and the plural form has both the plural, masculine/neuter prefix *ŋi-* and the plural suffix *-o*. For non-countable nouns, the agreement prefix shows whether the noun has a singular or plural form. Table 2.5, adapted from Dimmendaal (1983,

[1] See also the wide-ranging discussion in Acquaviva (2008).

p. 211, 234), displays examples of countable and non-countable nouns.

(9) a. e-kòrì
M.SG-giraffe
'giraffe'

b. ŋi-kori-o`
M.PL-giraffe-PL
'giraffes'

While non-countable nouns may be inherently plural (*ŋa-kipí* 'water') or inherently singular (*a-siɲoní* 'sand'), there is no compelling evidence that there is a semantic contrast in Turkana between inherently singular non-countable nouns and inherently plural non-countable nouns. In a later paper, Dimmendaal (2000, p. 230), discussing Nilo-Saharan languages in general, suggests that the number value of non-count derived nouns is largely a matter of etymology.

Countability Category	Singular	Plural	Gloss
Countable	a-mosiŋ´	ŋa-mósíŋ-o	'rhinoceros'
	a-rɪɛti´	ŋa-ríɛ́t-a	'desert'
Non-Countable	a-siɲoní	-	'sand'
	-	ŋa-kipí	'water'

Table 2.5: Countable and Non-Countable Nouns in Turkana

Countable nouns are further classified by the manner in which the morphological coding of singular and/or plural interpretations are achieved. Table 2.6 displays the three patterns in Turkana. The first one given is the singular/plural class where the plural is coded. The second line of the table shows the collective/singulative class, where the singular value is coded by a singulative marker. The third class is known as the "replacement" pattern (Dimmendaal, 1983), where both the singular and plural value are overtly coded. While the exact morphosyntactic and lexical details differ from language to language, Nilo-Saharan languages consistently distinguish these three types of singular/plural patterns.

As was the case for Welsh, those working on Nilo-Saharan languages have stated that there is a clear lexical semantic generalization underlying the collective/singulative class which contrasts with nouns of the singular/plural class. Dimmendaal summarizes:

Coding Type	Singular	Plural	Gloss
Singular/Plural	a-mosiŋ´	ŋa-mósíŋ-o	‘rhinoceros’
Collective/Singulative	ɛ-sɪkɪn-a´	ŋɪ-sɪkɪn´	‘breast’
Replacement	e-kúk-ut	ŋi-kúku-i´	‘chicken’

Table 2.6: Types of Countable Nouns in Turkana

> “... Nilo-Saharan words with meanings such as ‘bird’, ‘hair’, ‘leaf’, ‘louse’, ‘tooth’ are inherently plural, the corresponding singular expressing an individuated item from a collective or group, being marked with a singulative suffix. Entities congregated in large numbers or quantities tend to be morphologically unmarked in the plural in Nilo-Saharan languages. Likewise, words referring to items naturally occurring in pairs, such as ‘eye’ or ‘wing’, also tend to be unmarked in the plural, with the singular taking a singulative marker. Still other nouns have an unmarked singular form, taking a number suffix in the plural—a pattern that is much more common cross-linguistically. In Baale [a Nilo-Saharan language spoken in the region of the border between Ethiopia and Sudan], this set of nouns includes words such as ‘ax’, ‘tongue’, and ‘udder’, or names for animals not living in herds or groups, such as the word for ‘rhinoceros’.” (Dimmendaal, 2000, p. 229)

Although the Nilo-Saharan number systems have clearly developed independently from the number systems of Celtic languages, a comparison of the types of entities involved in the collective/singulative class in both languages reveals substantial overlap. Table 2.7 shows examples from the lexicon of Turkana for nouns of the same entity types found in Welsh, along with several other entity types which are not treated the same in Celtic languages.

While the collective/singulative class in Turkana bears much similarity that of Welsh, Turkana has an additional class of nouns where both singular and plural are coded for number by a suffix, the “replacement” pattern mentioned above. Dimmendaal observes that the replacement pattern is common for derived nouns, e.g. deverbals (Dimmendaal, 1983, p. 243). A comparison of different Nilo-Saharan languages indicates that the replacement pattern also appears to be a secondary development for nouns which notionally fit into

types of people: *e-sebey-it/ngi-sebey* 'Sebey person'/'Sebey people'; *e-turkana-it/ngi-turkana* 'Turkana person/Turkana people'

pair/grouped body parts: *e-pon-ol/ngi-pon* 'lip'/'lips'; *e-kyal-ai/ngi-kyal* 'tooth'/'teeth'; *a-ki-t/nga-ki* 'ear'/'ears'; *e-gec-ot/ngigece* 'wrist'/'wrists'

small animal and insects: *e-suro-t/ngi-suro* 'mosquito'/'mosquitos' ; *e-kur-ut/ngi-kur* 'maggot'/'maggots'; *e-lac-it/ngi-lac* 'bed louse'/'bed lice'

middle-sized animals: *a-taruk-ot/nga-taruk* 'vulture'/'vultures'; *a-toow-at/nga-toowa* 'duck'/'ducks'

vegatation/cereals/fruits: *a-tur-ot/nga-tur* 'flower'/'flowers'; *e-mar-et/ni-mare* 'bean'/'beans'; *e-nga-it/ngi-nga* 'blade of grass'/'grass'; *a-kuy-en/nga-kuui* 'leaf'/'leaves'

granular mass: *e-cok-et/ngi-cok* 'seed'/'seeds'; *e-tab-a/nga-tab* 'piece/pieces of tobacco'

Table 2.7: Sub-types of Singulative Nouns in Turkana (extracted from Ohta 1989)

both the collective/singulative and singular/plural classes. Dimmendaal (2000) provides evidence that some Nilo-Saharan languages have a three-way contrast between an uncoded collective form, a coded singulative, and a coded plural form. The example in (10) from the Shatt language shows that there is a natural group or set interpretation for *teeth* (10-b), which is designated by the unmarked collective form, which contrasts with the singulative reading (10-a) and (distributive) plural reading (10-c). In the closely related Shila language, this three-way contrast, where a collective interpretation is zero-coded, has apparently been simplified to a two-way contrast where only the coded singular and coded plural remain, as shown in (11).

(10) Shatt (Dimmendaal, 2000, p. 242)

a. nyix-te
nyix-SING
'tooth'

b. nyix
teeth
'(set of) teeth'

c. nyix-ke
nyix-PL
'teeth'

(11) Shila (Dimmendaal, 2000, p. 243)

a. nyir-te
nyir.SING
'tooth'

b. nyir-ke
nyir-PL
'teeth'

The three-way contrast recognized in Shatt also indicates, in parallel with the Welsh data, that the collective/singulative pattern cannot be reduced to an exotic version of a singular/plural pattern, but three distinct and contrastive values for grammatical number must be recognized as operative in these systems.

In sum, the Turkana grammatical number system, and more broadly those of Nilo-Saharan languages, appears to be making a three-way distinction between entity types that are not countable, those which typically come as multiple and related entities, and those which typically come as individual entities. This three-way distinction can be extrapolated from the categories of countability grammatically recognized, namely singular/plural, collective/singulative, and non-countable, as well as from three-way contrasts in interpretation of nouns, as was discussed concerning example (10). Additionally, the nouns which occur in the collective/singulative class are correlated with certain types of entities, similar to the situation in Welsh. These languages then demonstrate again that the countable and non-countable contrast must only be a partial description of the countability distinctions languages may express.

2.1.3 The Collective/Singulative in Maltese

Maltese also displays a tripartite number system, but differs in interesting ways both in terms of behavior with cardinal modifiers and in terms of the lexical semantic domains involved. Maltese is a Semitic language, yet has had little contact with other Semitic languages and Arab cultures for many centuries, but has had intense contact with European languages.

Maltese disposes of a countable/non-countable contrast which is witnessed by the usual core distributional properties. Countable nouns code a plural interpretation and allow modification by quantifiers implicating plurality and cardinal determiners. In contrast, non-countable nouns typically have only a single form, as in *arju* 'air', and do not permit modification by quantifiers implicating plurality or cardinal determiners. A further distinction is made for countable nouns: when modified by cardinals from 2 to 10, the plural form of the noun is required, while when modified by cardinals from 11 onwards, the singular form is required. Table 2.8 shows the paradigm for countable nouns.

As in Welsh and Turkana, Maltese possesses a collective/singulative contrast; however, the pattern of coding in Maltese differs. Unlike in Welsh and Turkana where the collective is formally a plural, the Maltese collective is zero-coded in the collective interpretation and formally singular and masculine. The unit interpretation is achieved by the addition of the

Singular	Plural	Indefinite Quantifiers	Cardinals 2-10
ˈskola	ˈskeyyel	ˈhafna ˈskeyyel	ˈerba˚ˈˈskeyyel
school.SG	school.PL	many school.PL	four school.PL
ˈrağel	irˈğil	ˈhafna irˈğil	ˈerba˚irˈğil
man.SG	man.PL	many man.PL	four man.PL

Table 2.8: Countable Noun Paradigm in Maltese

singulative suffix *-a*, which also renders the noun feminine yet still singular.

The combinatoric patterns with quantifiers add an additional layer of complexity to the collective's nominal paradigm. In contrast to countable nouns, which have just one plural form, collective nouns have an additional plural form when used with lower cardinal modifiers (2-10), known as the DETERMINATE PLURAL. For quantifiers of indefinite quantity, the collective form is used. The paradigm is given in table 2.9.[2]

Collective	Unit	Cardinals 2-10	Indefinite Quantifiers
hût	ˈhûta	ˈerba˚hûˈtit	ˈhafna ˈhût
fish	fish.SING	four fish.DETPL	many fish
ˈnemel	ˈnemla	ˈerba˚nemˈlit	ˈhafna ˈnemel
ant	ant.SING	four ant.DETPL	many ant

Table 2.9: Collective Noun Paradigm in Maltese

The types of entities which exhibit the collective/singulative paradigm in Maltese are given in table 2.10. There is substantial overlap with the types of entities seen in Welsh and Turkana—the core members of vegetation, insects and granular mass are present. Yet, there

[2]Although not regularly productive, many nouns that fall in the collective paradigm also dispose of an "internal" plural form which yields a 'different types of N' reading, as in (i-a) and/or a 'large quantities of N' reading, as in (i-b), both examples from Mifsud (1996). For more discussion of these types of interpretation of the plural, see Tsoulas (2008) for discussion of comparable Greek data.

(i) a. ˚ğîn 'pasta'
b. ˚ğeyyen 'different types of pasta'

(ii) a. dŭd 'worm(s)'
b. dwîd 'large quantities of worms'

are some striking differences—both the category of types of people found in Turkana (as well as in Arabic) is absent, along with mid-sized animals that typically come in groups, as seen for Welsh. If higher level animates are absent in nominal domains of the collective in Maltese, the categories of foodstuffs (*bread*, *cheese*) and materials (*iron*, *wood*), in contrast, are novel to the discussion so far in this chapter. That these entity types, which are typically non-count nouns in, for instance, English, are included indicates that the inventory of the collective class is different from what we have seen in Welsh and Turkana.

That the collective/singulative class spans entity types that are both countable nouns (*shoes*) and logically non-countable nouns (*iron*, *wood*) has often been noted in the literature on Maltese (see Mifsud 1996, p. 32, Sutcliffe 1936, and Borg 1980). These different entity types of the collective category can be further shown to manifest different behaviors in terms of their interpretation. They may either designate a set of individuals, whereby the singulative designates one member, or may designate a material, whereby the singulative designates a conventional portion. These different interpretations are shown in table 2.11, where for paired/collected objects, animals, and insects, the unit interpretation typically indicates a member of a group, but for foodstuff nouns, as well as for certain materials, the unit interpretation designates a conventional portion, e.g. the unit interpretation of 'bread' results in 'a loaf'. The countable interpretations of the entities in the foodstuff and material categories are reminiscent of interpretations attributed to the "Universal Packager" in English (see section 1.2.3), e.g. *a water* to designate a glass or bottle of water. In essence, it appears that the class of nouns which morphologically belong to the collective class is semantically heterogeneous.

Adjectival modification, which was discussed in relation to countability was discussed in section 1.1.2, provides a second way to establish that the collective category is comprised of two types of citizens, viz. logically countable objects and logically non-countable objects. Gil (1996) shows that adjectival modifiers which presuppose individuals, such as *big*, *square* or *fat* are, as expected, compatible with nouns that designate naturally countable entities, but incompatible with collective nouns which designate non-countable materials. For the logically countable nouns, the interpretation of the unit form is as expected—it is simply a large element in the set of things that the collective noun identifies. For material or substance nouns, *big* can only be interpreted when the noun takes the singulative or

paired objects: *zar'bûn* 'shoes'; *karkur* 'slippers' ; *tarağ* 'stairs' ; *bib* 'doors'

small animal and insects: *hut* 'fish' ; *tayr* 'fowls' ; *nemel* 'ant' ; *dub'b-in* 'flies'; *nahal* 'bees'

vegatation/cereals/fruits: *'ahm* 'corn' ; *eneb* 'grapes' ; *tin* 'figs' ; *lumi* 'lemon'

granular mass: *'ramel* 'sand' ; *trab* 'dust'

foodstuffs: *hobz* 'bread' ; *'ğobon* 'cheese' ; *cokkor* 'sugar'

materials: *ha'did* 'iron' ; *spag* 'string' ; *in'y-am* 'wood'

Table 2.10: Sub-types of Collective Nouns in Maltese (after Mifsud 1996)

Collective	Gloss	Unit	Gloss
tarağ	(a flight of) stairs	tagra	a step
zar'bûn	(a pair of) shoes	zar'bûna	a shoe
sigar	trees	sigra	a tree
baqar	cattle	baqra	a cow
nahal	bees	nahla	a bee
hobz	bread	'hobza	a loaf of bread
'ğobon	cheese	'ğobna	a cheese
sapûn	soap	sapûna	a bar of soap

Table 2.11: Two Interpretations of the Singulative in Maltese: Member or Conventional Portion

determinate plural.

(12) a. *Pietru għandu kartun kbir
Peter have.PRES:3:SG:M cardboard big.SG:M
Peter has big cardboard.

b. Pietru għandu kartun-a kbir-a
Peter have.PRES:3:SG:M cardboard-SING big-SG:F
Peter has a big cardboard.

c. Pietru għandu tmin kartun-iet kbar
Peter have.PRES:3:SG:M eight cardboard-DETPL big.PL
Peter has eight big cardboards.

(13) a. Pietru kiel ħut kbir
Peter eat.PAST:3:SG:M fish big.SG:M
Peter ate big fish.

b. Pietru kiel ħut-a kbir-a
Peter eat.PAST:3:SG:M fish-SING big-SG:F
Peter ate a big fish.

c. Pietru kiel ħut-iet kbar
Peter eat.PAST:3:SG:M fish-PL big-PL
Peter ate big fish.

For some nouns, which have sometimes been termed "flexible nouns" (Barner & Snedeker, 2005), the application of different plural markers brings out different interpretations, which again point to the dual status of these nouns as designating both stuff and pieces that are formed from that stuff. The example in (14) gives four possible forms for the Maltese word for 'thread'. According to Borg (1980), the indefinite plural permits quantification over quantities of material, "something like 'lengths of thread"' (p. 114), while the unit and determinate plural forms designate one or more than one pieces of material.

(14) Maltese (Borg, 1980, p. 114)

a. hajt
hajt
'thread'

b. hjût
hajt.PL [indefinite plural]
'lengths of thread'

c. hajt-a
hajt-SING
'a piece of thread'

d. haj-iet
haj.DETPL
'pieces of thread'

The upshot of this discussion is that the collective/singulative distinction in Maltese encroaches on the domain of typically non-countable nouns—a situation that differs from what has been seen so far.

The grammatical number system in Maltese, as in Welsh, shows every sign of being productive. Evidence from borrowing here too shows that entities in the world of the appropriate type are simply integrated into the collective category. Nouns borrowed from Italian and English are given in table 2.12, adapted from Mifsud (1996, p. 44-45). As was the case with borrowings into the collective paradigm in Welsh, here it is the plural form in the source language that is borrowed. In the case of the first two listings on the table, the plural suffix *-i* has been dropped, aligning these forms with native terms, which regularly end with the consonant.

Source Term	Collective	Singulative
ravioli (It.)	*ravˈyûl*	*ravˈyûl-a*
maccheroni (It.)	*mˈarˈrûn*	*mˈarˈrûn-a*
sandali (It.)	*sandli*	*sandl-a*
bricks (Eng.)	*briks*	*briks-a*

Table 2.12: Borrowings from Italian and English into the Collective Class in Maltese

Taken together, the data from Maltese demonstrate yet another instance where a three-way contrast is in effect between types of entities which are not countable, those which typically come in groups, and those which are typically individual objects. The lexical semantic domains of each category are not identical to the two systems previously reviewed, but the overlap is substantial. An additional wrinkle is provided by the use of the singulative suffix to derive individuated readings of non-countable nouns, in a manner similar to "packaging" in English.

2.2 Beyond Tripartite Number Systems: Dagaare

Dagaare, a Gur language spoken in Northern Ghana, presents an even more complicated grammatical number system. Dagaare, as spoken in the Central Dialect in and around the town of Jirapa, Ghana,[3] not only grammatically recognizes countable and non-countable nouns, but makes finer distinctions within both classes. Of particular interest is Dagaare's "inverse number marking" system for countable nouns, which I will discuss in detail. I begin by describing the countable/non-countable contrast and then turn to the different types of non-countable nouns and countable nouns.

Countable/Non-Countable Contrast The countable/non-countable contrast in Dagaare can be established by using some of the core distributional properties given in section 1.1. First, countable nouns display a singular/plural contrast while non-countable terms do not have a plural form. This is shown in table 2.13.[4]

Similarly, countable, but not non-countable, nouns combine with cardinal modifiers, as shown in (15).[5] The use of cardinal modification with non-countable nouns is not allowed even with shifts of meaning of the sort associated with "packaging" found in English, such as *two glasses of water*.

(15) dò-rí à-yí
pig-PL NHUM.PL-two

[3]Dagaare constitutes a dialect continuum which traverses the Upper-West and Northern regions of Ghana into Southern Burkina Faso. The Dagaare Language Commission (1982) established orthographical conventions and made broad dialect divisions into Northern, Central and Southern Dagaare. Central Dagaare, spoken around the area of the town Jirapa and Ullo, serves as the standard dialect for educational materials, church literature and radio broadcasts and is the dialect from which the data here is taken. Central Dagaare is also the basis for most linguistic analysis on Dagaare, including the early work of Wilson (1962), Kennedy (1966) and Hall (1973).

[4]Dagaare has two levels of tone: high (´) and low (`). Vowels in Dagaare distinguish whether advance tongue root (ATR) is present or not, viz. /i/ is +ATR while /ɪ/ is −ATR. The standard orthography for Dagaare collapses the representations for the vowels /o/ and /ʊ/ as well as for /e/ and /ɪ/, but I have followed Bodomo's (1997) orthographic system in which these distinctions are made. Using Bodomo's (1997) orthographic conventions is beneficial in that, for instance, it shows clearly that in the case of the number marker /ri/, which will be discussed in detail in this section, *-ri* and *-rɪ* are allomorphs which differ only due to ATR-harmony.

[5]Cardinal modifiers above *two* appear with agreement prefixes, which agree in number and distinguish human and non-human referents. The prefix *a-* agrees with non-human plural nouns.

Noun	Plural Form	Gloss
tìɛ́	tììrí	‘tree’
pìɛ́	pɛ̀rí	‘basket’
dʊ̀ɔ́	dòrí	‘pig’
kʊ̀ɔ́	-	‘water’
kãã̀	-	‘oil’
z͡ii	-	‘blood’
sáálá	-	‘charcoal’, ‘coal’

Table 2.13: Nouns with and without a Singular/Plural Contrast in Dagaare

‘two pigs’ (NHUM = Non-Human Prefix)

Although non-countable nouns do not have a plural form in general, they may combine with other number suffixes. First, both countable and non-countable nouns may optionally combine with a distributive plural suffix *-ree*, which designates ‘N in different locations’ or ‘different kinds of N’. For the term *kʊ̀ɔ́* ‘water’, the pluralized *kɔ̀nnɛ́ɛ́* designates ‘water in different locations’ or ‘different types of water’.

Noun	Dist. Pl.	Gloss
kʊ̀ɔ́	kɔ̀nnɛ́ɛ́	‘water/ (types of) waters’
mʊ́ɔ́	mʊ́ɔ́nɛ́ɛ́	‘grass/grasses in different locations’

Table 2.14: Distributive Plural in Dagaare

The distributive plural, however, is not able to combine with cardinal numbers, as shown in (16). Non-countable nouns then, while accepting a type of indefinite plurality, are not equivalent to canonically countable nouns which do permit cardinal modification, as was shown in (15).

(16) *kɔ̀nnɛ́ɛ́ à-yí
water.DISTPL NHUM.PL-two
‘two waters’

Other quantifiers in Dagaare that I have examined do not discriminate between countable and non-countable nouns in terms of their distribution. Still, different interpretations

of these quantifiers arise depending on whether they combine with nouns designating individuals as opposed to substances. For instance, *yágà*, designating "a great number/quantity of N", is compatible with both countable and non-countable nouns, and accordingly can be translated into English as either *much* or *many*, as shown in (17). Similarly, the quantifier *zàá* can be interpreted as equivalent to 'all', 'every' or 'each' depending upon the noun with which it combines. Thus, patterns of grammatical acceptability for such quantifiers do not distinguish between countable and non-countable nouns, although there is clearly a difference in interpretation: *yágà* with countable nouns refers to a great number of individuals, while with non-countable nouns it refers to a great quantity.

(17) a. nîì yágà
cow.PL QUANT
'many cows'

b. z͡ii yágà
blood.SG QUANT
'much/a lot of blood'

Varieties of Non-Countable Nouns Non-countable nouns can be further classified as to whether they permit a singulative suffix *-ruu*, generally designating 'a piece of'. Unlike the singulative morphemes discussed so far, the distribution of the singulative in Dagaare is very limited: it only combines with fewer than 100 nouns, and these nouns designate primarily granular aggregate terms or foodstuffs (*sɛ́rímáání* 'pepper', *kpɛ́ɛ́* 'malt', *múó* 'grass', kàmáánì'corn').[6]

Noun	Singulative	Dist. Pl.	Gloss
kʊ̀ɔ́	—	kɔ̀nnɛ́ɛ́	'water/–/(types of) waters'
mʊ́ɔ́	mʊ́ɔ́rʊ́ʊ́	mʊ́ɔ́nɛ́ɛ́	'grass/blade of grass/grasses'

Table 2.15: Number Marking Possibilities for Non-Countable Nouns in Dagaare

The distribution of *-ruu*, as shown in table 2.15, implicates that the non-count nouns

[6]There appears to also be a use of *-ruu* among some speakers to indicate a packaging reading, i.e. where *kpɛ́ɛ́rʊ́ʊ́* designates one sack of malt rather than a piece.

in Dagaare can be divided into those which accept *-ruu* and those which do not, a morphosyntactic division which corresponds to a difference in entity type: granular aggregates as opposed to liquids and substances.

Varieties of Countable Nouns Count nouns in Dagaare have singular and plural forms. Yet, whether the singular or plural is morphologically coded varies by noun, giving rise to a pattern known as "inverse number marking". The basic paradigm is given for the Dagaare words 'child' and 'seed' in table 2.16, showing a near[7] minimal pair where both nouns share the same stem, yet the morpheme *-ri* marks the plural interpretation for 'child' and the singular interpretation for 'seed'.

Gloss	Singular	Plural	Stem
'child'	bíé	bíírí	bì-
'seed'	bírì	bíè	bí-

Table 2.16: The Inverse Number Marking Paradigm in Dagaare

Additional examples of nouns similar to those in table 2.16 are given in tables 2.17 and 2.18. In table 2.17, nouns marked by *-ri* in the plural are shown on the right. In table 2.18, those marked by *-ri* in the singular are shown on the left. Tables 2.17 and 2.18 also display instances where *-ri* undergoes assimilation following nasals and liquids, resulting in the allomorphs *-nI* and *-lI*, respectively. (For vowels, I use capital letters as a variable for +/-ATR values, i.e. *-rI* ranges over *r*ɪ and *-ri*.)

The forms in tables 2.16, 2.17 and 2.18 can be shown to correspond to singular and plural interpretations by examining their interaction with other elements of the grammar that mark number. Examples (18)-(19) demonstrate distinct number agreement with the singular (*ŋâ*) and plural (*-má*) forms of demonstrative pronouns. In the plural, the agreement prefix further distinguishes between human (*ba-*) and non-human (*a-*) referents.

(18) a. bíé ŋâ
child.SG DEM.PROX.SG

[7]The pairs differ in tone and in that the plural of 'child' has a vowel which appears to have undergone lengthening. Anttila and Bodomo (2009) propose that Dagaare phonology makes use of lexically conditioned high vowel lengthening in such cases.

-V Singular	-rI/-nI Plural	Gloss
tìɛ́	tììrí	‘tree’
gbìé	gbèrí	‘forehead’
pìɛ́	pɛ̀rí	‘basket’
nàŋá	nànní	‘scorpion’

Table 2.17: Marked Plural Nouns in Dagaare

-rI/-nI Singular	-V Plural	Gloss
lúgrí	lúgó	‘prop, pillar’
nyágrí	nyágá	‘root’
fìlí	fìlɛ́	‘sore’
íílí	íílɛ̀	‘horn’

Table 2.18: Marked Singular Patterns

‘this child’

b. bíí-rí bà-má
child-PL HUM.PL-DEM.PROX.PL
‘these children’ (HUM = Human Prefix)

(19) a. bí-rì ŋâ
seed-SG DEM.PROX.SG
‘this seed’

b. bíè à-má
seed.PL NHUM.PL-DEM.PROX.PL
‘these seeds’ (NHUM = Non-Human Prefix)

Number words whose value is greater than one also show selection of plural forms and agreement. No agreement is visible, however, in the use of the word designating ‘one’, *yénì*, which acts as a modifier of the noun, directly attached to the noun stem.[8] Number words designating ‘two’ or more select for plural nouns and again take plural agreement prefixes. Examples of both nouns from table 2.16 combined with the number words for

[8]In Dagaare, modifiers such as adjectives are compounded with the noun stem and the adjective then supplies its own singular/plural suffix patterns.

'one' and 'two' are given in (20)-(21).

(20) a. bì-yénì (bì- + yénì)
child-one
'one child'

b. bíí-rí bá-yì
child-PL HUM.PL-two
'two children'

(21) a. bí-yénì (bí- + yénì)
seed-one
'one seed'

b. bíè à-yí
seed.PL NHUM.PL-two
'two seeds'

Much previous work on the nominal system of Dagaare has considered the pattern in table 2.16 from the perspective of a system of noun classes in Dagaare. Bodomo (1997), Kropp Dakubu (2005) and Bodomo and Marfo (2006) elaborate systems of noun classes for Dagaare based upon different singular-plural pairings of nouns.[9] The inverse marking pattern in these analyses is simply related to a distinction between two different singular/plural pairings: one is comprised of nouns ending in vowels in the singular and *-ri* in the plural, while the other is comprised of nouns ending in *-ri* in the singular and vowels in the plural.

Analyses of the different noun classes in Dagaare are clearly valuable from a diachronic and comparative perspective. As a Gur language, the nominal system of Dagaare stems from Proto-Gur, which possessed a highly developed noun class system (Miehe and Winkelmann 2007) and without a doubt, the pattern observed in table 2.16 historically derives from a noun class system. At the same time, accounting for the inverse marking pattern is worthwhile in its own right. If the pattern in table 2.16 were only a minor singular/plural pairing among many others, it would probably not merit much attention. Yet, in contrast to related languages, such as Gurenɛ (Nsoh 2002), which conserve more of the Proto-Gur

[9]This classification, although formed from singular-plural pairings, is not a gender system in the sense of Corbett (1991) or Güldemann (2000), where genders are established based on agreement classes. Dagaare has few concord phenomena, and the only three agreement classes that can be established are those seen in examples (18)-(19): singular, human plural and non-human plural.

system, the noun class system in modern Dagaare has largely decayed. The inverse pattern of table 2.16 has become the predominant pattern of nominal system, accounting for over 70% of the nouns in my current database which have singular and plural forms. The two other major singular-plural pairings are for nouns designating liquids and other typically non-countable nouns, as discussed, and humans.

A different perspective on singular and plural formation in Dagaare is provided by Anttila and Bodomo (2009), who provide detailed morphophonological analyses of the Dagaare nominal system. They uncover a range of regularities governing the morphophonology of Dagaare number inflection, which the account presented here is compatible with. At the same time, this pattern cannot be explained purely in phonological terms, as the near minimal pairs in table 2.19 demonstrate.

Gloss	Stem	Singular	Plural
'wild rat'	kù-	kúó	kúúrí
'hoe'	kù-	kùùrí	kùé
'granary'	bùg-	bùgó	bùgrí
'pillar'	lùg-	lúgrí	lúgó

Table 2.19: Near Minimal Pairs in Dagaare

In light of the examinations of the Welsh, Turkana and Maltese number systems, it is reasonable to think that the inverse number marking pattern is linked to the type of entities the nouns designate. A systematic examination of the lexicon shows that reliable asymmetries are visible across different nominal domains. Figure 2.1 shows results of counts in the lexicon for the relevant domains. The x-axis displays the domains while the y-axis displays the number of lexicon entries. The dark-shaded regions show the number of lexicon entries in a given semantic domain with the singular unmarked, while the light-shaded regions show the number which are unmarked in the plural and marked by *-ri* in the singular. For instance, the category of *mammal* shows 43 entries in the lexicon that are unmarked in the singular and 5 entries which are unmarked in the plural and marked in the singular by *-ri*. In these counts, I exclude derived forms, since they follow their own patterns, which tends to obscure any generalization.

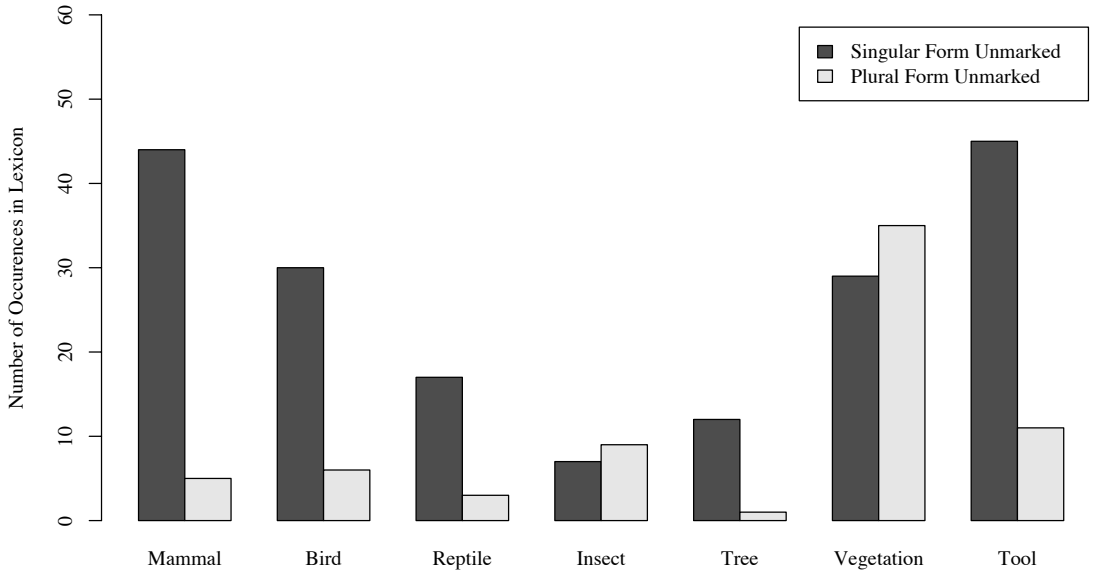

Figure 2.1: Number Marking Across Semantic Domains

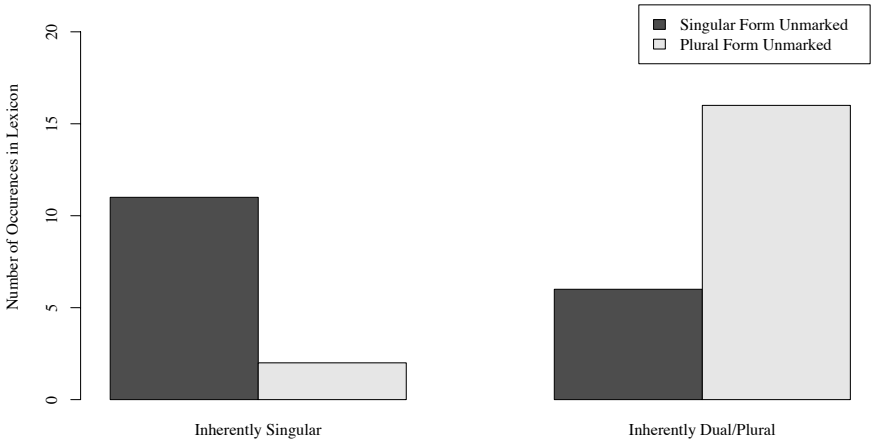

Figure 2.2: Number Marking and Inherent Plurality in the Domain of Body Parts

The overall pattens in the lexicon conform to the expectations one would have given the systems of Welsh, Turkana and Maltese. Nouns for higher-level animates, namely mammals, birds and reptiles are typically unmarked in the singular; however, the majority of nouns for insects have a plural that is unmarked. Similarly, nouns for trees are typically unmarked in the singular, while most nouns for vegetation are unmarked in the plural. By way of comparison, nouns for tools showed a strong tendency towards being unmarked in the singular.

Figure 2.2 shows that nouns for body parts which inherently come in pairs or groups are more likely to be unmarked in the plural while nouns for body parts that inherently come as singular items are more likely to be unmarked in the singular. The x-axis displays whether the noun is inherently singular, e.g. the term for *head* where canonically humans only have one, or inherently dual/plural, e.g. *eye* or *rib* where canonically humans have two and multiple of each, respectively. Again the y-axis displays the number of items in the lexicon for each category.

Thus, the system of Dagaare makes a distinction similar to what has been shown for Welsh and the other collective/singulative systems: nouns which are likely to co-occur

paired body parts: *dúmí/dúnó* 'knee'/'knees'; *gbérì/gbéè* 'leg'/'legs'; *núkúmí/núkúnò* 'fist'/'fists'; *yàgrí/yàgá* 'cheek'/'cheeks'

insects/reptiles/bugs: *dùndúlí/dùndúló* 'maggot'/'maggots'; *sòòrí/sòbó* 'locust'/'locusts'; *yʊ́ɔ́ní/yʊ́ɔ́mɔ́* (a type of small) 'frog'/'frogs'

vegatation/cereals/fruits: *kámáání/kámáánà* 'corn'; *lùnggúrì/lùnggúè* 'lemon'/'lemons'; *gúórì/góè* 'cola nut'/'cola nuts'

Table 2.20: Sub-types of Unmarked Plural Nouns in Dagaare

or form a natural pair or group are distinguished from nouns which are likely to be seen as individual entities. The manner of coding this distinction—through inverse number marking—is the novel aspect of the Dagaare system. Table 2.20 provides examples from these different types of entities.

In contrast to what was seen in Welsh and Maltese, this distinction does not appear to be productive any longer. Number marking for words borrowed into Dagaare appear to be influenced by phonetic similarity, as can be seen from the example *lɔ́ɔ́rì* (sg) / *lɔ́ɛ̀* (pl) 'truck; lorry', clearly a borrowing from the English word *lorry*. The Dagaare singular form *lɔ́ɔ́rì* corresponds to the English singular form *lorry*, and the Dagaare plural is formed on analogy with words for which the stem is vowel final and the singular is marked by *-ri*. A similar explanation can be found for the term *ŧìrí* (sg) / *ŧìɛ́* (pl) 'spoon', borrowed from Akan (Mark Ali, p.c.).

Altogether, the distribution of the different morphological markers implicates five nominal types that receive grammatical recognition in Dagaare: non-countable, granular aggregate, collective aggregate, individual and human. Once again, a purely countable/non-countable distinction would not capture all the nuances that Dagaare manages to express.

In summary, in this chapter I have reviewed grammatical number systems which display a large number of fine distinctions, many more than would be expected if grammatical number was only sensitive to a binary countable/non-countable distinction. There is a

persistent theme underlying these different distinctions: the grammatical category of collective/singulative nouns responds to a notional distinction among nouns concerning the propensity of their referents to appear together. The next chapter approaches these grammatical systems from the perspective of individuation, incorporating them in a general framework for countability.

Chapter 3

The Scale of Individuation

The preceding chapter set forth a range of different grammatical number systems, which have in common that they make a greater number of countability distinctions than English or other commonly considered Indo-European languages and that they display different variants of a collective/singulative class. Formally, the different collective classes across languages are similar in that the default nominal form, where no morphological coding is present, designates the collective value, while the coded form designates a single unit of the entity designated by the noun. The different languages show substantial variation in the inventory of entity types appearing in their collective/singulative classes.

This chapter first examines the relationship between the various entity types and grammatical number categories. I begin by observing that some types of entities appear to fall into a collective/singulative class more readily than others. The different entity types can be organized along a scale based on their likelihood to allow unit interpretations, but I show that only examining the entity type is insufficient. Section 3.2 examines evidence from the psycholinguistic literature bearing on abstract properties which may influence the categorization of nouns as countable or non-countable. The upshot is that ontological categorization, i.e. substance vs. object, is necessary to understand grammatical number categorization, but not sufficient. Several other additional properties, such as spatial configuration or whether an entity is associated with a function, also impact how entities are characterized in terms of countability. I show that the different factors related to individuation of an entity imply a set of individuation types which can then be organized in a scalar

fashion, giving rise to a scale of individuation.

With these foundations in place, section 3.3 explores the consequences of a scalar view of countability, elaborating a comparison of different language systems according to (i) which grammatical number classes are recognized, (ii) the manner of coding for each of the classes and (iii) the inventory of individuation types falling under each of the classes. The aim is to provide a method of comparing the organization of number systems in general. The picture that emerges is that languages structure their grammatical number systems in a way consistent with the scale of individuation. I explore the typological predictions of this scalar view on individuation, show how languages beyond those considered in chapter 2 align with the scale. I then consider the relation between the view proposed here and animacy in section 3.4, as well as frequency in section 3.5.

Having developed this broader perspective on the organization of number systems, section 3.6 proposes a variety of answers to some of the more recalcitrant issues surrounding the count/mass distinction. For instance, viewing individuation as a scalar phenomenon clarifies how languages may differ as to which nouns are countable or non-countable. Two languages may simply divide the scale of individuation up in different fashion. Variation among languages is expected, but the variation will generally be constrained to accord with the scale of individuation. Finally, I discuss the issue of nominal flexibility, its limits and how it accords with the view of countability defended here.

3.1 Comparing Grammatical Class, Coding, and Entity Type

During the examination of different grammatical number systems in the last chapter, it became clear that there was frequently a relation between certain entity types and certain grammatical number categories. For instance, the entity type “insects” is typically a member of the collective/singulative class. Yet, there is also much variation: the entity type “type of people” belongs to the collective/singulative class in Turkana but to the singular/plural class in Welsh. A first question to ask is whether a systematic picture of which entity types align with which grammatical categories can be developed, and further if this

can be formed in a manner that is predictive of which entity types are likely to fall under which grammatical categories and which entity types might show variation.

The relation between entity types and their manner of coding in languages with a singulative marker is presented visually in table 3.1. The entity types given here are simply ad hoc categories based on the discussion in chapter 2. The table pairs the different entity types discussed so far with their manner of coding in the four languages discussed. The rows of the table give the entity types, while the columns give the languages. For instance, the entity type "insects" has the singular as the coded value across all the languages, which is designated by *sing*. Shaded cells of the table indicate that for the language of that column the entity type is in the collective/unit class. (Dagaare has two categories of nouns for which the singular value is coded, zero-coded plurals with the singular marked by *-ri* and those nouns disposing of the optional singulative *-ruu*. These two categories are shaded in dark and light grey, respectively.)

The relation in table 3.1 between the different entity types and their coding clearly reveals that some entity types are coded via a collective category more often than others across the different languages. In particular, while insects and vegetation are uniformly treated as members of the collective category, higher animates are less often treated as such. On the bottom section of of the table, granular aggregates are treated across these languages differently than liquids, although languages differ as to how foodstuffs are coded. As the shaded region indicates, there appears to be a core set of entity types, namely insects and vegetation, where if a language grammatically recognizes a collective category, entities from these types will fall under it.

The pairings between entity type and morphological coding provide a method to establish an ordering on the entity types. Organizing the entity types on the basis of their coding preferences across these different languages, from zero-coding to singular-coding to plural coding, yields the scale in (1). This scale is not strictly ordered, as some entity types are ranked equally (as indicated by $\leq$).

(1) liquids < foodstuffs < granular aggregates < vegatation/cereals/fruits $\leq$ insects < small animals <pair/grouped body parts $\leq$ middle-sized animals < types of people < individuals

Entity Type	Coding			
	Welsh	Turkana	Maltese	Dagaare
types of people:	plur	sing	plur	plur
middle-sized animals:	sing	sing	plur	plur
pair/grouped body parts:	plur	sing	plur	sing
small animals and insects:	sing	sing	sing	sing
vegatation/cereals/fruits:	sing	sing	sing	sing
granular aggregates:	sing	sing	sing	0/sing
foodstuffs:	0	0	sing	0/sing
liquids:	0	0	0	0

Table 3.1: Entity Types and Their Number Coding in Welsh, Turkana, Maltese and Dagaare

The ordering of coding preferences in (1) appears to reflect the accessibility of the unit interpretation. The higher in the scale a noun type is, the more salient the unit interpretation becomes. The noun types lowest on the scale have zero-coded nouns, where the expression of a unit is either not coded directly or not uniformly coded. For instance, in Welsh, *llefrith* 'milk' does not directly code a unit, and a unit interpretation can only be brought about by using additional material, such as a measure phrase (e.g. *bottle of*). In contrast, *cacwn* 'hornet', a member of the insects entity type higher up the scale, has morphological means for regularly expressing a unit, viz. the singulative. In Dagaare, for some zero-coded nouns, a unit interpretation can be specified through the singulative morpheme, but this specification is not obligatory or regular. For singular-coded nouns, the expression of the unit value is regularly made available, but it is not the default. For plural-coded nouns, the expression of the unit value is the default.

As compelling as the scale in (1) might be, it falls short of being satisfactory. First, the fact that some members of an entity type belong to, e.g., the collective/singulative class does not indicate that they all do. As the lexicon counts for Dagaare in section 2.2 showed, these lexicalization patterns aligning entity type and grammatical coding are strong correlations, but not absolutes.[1] Further, the scale in (1) gives the impression that somehow the size of an entity is a relevant factor to its countability classification, yet we will see at

[1]I mention in passing that this is also the case for liquids: while many nouns which designate liquids are considered non-countable (*oil*), there are some, for instance those designating particular drinks (*martini*), which by virtue of their distributional properties are countable.

several points in this chapter that the influence of size on whether a noun is countable is an epiphenomen. Still, the picture that emerges from table 3.1 is that grammatical number appears to be responsive in some manner (i) to ontological distinctions among different entity types, such as substances vs. entities with discrete, and potentially countable parts and (ii) to the accessibility of units. This serves as an indication that grammatical number patterns with the meaning of the nouns.

The next section ties accessibility of the unit interpretation to the more abstract notion of INDIVIDUATION. As discussed in section 1.2, the hypothesis that countability marking reflects something about the world is one of the central points of controversy in the countability literature. Increasing evidence from psycholinguistic studies, which I will discuss in sections 3.2.1 and 3.2.2, indicates that speakers' use of grammatical number is bound up with how they interact with and view the world. I now turn to discussing some evidence which bears on this question, both that which supports such a relation and some evidence demonstrating that while such a relation may underlie countability, it is not always a simple relation.

A word of caution is in order before I proceed. The entity types discussed, and subsequently the ordering given in (1), is based on a limited number of languages. So far only four languages, along with English, have been presented. Several other systems will be discussed in section 3.3.4 which also accord with the generalizations that I will make in what immediately follows. Yet, what is truly required is a balanced and thorough typological examination of the relation between grammatical number coding and entity types, although that is beyond the scope of this work. I expect that the generalizations that I make below to carry over to a larger sample, but this of course cannot be known a priori. At the same time, all the data brought together here constitute a large expansion of the amount of data normally considered in the discussion of countability. At this point, it will provide a different perspective on countability phenomena, even if the final word about the relative countability levels of entity types such as "paired body parts" or "types of people" cannot be given here.

3.2 Individuation and Grammatical Number

The ordering on the scale in the preceding section appears to be linked to the propensity for the entity described by the noun to be such that it occurs as an individual. Exactly this property, the propensity for an entity to occur as an individual, known as individuation, has attracted equal parts interest and scrutiny in the psychological and philosophical literatures. Most broadly the thesis of individuation relates cognitive or perceptual qualities of objects to the grammatical realization of count and mass nouns.[2] The simplest hypothesis concerning the relation between entities in the world and the countability status of their names is that there is a direct correspondence between count nouns and reference to "individuals" in the world and non-count nouns and reference to "non-individuals". On a strong version of this correspondence theory, language users should "conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as non-distinct, uncountable, unindividuated things" (Wisniewski et al., 1996, p. 271).

This dichotomous view of individuation is discussed widely in the psycholinguistic literature (e.g. Bloom 1990, Bloom 1994; Wisniewski et al. 1996; Barner & Snedeker 2005). I will argue that, given the data from chapter 2, a dichotomous distinction is too coarse. Rather, a scalar approach to individuation leads to a proper account of countability.

Section 3.2.1 discusses some of the main evidence from the psycholinguistic literature that grammatical countability is related to the object/substance distinction, and thereby an individual/non-individual distinction, while section 3.2.2 discusses other psycholinguist work which indicates that individuation is related to factors beyond only the "object/substance" distinction. I then show that considering these factors together implicates a scalar approach to individuation. This sets up the machinery which will be used in section 3.3 to account for grammatical number systems.

[2] Another use of the term *individuation* in the philosophical literature relates to referentially individuating objects, and in particular to the sortal/non-sortal distinction and problems of identity (Brennan, 1988; Wiggins, 1980) Although countability is ultimately also bound up with referential aspects of language use, I will not pursue those connections here.

3.2.1 Individuation and the Object/Substance Distinction

Given that the most salient examples of countable and non-countable nouns are provided by nouns such as *dog* and *water*, respectively, much effort has been spent on disentangling the relationship between countable nouns and "objects" and non-countable nouns and "substances". Most researchers agree that there is a relation between the grammatical countable/non-countable contrast and the distinction between objects and substances, respectively, but the nature and direction of this relation has engendered significant controversy. In addition, there have been those who question that this relation delivers any substantive account of grammatical number categorization. In this section, I will discuss three of the more influential studies which provide evidence that a distinction between objects and substances is relevant to countability.

Most of the research on children's acquisition of countability reacts in some manner to Quine (1960), who put forth several far-reaching conjectures. According to Quine (1960), count nouns come with a "built-in manner of referring to individuals" (p. 91). Thus, count nouns are positively specified for individuation, an apparatus for individuating objects, viz. delimiting the relevant object from others and tracking its spatio-temporal identity. A second aspect of Quine's (1960) hypothesis concerns the direction of fit between language and the world. He claims that language provides the means for individuating objects: when children learn language, they simultaneously learn how to divide up the world into individual objects. According to this hypothesis, children in their pre-lingusitic state do not have ontological commitments to notions such as substance or object—it is only by way of language that these commitments come about.

The second part of Quine's hypothesis has led to much work in the language acquisition literature. Many researchers have found that children are able to pick up on perceptual cues that indicate a discrete solid object, and therefore, against Quine's hypothesis, children do seem sensitive to the object/substance distinction. The work of Soja et al. (1991), in particular, relates infants' perception of solid objects directly to questions surrounding acquisition of the count/non-count distinction, and makes the case that, contrary to Quine, infants apparently do have ontological commitments prior to learning English syntax.

To demonstrate that children have an association between solid objects and count nouns

at an early age, Soja et al. (1991) perform a novel word learning task, or "word extension task". First, the experimenters present the child with a novel item and instruct the child as to how it is named. Then, the child is presented with two other novel items which match the first on different perceptual properties. For the case at hand, an initial object would be either a novel solid object or a novel shapeless substance. The following two novel items consist of one which matches the initial object in shape and one which matches in material. It has been robustly attested that when a child is initially presented with a novel solid object, then the child will extend that name to another object of the same shape (Landau et al., 1988; Soja et al., 1991; Imai & Gentner, 1997). Yet, when the novel item is non-solid material, then the child does not tend to extend the name to the item with a similar shape, but rather extends it to the item formed from the same material.

From such evidence, Soja et al. (1991) argue that ontological distinctions underpin linguistic behavior, and not the converse, as Quine had conjectured. They conclude that "the present results show that presyntactic [with respect to the mass/count distinction] infants do see the world as composed of objects and non-solid substances (among other ontological types, presumably)..." (p. 206).

There is further evidence that this distinction is also relevant for speakers of languages other than English. Imai & Gentner (1997) provide evidence using the same experimental paradigm of the word extension task and show that Japanese speaking children also differentiate between objects and substances despite the fact that Japanese has no (straightforward) grammatical equivalent of the count/non-count distinction in English.[3]

The overall significance of this line of research is that ontological features of the world are relevant for understanding distinctions in language related to countability contrasts. In particular, the use of language in infants and language speakers at large appears to reflect, rather than create as Quine (1960) would have it, these different ontological properties related to countability.

If ontological distinctions are relevant, the next step is to determine to what degree they are relevant. Samuelson & Smith (1999) investigated in detail the correspondence in infants' vocabulary between solidity, shape, and material of the referent of a noun and the

[3]Much of this research is focused on isolating arguments for or against (some version of) the Sapir-Whorf hypothesis. This is not the focus here, although I will address it in passing when necessary to do so.

noun's categorization as countable or non-countable. Samuelson and Smith use solidity in a broad sense, following Soja et al.'s (1991) usage, to mean bodies that are "cohesive, bounded, spatio-temporally continuous and solid or substantial" (Soja et al., 1991, p. 183). Of particular interest is their finding that the correspondences between solidity and count syntax and non-solidity and non-count syntax, which had been the categories investigated in the previous studies, were real, but imperfect.

Samuelson & Smith (1999) used a corpus of 312 nouns that they took from the toddler version of the MacArthur Communicative Development Inventory, which serves as "a reasonably proxy for the nouns children learn early" (p. 4). The distibution of syntactic categories breaks down into very few non-countable nouns (31 nouns, 10% of the vocabulary), very many countable nouns (232 nouns, 74% of the vocabulary), as well as another portion of nouns which are judged ambiguous between countable and non-countable (49 nouns, 16% of the vocabulary). In examining this early vocabulary inventory of infants, they find that a general correspondence could be observed in the vocabulary between solidity and countable nouns and non-solids and non-countable nouns: "Count nouns comprise 74% of the nouns in the corpus as a whole but over 85% of the names for solid things. Mass nouns comprise less than 10% of the nouns in the corpus but more than 80% of the names for non-solid things" (p. 26). There were differences, however: solidity and shape were good predictors of count nouns, yet non-solids were less good predictors of non-count nouns. The general finding is that categorization as a countable noun is significantly correlated with shape and solidity, but material-based categorization, non-solidity and non-countable syntax are only weakly correlated.

The studies in Samuelson & Smith (1999) demonstrate that ontological features of entities, such as objects which are solid and/or have rigid shape, are most likely important for acquisition of grammatical number categories, and by implication, ontological distinctions are not to be lightly dismissed when investigating the basis of grammatical number. At the same time, their studies also indicate that whatever the ultimate explanation of the categories of count/non-count is, it cannot be reduced to a simple ontological distinction between substance and object.

A third set of experiments by Prasada et al. (2002) with adult speakers provide evidence that there are multiple factors at play which determine whether speakers are likely to treat

a given object as associated with a countable or non-countable noun. Further, they argue that classification of nouns as countable or non-countable is not determined by sets of immutable ontological properties of entities, but rather their classification is sensitive to speakers' construal of an entity as possessing non-arbitrary structure or not, respectively.

Prasada et al. (2002) examine whether factors such as regular shape or association with a function influence whether speakers name an entity as a countable or non-countable noun. For instance, they present participants with items that have either a regular or irregular shape. They then ask the participants to choose between two sentences describing the item which differ in the use of count or non-count syntax, i.e. 'There is a blicket in the tray.' vs. 'There is blicket in the tray.' The results showed that participants choose countable nouns more often for the regularly shaped items (73% of trials) than for the irregularly shaped ones (23% of trials). This difference is highly significant ($p < .0001$). Of particular importance is that both items are discrete objects: thus countability categorization is unlikely to reduce to a simple distinction between elements of the world which are discrete as opposed to non-discrete.

Subsequent experiments show that construing entities as countable or not is not strictly linked to whether their shape is irregular or not. Using the irregularly shaped items from the previous experiment, Prasada et al. (2002) show that when presented with multiple instances of the same irregularly shaped item, speakers are more likely to construe that item as corresponding to a countable noun. In one experiment, they showed participants either a set of items all with the identical irregular shape or only one item with that shape. In a second experiment, they showed participants either a set of items all with the identical irregular shape or a set of items with different irregular shapes. In each case, the participants then would choose a sentence containing either a countable or non-countable noun to describe the items. For both experiments, Prasada et al. (2002) found that the repeatability of structure was a significant effect ($p < .01$): participants more often selected sentences with countable nouns to describe the set of items whose structure was repeated as opposed to single items or sets of items with different irregular shapes.

A final study investigates whether a function associated with an item influences whether participants choose to assign a countable or non-countable noun as a description of the item.

Prasada et al. (2002) distinguish between "structure-dependent functions" and "structure-independent functions". Structure-dependent functions are those dependent on the entity's structure, e.g. one item, a piece of plaster board, was such that it fit into a slot in a mechanism and turning the item caused a bell to ring. The structure-independent function of the same item was that the bell could be rung by hitting it with the item—which clearly did not rely on the item having a particular structure. Participants viewed video tapes of the items being used in their structure-dependent or structure-independent function. The results showed that if an item is seen used in its structure-dependent function, it is more likely to be described by a countable noun (p <.03).

Taken together, the results from Prasada et al. (2002) provide compelling evidence that countability categorization does not reduce to a strictly ontological categorization between objects and substances in some objective sense. Instead, speakers, at least of English, categorize these entities in terms of countability according to whether they *construe* an entity as an object rather than as a substance. The countability status attributed to entities does not arise exclusively from their physical properties: an item with an irregular shape is considered less countable when contrasted with an item of regular shape, but when multiple instances of the item are present, it is judged more countable. I will return to the construal of entities in the world and its importance for countability in section 3.6 when discussing the common critiques of an individuation account of countability. I now turn to other studies which show that categorization via count and non-count syntax is in effect in tasks which are not reducible to a binary distinction between substance and object.

3.2.2 Beyond Substances and Objects: Other Factors Relevant to Individuation

As the foregoing discussion has shown, the contrast between countable and non-countable nouns is related to, but not exhausted by, a distinction between objects and substances. This naturally leads to the question of what other principles may be in effect. One answer to this is proposed by Wierzbicka (1988), who argues that there are multiple principles of categorization underlying the count/non-count distinction, including distinguishability of elements, divisibility, heterogeneity, and the manner in which people interact with an

entity.

Wierzbicka (1988) defends the idea that the morphosyntax of countability is closely related to the manner in which language users conceptualize the world, arguing that "form-classes are semantically motivated, and that difference in grammatical behavior reflect iconically differences in meaning" (p. 501). Rather than reducing the conceptual distinction to a simple individual/non-individual distinction, she articulates a set of factors that are relevant for categorization of nouns into different countability categories. I will discuss two factors in detail, DISTINGUISHABILITY and MANNER OF INTERACTION, which have both been examined experimentally by Middleton et al. (2004).

Wierzbicka argues that nouns designating entities for which the constituents are more easily distinguishable are more likely to be used as count nouns, while those nouns designating entities for which the constituents are not easily distinguishable will be used as mass nouns. For example, she argues that *beans* is more likely to be a count term than *rice* since individual beans are in principle easier to distinguish than individual grains of rice. Middleton et al. (2004) examine this hypothesis experimentally, devising a task in which subjects had to match a nonce count or mass term with one of two graphical displays of novel aggregates which varied in distinguishability. The graphical displays of novel aggregates were sets of 40 elements where "each element was a simple shape with a black-to-white gradient that appeared slightly 3-dimensional and did not obviously resemble the constituents of any familiar aggregate" (p. 382). They then presented subjects with pairs of aggregate displays which varied along the dimensions of spatial proximity to other elements (*Close* versus *Apart*) and size of elements (*Large* versus *Small*). For example, a subject would see two sets of an element where for one set, each element was large and spatially separated from the other and for the other set each element was small and spatially contiguous with other elements. The subject would then decide which picture aligned with a phrase such as "This is worgel." The general results were that subjects' choices of countable or non-countable terms were significantly influenced ($p < .001$) by spatial proximity, but not by the size, of the elements.

The second factor argued for by Wierzbicka (1988) is the canonical manner of interaction with a given entity. She exemplifies this with examples such as the naming of berries in Polish, generally count terms because, she claims, people interact with them one by

one, viz. picking/eating them, while farmers selling berries typically use mass syntax to describe berries since they interact with them in quantities rather than individually. This factor was investigated via novel objects, again by Middleton et al. (2004). They presented subjects with a novel aggregate, "yellow decorative coarse-grained sugar", in a cardboard box, which the subjects then needed to match to one of two phrases presented in count and mass syntax (e.g. "This is worgel/These are worgels"). The experimenters manipulated the mode of interaction with the aggregate. In the baseline condition, the subjects simply observed the material and then were presented with a response sheet to decide which phrase was appropriate. In the interaction condition, the experimenter and the participants used a thin paper-clip implement to scoop up individual grains of the material and insert each grain into a hole in a board distinct from the box containing the material. The participants then were presented with the response sheet to decide which of two phrases was appropriate, one with mass and one with count syntax. The responses for the baseline and interaction conditions are inversely related: a majority of participants in the baseline condition (69%) selected a mass phrase while a majority of participants in the interaction condition (61%) selected a count phrase.[4] While this result is not definitive, it would appear that the mode of interaction with an aggregate can affect the manner by which it is referred to.

These results show in different ways that individuation is not only related to intrinsic features of objects, such as shape, but also related to whether elements are recognized as sufficiently independent from one another, either spatially or functionally, to be construed as individuals. Again, these factors are not due to objective facts about things in the world, but involve how entities are construed by speakers.

Independent supporting evidence for this view of individuation comes from the semantic shift of particular words. Zwicky (2001) discusses nouns which while initially categorized as non-countable, due to changes in circumstances of their use, are re-categorized as countable nouns. One instance is provided by the word *chad*, which designates the tiny pieces of paper left over from punching punch-cards. This word underwent a shift in countability status during the tight 2000 U.S. presidential election. It became critical after the election to count each vote through determining whether the ballot had been completely

[4]This distribution is significantly above chance (p <.05).

punched through, and accordingly to examine the pieces of paper designated by *chad*. Previous to the election, most of those working with punch-card ballots typically used the word as a non-countable noun, but during the election, as individual votes were scrutinized, *chad* was also used as a countable noun. This is not surprising if interaction with entities is related to countability. In earlier circumstances, *chad* was equivalent to left-over paper from punch-cards, but in the election, each chad became significant. A similar shift can be observed for the word *email*. Zwicky (2001) reports that for many speakers *email* had only a non-countable use when the term was first introduced, which as of current writing in 2012, appears to have given way to the countable form also being widely accepted. Thus, the factors discussed in this section are not restricted to experimental settings, but provide plausible routes of explanation for shifts in meaning, routes of explanation which are not available for a strictly grammatical account, for instance. I now consider the combined effect of the different factors considered so far.

3.2.3 Individuation: A Scale or Dichotomy?

As the preceding discussion has indicated, the evidence and arguments concerning the relation between the individual/non-individual contrast and the grammatical countable/non-countable contrast is complex and involves multiple factors. The study of Samuelson & Smith (1999) shows that for natural concrete nouns there is a tight relation between countable nouns and nominal referents designating solid objects with a recognizable shape; yet the relation is not so straightforward between non-countable nouns and nominal referents which are non-solid objects whose categorization is based on material. Prasada et al. (2002) show that whether shape leads to classifying a noun as countable depends on whether the shape is regular or, more importantly, repeated. At the same time, all the novel entities present in the Middleton et al. (2004) experiment have a distinctive and repeated shape, yet the subjects judge the grammatical countability status of the novel nouns in terms of the relation among the elements, viz. the degree of contiguity. Here the implication is that even having a regular and repeated shape is not sufficient for classification as a countable noun, the entity must also be independent, either spatially or functionally. In sum, while there is solid evidence of a relation between different facets of individuation and the grammatical

realization of the number properties of nouns, the complexity of the different factors is not captured by a simple binary feature +/− *individuated*. This section proposes that many of the shortcomings of the individuation account can be met by viewing individuation as a scalar, rather than binary, phenomenon.

Taken together, the various types of experimental evidence reviewed indicates that there are several distinct ingredients to classification in terms of countability that need to be taken into consideration for a full account. Future work may be able to tease apart the interrelations among the factors and their relative importance, but for now, I simply consider the minimal set of distinctions needed. (As the focus is on natural concrete entities, I leave aside factors related to function as these are more pertinent to artifact nouns.) First, the most clearly needed is a distinction between whether the referent designated by the noun has perceptible[5] minimal units. Whether these minimal units are sufficiently individuated to merit being described by a countable noun depends (at least) on the various factors examined in Prasada et al. (2002), regularity and repetition of shape. This distinction between presence or absence of perceptible minimal units is clearly related to several of the distributional properties discussed in section 1.1.2, such as the felicity of nouns as the complement of *to count* as well as the observed behavior with the comparative. Second, as the experiments in Middleton et al. (2004) vividly demonstrate, the contiguity between elements that a noun designates influences the countability status of the noun. Unlike the presence or absence of perceptible minimal units, which makes a categorical distinction between two classes of objects, the contiguity between elements is a matter of degree.

Given the various distinct factors that are relevant for countability, a view on individuation which divides entities into those which are individuated from those which are non-individuated would appear to be an over-simplification. Even though English is often considered to have a countable/non-countable distinction, the differential behavior of nouns across a variety of diagnostics has demonstrated that there are, as pointed out by Allan (1980), "degrees of countability" that are grammatically detectable. Thus, from different perspectives, it is plausible that countability, both in terms of grammatical classes

[5]I use the term PERCEPTIBLE here in order to steer the discussion away from the relation between, for instance, the extension of *water* and individual H_2O molecules which has sometimes been taken to be a worry in the literature (Quine, 1960; Landman, 2011). I will however briefly return to this question in chapter 4.

that can be isolated as well as the factors relevant to individuation, is better viewed as a scalar phenomenon rather than as a binary contrast.

A scalar view on countability in terms of the factors discussed so far would provide a straightforward interpretation for the different entity types discussed in section 3.1 in terms of different levels of individuation. Least individuated are entities without any perceptible elements—the standard description for substances. Those entities with perceptible units, but which are in no way separated from one another, as is the case for granular aggregates such as *rice* or *sand*, are more individuated than substances, but still quite low-ranking in their level of individuation. Those entities with perceptible units that are separable from one another but still connected in some fashion, whether spatially near or functionally united, as is the case for what I will term collective aggregates such as *ants* or *cherries*, are then again more individuated. Those entities whose elements are independent from one another, not connected in a regular manner to other elements of the same class, are the most individuated. I will refer to these different classes as INDIVIDUATION TYPES, and to their scale, given in (2), as the SCALE OF INDIVIDUATION.

(2) substance < granular aggregates < collective aggregates < individuals

Countability has, in fact, been related to scalar structure in prior work, although the set of facts that were under examination were different. Lucy (1992), and later Gentner & Boroditsky (2001), propose that individuation is a continuum between amorphous substances and highly individuated items.[6] The scales proposed by Lucy (1992) and Gentner & Boroditsky (2001) are given in (3) and (4), respectively.

(3) Animacy Continuum adapted from Lucy (1992)
stuff < objects < animals < humans

(4) Individuation Continuum in Gentner & Boroditsky (2001)
amorphous < large simple objects < complex structurally cohesive objects < small mobile objects < vehicles <animals < humans

[6]Another proposal for a individuation continuum is put forth in Contini-Morava's (2000) analysis of the Swahili noun class system.

Both scales given are elaborations of the animacy hierarchy; thus, for both Lucy (1992) and Gentner & Boroditsky (2001) the notion of individuation and animacy distinctions can be conflated into one linear scale. As will be discussed in section 3.4, grammatical number is highly correlated, yet independent from animacy, although this was not generally recognized at the time of Lucy's proposal. Another observation is that the scales in (3) and (4) do not provide a middle ground between substances and objects which was seen to be crucial for entities typically coded as collectives. In sum, the larger cross-linguistic data now assembled imposes a more nuanced view on the relation between notional categories and grammatical number. I now turn to applying the scale of individuation in (2) to the grammatical number systems discussed so far.

3.3 The Scale of Individuation and Grammatical Number Systems

This section argues that the scale of individuation just developed underlies the structure of grammatical number systems. Laying out the interaction between entity types and coding types across the different languages of interest here will demonstrate that grammatical number systems can be viewed as coherent systems whose morphosyntactic operations respond to the lexical semantic needs of different entity types.

3.3.1 Relating Entity Types and the Scale of Individuation

I first relate the scale of entity types developed in section 3.1, repeated in (5), to the scale of individuation, repeated in (6). This will help to clarify the relation between entity type and individuation type, and fix intuitions about the sorts of elements which fall under the different individuation types.

(5) liquids < foodstuffs < granular aggregate < vegatation/cereals/fruits ≤ insects < small animals <pair/grouped body parts ≤ middle-sized animals < types of people < individuals

(6) liquids/substances < granular aggregates < collective aggregates < individuals

The scale of entity types in (5) was elaborated based on the morphological coding across the different languages examined in chapter 2. Yet, this scale based purely on entity types would not be predictive of the structure of grammatical number systems. The problem is that different entities which may be similar in kind, such as different types of plants *rose* and *ivy*, may differ in their countability coding preferences. (This mismatch has traditionally been given as an argument for the alleged arbitrariness of countability classification.)

The last section, however, developed the scale of individuation in (2) based on individuation properties, which cross-cut different types of entities. As noted in the discussion of the data from different languages in chapter 2, there is not a determinate relationship between an entity being a member of a given entity type and the noun designating that entity being a member of a given grammatical number category. For instance, in Dagaare, a noun designating an insect is likely to be a member of the unmarked plural class, but this is not an absolute rule. Rather, it is the behavior of the entity in terms of individuation properties which leads to its classification in one grammatical number category or another. This is again the situation in the other languages. For instance, not all nouns designating small-sized animals in Welsh fall under the collective class, only those whose "characteristic way of living together [is] in swarms, flocks, herds and shoals" (Stolz, 2001, p. 65). This indicates that it is in terms of individuation properties, and accordingly the scale of individuation, that generalizations about the relationship between meaning and grammatical number categories can be established.

While there is no direct relation between the scale of entity types in (5) and the scale of individuation in (6), the scale of entity types does contain information about the likely members of individuation types in a given language. For instance, if a language treats mid-sized animals as falling in the collective class, then it is likely that it will also treat small animals as falling in the collective class.

From the comparative data that is available, it appears that languages are comparable in how they classify certain entity types under certain individuation types, such as large animals being classified as individuals and cereals being classified as granular aggregates, no doubt due to the causal regularity of the world (see Davidson 1973). Yet, there is much room for disagreement among languages and cultures as to whether a term should be viewed in one fashion or another. The discussion of the difficulties surrounding the

translation of the collective/unit distinction in Welsh showed exactly that. This all is to say that wide comparison across different languages will surely reveal similarities as well as differences. The question to be answered by future typological work concerns the degree to which cross-linguistic variation in grammatical number classification of different entity types, i.e. whether the entity types are more or less countable, reliably correlates with differences in construals of the entities as more or less individuated in those language communities.

3.3.2 Relating Individuation Types and Morphological Coding

Having proposed that individuation types are the proper level from which to examine the underpinnings of grammatical number systems, I now proceed to examine the relation between individuation type, coding and grammatical class across the languages under discussion. Although Welsh, Turkana and Maltese all differ in terms of their manner of grammatical expression and the entity types which fall into their grammatical categories of number, at the level of individuation types, they are comparable as they all possess tripartite number systems which treat *collective aggregates* and *granular aggregates* as falling under a single grammatical category. As discussed in chapter 2, this category is morphosyntactically distinct in comparison with the grammatical categories which *liquids/substances* and *individuals* fall under.

Table 3.2 displays the individuation types, ordered according to the scale of individuation, along the top. The language is named in the leftmost column. Here I use Welsh as a representative of the tripartite systems. The table displays each of the grammatical categories in Welsh as a shaded region spanning the relevant entity types. Each of the shaded regions also indicates the grammatical category's manner of coding: (i) no number coding (*0*), (ii) zero-coded non-unit interpretation with a coded unit (*-yn*), (iii) or zero-coded unit reading with a coded plural (*-od*). The table shows how a tripartite number system partitions the entity types into different grammatical categories.

The grammatical number system of Dagaare is represented in table 3.3. Here, the four individuation types each manifest a distinct grammatical category of number, in contrast to the tripartite systems.

Language	liquids/ substances	granular aggregates	collective aggregates	individual entities
Welsh	0	0/Singulative (*–yn*)		0/Plural (*–od*)

Table 3.2: The Grammatical Number Categories of Welsh along the Scale of Individuation

Language	liquids/ substances	granular aggregates	collective aggregates	individual entities
Dagaare	0	0/Singulative (*–ruu*)	0/Singular (*–ri*)	0/Plural (*–ri*)

Table 3.3: The Grammatical Number Categories of Dagaare along the Scale of Individuation

Table 3.4 compares Dagaare, Welsh and English in terms of the scale of individuation. Each of these systems differs substantially from the others, yet they all still are comparable: they all partition the scale of individuation into a discrete set of grammatical categories. English's grammatical number system does not make the more fine-grained distinctions that Welsh and Dagaare do, which can be seen visually in that the class of nouns which are non-countable in English is larger than in the other two languages.

The general picture that arises from this table is that morphosyntactic classes appear to be carved out of the scale of individuation. In this sense, the table represent both the grammatical and meaning-based facets of countability. Morphosyntactic categories of number are grammatical phenomena—different languages have established different categories. But these categories are based in individuation types, or coherent combinations of individuation types, and therefore are also grounded in the meaning of nouns. Languages may carve up the space defined by the scale in different fashions, but all of these languages are responding to a similar functional need: to provide means of quantifying individuated entities, while not being burdened with count morphology for nouns which are not (sufficiently) individuated. In the remainder of the section, I examine the scale's typological implications and how several languages of different types than those investigated in chapter 2 fit with the scale of individuation.

Language	liquids/ substances	granular aggregates	collective aggregates	individual entities
Dagaare	0	0/Singulative (*–ruu*)	0/Singular (*–ri*)	0/Plural (*–ri*)
Welsh	0	0/Singulative (*–yn*)		0/Plural (*–od*)
English	0		0/Plural (*–s*)	

Table 3.4: The Scale of Individuation: Dagaare, Welsh and English

3.3.3 Typological Predictions

In this section I will set out some of the predictions from the scale of individuation. First, the core prediction is that the grammatical number system of a given language will respect the structure of the scale. In particular, there should not be systems where a category of grammatical number spans two disconnected segments of the scale. An example of a system which would violate this condition is shown in table 3.5 where individuals and granular aggregates both belong to a singular/plural class while collective aggregates form a distinct class, which results in the singular/plural class as being discontinuous along the scale of individuation. As a result, grammatical number systems should partition the semantic space of the scale into only as many segments as the language has categories of grammatical number. These predictions will be refined in section 3.4 when the influence of animacy is also considered.

Language	liquids/ substances	granular aggregates	collective aggregates	individual entities
Bad System	0	0/Plural	0/Singulative	0/Plural

Table 3.5: Hypothetical Grammatical Number System Violating the Prediction of the Scale of Individuation

A further prediction concerns the coding of the different number categories, or "markedness" patterns. There is a choice to be made as to which countability value is the zero-coded, or "unmarked", value. For instance, for countable nouns in English the zero-coded value is the singular, while the coded, or "marked", value is the plural. In contrast, for the collective class in Turkana, the plural value is the zero-coded value while the singular is the coded or "marked" value. The prediction that the scale makes is that the higher the level

of individuation of a grammatical class, the more likely the designation of singular entities will be the default, while the lower the level of individuation of a grammatical class, the more likely the designation of multiple entities will be the default. The limiting case is the portion of the scale for which there is no countability contrast.

3.3.4 Other Types of Number Systems

I now briefly show how the scale of individuation aligns with other grammatical number systems beyond those discussed in chapter 2. The finding will be that although these systems have very different means of manifesting grammatical number, e.g. classifiers, the categories of nouns that can be distinguished accord with the scale of individuation.

Miraña Miraña, a highly endangered Amazonian language spoken in southern Columbia, possesses both nominal classifiers and inflectional number. My discussion closely follows Seifart (2009) who carefully distinguishes different types of nouns through the application of different distributional properties. First, there is a primary division between countable and non-countable nouns, detectable by whether a noun may combine with inflectional number markers. The examples in table 3.6 show the differing behavior of nouns according to this distributional property. Non-countable nouns do not permit number inflection, while it is obligatory for countable nouns when expressing multiple entities.

Countability Category	Base Form	Plural Inflected Form	Gloss
Countable	kaːni	káːni-mɯ	‘father’
	ɨːβa	ɨːβa-mɯ	‘macaw’
Non-Countable	kaː	*ka-mɯ	‘ants’
	íːnɯ	*íːnɯ-ːnɛ	‘earth’

Table 3.6: Countable and Non-Countable Nouns in Miraña

Non-countable nouns instead combine with class markers to refer to single objects and these forms then permit inflectional number (as well as cardinal numbers). This is shown in examples (7)–(8).

(7) a. íːnɯ
earth
'earth'

b. íːnɯ-ba
earth-SCM.3DIM
'piece of earth'

SCM = Specific Class Marker

c. íːnɯ-bá-ːkɯ
earth-SCM.3DIM-PL
'pieces of earth'

(8) a. kaː
ant
'ants'

b. ka-ʔba
ant.SCM.3DIM
'an ant'

c. ka-ʔbá-mɯ
ant.SCM.3DIM-PL
'some ants'

Among nouns which are non-countable as determined by their inability to directly take inflectional number, there is a further distinction to be made for nouns which in their bare form designate a collection. Seifart observes that many nouns, mainly animate but not human, differ from nouns designating inanimate objects or stuff: "even though the non-unitized forms do not take number inflection, the singular-plural distinction is relevant for them in the sense that the underived form refers to groups of animals and the unitized form to single animals" (p. 39). This can be seen with the word for ant in (8), where the bare form refers to groups of ants, and the application of a class marker results in reference to single ants. The plural form, according to Seifart, is referring to some smaller number of entities, in this case glossed by "some ants". Thus, this class of nouns in Miraña, where the reference to collections is the default, aligns with the collective/singulative classes that have been explored.

In sum, although the grammatical means employed differs sharply from the systems examined in chapter 2, Miraña makes divisions between non-countable nouns, aggregate nouns and countable nouns, which align with the predictions of the scale of individuation.

This is summarized in table 3.7. Another noteworthy point is that the types of entities at issue in Miraña's aggregate category partially overlap with what was expected from the discussion of entity types in section 3.1, namely it includes insects, but many other animates, such as rats, are also included.

Language	liquids/ substances	granular aggregates	collective aggregates	individual entities
Miraña	0		0	0/-mɯ

Table 3.7: The Grammatical Number Categories of Miraña along the Scale of Individuation

Yudja I now turn to a language which serves as a limiting case. Yudja, another Amazonian language (Juruna family, Tupi stock, spoken in Brazil), is interesting in that a countable/non-countable distinction is only very weakly present in the language. My discussion follows Lima (2010). Nouns in Yudja are allowed to appear bare, unspecified for number or definiteness, as shown in (9), from Lima (2010, p. 159).

(9) ali ba'ï ixu
child paca eat
'The/a/child(ren) eat(s)/ate the/a paca(s)'
Literal: an undefined number of children eat(s)/ate an undefined number of pacas.

Lima (2010) reports that the distribution of numeral modification in the language is unrestricted, finding that numerals are able to combine freely with nouns designating substances or individuals; thus, they do not serve to distinguish between countable and non-countable nouns. Some examples are given in (10) and (11) (Lima's (39) and (37)), which are reported as acceptable either in conventionalized contexts, e.g. containers of sand/blood, or ad hoc, unconventional contexts, e.g. clumps of sand or drops of blood. Lima (2010) further claims that in Yudja "all nouns can be combined with all quantifiers without restriction" (p. 160).

(10) Yauda ali eta apapa
two child sand drop.REDUPL
'Children drop two sand(s)'

(11) Txabïu apeta ipide pepepe
three blood on the floor to drip.REDUPL (three events)
'Three bloods dripped on the floor'

The sole discriminating piece of number morphology is an optional plural morpheme *-i* which is restricted to human nouns (Fargetti, 2001). Lima further notes that when a human noun has plural reference, use of *-i* is preferred. Clearly, Yudja provides an extreme case in which the grammatical traits of a countable/non-countable distinction are hardly present. This system is, however, still consistent with the scale of individuation: only very highly individuated entities (humans) manifest grammatical number, while everything lower on the scale is unspecified.

Language	**liquids/ substances**	**granular aggregates**	**collective aggregates**	**individual entities**
Yudja			0	0/Plural (*–i*)

Table 3.8: The Grammatical Number Categories of Yudja along the Scale of Individuation

Kiowa The application of the scale of individuation also extends to much more complicated systems. One of the most complicated systems known is that of Kiowa (Kiowa-Tanoan, spoken in Oklahoma). Like Dagaare, Kiowa has an inverse number marking system, but also makes a three-way distinction between singular, dual and plural number values. The complexity of the system prevents giving a thorough treatment here, but I will simply remark that the detailed treatments of Watkins (1984) and Harbour (2008) point out that the different classes of grammatical number correspond to natural lexical semantic classes which align with what one might expect with regard to the scale of indiviudation.

Harbour (2008), expanding upon Watkins (1984), establishes nine classes of nouns, including five major ones: (i) animates or inanimates which are capable of motion, (ii) "plants and plant material, natural and man-made objects and a small number of body parts" (Watkins 1984, p. 85), (iii) vegetation forming natural collections or other nouns disposed towards referring to a cohesive group, (iv) pluralia tantum, abstract nouns, and for many speakers granular aggregates, and (v) substances such as milk or honey. There are many interesting nuances, such as the fact that many nouns belong to both class (ii)

and (iii), which then highlight distributive and collective readings, respectively. For the moment, it suffices to note that for the major classes, they align rather well with what the scale of individuation would predict—ranging from highly individuated entities (class i) to vegetation and general objects (class ii) to collectives (class iii) to pluralia tantum and granular aggregates (class iv) and finally to substances (class v). To fully connect this system with the scale of individuation, the scale needs to be able to represent both animacy and the dual. I treat animacy in section 3.4, but leave the dual for future work since a comprehensive treatment would take us too far afield.

Summary The different languages examined here exhibit vastly different grammatical number systems in terms of their morphology and degree of complexity. Yet, each can be seen as reflecting the organization of the scale of individuation in their own way. It is worth contrasting these results with a view which proposes that countability reduces to a binary +/−*individual* distinction. Such a view would confront difficulties in the face of Kiowa's wealth of classes, as well as be forced to conclude that what speakers of Yudja consider to be individuals is very restricted compared with what speakers of other languages do. Treating individuation as a scalar phenomenon does not lead to these difficulties. A scalar view on individuation provides a way to address the richness of different number systems. Further, it takes the notion of individual to be a graded concept, and thus it is expected that some languages, such as Yudja, might restrict their expression of grammatical number to only the types of individuated entities which are highest in the scale. I now turn to exploring the relation between the scale of individuation developed in this section and animacy, followed by its relation with frequency.

3.4 Countability and Animacy

The discussion so far has been devoted to untangling the different distinctions grammatical number systems make, focusing on the classification of entities as collectives and how this interacts with grammatical number systems in general. I have been arguing that a scalar structure, based on individuation, underlies countability phenomena; however, a look at a wider range of data quickly shows that there are other potential factors that contribute to

the phenomena to consider as well. This section addresses the influence of ANIMACY, which is another important factor in nominal countability.[7]

Animacy has been reliably tied to the manifestation of plural marking in a large number of languages. Yet, the relation between animacy and other types of grammatical number marking, such as the collective/singulative or the dual, has proved to be something of a puzzle. I will show in this section that viewing the manifestation of grammatical number as grounded in individuation also leads to understanding how the factor of animacy influences to the inventory of nouns which fall in the collective/singulative category in a given language.

The pioneering typological study by Smith-Stark (1974) demonstrates that the degree of animacy of a noun's referent is correlated to the likelihood of that noun being able to express plural marking. Smith-Stark (1974) develops the scale in table 3.9, which he considers more properly as a scale of the likelihood of participation in the speech event. The associated claim is that if a language expresses pluralilty for noun types in a given position on the hierarchy, then the language will also be able to express plurality for any noun type which is higher on the hierarchy.

inanimate < animate < human < rational < kin < addressee < speaker

Table 3.9: The Smith-Stark Hierarchy

An instantiation of this prediction is provided by Tamil, as shown in table 3.10. The shaded area of the animacy hierarchy in table 3.10 indicates the portion which manifests a singular/plural contrast. According to the discussion in Smith-Stark (1974, p. 662), nouns designating rational beings typically manifest plurality while those designating irrational (non-caste) nouns, including *child* and *infant*, are rarely pluralized. Smith-Stark's claim is that any noun type located higher on the hierarchy than rational nouns would also have obligatory number marking which is the case with, say, pronouns which all manifest a singular/plural distinction.

[7]Less studied, but equally intriguing, is the interaction between referentiality and/or definiteness and number marking. Although pursuing this question is beyond the scope of this work, this is a fruitful avenue for future work. See discussions of the interaction between number and referentiality from very different perspectives in Corbett (2000); Gil (2003); Schwarz (2010); Grimm (to appear (b)). See also Behrens (1995,

inanimate < animate < human <	rational < kin < addressee < speaker

Table 3.10: Tamil's Grammatical Number System along the Smith-Stark Hierarchy

While the hierarchy in Smith-Stark (1974) has proven remarkably robust for predicting the occurrence of plural marking, its relation to other grammatical number categories has remained opaque. Corbett (1996) discusses data pertaining to what he terms "minor number" categories, which are essentially any number category that is not the singular/plural category, such as dual, paucal, collective or mass. Corbett (1996) provides data from eight different languages which proves problematic for the Smith-Stark hierarchy. For all the examples Corbett considers, the number categories, such as the collective, apparently are not sensitive to the animacy level of the noun.

A case in point is the collective in Maltese, which, as Corbett points out, contains a mixture of animate ('flies') and inanimate ('corn','shoes') nouns, but does not pertain to, for instance, 1st and 2nd person pronouns, violating the predictions of the Smith-Stark hierarchy. This is shown in table 3.11, where the shaded region indicates for which part of the hierarchy the collective/singulative is expressed.

inanimate < animate <	human < rational < kin < addressee < speaker

Table 3.11: Smith-Stark Hierarchy Violated by Maltese Collective

The question arises, then, what precisely the relation is between number categories such as the collective and the animacy hierarchy. From the data considered so far it would seem that the two are independent—the use and function of a number category such as the collective is picking up on attributes of entities that are not directly related to animacy. This is the conclusion that Corbett comes to: while the Smith-Stark hierarchy governs the portion of the lexicon which grammatical number generally applies to, there may be "patches" of the grammar where minor number applies. In these cases, a very particular semantics is at play. Indeed, for some of the examples Corbett considers, such as the collective plural in Budugh (Lezgian; spoken in northeastern Azerbaijan) which has a very

p. 57 ff.) who argues for several other dimensions to the countability distinction.

limited distribution, being restricted to just five nouns referring to paired body parts, much more does not need to be said. Yet, for the systems considered in chapter 2, which have smaller categories of number which are nonetheless productive, one would hope something more general could be said about the relation between number categories, individuation types and the animacy hierarchy.

Viewing grammatical number as related to individuation provides the key to understanding the behavior of the collective/singulative class in relation to the animacy hierarchy. In order for nouns to be members of the collective/singulative class, the tendency to view the entity as habitually occurring in a group must be greater than the tendency to view the entity as habitually occurring singly. Yet, the higher one ascends in the animacy scale, the more individuated the entities are, and the greater the tendency to view them as occurring singly as opposed to coming in groups. Thus, from the viewpoint of individuation, the occurrence of the collective/singulative category should be *inversely* related to the animacy hierarchy. In other words, if a language disposes of a collective/singulative class, the higher a noun rates on the animacy hierarchy, the less likely it will fall in the collective/singulative class. This is the inverse of the claim Smith-Stark (1974) makes for plural marking.

This discussion can be put in a more precise form by representing the different possibilities that arise when the scale of individuation interacts with the animacy hierarchy. A simple technique for multiplying different linear scales was demonstrated by Aissen (2003), and I will refer to it in what follows as an Aissen lattice. In our case, it is necessary to take the product of the countability scale and animacy hierarchy. I will modify the animacy hierarchy from the form given by Smith-Stark (1974). First, I will not represent the categories above *human*, as they are not relevant for the distribution of the collective. Second, it has been noted that many languages do not treat all members of the category *animate* equally in terms of number marking. Haspelmath (2005) notes that the category of animates is often subcategorized into "higher" and "lower" animates, which I will adopt here. Such a distinction was already seen in effect with regard to Miraña, which categorized salient animates, such as those of high cultural value ('macaw'), with humans in the singular/plural class, as opposed to other animates whose nouns in their base form had collective reference.

The lattice in figure 3.1 exhausts the combinatoric space of the two scales. For the segment of the individuation scale from *collective aggregate* and higher, the entirety of the

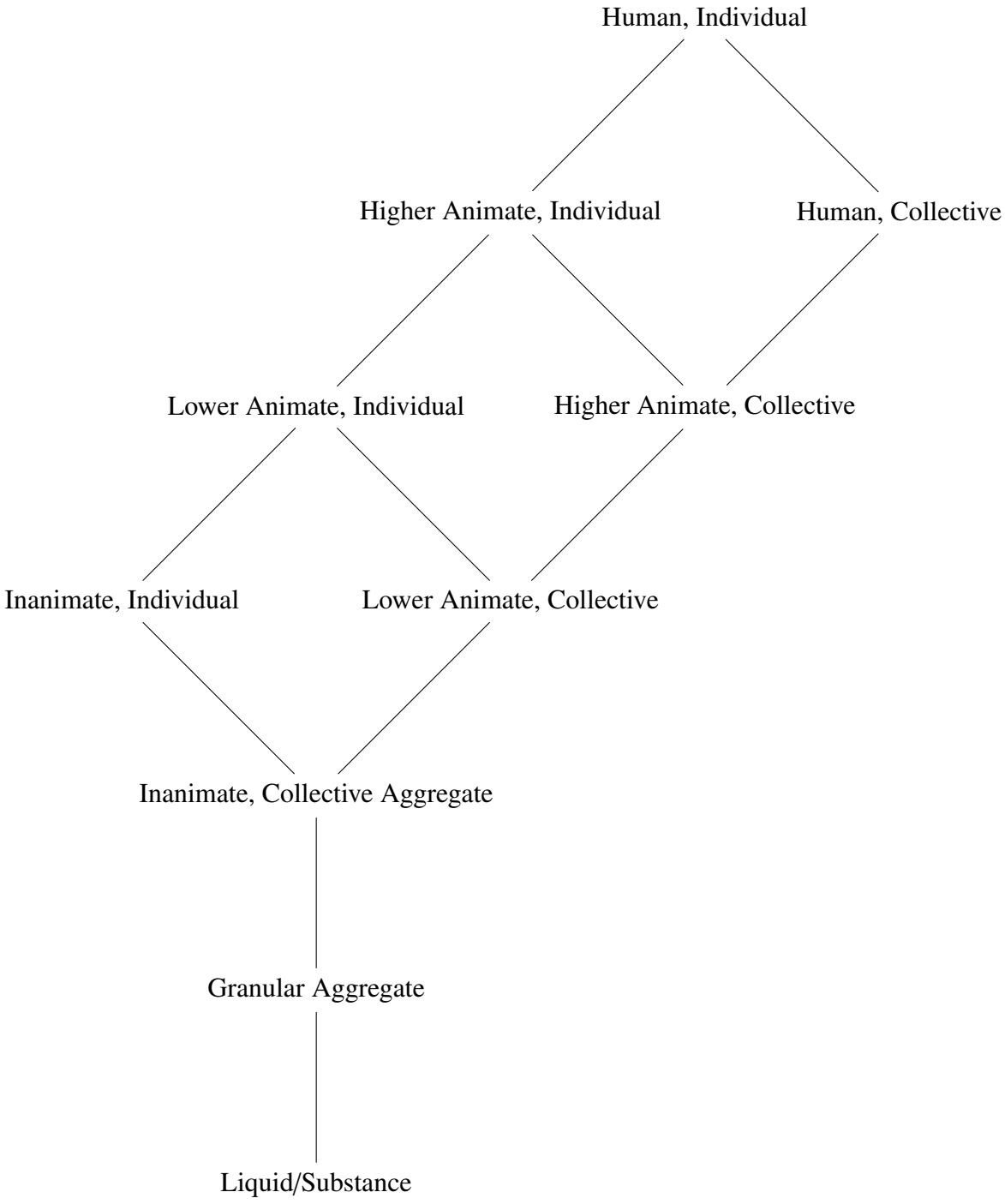

Figure 3.1: The Lattice of Animacy and Individuation

(simplified) animacy hierarchy is attached to each node. I consider that *granular aggregate* and *liquid/substance* are necessarily inanimate, whereby incompatible with the rest of the animacy hierarchy, so they are just represented by single nodes.

In figure 3.2, the systems of Welsh, Turkana, and Maltese are represented on the combined animacy/individuation lattice. The extent of the collective/singulative class in each language is indicated by shading which covers the relevant nodes of the lattice. While at the abstract level of the individuation hierarchy, these three languages are highly similar, through this mapping we can see differences across the different systems.

Returning to the relation between the animacy hierarchy and minor number categories, it appears that there is an interaction between the distribution of the collective categories in these languages and the animacy hierarchy. Yet, unlike plural-marking across languages, which differs in how far it descends the animacy hierarchy, the collective/singulative classes in these languages differ in how far they *ascend* the animacy hierarchy. For instance, the Welsh collective/singulative class contains inanimate nouns, insects, small- and mid-sized animals. Turkana's contains those and in addition human nouns, thus it includes a larger, but still connected section of the hierarchy. The collective/singulative category in Maltese, however, spans a smaller segment of the animacy hierarchy. Maltese has a restricted number of animate members in its collective/unit category, essentially limited to insects. This restriction is somewhat surprising given that this category historically developed from the Arabic collective, which, like Turkana's, contained nouns designating larger animals, such as cows, and collections of humans, such as names of nationalities or professions. Mifsud (1996) notes that there has been a historical shift whereby in Maltese these nouns were integrated into the singular/plural category. The important point for the moment is that this decrease in membership seems again to proceed along the lines of the animacy hierarchy. What does not appear to be attested is a collective/singulative class which includes many nouns referring to humans, and many referring to inanimates, but none to animates, i.e. one that would be discontinuous.

Given this evidence, the conjecture about the structure of grammatical number categories can be refined with respect to animacy. A grammatical number category should not be discontinuous across either segments of the individuation scale or segments of the animacy hierarchy. Stated in the converse fashion, grammatical number categories should

occupy connected regions of the animacy-individuation lattice.

From this perspective, the "minor" number categories do actually respect the animacy hierarchy, but simply in a different fashion than plural marking does. The singular/plural distinction occupies the upper regions of the animacy hierarchy, and any extension into the lower categories is predicted to occur in a continuous fashion. The collective/unit distinction occupies the middle region of the animacy hierarchy, and any extension into the upper/lower categories is predicted to occur in a continuous fashion. The functional grounding for this division of labor is evident: higher animate entities, and certainly speech act participants, are clearly individuated. Thus, if anything will be high on the individuation hierarchy, it will be such entities.

3.5 Individuation, Grammatical Coding, and Frequency

I will now briefly consider the relationship between the proposed individuation account and frequency. Two notions of frequency are relevant here: non-linguistic frequency, i.e. frequency in relation to elements in the external world, or linguistic frequency, e.g. frequency in relation to occurrences of words or morphemes.

As discussed in chapter 2, the singular/plural and collective/singulative classes display coding asymmetries: one form, e.g. the singular, is zero-coded (or "unmarked") while the other form, e.g. the plural, is overtly coded. Frequency has become increasing relevant in typological studies since coding asymmetries such as those found in grammatical number systems have sometimes been attributed to asymmetries in frequency (Haspelmath, 2008). I first discuss the notion of frequency of elements in the external world, then turn to textual frequency. Finally, I discuss the view wherein economy, i.e. using less coding for more frequent forms, and learnability are two contrasting forces which shape language systems and how that view aligns with the patterns found in the different grammatical number systems of chapter 2.

Reference and Frequency It is tempting to think that raw frequency of appearance of the referents of a noun would be causally related to how nouns are categorized in terms of countability, as suggested by Haspelmath (2008). For instance, it is reasonable to suppose

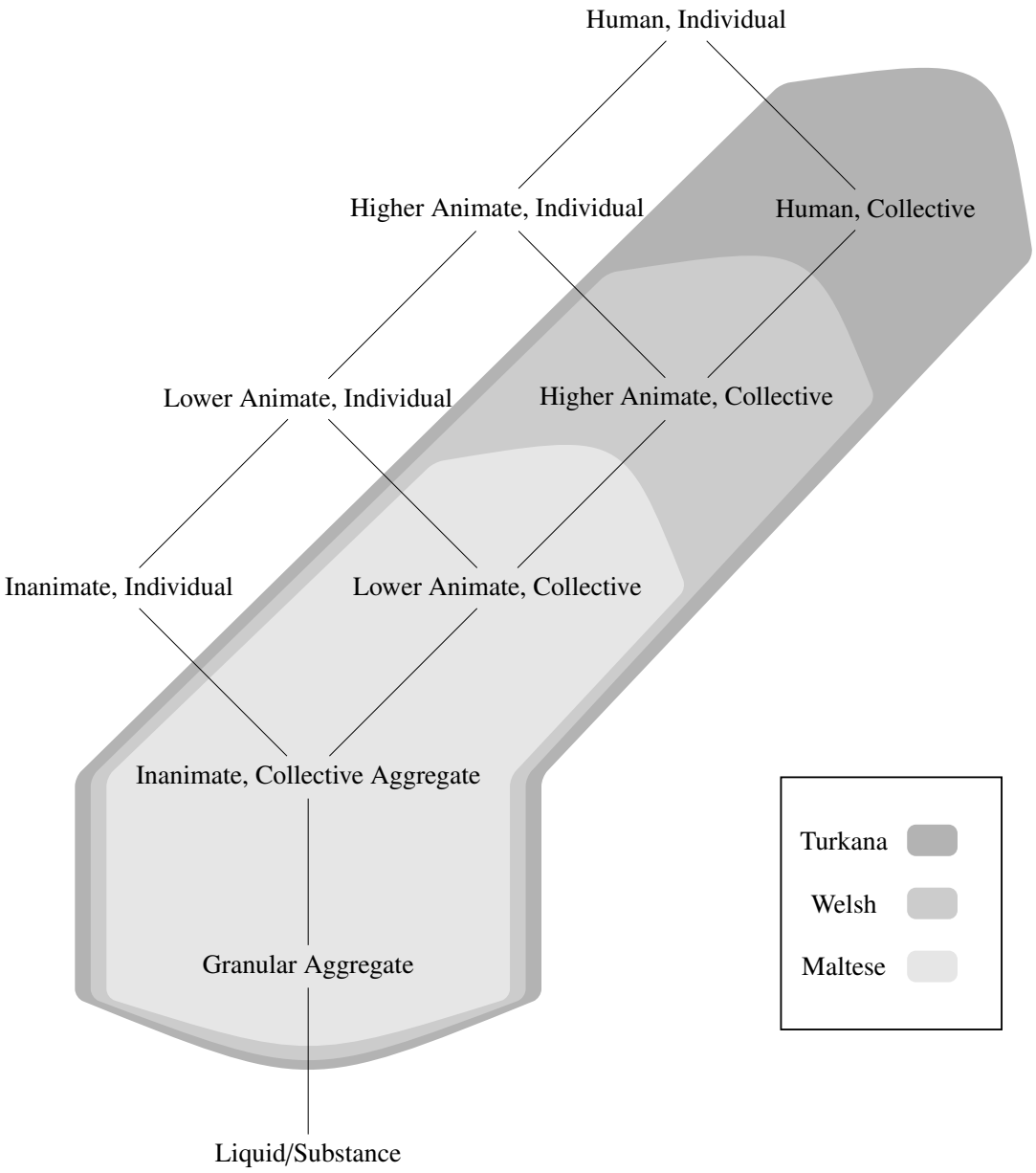

Figure 3.2: The Collective/Singulative Classes in Turkana, Welsh and Maltese on the Lattice of Animacy and Individuation

that since ants typically appear in great numbers, it is more economical to use a word-form for which the plural interpretation is the default. Yet, from the sources consulted for the various languages in chapter 2, frequency of co-occurrence appears to be, at best, a necessary, but not sufficient condition for a noun to fall into a collective/singulative class.

In his examination of Welsh, Stolz (2001) argues explicitly that whether a noun is assigned to the collective/singulative class is not "a matter of quantity" (p. 65). In part, it depends on the behavior of the entities, for instance, for animate entities it is "the characteristic way of living together in swarms, flocks, herds and shoals" (p. 65). Stolz further argues that gregariousness of entities is also not in itself sufficient, providing several counter-examples such as the Welsh term for *cow*, an animal which does come in herds, but is a member of the singular/plural class in Welsh. Rather, it is whether the entities are perceived as habitually coming together[8] which is determinate, according to Stolz.[9] In the discussion of Turkana in section 2.1.2 as well, I provided a quote in which Dimmendaal observed "living in herds or groups" (Dimmendaal, 2000, p. 229) is the relevant characteristic for nouns designating animate entities which are zero-coded in the plural. Further, as observed in the discussion of Welsh and Turkana, there is a meaning contrast between collective and plural values, and therefore it seems unlikely that nouns are categorized in the collective/singulative class simply by virtue of the noun typically referring to multiple entities. Rather, the noun must refer to entities habitually appearing in some number, which additionally must be viewed as linked together, for instance, through collective living, or related to a common source, e.g. multiple berries related by growing on the same branch of a bush. In sum, if an entity appears in multiples, that is a precondition, rather than the ultimate explanation of why a noun may be classified in the collective/singulative class.

[8]Given the discussion in the preceding sections, one could also speculate that the type of interaction also has an influence—namely, since there is more interaction with individual cows (e.g. milking) than with pigs or rabbits, cows would be more likely to be treated as individuals and be categorized in the singular/plural class.

[9]Stolz (2001) is similarly critical of the factor of size, noting that while there is a high proportion of nouns which designate small entities or animals, this does not cover all the cases, and appears to be secondary consideration compared to swarming or herding behavior. This would appear to align with the results of the first experiment of Middleton et al. (2004), presented in section 3.2.2, where size is not a significant factor.

Textual Frequency Coding asymmetries, and the notion of "markedness", have been related to textual frequency, as in Greenberg (1966), where zero-coded forms are shown to have greater textual frequency than overtly coded ones. The frequencies relevant here are the frequency with which an entity is spoken of as referring to multiple referents as compared to the frequency with which it is spoken of as referring to a singular referent. By the same logic as in the preceding paragraphs, textual frequency is not enough to fully explain how nouns are categorized in terms of countability. Still, this section demonstrates that there is a correlation: nouns which are typical members of a collective/singulative class are also observed to have greater textual frequency in the plural.

If the semantic domains relevant to "unmarked plurals" in languages like Welsh or Dagaare are such that speakers more frequently refer to multiple referents than singular referents, then languages which do not possess morphologically unmarked plurals should still display asymmetries in the same semantic domains in terms of textual frequency. In order to evaluate this prediction, I examined frequencies for nouns in the semantic domains of animal and insect from the COBUILD corpus (18 million words) provided by CELEX (Baayen et al., 1996). Using basic terms and terms which had correspondents in the vocabulary of Dagaare, I calculated the plural-to-singular ratio for these two domains, shown in figure 3.3, where the x-axis represents the ratio of the token frequency of plurals to the token frequency of singulars and the y-axis represents the number of lexical items. The graph indicates that there is a clear trend for insect terms to have a plural/singular ratio greater than 1, i.e. insect terms occur more frequently in the plural, while animal terms tend to have a plural/singular ratio less than 1, i.e. animal terms occur more frequently in the singular.

Similar evidence is provided by Baayen et al. (1997) who investigate grammatical number morphology and the phenomenon of "local markedness" (Tiersma, 1982) in Italian from a psycholinguistic perspective. They first isolate a set of, what they term, "plural dominated" nouns in Italian (*dent-e* 'tooth', *capell-o* 'hair', *pied-e* 'foot', *gamb-a* 'leg', *scarp-a* 'shoe') which they contrast with "singular dominated" nouns (*nas-o* 'nose', *piazz-a* 'square', *ombr-a* 'shadow', *region-e* 'region', *pont-e* 'bridge'). The examples of plural dominated nouns cited already suggest that these will align with the nouns at issue in the collective/singulative classes discussed in chapter 2. The Italian data thus provides further

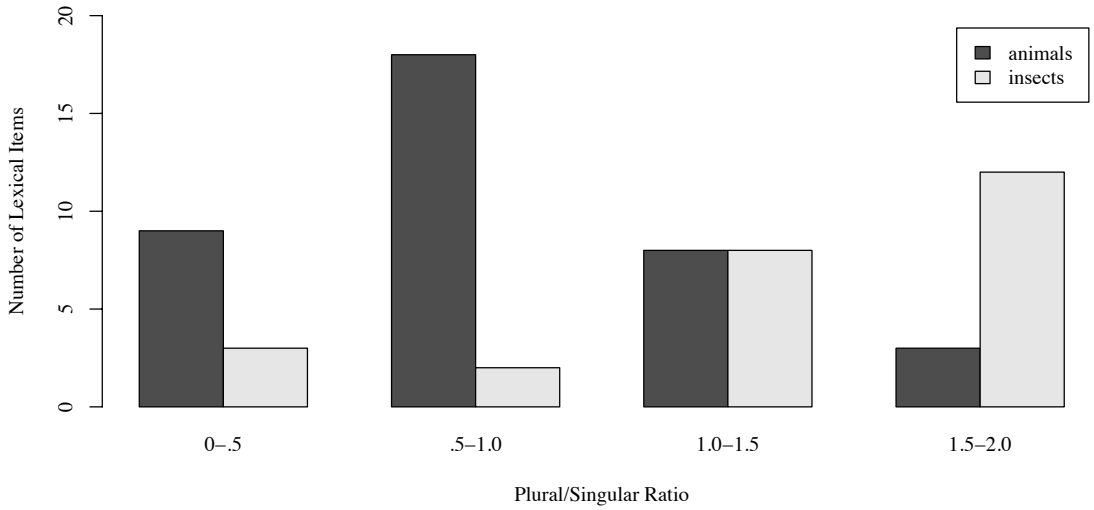

Figure 3.3: Number Marking Frequency Patterns in English for *animals* and *insects*

support for the connection between entity types that show up as morphologically zero-coded in some languages and frequency of expression in the plural form. Baayen et al. (1997) further argue, based on evidence from a lexical decision task, that plural forms of plural dominant nouns are stored separately in the mental lexicon. These results indicate that there is a testable difference between ordinary plurals and plural dominated plurals, which in terms of frequency corresponds to the difference in morphological patterning that was observed in chapter 2.

Economy and learnability Given the correlation between the entity types which fall into collective/singulative classes and textual frequency, it stands to reason that zero-coded plurals or collective countability classes are beneficial from the point of view of economy of expression.[10] In English, for many words such as the insects in figure 3.3, one must more often than not pronounce an *-s*, an option less economical compared to treating multiple reference as the default. In languages such as Dagaare or Welsh, the default form simply

[10]I would like to thank Paul Kiparsky for leading me to the connections in this section.

corresponds to the more frequently used form.

The trade-off for increased economy gained through default forms which refer to multiple referents is increased complexity of what must be learned. While in English, children must only learn two categories of nouns with respect to number: non-countable nouns and those with a singular/plural contrast. In Welsh or other languages with a collective/singular class, there is an extra category which must be learned. In fact, this appears to require some effort on the part of children acquiring such systems. Although research on acquisition of such grammatical number systems is still largely unexplored, some results indicate that collective categories are more complex to master. Ravid & Hayek (2003) have investigated the acquisition of collective forms in Palestinian Arabic in children from ages 4-8, where their results indicate that collective nouns in Arabic are quite difficult to acquire. While the oldest children performed with around 85% accuracy for duals and plurals, they only had, at best, 50% accuracy for collective forms.

In sum, while frequency alone is not sufficient to explain the coding patterns of the different grammatical systems discussed here, there is an intimate link between which word-forms correspond to the default situation in the world and textual frequency. In turn, this supports viewing languages such as Dagaare as providing an economical grammatical number system.

3.6 Resolving Challenges to Meaning-Based Account of Countability

This chapter has made the case that countability should be understood in light of several, interrelated, elements: grammatical number categories, individuation properties and entities in the world. As discussed in chapter 1, this position is not to be taken for granted—many researchers have argued that countability does not involve anything beyond a grammatical classification or that countability distinctions are simply epiphenomenal. For these researchers, the distinction between countable and non-countable nouns is either an (ultimately) arbitrary fact about the grammars of different languages or, more simply still, an empty notion, respectively. I have been arguing that the distinction between countable and

non-countable nouns is a substantive one—entities which are construed as highly individuated according to various properties are likely to be named by countable nouns, while entities which are not construed as individuated are likely to be named by non-countable nouns. This categorization is not always clear-cut, but rather based on multiple nominal properties and may be realized in different fashions depending on the possibilities of the grammatical number system at issue.

This section revisits some of the challenges for giving an account of countability which invokes nominal meaning. I first examine cross-linguistic and lexical variation in countability classification. I then examine the arguments of Barner & Snedeker (2005), who have claimed that the relation between grammatical number categories and the presence of individuals in a noun's denotation is contrary to the predictions of an individuation account. Finally, I explore the phenomenon of nominal flexibility, viz. the fact that *apple* can appear in countable contexts (*three apples*) and non-countable contexts (*apple in the salad*), which has sometimes been taken as an argument against a meaning-based account of countability.

3.6.1 Cross-linguistic Mismatches and Lexical Doublets

Two prevalent arguments for dissociating countability from nominal meaning concern (i) cross-linguistic variation in how different entities are categorized in terms of countability and (ii) nouns in the same language which while near-synonyms, differ in countability categorization. I repeat the quote, given in section 1.2, from Palmer (1971) about the separability of grammatical and semantic distinctions:

> "It is easy enough to show that grammatical distinctions are not semantic ones by indicating the many cases where there is not a one-to-one correspondence. ... examples are to be found in *foliage* [mass] vs. *leaves* [count], in English *hair*, which is singular, vs. French *cheveux*, plural. These distinctions are grammatical and do not directly correspond to any categories of meaning." (Palmer 1971, p. 34–35)

This is the standard argument against basing countability in the meaning of nouns. For instance, Chierchia (2010a, p. 150–153) (or Rothstein (2010, p. 346–348) following him)

gives the same line of argumentation as quoted here, before going on to argue for a grammatical account of countability.

> "The world can be well made of substances and discrete entities ... But the mass-count distinction is something else. For one thing, the two distinctions simply do not coincide. Moreover, languages appear to have some freedom in how they classify their nouns. We must conclude that the mass-count distinction does not appear to be readily and completely reducible to any known extralinguistic one." (Chierchia, 2010a, p. 153)

This critique then is of a semantic theory of countability where there is a direct, one-to-one correspondence between, for instance, discrete entities in the world and countable nouns and substances and non-countable nouns. These authors are certainly right in pointing out that such a proposal would not work—for such a simple distinction would be inadequate to address all the facets of countability that have come up so far. Further, as pointed out in Bloom (1990, p. 102), this simplistic view where countability is directly based on the external world has, however, only very rarely been actually proposed. As far as I am aware, this view was only seriously considered in quite early discussions in the philosophy of language literature (Grandy, 1973; Moravcsik, 1973; Cheng, 1973). Aside from that, this view is mainly found as a foil in critiques of meaning-based approaches.

These arguments given against a direct, one-to-one relation between entities in the world and countability categories succeed against that particular view, but certainly do not rule out all accounts of countability based in nominal meaning. In fact, a view relating countability to individuation entirely agrees that the explanatory burden should not be placed on the external world, rather on *construals* of items in the world. For instance, the work of Prasada et al. (2002) explicitly sets out to determine which factors are relevant to construing objects in the world as countable or non-countable nouns (and does not set out to determine which objects in the world are countable or non-countable). Likewise, the experimental findings of Middleton et al. (2004) cannot be attributed to the extensional level—for the entities did not change, only their configuration or participants' interaction with them. Most work on countability which invokes the notion of individuation has clearly distinguished construals of objects from objects in the world. (See McCawley 1975; Mufwene

1984; Bloom 1990; Wisniewski et al. 2003, inter alia). Even the early proposal of Quine (1960) shows awareness of this and argues that "The contrast lies in the terms and not in the stuff they name ... 'shoe' ... and 'footwear' range over exactly the same scattered stuff" (p. 91).

In sum, these critiques should be recognized as pointing out data that a meaning-based account must cover, which I now turn to; however, they do not force one to abandon a meaning-based account and accept a purely grammatical account.

Reconciling Cross-Linguistic Mismatches The cross-linguistic mismatches demonstrate that the grammatical realization of these semantic categories is not identical across languages. An initial comment is that this would appear to be a common situation. A large number of grammatical phenomenon with semantic content are recognizably comparable across large numbers of languages, say, for instance, the use of perfect tense or modals, even while the precise contexts of felicitous use may vary from language to language. The fact that there is variation in how particular words are categorized in different languages does not invalidate that there is clearly a core phenomenon that is in need of accurate description and theorizing.

The general view reached here provides a useful perspective, which if adopted, reveals the problems related to mismatches to be only apparent. First, it is useful to unpack all the elements that have come up in the discussion so far. Figure 3.4 displays four different levels that have been discussed: things-in-the-world, lexical nouns and their properties, individuation types and grammatical classes.[11]

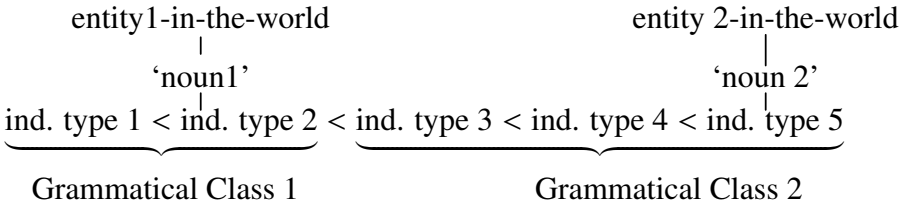

Figure 3.4: Mapping between Things-in-the-World, Lexical Nouns, Individuation Types and Grammatical Classes

[11]This separation of different levels is implicit in much work which emphasizes individuation, and is also explicitly laid out in a similar fashion in Bloom (1990, p. 107)

Displaying the different elements in this fashion allows us to clearly see that there is room for indeterminacy in how an entity in the world may be categorized in a grammatical number category. Considering the relation between grammatical categories and individuation types, i.e. the bottom two levels of figure 3.4, there is indeterminacy present in how different languages realize different individuation types. This indeterminacy was visible in table 3.4 where the individuation type of *granular aggregates* was realized in three different manners: non-countable (English), collective/unit (Welsh) and uncountable with optional singulative (Dagaare). It appears clear from the comparative work of the preceding chapter that grammatical number systems put together different individuation types in different manners. Further, these grammatical number categories are not incomparable, but rather simply differ in the number and locations of the divisions made along the scale of individuation.

Mismatches across languages, such as Palmer's example of *hair* and *cheveux*, may further indicate either that categorization via the relevant factors is less determinate for some portions of the lexicon than for others or that there are differences in how the entities are construed across the different languages. Precisely this point has been made in work comparing the categorization of spatial relations across languages:

> "All else being equal, cross-linguistic agreement in semantic categorization suggests relative uniformity in the way that people readily conceptualize the domain, while disagreement suggests that the domain is more open to alternative conceptualizations and so more in need of language-specific learning." (Gentner & Bowerman, 2009, p. 468)

From this perspective, cross-linguistic mismatches do not pose a problem for the claim that a grammatical category is based on semantic categorization—this is actually the expected state of affairs.

Reconciling Countability Mismatches in Lexical Doublets The second type of mismatch illustrated in the Palmer quote occurs with what I will call lexical doublets (*leaves* vs. *foliage*). It appears to be a paradoxical situation: how could two words which refer to the same things in the world belong to two different countability categories? I will show

that the apparent equivalence between these words does not hold up to closer scrutiny. More generally, there is also indeterminacy in mappings from things-in-the-world to nouns: certain entities may permit multiple lexicalizations, each with distinctive properties, which in turn potentially have different effects on each noun's countability categorization.

I first note that in a given situation, such as the sentences in (12), *leaves* and *foliage* may be interchangeable.

(12) a. The leaves have turned red.
b. The foliage has turned red.

Yet, this apparent synonymy does not hold in general. (A similar point holds for many pairings of artifactual nouns, such as *coins* and *change*. See discussion in Clausen et al. 2010.) Lexicographers are careful to note the differences between the two terms. For instance, (13) and (14) give the dictionary entries for *leaf* and *foliage*, respectively, from the Oxford English Dictionary (OED).

(13) *leaf*: "An expanded organ of a plant, produced laterally from a stem or branch, or springing from its root; one of the parts of a plant which collectively constitute its foliage." (OED)

(14) *foliage*: "The leaves (of a plant or tree) collectively; leafage." (OED)

These meaning differences can be further brought out by finding contexts where the two senses are distinguished, as in the sentence given in (15).

(15) I raked the leaves/#foliage into a pile.

Demoting the *leaves/foliage* doublet from total- to near-synonyms does not fully address the problem that Palmer is getting at, but points the way to its resolution. First, it is useful to distinguish the extension of a noun, the things-in-the-world that the noun picks out, from the noun itself, which is a *description* of a type of entity. Here I am only making a distinction between entities and entity descriptions, which echoes the distinction often emphasized in the aspect and event semantics literature between *events* and *event descriptions* (Krifka, 1992). Continuing one step further, a particular noun meaning can be viewed as a

collection of attributes which pick out things in the world. These are the attributes that must be shared by those entities in the world to which the noun can refer. A noun then presents a *construal* of the entity in the world it is describing by selecting the attributes used to single out this entity. Yet, the entity may have other attributes beyond those specified by one particular noun, which in turn may be relevant for a different noun which describes the same entity. With these remarks in mind, the problem brought about by the near-synonyms can be recast in terms of the two terms presenting distinct construals of the entities. In this case, the different construals are associated with different individuation properties. In the case of *leaves*, the focus is on the individual leaves, while for *foliage*, the focus is on the collectivity of leaves and other associated vegetation.

3.6.2 Mismatches Between Countability Categories and the Presence or Absence of Individuals

I now turn to a different argument for skepticism towards connecting grammatical categorization in the form of a count/non-count distinction to individuation, or a contrast between individuals and non-individuals. A compelling argument is put forth by the study of Barner & Snedeker (2005), who focus on the quantificational behavior of nouns such as *furniture* or *silverware*. The argument is simple and elegant, and although it involves nouns designating ‘s, which are not the focus here, it is worthwhile to address some of the points that surface in Barner & Snedeker (2005), for they serve as a reminder of how complicated a full account of the facts must be.

Theories which link countable and non-countable realizations of nouns to a binary distinction between individuals and non-individuals make clear predictions for comparisons of quanity, namely in expressions such as *Who has more X?*, which I will describe as the quantity query. Barner & Snedeker (2005) provide a simple and clear experimental design, named the “quantity judgement task” to test how speakers judge quantities of different nouns. They presented participants with pictures showing two different characters who had different quantities of the entity at issue, and participants then had to decide which character had more. One of the characters possessed only one object, but which was very large, while the other character possessed three small objects. The stimuli were constructed in

such a way that the three smaller objects had a smaller combined volume and surface area than the single large object, which allowed the experimenters to determine whether the participants made judgements according to number as opposed to mass or volume.

The results were particularly clear. When speakers use a quantity query with count nouns, inasmuch as count nouns are individuals, the quantity judgement should be based upon individuals. Barner & Snedeker (2005) show, as expected, participants judge *Who has more chairs?* in terms of who has more individual chairs. When speakers use a quantity query with non-count nouns, as non-count nouns are not individuals, the quantity judgement should be based upon quantity of stuff/material. Again as expected, participants judge *Who has more toothpaste?* in terms of who has greater quantities of toothpaste material. The problem arises with what Barner & Snedeker (2005) name "object-mass" terms, i.e. functional aggregates such as *mail* or *furniture*. As these nouns have non-count syntax, the expectation is that participants will resolve *Who has more furniture?* in terms of which picture has more 'furniture-stuff', yet participants instead resolve this question by quantifying over the number of pieces of furniture. Barner & Snedeker (2005) argue from this result that there is no one-to-one correspondence between syntax and semantics. They propose the difference in countability behaviors of different lexical items stems from linguistic features, which have a less direct connection (or, under some interpretations, no connection) with, e.g. perceptual, qualities of entities in the world. Accordingly, they argue that nouns such as *furniture* are lexically specified as non-countable nouns but denote individuals.

Barner & Snedeker (2005) make a compelling case that grammatical countability is not identical to whether entities are logically countable. Although Barner & Snedeker (2005) situate their discussion of the relation between individuation and countability in terms of the performance of nouns within the quantification judgment task, as was shown in section 1.1.2, many other contexts demonstrate that nouns which are grammatically non-countable may designate entities which are constituted by countable individuals, as shown in (16).

(16) Ed counted the furniture/sand.

More generally, that a noun's countability behavior in some respects (quantifiers, pluralization) does not match its behaviors in others (quantification judgement) is actually often the case, as shown in chapter 1 and as discussed by Allan (1980).

Accepting the results of Barner & Snedeker (2005) again does not force one to abandon a line of explanation of countability involving individuation. Granting that mismatches between grammatical number classification and logical countability occur, it still does not follow that grammatical number and individuation are not related. The results of Barner & Snedeker (2005) are certainly a good argument against a *simple*, *binary* relation between syntax (+/−count) and semantics (+/−individual), but are not sufficient to exclude any account where syntax is responsive to semantics or individuation has a part to play. In point of fact, given the variety of factors that are relevant discussed in section 3.2, the relation between an entity and its countability classification appears to be anything but simple.

I conclude this discussion by considering how the defenders of a theory of countability based in individuation would counter the arguments of Barner & Snedeker (2005) for natural concrete entities such as *sand* or *rice*, which also allow access to individuals yet are grammatically non-countable. An account of individuation, in order to be consistent with the evidence from, e.g. comparatives, would need an explanation of how nouns such as *sand* can both be less individuated than count nouns and yet provide access to individuals. Yet, the evidence put forth in section 3.2.2 provides just such an explanation: individual elements may be present, but the configuration of these elements is such that the individual elements are not salient, as would be the case for *rice* and other granular aggregates which have a high degree of spatial contiguity.

3.6.3 Flexibility of Nominal Interpretation

Up to this point, I have been providing a treatment of countability which relies on identifying the typical designation of different nouns, and then explaining their countability status as determined by the possibilities provided by that designation. As discussed in chapter 1, it has been observed that nouns may be interpreted differently depending on their context. This phenomenon, which I will term NOMINAL FLEXIBILITY (also termed "elasticity" in Chierchia 2010b), has generated a large amount of discussion in the literature. Thought experiments such as the "Universal Grinder", as shown by the example in (17) repeated from chapter 1, clearly bring out this flexibility. The counterpart to the Universal Grinder is the "Universal Packager", which given a non-countable noun produces a countable noun, as

shown in example (18), also repeated from chapter 1. These examples at first sight appear to go against the very idea of a grammatical division between countable and non-countable nouns.

(17) There is dog all over the road.

(18) Two beers, please.

Such data has been used to motivate several different arguments impinging on the nature of countability. The primary fact that follows from such nominal flexibility is that accounting for countability by positing two syntactic categories, e.g. +*count* and +*mass*, is not sufficient. It is simple to see that should *dog* be labelled in the lexicon as +*count* and nothing more is said (an analysis found in Chomsky 1965), an explanation for the sentence in (17) would be lacking.

It is now commonly accepted that nominal flexibility presents an insurmountable problem for a syntactic feature approach to countability, but several researchers have argued that this data has much more radical consequences. Often, this data has been taken to motivate the position that nouns are not differentiated in terms of countability, but treated uniformly, for instance as completely underspecified as to countability properties. This stronger position contrasts with a proposal where nouns have a canonical designation, e.g. *dog* refers to individual dogs, and not dog-stuff. In an approach where nouns have a canonical designation, an actual difference between *dog* and *water* is asserted, at least in the type of objects that they standardly designate. Contexts such as "X all over the road" serve to coerce nouns to designate something different, but related, to the canonical designation of the noun. The more radical position claiming that nouns are underspecified with regard to countability denies that one sense of a noun is privileged.

An important assumption of the underspecification position is that nominal flexibility is absolute, in other words the "Universal" in the names "Universal Grinder" and "Universal Packager" is to be taken quite literally. While some have pointed out difficulties with the flexibility of particular nouns, such as Ware's (1975) observation that *hole* and *pore* are resistant to grinding contexts, it has remained a prevalent view that nominal flexiblity is

absolute. Pelletier, in particular, has been a proponent of this position across a variety of papers (Pelletier, 1979, 1991; Pelletier & Schubert, 2004; Pelletier, to appear), as exemplified in the following quote:

> "*Every* noun—even *hole* and *pore*—sometimes occurs in noun phrases which we would intuitively call +mass. And every noun sometimes occurs in noun phrases we would intuitively call +count." (Pelletier & Schubert, 2004, p. 270)

If this claim were accurate and any noun could find itself as a countable or non-countable noun depending on the context, then lexically specifying nouns for a particular countability status, whether syntactically or semantically, would be an inelegant treatment. If *dog* can equally refer to individual dogs or dog stuff, and similarly *water* can refer to individual portions of water as well as water stuff, then there would appear to be little basis for saying that an intrinsic facet of *dog* is that it is semantically interpreted as designating individuals, while it is not an intrinsic facet of *water* that it designates individuals. A more appropriate treatment would then be that all nouns would be either completely underspecified for countability or would be of the same type, e.g. underlyingly designating stuff, which then could be overwritten by the noun's eventual surrounding context. Both positions have been proposed, in Pelletier (to appear) and Borer (2005), respectively.

The influential account of Borer (2005) follows the latter path and provides several arguments why countability should be taken as grammatically determined and unrelated to the designation of any given noun. Some of the arguments hinge upon theory internal issues such as avoiding redundancy of features and the formal simplicity provided by underspecification. Of greater interest here are the empirical foundations of her argument: (i) nominal interpretations are flexible with respect to countability and (ii) a typological claim that the non-countable interpretation is basic. While her framework puts forth an elegant treatment of nouns, I will argue that its empirical basis is ill-founded—nouns are not permissive without limit, nor is this flexibility typologically universal. Further, a closer examination shows that not recognizing the constraints on nominal flexibility leads to several missed generalizations.

The core of her view is that listemes (elements of the lexicon) have no formal properties, but correspond to "'stuff' which is poured into the structural mould to be assigned

grammatical properties" (p. 108). All nouns start as non-countable nouns, i.e. have a default "mass" interpretation, wherefrom a count interpretation may be licensed by the structure of the noun phrase. Borer illustrates her claims with Chinese; she, along with many others, argues that "all nouns in Chinese have extensions which are mass," a position which has been widely proposed (Allan (1977), Sharvy (1978), Lucy (1992, 89), Krifka (1995), Chierchia (1998)). Borer, like Sharvy (1978), extends this view to all languages. Borer argues against lexical treatments of countability, which assert different countability statuses for different nouns. She states that, although in their favor they are able to account for idiosyncrasies, lexical accounts are not capable of capturing the generalization which applies to "just about all nouns and which cuts across languages" (p. 103).

In order for this position to be maintained, the accuracy of these wide-ranging claims must be demonstrated. Unfortunately, no significant number of nouns is discussed nor is further typological support displayed beyond the discussion of Chinese of languages for which all nouns have "extensions which are mass". The next two subsections examine, and provide substantial sets of counter-examples to, these claims.

3.6.4 Types of Gaps in Nominal Flexibility

A wide range of counter-examples to viewing nouns as absolutely flexible have in fact been pointed out at various places in the literature. As shown in (19), artifactual nouns such as *book* are difficult to grind (Chierchia, 2010b). Also, as mentioned above, Ware (1975) points out that nouns designating orifices (*hole*, *mouth*) do not grind well, as in (20).

(19) ??There is book all over the floor.

(20) ??There is hole all over the street.

Those who support absolute flexibility typically respond to such critiques by finding different contexts which support a "mass occurrence". For instance, Pelletier & Schubert (2004, p. 270) provide contexts which arguably support a "mass interpretation" for the putative counter-examples to the grinder, *book* and *hole*, given in (21) and (22), respectively.

(21) He has more book than bookshelf.

(22) This site has more hole than building.

The presumed readings are that in (21) the quantity of books is more than the space available in the bookshelves and in (22) there is more quantity of empty space at the building site than completed building. These examples are effective counter-examples against an analysis where countability is treated through syntactic labeling, as Pelletier and Schubert point out in their discussion of these examples. Yet, finding contexts where non-countable uses of these nouns are tolerable[12] is not tantamount to a positive argument for treating all nouns uniformly, either as underlyingly underspecified or underlyingly non-countable. To the contrary, they show how fragile such an approach is: if all nouns were, for instance, underlyingly non-countable, it is then surprising that these reading emerge only in such uncommon examples.

For the opposing view, in which nouns are associated with canonical designations and different contexts may lead to re-interpretation, the context *more X* simply provides another context which may serve to coerce the standard meaning of a noun in a particular way. The difference in acceptability of a non-countable occurrence of, e.g., *hole* in grinding and comparative context follows from the fact that the two different contexts coerce nouns in different fashions. In the grinding context, what is at issue is the *stuff* related to the entity named by the noun. In the comparative contexts, what appears to be at issue is the *amount* of what is named by the noun. While there is no obvious way to associate *hole* with material stuff, it is much easier to think measure a hole's volume, which allows interpreting *more hole*.

While much effort has gone into pointing out instances of nominal flexibility, less attention has gone into understanding the nature of the counter-examples to the Grinder. If the distribution of the counter-examples were idiosyncratic, then it would indeed be sensible not to pay much attention to them. Yet, the different counter-examples can be arranged into semantically coherent groupings, indicating that there are generalizations to be had. I have grouped different sets of counter-examples in table 3.12, which is by no means an exhaustive listing. If a member of these groups has been previously pointed out in the literature, I

[12] It has been noted in a variety of places that most of the sentences the undergo coercion have a jocular feel (Bloom, 1990, p. 130, fn. 22). In general, the acceptability of many of these sentences as "interpretable" is not to be taken for granted, but often depends heavily on the patience and good humor of one's interlocutor.

have acknowledged the source.

Resist grinding (no substance interpretation)	**Group nouns** (Bloom, 1990) *committee, team, flotilla*
	Abstract shapes (lower dimensional) *triangle, square, line, point*
	Units of measurement *hour, mile, second, day*
	Nouns of negative space (Ware, 1975) *hole, mouth*
	Event nouns (Ware, 1975; Brinton, 1998) *trick, act, arrival, blink, smile, run*
Resists packaging (no unit interpretation)	**Functional aggregates** (Bale & Barner, 2009; Borer, 2005) *furniture, change, foliage, mail*
	Granular aggregates *sand, foliage, barley, dirt*

Table 3.12: Classes of Nouns Resisting Grinding and Packaging

The status of these gaps differs: some of the gaps are accidental (*sand*) while some would appear to be obligatory and due to the nature of what the noun designates (*triangle*). The accidental nature of some of the gaps can be seen by tracing the history of the use of a noun. For instance, historically some of the granular aggregate nouns had designations which are individuated in certain circumstances—e.g. an earlier nautical use of *a sand* designated a sand bar (Oxford English Dictionary, 2012).

Other gaps such as the lack of non-countable interpretations for nouns of negative space or abstract shapes would seem to be related to the nature of what the noun designates. In other words, the inability of *triangle* to designate "triangle stuff" follows from the fact that a triangle is not a solid object. An attempt to put abstract shape nouns into a suitable grinding context is given in (23). Similarly, nouns which desigate units of measurement such as *meter* fail, as demonstrated in (24).

(23) a. ?There is triangle all over this chalkboard.
b. ?There is point all over the computer monitor.

(24) a. ?There is meter all over this tape measure.
b. ?This sidewalk has a lot of meter.

Both of these classes of nouns also are anomolous in the comparative construction as well, as shown in (25). The explanation appears clear: the members of these classes are not gradable in a way that is compatible with the comparative.

(25) a. ?This fabric has more triangle than square.
b. ?This clock has more second than hour.

These types of nouns stand as clear counter-examples to the claim that all nouns could be underspecified or have a default "mass" interpretation. These are classes of nouns which, by virtue of what they designate, designate individuals and do not provide any flexibility.

In the other direction, the class of functional aggregates, e.g. *furniture*, whose members are non-countable, adamantly resist being re-interpreted as countable. Attempts to construct unitized or sort readings are given in (26) and (27), respectively.[13]

(26) Hand me those three waters/#furnitures.

(27) Ed collects wines/#furnitures from Italy.

This is another instance where nominal flexibility, while wide-ranging, displays its limits. A full account of countability of course must not only explain the underlying reason for nominal flexibility, but also provide an account of the counter-examples.

3.6.5 Typological Variation in Nominal Flexibility

The accounts which would deny intrinsic countability distinctions between nouns, e.g. *dog* vs. *water*, lead to a typological prediction that languages other than English will behave in a comparable fashion. In other words, other languages should (i) demonstrate underspecification or uniformity of countability status across all nouns and (ii) operations such

[13]The difficulty posed by these nouns seems to be generally recognized, e.g. Bale & Barner (2009, 229) for some discussion. Borer also recognizes that these nouns are problematic in English for her proposal (p. 103, footnote 13), but does not provide an explicit account of them.

as grinding and packaging should be available. This section examines some of the relevant evidence that has been put forth for these claims.

I will first consider the typological claim of Borer (2005), who, having argued that all nouns are flexible, argues that all languages are such that nouns have a default non-countable interpretation, akin to the purported situation in classifier languages such as Chinese. There are two claims at work here (i) all nouns in Chinese are underlyingly non-countable and (ii) other languages pattern similarly, although they differ in the surface details.

Although the claim that all nouns in Chinese are underlyingly "mass" has been widely proposed, as mentioned above, it has been increasingly recognized that Chinese is more complicated than has been assumed. First, although not all of the standard diagnostics for countability that function in European languages, such as number morphology, are directly transferable to Chinese, Yi (2009) shows that countability differences are manifested in several ways. First, there are several classifiers which combine only with what appear to be countable nouns, including the general classifier *ge*, which Yi states approximates the term *individual*, as well as the classifier counterparts to terms such as *pair* or *dozen*. Additionally, some of the diagnostics applicable to English also distinguish countable nouns in Chinese. The distributive quantifier *geh* 'each', as exemplified in (28) (Yi, 2009, p. 221), is only licit with countable nouns.

(28) Niu (dou) geh you changchu he duanchu.
cow (all) each have strength and shortcoming
'Each cow has strengths and shortcomings.'

The behavior of size adjectives, which distinguishes countable and non-countable nouns in English (Bunt, 1985), is yet another diagnostic that is applicable in Chinese. The size adjectives *da* 'big' or *xiao* 'small' are only applicable to countable nouns and infelicity arises for the Chinese equivalents of *big water*, just as in English. In sum, Chinese, the parade example of classifier languages, does not lend support to the proposal that all nouns are uniform with regard to countability; rather, here too different classes of nouns are detectable.

Turning to the typological validity of the nominal flexibility claim, this too receives less than adequate support. It is beyond the scope of this work to provide a balanced typological

survey of the availability of operations such as grinding or packaging, but there has been sufficient discussion in the literature which as a whole illustrates the difficulties that the nominal flexibility hypothesis faces.

A finding across several languages has been that the operation of grinding does not function in the expected manner, counter to what the nominal flexibility hypothesis would predict. Cheng et al. (2008) discuss the Chinese data given in (29), where the provided context should replicate the effect of grinding. Yet, instead of the expected substance reading of 'dog', Cheng et al. (2008) report that the only reading available for (29-a) describes "a situation in which a wall has been decorated with numerous little dogs" (p. 52). To achieve the substance reading, the noun must be modified as in (29-b). True substance nouns, such as *water*, are however felicitous in this context.

(29) a. qiáng-shang dou shì gǒu.
wall-top all COP dog
'There are dogs all over the wall.'
NOT: 'There is dog all over the wall.'

b. qiáng-shang dou shì gǒu-ròu.
wall-top all COP dog-flesh/meat
'There is dog(meat) all over the wall.'

c. dì-shang dou shì shu?
floor-top all COP water
'There is water all over the floor.'

Similar restrictions on grinding are found in other languages. For instance, as discussed in Nunberg & Zaenen (1992), J. Sadock reports that Greenlandic Eskimo does not permit grinding to designate meat associated with an animal, as in the standard grinding examples in (17). Greenlandic Eskimo does, however, permit something like grinding with names of trees, through which the type of wood can be designated. In my own fieldwork on Dagaare, whose number system was discussed in chapter 2.2, I found that speakers also reject grinding. Translations of the canonical grinding contexts, such as given in (30-a), are unacceptable, instead requiring modification, as shown in (30-b)

(30) a. ??baa paale la a sori (zaa)
dog full PART DEF road (all)

'There was dog all over the road.'

b. ba-nɛnɪ paale la a sori (zaa)
dog-meat full PART DEF road (all)
'There was dog meat all over the road.'

While I have focussed on grinding so far in this section, there appears to be similar variation in the availability of packaging. The discussion of Yudja in section 3.3.4 provided an example of a language where packaging was unrestricted. Dagaare represents the other end of the spectrum, where, according to my informants, packaging is essentially never allowed.[14] Wiese & Maling (2005) provide a study showing that even the closely related languages English, Icelandic, and German show many different manners of and restrictions on packaging.

In sum, the behavior attested in these languages indicates that nominal flexibility is not a universal property of nouns, but rather is subject to significant cross-language variation and in some languages is quite restricted. In turn, this indicates that the predictions of accounts which place a high value on the phenomenon of nominal flexibility are not so easily validated. Yet, the variation seen aligns with what would be expected on an account which appeals to nominal meaning. Restrictions on grinding and packaging are found, but the opposite has not been observed, i.e. languages where there exist lexical items only referring to, say, dog-stuff and packaged waters, but which lack a way to refer to individual dogs or to water as a substance.

This concludes the discussion of various objections made against a semantic account of countability. I now turn to developing a formal account of countability.

[14] As far as I was able to discern, only one noun *dãã*, referring to the local alcoholic beverage, was able to be packaged. All other nouns were strongly rejected.

Chapter 4

A Formal Model of Countability

The preceding chapter established a global view on the typological possibilities for the grammatical encoding of countability contrasts. This chapter focuses on the formal representation of countability contrasts within a linguistic theory.

This chapter will first discuss some basic tools employed in representing countability and nominal semantics, and then review some of the seminal formal semantic accounts of countability contrasts. The primary technique that has been used is structuring the domain of discourse through part-of relations, implemented through mereologies or equivalent mathematical structures, i.e. Boolean algebras or semilattice structures. Accounts often relate their explanations of countability to properties defined in terms of the part-relation, such as *cumulativity*. While these different accounts are comparable in that they use similar modeling techniques, they also differ, rather substantially, in the details of implementation and, perhaps more crucially, in exactly which distinctions they attempt to model in the first place. In other words, although these proposals may be related in terms of their algebraic spirit, they differ as to what the countable/non-countable distinction signifies. In section 4.1, I first lay out the core axioms and definitions of mereology, then discuss the higher-order properties (*cumulativity*, *atomicity*) that can be defined over part structures, and proceed to describe some of the more influential accounts, and note some shortcomings.

I then motivate an extension of these accounts to address the phenomena discussed in chapters 2 and 3. In section 4.2, I set out the foundations necessary for the alternate account.

I enrich standard mereology with topological relations, which permit formal recognition of whole objects as well as of different types of connections that may hold between entities. Section 4.3 demonstrates how the addition of connectedness relations allows for a richer set of denotation types for nouns. While building upon previous accounts, the explanatory burden here is no longer solely on how an entity is related to its parts, but also on how the referents of a noun are related to one another in terms of the topological relation of CONNECTEDNESS. In short, nouns whose referents tend to come in connected clumps will be less countable than those whose referents come separated from one another, i.e. individuated. The remainder of the chapter relates this view to the generalizations from chapter 3 and then provides an account of the grammatical number systems of English, Welsh and Dagaare, as discussed in chapter 2.

4.1 Nominal Semantics and Part Structures

The central problem facing theorists who would like to account for mass as well as plural terms is that the traditional logical tools for analyzing language, namely set theory and predicate logic, assume that singular entities are under discussion. Predicate logic and set theory assume that individuals are predefined, and because of this, it is difficult to integrate nouns designating, for instance, liquids. For example, within a predicate logic representation of natural language, in the same way that $dog(x)$ would be true of the individuals in the world which are dogs, $water(x)$ would be true of the individuals in the world which are water; however, it is not clear that it is coherent to have the model treat individuals in the world which are water in the same fashion as it treats about individuals in the world which are dogs. The basic challenge is to develop a framework in which substances (and plural) entities are given equal standing as singular entities.

The most widely adopted view is to model plural and non-countable terms using MEREOLOGY. Mereology, derived from the Greek $\mu\epsilon\rho o\varsigma$ meaning 'part', is the theory of parthood. The study of parthood can be traced back to the Pre-Socratics (Varzi, 2011), but mereology in its modern, and formalized, incarnation stems from the works of Leśniewski (1916) and was further developed under the title of "calculus of individuals" in the middle of the twentieth century by Leonard & Goodman (1940) and Goodman (1951).

Mereology has figured in a wide range of philosophical analyses, but more relevant to our purposes, analyses of countable/non-countable distinction have mainly made use of a mereological framework. Quine's (1960) discussion of non-countable nouns set the standard for using mereological concepts as a way to think about the designations of countable and non-countable nouns. In particular, his conception of non-countable nouns as "scattered individuals" is a distinctly mereological proposal.

In this section, I will first discuss why mereological systems have been viewed as advantageous, and then outline a standard version of a mereological system which will serve as a foundation for the extensions in section 4.2. This discussion of standard mereology will be limited to setting out the facets of mereology which are needed to develop formal approaches to countability.[1]

4.1.1 Standard Mereology

A central advantage of using mereological systems in linguistic applications is mereology's liberal view of what an individual designates. Indeed, in the early formulations of mereology, "individual" had a technical meaning, simply designating what the lowest logical type used in the system represents. The ontology of Leśniewski (1916) eschewed any bias towards singular individuals—a mereological "individual" may correspond to a number of entities in the world which are disjoint and even widely spatially separated. Nor does an individual necessarily have any boundaries—from the viewpoint of mereology, "an individual is simply a segment of the world of experience, and its boundaries may be complex to any degree" (Goodman, 1951, p. 42). Thus, an individual can be any section of the world, or combinations thereof. At this point, it is worth warding off terminological confusion concerning the use of *individual*. Its use in the mereological literature is clearly different from the pretheoretic sense in which I have used *individual* in the preceding chapters, where it referred to well-defined physical objects in opposition to stuff. To eliminate any confusion, I will refer to individuals in the mereological sense as M-INDIVIDUALS and retain INDIVIDUAL for the pre-theoretical use.

[1]For in-depth discussion of the foundations of mereology and the variants on the basic system from a philosophical perpective, see the treatments in Simons (1987) and Varzi (2011). For considerations on connecting mereology and linguistics, see Piñón (1995) and Champollion (2010).

I now present a standard axiomatization of mereology, largely following the discussions in Casati & Varzi (1999) and Varzi (2011), as well as Champollion (2010), who provide a detailed discussion of the standard assumptions and axiomatizations of mereology. The lexical core of a mereological theory is provided by a treatment of the "part-of" relationship, $\leq$. There is wide consensus that to conform to its intuitive use, the part relation must be reflexive, antisymmetric and transitive. These restrictions are given as the first three axioms.[2]

M 1. $x \leq x$ *(Reflexivity)*
(Every thing is part of itself.)

M 2. $x \leq y \wedge y \leq x \rightarrow x = y$ *(Antisymmetry)*
(Two distinct things cannot both be part of each other.)

M 3. $x \leq y \wedge y \leq z \rightarrow x \leq z$ *(Transitivity)*
(Any part of any part of a thing is itself part of that thing.)

Given the part relation as primitive, other useful ancillary relations can be defined. The definitions in D1–D3 provide the relations OVERLAP, UNDERLAP, and PROPER PART.

D 1. $O(x, y) =_{def} \exists z(z \leq x \wedge z \leq y)$ *(Overlap)*
(Two things overlap when they share a part.)

D 2. $U(x, y) =_{def} \exists z(x \leq z \wedge y \leq z)$ *(Underlap)*
(Two things underlap when they are both part of another thing.)

D 3. $x < y =_{def} x \leq y \wedge \neg(y \leq x)$ *(Proper Part)*
(x is a proper part of a thing if it is a part of a thing which itself is not part of x.)

Through selecting different axioms and relations, a variety of different mereological systems can be devised (see Simons 1987 and Varzi 2011 for detailed discussion). I will discuss the axiomatization of the standard version which has been assumed in canonical works on different topics of countability in the formal semantic tradition, such as Link (1983), Krifka (1989), and Landman (1989). This standard framework has been argued for

[2] All the axioms given here are universally quantified.

Figure 4.1: An Object, a, with a Solitary Proper Part

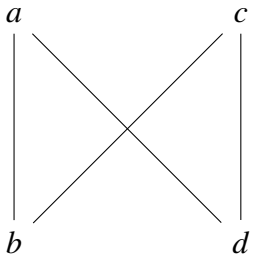

Figure 4.2: Two Different Objects, a and c, Sharing All Parts

on philosophical grounds, but the fact that it has a well-understood algebraic structure has certainly added to its appeal. In particular, there is a deep connection between models of mereology and Boolean algebra structures with the null element removed.[3]

While the part relationship and some dependent concepts have now been given, without saying anything more, we would allow structures in our models which have traditionally thought to be in disaccord with what a theory of parthood should represent. For instance, figures 4.1 and 4.2 show unwanted types of structures that are compatible with the axioms laid out so far. In a graph such as 4.1 which represents a mereological model, when two elements are connected by a line, the lower element is taken to be part of the higher element. Two types of extensions limit allowable models, which exclude models such as those given in figures 4.1 and 4.2.

The first extension involves restricting the manner in which an m-individual can be decomposed into different parts, in particular ruling out the situation, shown in figure 4.1, where an m-individual (a) is decomposed into a single proper part. An axiom known as either SUPPLEMENTATION or the REMAINDER PRINCIPLE is added in order to ensure such models are excluded.

M 4. $x < y \rightarrow \exists z(z \leq y \wedge \neg O(z, x))$ *(Supplementation)*

[3]This result is due to Tarski (1935), who demonstrates that for this axiomatization of mereology, the parthood relation has equivalent properties to the set-inclusion relation in standard set theory.

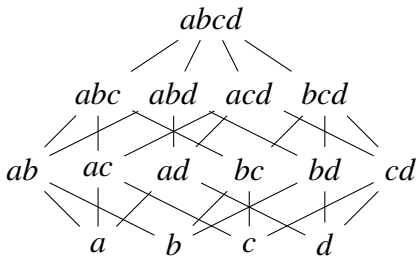

Figure 4.3: Example of a Licensed Mereological Structure

(If x is a proper part of y, there is an additional part of y which is distinct from x)

This effectively rules out the model in 4.1, since there is no additional part of a distinct from b. Yet, it is not sufficient to rule out 4.2. In this model, a has two proper parts and d may count as the supplement of b (or vice versa). Similarly for c, where b counts as the supplement of d (or vice versa). Since a and c are not parts of one another, M4 is trivially satisfied. In order to rule this model out, a second extension is necessary.

This second type of extension to the axioms of mereology concerns adding individuals together. The definition of the sum operation is given in D4. I also add notation for a binary sum operator and a generalized sum operator in D5 and D6, respectively. The traditional principle of mereology, again not uncontroversial as will be discussed further in section 4.2, is that for any two individuals, there is also a sum of those two individuals. Another way to state this is that the part structures that are of interest are those which are closed under sum formation, which is ensured by requiring every set of m-individuals to have a unique sum, for which the axiom is given in M5. By requiring unique sums, the model in 4.2 is now ruled out, since, for instance, b and d have both a and c as a sum, violating uniqueness.

D 4. $sum(x, P) =_{def} \forall y[P(y) \rightarrow y \leq x] \wedge \forall z[z \leq x \rightarrow \exists z'[P(z') \wedge O(z, z')]]$
(A sum of a set P is a thing that contains everything in P and whose parts each overlap with something in P.)

D 5. $x \oplus y =_{def} \iota z\ sum(z, \{x, y\})$

D 6. $\bigoplus X =_{def} \iota z\ sum(z, X) where X is any set$

M 5. $\forall P[P \neq \emptyset \rightarrow \exists! z sum(z, P)]$ *(Uniqueness of Sums)*
(Every nonempty set has a unique sum.)

The joint effect of these additions is to restrict the allowable models to those such as in figure 4.3. While each pair of m-individuals has a sum, not every pair of m-individuals is required to share a part, or in other terminology have a product. While *ab* and *ad* in 4.3 share a part through *a*, there is, for instance, no individual such that it is a part of *a* and *b*. In particular, in standard mereology there is no "null individual" which belongs to all other individuals in the way that the empty set is a member of all other sets in set theory. This position is typically held for philosophical reasons—as belief in a null individual is rather counterintuitive. This has the effect that the structures that are licensed here are a special type of lattice, a semilattice, the "semi-" indicating that the structure is closed under only one operation, here sum formation, as opposed to lattices proper, which are closed under sum and product operations.

The theory arrived at through these different axioms is the standard version of mereology used in philosophy and semantic theory, known in the literature as General Extensional Mereology. This system is expressive enough to represent the standard Boolean operators, such as difference and complement, given in D7 and 8, respectively.

D 7. $x - y =_{def} \iota z \forall w(w \leq z \leftrightarrow (w \leq x \wedge \neg O(w, y)))$ *(Difference)*

D 8. $-x =_{def} \iota z \forall w(w \leq z \leftrightarrow \neg O(w, x))$ *(Complement)*

I now turn to how such structures have been argued to provide an explanation of different countability phenomena.

4.1.2 Higher-Order Properties

In addition to being tolerant of different types of entities (plurals, substances), mereology provides a natural framework in which to elegantly model properties of entities through their behavior in terms of the part-relation. As discussed in chapter 1, a variety of higher-order properties have been defined with the goal of characterizing facets of entities which in turn correspond to their countability status. Many researchers have proposed that properties such as cumulativity, divisiveness, or atomicity serve to characterize the countability status

of different predicates. Divisiveness and cumulativity are the properties often claimed to be relevant for the denotation of non-countable nouns,[4] while atomicity is felt to be the relevant property for count nouns. Much discussion in the literature however has shown that all of these properties are controversial in terms of their capacity to shed light on the nature of countability. In this section, I will discuss these different properties and some of the associated criticisms.

CUMULATIVE REFERENCE, discussed early on in work in countability by Quine (1960, p. 91),[5] characterizes predicates which if they are true of two m-individuals separately, then also hold of the m-individuals together. This property, whose definition is given in D9, licenses inferences such as in (1). Cumulativity does not only characterize non-countable predicates such as *water*, but also plural predicates, as shown in (2); however, it fails in the case of singular predicates as in (3).

D 9. *Cumulative(P)* = $[P(x) \wedge P(y) \rightarrow P(x \oplus y)]]$

(1) A is water and B is water
→ A and B together are water.

(2) A are dogs and B are dogs
→ A and B together are dogs.

(3) A is a dog and B is a dog
↛ A and B together are a dog.

A predicate P has DIVISIVE REFERENCE if for an x falling under the predicate, then P is true of any part of x. This is stated in D10. For instance, given a quantity of mud, and assuming that *mud* is a divisive predicate, then it follows that every part of this mud is again mud. One of the first to argue for the importance of divisiveness as a characterization of non-countable nouns was Cheng (1973), and it is sometimes known as *Cheng's Condition*.

D 10. *Divisive(P)* = $\forall x[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$

[4] As pointed out by Koslicki (1999), while these properties have been a mainstay in the modern discussion of countability, they date back at least to the Pre-Socratics. Aristotle, for instance, discusses 'homoiomerous' substances, describing entities such as *flesh*, *blood*, or *marrow*.

[5] Quine attributes this property to Goodman (1951).

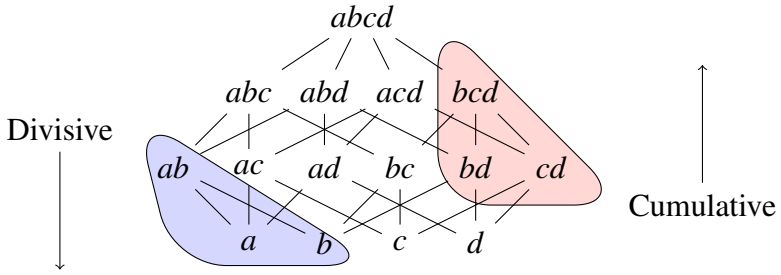

Figure 4.4: Cumulativity and Divisiveness as Closure Properties

If predicates such as *mud* are divisive, then it is clear why they are non-countable: these predicates do not provide a stable unit for counting. For an entity that could be described as *one mud* could by the same rights be described as any arbitrary number of muds, since all its parts are mud as well. This property also connects with the observation, put forth in Cartwright (1975), that non-count nouns correspond to entities that do not have an individuating standard. Unlike *dog*, which has a standard as to what counts as an individual, entities such as *mud* have no such standard.

In terms of the part structure, both cumulativity and divisiveness are "closure" conditions: cumulativity is closure under sum formation, while divisiviness is closure under part-taking. In other words, if a predicate is cumulative, it permits going upwards in the semilattice, and if it is divisive, it permits going downwards in the semilattice. The effect of the two properties is shown in figure 4.4. This figure also demonstrates the naturalness of these properties in a mereological framework: mereology allows one to discuss what happens along the vertical dimension of the semilattice, viz. what element is part of another or contains another element and these properties state in which cases reference may spread along the vertical dimension of the semilattice.

Predicates for which cumulativity and divisiveness obtain are often known as having HOMOGENEOUS REFERENCE.[6] For instance, Bunt (1985) and Lønning (1987) both take homogeneous reference as the hallmark of non-countable nouns.[7]

The third property, ATOMICITY, is often intended to characterize discrete individuals. It differs from the other two properties as it is not a closure condition but a restriction on part-taking. There are different ways atoms have made their appearance. From a mereological point of view, an atom is an individual which has no proper parts, as stated in D11. Some approaches simply have models that are atomistic (Link, 1983; Chierchia, 1998a). This is implemented by adding an axiom requiring everything to be ultimately composed of atoms, as in M6.

D 11. $Atom(x) \leftrightarrow \neg\exists y(y < x)$ *(Atom)*
(An atom is an individual which has no proper parts)

M 6. $\forall x\exists y(y \leq x \wedge \neg\exists z(z < y))$ *(Atomicity)*
(For any element, there is a part for which there doesn't exist a proper part)

A different approach defines atoms relative to a property, as given in D12.

D 12. $Atomic(x,P) = P(x) \wedge \neg\exists y[y < x \wedge P(y)]$ *(Atomic relative to a property)*

Given this definition, one can further define what it means for a predicate to be atomic, as in D13 from Krifka (1989), where $Atomic_P$ indicates atomic relative to P.

D 13. $Atomic(P) = \forall x[P(x) \rightarrow \exists y[y \leq x \wedge Atomic(y, P)]]$ *(Atomic predicate)*

Singular countable predicates are thought to be atomic (in one of these senses), given the inferential behavior demonstrated in (4), in contrast to non-countable predicates such as *water*, shown in (5). Note that, when an entity has parts, atomicity and divisiveness are mutually inconsistent—in other words, a predicate P cannot simultaneously require for an x that all of its parts satisfy P (divisiveness) and that there is no part satisfying P (atomicity).

[6] This is an unfortunate terminological choice as "homogeneous" has other well established uses, both pre-theoretically and technically.

[7] The characterization of predicates as homogeneous is sometimes extended from nominal predicates to adjectives or verbs (Quine, 1960; Lønning, 1987), although this extension has not been widely accepted.

(4) A is a dog and B is part of A
$\rightarrow$ B is not a dog.

(5) A is water and B is part of A
$\nrightarrow$ B is not water.

These properties, taken together, have been thought of as classifying nouns into three types: *singular*, *plural*, and *non-count*. Yet, they also give rise to other possibilities. Assuming all three properties are valid, the possible combinations are laid out in table 4.1. The third and fourth lines of the table show the two routes available to characterize non-countable nouns, depending on whether they are taken to be divisive. A natural question is if other categories, which would be predicted by this constellation of properties, are attested: predicates which are cumulative, but neither divisive or atomic (as on the fourth line) or predicates that are divisive but not cumulative (as on the fifth line). To my knowledge, such examples have not been proposed. In sum, while these three properties classify nouns into three types, it is not an economical classification system, and also one that also leaves some questions unanswered.

Table 4.1: Nominal Types Classified by Mereological Properties

Nominal type	Cumulative	Divisive	Atomic
singular (*dog*)	No	No	Yes
plural (*dogs*)	Yes	No	Yes
non-count (*water*)	Yes	Yes	No
absent	Yes	No	No
absent	No	Yes	No
absent	No	Yes	Yes
absent	No	No	No
absent	Yes	Yes	Yes

In addition, the connection between these properties and the classification of nouns as countable or non-countable is controversial for each property. Cumulativity is generally accepted as being true of non-countable and plural predicates, but that of course does not distinguish countable from non-countable nouns and so it is unclear if this property by itself

has much to do with countability per se.[8]

Turning to atomicity and divisiveness, these properties have garnered much more attention in the literature—for when these properties are applied systematically across the lexicon, many counter-examples arise. While most countable nouns can be reasonably treated as atomic predicates, there are still several classes of problematic nouns, discussed at length in Zucchi & White (2001) and Rothstein (2010). One such class of nouns is *twig*, where a part of a twig may again merit being designated as a twig. Geometrical terms such as *line* or grouping terms such as *sequence* show the same behavior. Terms of indefinite reference, such as *event*, *thing*, *a quantity of X* also fail to be atomic.

Divisiveness is the most controversial of the properties that have been argued to be relevant for non-countable predicates and there is wide disagreement as to the usefulness of the property. Adopting divisiveness leads to a dilemma, different parts of which have been recognized at different places in the literature. Divisiveness in its full generality, as presented in D10, turns out to be false—for many non-countable, and arguably all, nouns, attributing unrestricted part-taking as part of their nature is simply a false characterization. Quine (1960, p. 99) already pointed out that "there are parts of *water*, *sugar* and *furniture* too small to count as *water*, *sugar* and *furniture*." This is known as the "minimal parts problem".

For nouns designating materials, the minimal parts problem arises at the subatomic level: *water*, some argue, can be thought of as being composed of minimal bits, i.e. H_2O molecules, but the subatomic parts, such as lone hydrogen atoms, are not *water*.[9] A natural response to this worry is to deny, rather plausibly, that the subatomic level plays any role in the intuitions of language speakers. But even excluding these cases, there are several other problematic cases for divisiveness. Artifactual non-count nouns such as *furniture* also fail: a part of an element falling under the predicate *furniture*, say a leg of a table, is not again

[8] As pointed out by Grandy (2007), complex NPs can be constructed which are problematic for cumulativity. He considers the two expressions *object with mass more than 2 kgs*, which satisfies cumulativity, and *object with mass less than 2 kgs*, which does not satisfy cumulativity. Grandy (2007) states, "It is difficult to see any important respect in which the two expressions divide their reference differently."

[9] Although some have claimed that individual H_2O molecules are instances of *water*, this is not generally the position of researchers in the physical sciences. Anderson (1972), in an influential paper titled "More Is Different", shows many examples where the properties of the aggregate differ from those of the minimal part. One example would be that water has the property of being wet, but individual H_2O molecules do not.

furniture. Those who are committed to the usefulness of divisiveness may again counter that *furniture* seems to be an atypical non-countable term. Yet, there are further examples of canonical non-countable terms which have clear minimal parts and do not conform to the predictions of divisiveness. The most effective illustration comes from predicates such as *soup*.[10] If *soup* is divisive and *soup* is true of some a, then any part of a, say a_1, is also soup. Yet, should a carrot be in the soup, the part of the soup m which consists of some interior chunk of the carrot, does not count as an instance of *soup*.

One horn of the dilemma then is taking divisiveness at face-value, as it then just delivers the wrong results. The other horn is adopting a weaker version of divisiveness, but this has typically led to injecting vagueness into the workings of the property, which in turn makes its success difficult to assess. One technique to weaken divisiveness is to claim that it is effective "up to minimal parts", whereby *water* is divisive up to H_2O molecules, *furniture* is divisive up the individual chairs, tables, etc. One obvious difficulty is that, if this is legitimate for nouns such as *furniture*, there is no way then to block divisiveness from successfully applying to plural nouns either, for *boys* can be divided up into the individual boys. But this of course eliminates the explanatory power of the property which made it attractive in the first place.

A different way of weakening divisiveness is put forth in Koslicki (1999):

D 14. *Weakened Divisibility [Divisiveness]: A predicate P is weakly divisive iff P permits a myriad of unprincipled divisions of what it applies to into parts.*

This weakened version of divisiveness of course is not valid, as Koslicki (1999) notes, for predicates such as *furniture*, which while non-countable, still do not permit unprincipled divisions. It certainly performs better for capturing divisibility of predicates which designate material than the unrestricted version does, for this definition would apply equally well to *water* or *soup*. At the same time, the definition hinges upon notions such as "unprincipled" and "myriad", which, from a theory-internal perspective, is problematic as there is no obvious way to implement such notions in standard mereology. Nothing in the standard framework tells us what a principled or unprincipled division is. Koslicki points out yet other difficulties not tied to the use of mereology, namely that some countable nouns

[10] Examples of this type were, to my knowledge, first discussed in Taylor (1977).

apparently satisfy the weakened version of divisiveness:

> "However, there is no a priori guarantee that all stuff-like entities are reflected in language in the form of mass predicates. For example, clouds are quite stuff-like, but the noun 'cloud', in English, standardly only has count-occurrences. For this reason, more fine-grained mereological generalizations, in addition to their more restricted scope, can also sometimes be expected to cut across linguistic lines. However, such generalizations may nevertheless be quite useful in characterizing the paradigm cases of what we count and what we do not count." (p. 64)

In sum, while this weakened version of divisiveness is the best version that I have seen in the literature, its success is still partial.

Another way of attenuating the criticisms of divisiveness has been to weaken the relation between the application of the property and objects in the actual world. Authors such as Cheng (1973) make it clear that they believe that divisiveness does characterize objects in the world, but there is a second approach: Bunt (1985) argues that "mass nouns provide a way of speaking about things *as if they did not consist of discrete parts*" (p. 45). In other words, for Bunt (1985), the criticisms of divisiveness carry no force, since the properties determining countability have no necessary connection to the physical world, but only to the semantic model.

Bunt (1985) defines "mass" nouns as those which adhere to the *homogeneous reference hypothesis*: "Mass nouns refer to entities as having part-whole structure without singling out any particular parts and without making any commitments concerning the existence of minimal parts" (Bunt, 1985, p. 46). It is worth pointing out that the homogeneous reference hypothesis is a negative definition: it only tells us what non-count nouns do *not* do, but gives little information as to how to recognize non-count through positive characteristics.

Bunt later explains that there is a clean separation between the actual world and the application of these properties:

> "... the use of a mass noun constitutes a way of referring to something as if it is a homogeneous mass, as opposed to a discrete collection of objects. The

> linguistic evidence supporting this view, summed up in the homogeneous reference hypothesis, is the same for all mass nouns: the count/mass distinction is a *formal* phenomenon, not a referential one. Although we may have different actual beliefs about the 'homogeneity' of the referents of such words as 'furniture', 'luggage', 'computing equipment', 'shoe polish', 'time', or 'rice', from a formal semantic point of view these mass nouns should all be treated alike. ... Actual beliefs about the world simply play no part at the formal level." (Bunt, 1985, p. 129-130)

This view, then, asks us to believe that a property that is neither responsive to physical reality nor to our beliefs in the world should be held responsible for a rather widespread form of grammatical classification. Free from such constraints, the application of divisiveness of course runs into no problems. The crucial question, left unanswered, is what positive reasons can be given for adopting this view. In fact, there has been little compelling positive evidence in favor of adopting the properties of divisiveness and cumulativity as a characterization of non-countable terms, beyond their initial plausibility.

In summary, the wide adoption, and distinguished pedigree, of these properties is an indication of their intuitive appeal. It is also clear that these are sensible properties to propose from the mereological perspective as they all exploit the part-relation in a natural way. At the same time, decades of research have shown that the correspondence between these properties and countability phenomena is imperfect. Such a misalignment indicates that alternate properties are worth considering, as I will do in section 4.3. Now, I turn to discussing two implementations of mereological accounts of nouns.

4.1.3 Implementations of Nominal Semantics

Wide agreement on using a mereological, or equivalent, system and characterizing the higher-order propereties has still left much room for a variety of different proposals. I will discuss the proposals of Link (1983), Krifka (1989), and Chierchia (1998b). This survey cannot hope to be exhaustive, but these proposals have not only been influential, but they provide distinct manners of implementing the concepts discussed in sections 4.1.1 and 4.1.2, and thus provide a good coverage of the different routes that have been taken in

implementing countability.

Link (1983): A sortal distinction between countable and non-countable nouns

Link (1983) is *the* seminal work on countability—it achieves a large number of advances in surprisingly few pages and the bulk of the countability literature that follows the formal semantic tradition reacts to it in one way or another. I will limit the discussion here to its relevance for the countable/non-countable distinction, and not discuss other elements such as his account of the distributive/collective distinction.

Recognizing the difficulties posed by plural and non-countable terms, Link argues that to provide a successful account of plural and "mass" terms, it is not necessary to give up the set-theoretic treatments that had been successful before, but just to extend them. The proposal is that individuals must be viewed within a context of a lattice structure, which specifies the relations holding among the individuals. To treat plural reference, Link proposes that the domain of discourse should not only include the set of entities that are the elements of the discourse, but also include all the combinations of these entities. Thus, if the entire universe consists of a table a and two chairs, b and c, then we need to represent not only these three entities but all of their combinations, such as $a \oplus b$. For any predicate P, the set of all its sums is defined as P^*, as given in D15. The totality of the entities and its combinations generates a complete join semilattice, viz. the structure licensed by the standard form of mereology given at the end of section 4.1.1. Link considers the domain of individuals E to be atomic, thus atomicity is an assumption of the model. The subset of atomic elements is designated by A.

D 15. $P^* =_{def} \{x|\exists P' \subseteq P[x = \bigoplus P']\}$
(Algebraic Closure) (P^* is the set that contains any sum of things taken from P.)[11]

Turning to the countable/non-countable distinction, Link (1983) advocates a clean ontological distinction between individuals and the material substance of the world.

> "Our guide in ontological matters has to be language itself, it seems to me. So if we have, for instance, two expressions a and b that refer to entities occupying

[11] I am using here the more perspicuous notation of Champollion (2010, p. 15)

> the same place at the same time, but have different sets of predicates applying to them, then the entities referred to are simply not the same. From this it follows that my ring and the gold making up my ring are different entities; they are, however, connected by what I shall call a *constitution relation*. ... individuals are created by linguistic expressions involving different structures even if the portion of matter making them up is the same." (p. 13)

These ontological categories of individuals and stuff correspond in the model to two different domains of entities, the set of individual entities E and a second set D, representing portions of matter, respectively. The domain of material substance, D, is also ordered by the part-relation, again generating a complete join semilattice. Countable nouns have their extensions in E while non-countable ("mass") nouns have their extensions in D.

Link then specifies a relation between the members of E and D: the matter in D is considered as a subset of the atoms in E, namely of the set A, whereby for every individual, the matter that makes up that individual is represented. Formally, this relationship is achieved by defining a homomorphism between the two lattices.

To illustrate, consider a simple model where the universe consists of one table. For the individual a, the table, belonging to the domain E, there is a corresponding element of the set D, a^m which is the material portion which constitutes a. In this case, a^m is the wood which materially constitutes the table. This relation is given by the homomorphism: $h(a) = a^m$. This relationship is completely general: every predicate which has its denotation in E also has a "mass term correspondent", P^m whose denotation is in D.

This account is not only elegant, but yields several desirable results. This system represents the distinction between countable objects and non-countable substances—the different nominal types as different types of elements in the model. Count nouns belong to E, with singular nouns as atoms and plural nouns as sums, while non-count terms are elements of D. Thus, countability here completely aligns with an ontological distinction between material substances and individuals.

Additionally, the homomorphism between the two domains easily accounts for the existence of contextual countability shifts. Link gives the following examples, where P represents *is an apple*, Q represents *in the salad*, and P^m is the mass term correspondent *is apple*.

(6) a. There is an apple in the salad.
$\exists x(P(x) \wedge Q(x))$
b. There is apple in the salad.
$\exists x(P^m(x) \wedge Q(x))$

Link also represents some of the higher-order properties, although they arise in different fashions. (Link considered divisiveness not ready for implementation, and refrains from treating it.) Atomicity, as discussed, is simply built into the model. Cumulativity, however, is a theorem that can be derived in this system—for both predicates of mass terms and plural predicates. Thus, cumulativity comes about because of how entity types are structured, namely mass and plural predicates give rise to structures for which cumulativity is satisfied.

The common critique of Link (1983), however, is that the countability facts that need explanation cannot be reduced to a sortal distinction between objects and substances. For instance, Lasersohn (2011) states "the proposal makes sense only under the very narrowest construal of the term mass noun, under which it refers only to those nouns which can function as names of physical substances" (p. 1138). Standardly, *furniture* or similar nouns serve to illustrate this point, but equally important are the variety of aggregates which came to light in chapter 2, which have a different countability status than a strict binary distinction between substances and objects would predict.

Krifka (1989): Properties of nominal meaning

The work of Krifka (1989, 1992, 1995, 1998) develops the core ideas of an algebraic approach to countability in several novel directions. Following Link (1983), Krifka structures the domain of objects in a way that is consistent with the basic principles of mereology set out above, and models nominal meaning as complete semilattice structures lacking the null element. While Link's (1983) account of countability came down to the substance/object distinction, which was explained as a sortal distinction in the model, Krifka's theory expresses the difference between countable and non-countable entities relative to predicates.[12] The end effect of this view is that a particular element in the semilattice may be designated

[12]Krifka does appeal to using different sorted domains to give a solution to the problem of nouns such as *twig*, but what is important for the moment is that this distinction is not what serves as an explanation of countability behavior of predicates.

by a variety of different predicates. Unlike in Link's system, where individual entities and portions of material inhabit different, although related, semilattices, here the same element of the lattice may, for instance, be described both by a predicate designating atomic m-individuals (*table*) and a predicate designating substances (*wood*).

In this system, there are two differences between countable and non-countable nouns at the lexical level: (i) the satisfaction of the properties of cumulativity or quantization and (ii) whether there is a built-in natural unit measure function. Higher-order properties such as cumulativity are then specified for particular natural language predicates rather than as properties of the model. For instance, the predicate *gold* is translated as a cumulative predicate $\mathbf{CUM}_o$**(gold)**. While non-countable nouns are assumed to be cumulative, Krifka claims that countable nouns satisfy the property of being QUANTIZED, given in D16. (This is a stronger condition than atomicity, and actually implies atomicity.)

D 16. $Quantized(P) =_{def} \forall x \forall y[P(x) \wedge P(y) \rightarrow \neg y < x]$
(A predicate P is quantized if and only if no P-entity can be a proper part of a P-entity.)

The second aspect of Krifka's account is the use of measure functions to model expressions such as *two liters* or *five ounces*. He further claims that measure functions are relevant for the distinction between countable and non-countable nouns. Partly based on an analogy with classifier languages, Krifka considers countable nouns, e.g. in English, to have a built-in measure function that counts "natural units" of the noun. Countable nouns under this account are two-place relations between numbers and entities, while non-countable nouns are one-place entities, as exemplified by contrasting the lexical entries for *cow* and *gold* in (7-a) and (7-b), respectively. Krifka then explains that there are two criteria of applicability at work: one which is "qualitative" and represented by the nominal predicate, viz. *gold* or *cow*, and a second which is "quantitative", represented by a "natural unit" (NU) measure function.

(7) a. $\lambda n \lambda x$[COW(x) $\wedge$ NU(COW)(x)=n]
b. λx[GOLD(x)]

In sum, by treating countability distinctions at the level of predicates rather than as a matter of ontology, Krifka's system is able to treat nouns such as *furniture* or *cattle* more plausibly

than Link—for these predicates are descriptions of entities which adhere to additional properties such as cumulativity, rather than necessarily distinct in any ontologically significant way.

Chierchia (1998): Non-countable nouns as plurals

I briefly describe the proposal of Chierchia (1998a) as it explores yet another logical possibility within the tradition of using part structures to model the nominal domain. The innovative part of the proposal is to consider the entire domain to be atomic. Thus, all nouns have minimal, atomic elements. Chierchia (1998a) puts particular emphasis on nouns such as *furniture*, which would appear to indeed have minimal atomic elements, to support this view. Non-countable nouns such as *water* or *rice*, according to Chierchia, also have atomic minimal elements, but these are "vaguely specified" (p. 68).

Like Krifka (1989), countability distinction are specified relative to predicates. The difference between the predicates *table* and *furniture* comes about from the lexical entries. *Table* is "associated ... with individual atoms", whereas "*furniture* does not single out a set of atoms, but a whole ... sublattice" (p. 68). Thus, non-countable nouns are treated in a fashion akin to plurals.[13]

The many parallels between non-countable and plural nouns lend plausibility to Chierchia's (1998a) proposal. Yet, there have also been criticisms that this proposal does not adequately distinguish substance non-countable terms (*water*) from nouns which are lexical plurals (*police*, *cattle*) (Lasersohn, 2011).

4.1.4 Desiderata

This section lays out some goals for what a formal system should represent, based on the discussion so far. First, the distinction between objects and substances should be represented. The evidence presented in chapter 3 indicated that the distinction between entities which are viewed as coming in minimal units as opposed to not coming in such units is

[13]Chierchia (1998a) distinguishes them in that plural nouns are actually true of groups but not of individuals, while non-countable nouns are true of both.

foundational for the countable/non-countable distinction, as ordinary intuition would suggest. Therefore, this distinction should find a place in the account. While the properties of atomicity and divisiveness were designed to capture just this distinction, they face serious challenges as reviewed in the previous section.

A different problem is posed by the distinction between aggregates, as discussed in chapter 2, and plurals. Standard mereology does not provide a way to distinguish sums which correspond to plural individuals (*boys*) from those which correspond to aggregates (*ants*)—both are simply represented as sum individuals. Given the central role that aggregates play in certain grammatical number systems, the account should be able to distinguish the two types of entities.

Finally, the analyses reviewed in section 4.1.3 aim to capture a binary distinction between countable and non-countable nouns, and thus those analyses rely on binary contrasts, e.g. +/− *atomic*. A broader view of the data, as given in chapter 2 and chapter 3, has shown that a binary distinction is insufficient. The formal system should be capable of representing gradations of individuation, and provide an account of the scale of individuation.

I now turn to presenting an extension of mereology, which then will be used in section 4.3 to model objects, substances and aggregates, then in 4.4 to provide a semantic account of several of the grammatical number systems presented in chapter 2.

4.2 Extended Mereology: Part-Whole Structures and Connectedness

This section presents an extension of the standard mereological framework by adopting various notions from topology. There are several motives for such an extension. In addition to the difficulties, both well-known and less well-known, that I have discussed regarding the application of mereology to countability phenomena, mereology has been criticized both with respect to its adequacy for other applications and with respect to its foundations. If one uses mereology to model objects in the world, there are significant gaps between intuitions about how objects stand in the world and what mereology delivers. Two prominent, and related, shortcomings of mereology are its commitment to unrestricted sum formation and

its failure to account for what a whole entity is. Given this, many philosophers have argued that mereology is, as a theory, not sufficient for modelling tasks it is usually assigned, viz. a theory of objects.

Recent developments in the mereology literature have indeed been attempting to push beyond only using the part relation and tend to develop an account of other prominent relations that entities may stand in, in particular adding relations which do service to topological notions. What is most relevant here is that some of the shortcomings of the accounts of countability phenomena based in part structures may be a result of, or related to, the general shortcomings of the mereological framework that is used. I will argue that some of the new techniques that have been developed to aid mereology better represent the world in turn help in accounting for the outstanding desiderata of countability phenomena.

As mentioned in section 4.1.1, the requirement of unrestricted sum formation, i.e. that for each element in the domain, there is a corresponding sum composed of those two elements (see M5), has often lead to uneasiness. The question often turns upon what the status of this sum "individual" is. Does the sum of my office chair and computer monitor qualify as an "individual" in the same way as the sum of the two halves of my desk? The two types of entities do indeed seem to be quite different in nature—one sum, that of my office chair and computer monitor, is of two accidentally related entities while the other sum, that of the two halves of my desk, is of parts of a whole. Some theorists have seen this as one cause for abandoning the mereological enterprise. On the other hand, defenders of mereology argue that accepting sums of diverse entities, while perhaps not resulting in natural entities, is not in itself a harmful practice. Most famously David Lewis (Lewis, 1991, p. 81), has argued that these sums are "ontologically innocent" (p. 81), stating: "In general, if you are already committed to some things, you incur no further commitment when you affirm the existence of their sum" (p. 81).

If the existence of arbitrary sums can be countenanced, it still does not mean that mereology provides a satisfying account of ordinary objects, as opposed to more technically motivated m-individuals. The debate over the acceptance of unrestricted sum formation points to a true deficiency of mereology. Mereology is often said to be a theory of the part-whole relation, but in truth it is only a theory of parthood. Parthood is the primitive relation assumed from which all else is constructed. The sum of two entities $a \oplus b$ is often casually

equated with the "whole" to which the parts *a* and *b* belong. Yet, sums simply cannot be equated with wholes—the sum of the Empire State Building and the letter 'z' exist by the same rights as the sum of two halves of an apple making a whole apple. As Casati & Varzi (1999) state, "it is impossible ... to draw a distinction between 'good' and 'bad' wholes" (p. 51). Casati & Varzi (1999) provide a clear-sighted discussion of this deficiency:

> "What exactly is the difference between the cup and the broken glass? What is it that makes the cup one thing, as opposed to the many pieces of the broken glass? The difference cannot be a purely mereological one. Mereologically, for every whole there is a set of parts, and to every set of parts (that is, every arbitrary collection of objects) there may in principle correspond a complete whole, viz. their mereological sum or fusion. One could argue that not every sum is legitimate—that not every sum is a good whole. But there is no way, mereologically, to draw a distinction here; there is no way to rule out 'bad wholes consisting of scattered or ill-assorted entities (the whole consisting of the pieces of the broken glass, or the whole consisting of the broken glass, the unbroken cup, and your favorite Chinese restaurant) by thinking exclusively in terms of parthood. If we allow for the possibility of scattered entities, then we seem to lose the possibility of discriminating them from integral, connected wholes. Yet we cannot just keep integral wholes without some means of discriminating them from wholes that come in pieces." (Casati & Varzi, 1999, p. 11)

In sum, with only a part relation in the theory, there is no way to do justice to the part-whole contrast. These criticisms of mereology's expressive power are in turn relevant for an adequate treatment of the countability distinctions of different nominal predicates.

At its most intuitive, the distinction between countable and non-countable nouns is tied to the distinction between objects and substances. (This distinction is obviously too coarse for, e.g. non-countable nouns in English, but at least this must be covered.) Objects, in turn, are intuitively whole objects, as Soja et al. (1991) and Prasada et al. (2002) have shown for the developmental literature. Yet, from the proceeding discussion, it appears

that mereology doesn't provide a satisfactory treatment of what it is to be a whole object.[14] If the criticisms of mereology are well-placed, then it is clear from where, at least some of, the problems modeling countability stem—mereology is not suited to represent (at least) one side of object/substance divide.

It also becomes clear then why higher order properties such as *atomic* or *quantized* fail as characterizations of objects—they are an attempt to characterize wholes in a theory which is unable to carry out that task. The property of divisiveness enforces that the reference of a predicate descends the semilattice. The property of atomicity restricts the reference of a predicate from descending in the semilattice. But as the part-relations do not say anything about what a whole object is, these properties, based on the part-relation, do not provide a characterization what a whole is either. Instead, these properties demonstrate the awkward fit between the explananda and the characterization in terms of parts. Atomicity/quantization leads to finding parts when one doesn't want them, as in the *twig* case. On the other hand, divisiveness leads to *not* finding parts where one needs them, e.g. according to divisiveness, one should (but doesn't) find increasing smaller parts of furniture which count as *furniture*.

Much recent work in philosophy and ontological modeling (Smith 1996, Casati and Varzi 1999, Varzi 2006) has explored extending standard mereology with connectedness relations. These various theories are known as "mereotopology". Early studies of mereotopology can be found in Whitehead (1920, 1929); however, Whitehead's proposals were not followed up on until work in artificial intelligence began to explore mereotopological formalisms in order to develop fine-grained representations of spatial relations—knowledge of relations such as NEAR, INSIDE or ON TOP OF require a richer set of distinctions than standard mereology furnishes. While most of the discussion of developing mereotopology has taken place in philosophy (Roeper, 1997), artificial intelligence (Clarke, 1981) and ontological modeling (Smith, 1996), the connection to natural language semantics has also been pointed to, if not followed up on systematically by linguists. For instance, Smith (1996), an early proposal for mereotopology, explicitly argues that mereotopology would be useful for those working in the cognitive grammar tradition, where many distinctions

[14]This is one of the criticisms put forth by Moltmann (1997) in relation to the countable/non-countable distinction, an important precedent for the view I argue for. Moltmann also proposes that whole objects must be formally recognized, although her account rejects mereology altogether.

such as "boundedness" are proposed, which lacks a formal framework in which they can be systematized.

In passing, it is worth noting that the machinery I will propose for countability, while a departure from the standard mereological proposals for nominal meaning, is most likely necessary independently for a full treatment of spatial relations in linguistic theory. Topological notions regularly are used in the literature on locatives (Kracht, 2002), while many current theories of prepositional semantics assume that topological notions such as *convexity* are available within semantic theory (see Zwarts & Winter 1997). The following section presents an extension of mereology which will then be put to use in section 4.3 to give an account of both the richer data concerning collectives, elaborated in chapter 2 and 3, as well as the core data of objects and substance which has proved problematic to model over the last decades.

4.2.1 Adding Topological Axioms

Different techniques have been explored to relate mereology and topology. Some have taken mereological relations as basic while others have taken topological relations as basic. Here I will be following a technique which simply extends mereology with topological relations, although nothing crucial will hinge on this choice, and it is mainly chosen for perspicuity, and ease of comparison with other proposals. Further, I will be following in particular the discussion of strengthening mereology with topology found in Casati & Varzi (1999).

The fundamental change is adding the relationship of CONNECTEDNESS, which then interacts with the different definitions and axioms of standard mereology. The intuitive definition is that two entities are connected if they share a common boundary. An alternate characterization is that if two entities are disjoint, then they are SEPARATED. Accordingly, any entities that are not separated are connected. It is easy to illustrate with some examples from the physical world. My chair is touching the floor, and therefore qualifies as connected to the floor. The upper two-thirds of my mug is connected to the lower two-thirds, in fact they overlap, sharing a common part in the middle third. My mug and the Taj Mahal are, however, separated (although they do have a mereological sum). Figures 4.5 and 4.6 show

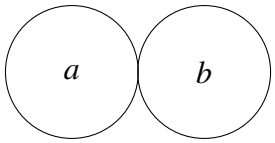

Figure 4.5: External Connection

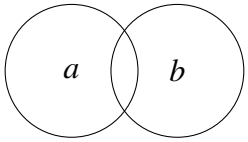

Figure 4.6: Strong Connection

two varieties of connection between two circles: on the left, the two circles are connected by touching, while on the right, the two figures are connected and overlap.

Some basic requirements are that the relation C, CONNECTED, is reflexive and symmetrical. These are added as axioms in T1 and T2.

T 1. *C(x,x) (Reflexivity)*
(Everything is connected to itself)

T 2. *C(x,y)* $\rightarrow$ *C(y,x). (Symmetry)*
(If x is connected to y, then y is also connected to x)

There are some further intuitive interactions with the mereological relations *part*, $\leq$, and *overlap*, O, that Casati & Varzi (1999) note. I discuss two interactions. First, the axiom in T3 ensures that parthood implies connectedness.

T 3. $x \leq y \rightarrow \forall z(C(x,z) \rightarrow C(z,y))$
(If an m-individual is part of another, whatever is connected to the first is connected to the latter)

From T3 the relation in (8) follows, whereby overlap implies connectedness.

(8) $O(x,y) \rightarrow C(x,y)$
(*if two m-individuals overlap, they are connected*)

With these additions, one can define mereotopological properties which make fine-grained distinctions among different configurations that objects may be in, such as INTERNAL PART (IP), INTERNAL OVERLAP (IO) or TANGENTIAL OVERLAP (TO) which are illustrated in figures 4.7, 4.8, and 4.9, respectively.

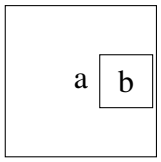

Figure 4.7: Internal Part

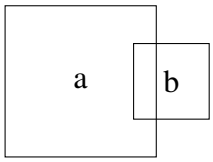

Figure 4.8: Internal Overlap

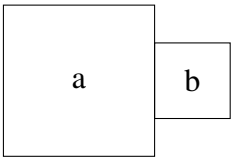

Figure 4.9: Tangential Overlap

D 17. $IP(x, y) =_{def} x \leq y \wedge \forall z(C(z, x) \rightarrow O(z, y))$ *(Internal Part)*

D 18. $IO(x, y) =_{def} \exists z(IP(z, x) \wedge IP(z, y))$ *(Internal Overlap)*

D 19. $TO(x, y) =_{def} O(x, y) \wedge \neg IO(x, y)$ *(Tangential Overlap)*

Standard notions of topology can also be given expression within the framework. By way of the generalized sum operator, the notion of the INTERIOR of an entity can be defined as in D20, which is the sum of an m-individual's internal parts. Being able to specify the interior of an object will be critical in defining an entity which forms a whole. Further topological operators such as the *exterior*, *closure* or *boundary* of a given entity can then be defined. Figures 4.10-4.12 illustrate the interior, exterior and closure of a circle. For example, the shaded part of figure 4.10 is the interior of the circle, where the dashed line indicates that the interior does not included the boundary of the circle.

D 20. $ix =_{def} \bigoplus X$ *where* $X = \{y : IP(y, x) = True\}$ *(interior)*

D 21. $ex =_{def} i(-x)$ *(exterior)*

D 22. $cx =_{def} -(ex)$ *(closure)*

D 23. $bx =_{def} -(ix \oplus ex)$ *(boundary)*

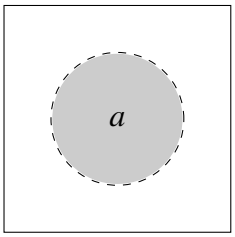

Figure 4.10: Interior

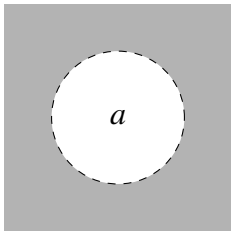

Figure 4.11: Exterior

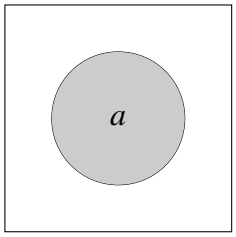

Figure 4.12: Closure

This concludes the discussion extending mereology with topological relations. What has been developed so far is sufficient for representing the notion of an integral whole as well as different degrees of connectedness, which in turn will be used in section 4.3 to model the countability distinctions.

4.2.2 Defining Wholes

Returning to one of the primary motivations for adding topological relations, it is now possible to distinguish entities which form one piece as opposed to arbitrary sums. This is achieved through recognizing "self-connected" (SC) entities—individuals which cannot be divided into two separated parts. The basic definition is given in D24.

D 24. $SC(x) =_{def} \forall y, z(\forall w(O(w, x) \leftrightarrow (O(w, y) \vee O(w, z))) \rightarrow C(y, z))$
(An entity x is self-connected if any two parts that make up the whole of x are connected to each other)

We can now distinguish between objects which come in one piece as opposed to scattered objects. For instance, suppose our domain of discourse contains four individuals, a and b are two halves of a sphere, and c and d are my left shoe and the Eiffel Tower. Then

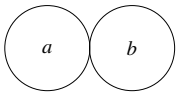

Figure 4.13: The Sum of a and b Is Self-Connected, but not Strongly Self-Connected

we can distinguish the sum $x = a \oplus b$ from the sum $y = c \oplus d$. x is a whole, i.e. SC(x) is true, while the "scattered individual" y is not a whole, and SC(y) is false.

Yet, this definition does not rule out a sum of entities which just "touch", as in $a \oplus b$ in figure 4.13. A stronger notion is needed to do justice to the notion of a whole. Casati & Varzi (1999) supply two stronger self-connectedness properties. First, a stronger requirement is that an entity's interior must be connected, which correctly rules out configurations as in figure 4.13.

D 25. $SSC(x) =_{def} SC(x) \wedge SC(ix)$ *(Strongly Self-Connected)*

If an individual has the property of being strongly self-connected, that still does not guarantee that it is also a whole in the intuitive sense. Consider the m-individual, i.e. "the segment of the world of experience" according to Goodman, that corresponds to exactly one half of a sphere. The half-sphere qualifies as strongly self-connected, but it does not intuitively qualify as a whole, for that would be the entire sphere. Since we are interested in unity, we can define being *maximally* connected relative to a property, which provides our final definition of what it means to be a whole. If an entity satisfies Maximally Strongly Self-Connected (MSSC) as given in D26, then that entity is the largest entity satisfying that property which is self-connected.

D 26. $MSSC(x, P) =_{def} P(x) \wedge SSC(x) \wedge \forall y(P(y) \wedge SSC(y) \wedge O(y, x) \leftrightarrow y \leq x)$
(Maximally Strongly Self-Connected relative to a Property)
(An m-individual is Maximally Strongly Self-Connected relative to a property if (i) every (interior) part of the individual is connected to (overlaps) the whole (Strongly Self-Connected) and (ii) anything else which has the same property and overlaps it is once again part of it (Maximality))

Such distinctions may seem subtle, but actually have very welcome consequences for countability as I will discuss in section 4.3. To anticipate, a fence actually contains other

m-individuals which also satisfy being a fence in the material sense, as discussed in section 4.1.2, but the only one that satisfies being a MSSC individual is the largest fence that contains the others.

4.2.3 Varieties of Connectedness

The basic connectedness relation C is taken as primitive, from which different degrees of connectedness can be defined. The two primary types of connectedness turn upon whether the two connected entities overlap in a substantive sense or only tangentially.

Strongly Connected is defined between two individuals when they overlap substantively. This can be succinctly defined in the full theory of mereotopology as in D27. Recalling the notion of strong self-connectedness (SSC), a relation of strong connection can be defined between two m-individuals if their sum is strongly self-connected. Note that defining strongly connected in terms of simple overlap leaves something to be desired, as that would include cases of tangential overlap as being considered strongly connected. Here the interiors of the two individuals are constrained to be overlapping, a stronger condition.[15]

D 27. $StrongC(x, y) =_{def} \exists w \exists z(w \leq x \wedge z \leq y \wedge SSC(w \oplus z))$

In contrast to Strongly Connected, one can define the relation that holds between individuals which are only tangentially connected, or externally connected (ExtC), where two m-individuals are connected, but their interiors do not overlap, as given in D28.[16]

D 28. $ExtC(x, y) = C(x, y) \wedge \neg C(ix, iy)$

Many other variations of connectedness have been explored in the literature. Varzi (2007) discusses relations such as by-connection (BC), a three-place relation, which can state that, for instance, two cups on a table are connected by virtue of both being on the table. This leads to a binary mediately connected (MC) relation. It is possible that such notions will be useful for describing the connectedness conditions holding for objects such as *eyes* or *fingers*, which in some languages are treated as inherently plural. For instance,

[15] This property is referred to in Varzi (2007, p. 1003) as 'firm connection'.

[16] An equivalent to the external connection relation is used by Krifka (1998) to define "adjacency structures" which are used to model paths.

my thumb and index finger are not directly connected, but are mediately connected via my hand.

D 29. $BC(x, y, z) =_{def} C(x, z) \wedge C(z, y)$

D 30. $MC(x, y) =_{def} \exists z BC(x, y, z)$

To these variations, we can add, in a slight abuse of terminology, a notion of PROXIMATELY CONNECTED (ProxC), where two entities are not contiguous, but "sufficiently near". (To foreshadow, this is an appropriate connectedness relation to describe swarming insects for instance.) To establish this relation, a definition of distance must be given. I assume a standard distance function d, which given two entities, returns the distance between them.[17] A relation which holds of entities which are 'proximate' to one another can be defined as in D31. Of course, the aim is not to characterize nearness in any absolute sense, but relative to the entities under discussion. I assume that the relevant value of n will be determined with respect to the predicate.

D 31. $ProxC(x, y) =_{def} d(x, y) \leq n$

This concludes the addition of different varieties of connectedness that will be useful for characterizing different entity types with respect to their relation to different countability classes. It will be convenient in what follows to continue to use lattice representations, but enrich them with regions representing which elements are connected. For instance, the example of four m-individuals a and b, two halves of a sphere, and c and d, my left shoe and the Eiffel Tower, can be represented in 4.14. The four m-individuals generate the lattice and individuals which are connected are covered by a shaded region.

I now turn to applying these different distinctions to the problems of countability.

[17] More precisely, a distance function d is from the Cartesian product of a set with itself into the nonnegative real numbers: $D : S \times S \rightarrow \mathcal{R}^+$. Four further conditions are standardly specified:
(i) d(x, y) = 0 iff x = y
(ii) d(x, y) > 0
(iii) d(x, y) = d(y, x)
(iv) d(x, z) ≥ d(x, y) + d(y, z)

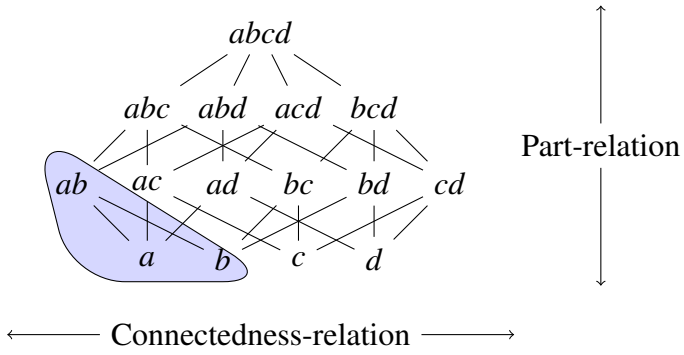

Figure 4.14: Lattice Decorated with Connectedness Relations

4.3 Modeling Countability in Terms of Mereotopological Properties

The previous section provided an extension to the mereological notions common in semantics, leading to definitions of different concepts related to parthood, wholeness and connectedness. This section uses this richer framework to address the semantic underpinnings of countability distinctions. In discussing the elements necessary to generate the scale of individuation, I argued that to treat the variation in predicates with regard to countability, both the presence/absence of minimal units as well as the degree of contiguity that typically holds among those minimal units must be recognized. The previous section has provided ways to express these critical contrasts. The property of Maximally Strongly Self-Connected characterizes individual wholes. Different forms and degrees of connectedness permit characterizing the different spatial configurations in which entities may stand. Thus, the formal machinery developed in the preceding sections provides a framework which address the differences seen across entity types and across languages at the end of chapter 3. Further, as I will show in detail in section 4.4, the framework provides a natural account of the grammatical number systems discussed in chapter 2.

After a discussion of the theoretical assumptions about the semantic representation of nouns, I will proceed in sections 4.3.2–4.3.3 to establish a set of denotation types for the different types of nouns discussed in chapter 3. This account will proceed much in the same way that previous accounts have fixed nominal types according to properties such as,

for example, "cumulative". The difference simply lies in using mereotopological notions: different semantic classes of nouns can be distinguished by virtue of which connectedness relations may or must hold among the individuals in their denotation.

I will assume, following much work in the literature on nominal semantics and generics (Carlson, 1980; Krifka, 1995; Zamparelli, 1999; Müller-Reichau, 2006), that nominal meaning includes both reference to objects and reference to kinds, or more broadly to *concepts*, as proposed in Krifka (1995). These different elements are related by a realization relation R between *concepts* and the instances of the concept at the level of objects. The referential use of *dog* then is tied to the realizations of the concept *dog*, or as I will sometimes say, the instances of a concept. A second relation, discussed in Krifka (1995), is a taxonomic relation T holding between *kinds/concepts* and their subkinds or subconcepts, where the subkind reading of *dogs* would correspond to "different types of dogs", such as *beagle* or *terrier*. In the system of Krifka (1995), which I will largely follow here, the basic meaning of a noun is given by (9),[18] which gives "the property of being a specimen or subspecies, or an individual sum of specimens or subspecies" (Krifka 1995, p. 399). In other words, one way an entity may satisfy the predicate *dog* is if it is an individual dog, or a plural individual composed of dogs, which are objects related to the kind *dog* by the realization relation R. In (9), variables ranging over object-level entities are subscripted with *o*, while variables ranging over kind-level entities are subscripted with *k*. Yet another way to satisfy the predicate is on the taxonomic reading, where the entity must be an individual subkind of dog (*terrier*) or sum of subkinds, which are related to the kind *dog* by the taxonomic relation T. I will not treat the taxonomic readings of nouns, so will simplify in what follows by only considering the R relation.[19]

(9) $\lambda y_k \lambda i \lambda x_o [\mathrm{R_i}(x_o, y_k) \vee \mathrm{T_i}(x_o, y_k)]$

In this section, I will be considering the denotation of nouns relative to their property-level denotation. In the following section, I will consider the meaning of nouns at the level of

[18] Krifka (1995), as in Krifka (1989) discussed in section 4.1.3, also includes a cardinal measure function, *NU(x)*, as a primitive element of noun meaning. As that is not necessary for the account here, I will not include it in the representation here, although there is nothing that I will present that would be incompatible with that approach.

[19] I will also simplify in what follows by giving only the extensional variant.

their lexical entries, i.e. when *dog* in English designates the singular form.

I will also assume that the representations developed in this section are compatible with a cognitive view of meaning. Model-theoretic approaches often have a reputation for being not connected to notions such as "conceptualization" which was given much consideration in chapter 3. Yet, the two approaches are not incommensurate, and the approach that I will take here is one that would accommodate both. This is in accordance with other views on semantics, such as expressed in this quote from Krifka (1998):

> "Model-theoretic semantics in the tradition of Montague, Lewis and Cresswell has often been seen as opposed to cognitive approaches as developed by Lakoff, Langacker, Wierzbicka, Jackendoff, Bierwisch, and others. It was believed that model-theoretic semantics is forced to a 'realistic' view, in which natural-language expressions are interpreted by real entities, like objects and possible worlds, whereas cognitive semantics is concerned with cognitive models of reality. I don't see that model-theoretic semantics has to be realistic in this sense. We can make use of the techniques developed in the model-theoretic tradition and assume that expressions are interpreted by elements of conceptual structures that in turn are related to real entities by some extra-linguistic matching. This is how I would like to understand the algebraic structures discussed below: As attempts to capture certain properties of the way we see the world, not as attempts to describe the world how it is." (Krifka, 1998, p. 198)

Having set out the basic preliminaries and spirit of the analysis, I turn to adapting the mereotopological system developed in section 4.2 to modeling the countability distinctions discussed in chapter 3.

4.3.1 Individuals

I treat predicates which refer to individuals (in the pre-theoretical, non-mereological sense) as those predicates which are restricted to designating whole objects or sums thereof. Very little must change about the way that we represent the meaning of count nouns, except for the substitution of the property of being a whole object, viz. MSSC, for the property

atomicity. For instance, the condition given in (10) of an INDIVIDUAL-predicate is an adaptation of the definition of an atomic predicate in D13. This condition prevents strongly connected (overlapping) individuals from being in the extension of an individual predicate. The singular property of, for instance, *dog*, given in (11) will simply pick out individual whole dogs.[20]

(10) INDIVIDUAL(P)$\rightarrow \forall x[P(x) \rightarrow \exists y[y \leq x \wedge MSSC(y, P)]]$
(*A predicate is* INDIVIDUAL *if any m-individual that satisfies the predicate has a (not necessarily proper) part which is a MSSC m-individual.*)

(11) $[\![dog]\!] := \lambda x_O[\mathrm{R}(x_O, \mathrm{Dog}) \wedge \mathrm{MSSC}(x_O)]$

Using the full strength of maximally strongly self-connected (MSSC) m-individuals overcomes the problems for atomicity related to predicates such as *fence* or *twig*, discussed in section 4.1.2. As a result of the maximality condition, the problems with predicates such as *fence* simply do not arise: INDIVIDUAL-predicates restrict their reference to pick out only MSSC m-individuals, which by definition will be in each case the largest individual satisfying, for instance, *fence*. A part of a fence may have all the requisite features to satisfy the predicate *fence*, but it will not be a MSSC-individual. Thus, when *a fence* is uttered, this will pick out only the whole object, although there are a variety of ways to pick out portions of the fence, which, if considered in isolation, would also qualify as an instance of the kind *fence*.

This explanation extends to other examples problematic for atomicity, such as *twig* and *line* or even the more fanciful example of the Pope's crown (Wiggins, 1980, p. 73), which is problematic since it contains three smaller crowns.[21] When using MSSC as the foundational property of countable nouns, these problems simply don't arise.

[20] I will often simplify the representation, as in (11), by leaving out the property relative to which the entities are MSSC when it is clear.

[21] One class of predicates, exemplified by *bouquet*, that have been suggested as being "homogeneous" nouns would appear to not admit of this explanation. Rothstein (2010) provides the following comments:

> "... suppose that I have a bouquet of flowers that I split, giving half to my daughter and half to her friend. Then, either there is a single bouquet that has been split so that each girl has half a bouquet, or each girl has a bouquet of flowers (albeit smaller than the original one)." (Rothstein, 2010, p. 356)

4.3.2 Substances

I now turn to the designation of substance terms such as *water* or *iron*. Part of the intuition behind the property of divisiveness, as well as Schwarzschild's (to appear) or Chierchia's (1998b) analysis of non-countable nouns, is that pure substances such as *oil*, *water* or *wood* have the trait that when one refers to an m-individual, then simultaneously one refers to multiple instances. Yet, as reviewed in section 4.1.2, modeling this intuition using the property of divisiveness leads to the unwelcome consequence of the "minimal parts problem". The other route, collapsing substance terms with plural entities, captures this intuition but violates other intuitions, viz. that substances are simply different from a collection of entities.

Within the framework developed here, the intuition that substances involve multiplicity in a certain sense can be satisfied, yet something still stronger can be said. Substances manifest the property that a given instance of, say, water, will overlap with other instances of water. In terms of the framework just developed, instances of a substance come *strongly connected* with other instances. In this sense, the presence of an instance of water implies the presence of another instance of water. Recalling that the converse of *connected* is *separated*, another way to fix the intuitions for this property is that no instance of a liquid/substance is separated from all the other instances of that liquid/substance. The relevant property is that each occurrence of water is *locally strongly connected*.[22]

The condition given in (12) states that substance predicates require their extension to be comprised of m-individuals which are strongly connected to other m-individuals of the same substance. This will be satisfied, for instance, by a section of a pool of water—it overlaps other sections of the pool of water, which are again water. Also, since connectedness is implied by parthood, the whole pool of water is strongly connected to its parts which are

Here, after the division of the flowers, each set of flowers will be treated as an MSSC m-individual, and resulting in two bouquets. This example points to an interesting fact about these predicates that their designation is temporally dependent. Yet, unlike Rothstein, I do not think that this implicates that countability in general needs to be relativized via a contextual parameter, but simply that these particular predicates are context dependent. Further discussion of similar predicates which present a temporary grouping of a group of objects, such as *flotilla*, is provided by Mihatsch (2007).

[22] 'Locally' connected indicates that not every instance of water is connected, since clearly water may appear parceled out into puddles, rivers, glasses and so on; but locally, an instance of water is always connected to another instance of water, where the second instance may contain or be contained in the first.

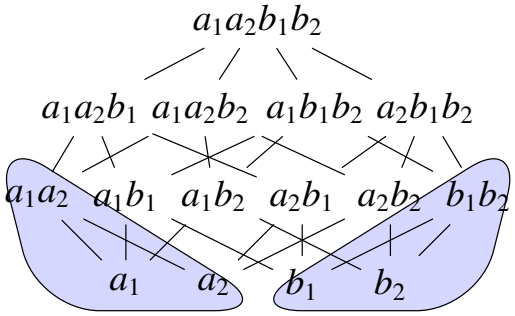

Figure 4.15: The Connected Components of a Sum of Two Portions of a Substance

water.

(12) SUBSTANCE(P) $\rightarrow [\forall x[P(x) \rightarrow \exists x'[R(x', P) \wedge x' \neq x \wedge \text{StrongC}(x, x')]]$
(*If P is a substance predicate then all m-individuals that satisfy P are strongly connected to a distinct m-individual of the same substance.*)

This condition will also be satisfied when taking the sum of two portions of water, say two pools—here supposing that an instance of water was the sum of two portions of water $a \oplus b$. Then our condition tells us that there are additional strongly connected m-indiviudals for each element of the sum-individual. In the lattice representation in figure 4.15 these are given by a_1,a_2 and b_1,b_2, where the shading again indicates connectedness in the relevant sense.

This condition on the semantics of substances has several advantages over the properties used by previous approaches. First, unlike cumulativity, the condition in (12) distinguishes substance predicates from nouns with plural denotations. Plural nouns such as *boys* do not satisfy the condition of having their referents being (locally) strongly connected.

Second, the difficulties facing divisiveness are avoided under this account. The difficulty was that many substance nouns do not lend themselves to being 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 what divisiveness predicts. The characterization of substance terms given in (12), 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. Yet, there is no commitment to infinite part-taking, the problematic aspect of divisiveness.

A final remark concerns the observation that there is no (natural) individuating standard for substance entities (Cartwright, 1975; Quine, 1960). It turns out that in the system developed here establishing a "minimal instance" for a substance term leads to contradiction, which is a welcome result. I take a minimal instance to be one which (i) does not contain another entity of the same sort (which is the atomicity property) and (ii) is able to be isolated from other instances of the entity (which is the property of being separated from other instances, in other words not being connected to other instances). A predicate such as *boy* will clearly satisfy having a minimal instance. Yet, an entity falling under a substance predicate, assuming that it obeys the condition given in (12), cannot satisfy these conditions. Suppose there was a minimal instance of water. In order to satisfy (12), this instance must be strongly connected to another instance of water. Assuming that this other instance is strongly connected to another instance by containing it violates the first condition of being a minimal instance, whereas assuming that it is connected in any other manner (viz. overlap) violates the second. Thus, assuming that a substance term has a minimal instance leads to a contradiction.

In sum, treating substance terms as designating entities which must be strongly connected to like entities avoids some of the classic problems facing divisiveness, yet still represents the core intuitions that one would want the semantic theory to represent.

4.3.3 Aggregates

Aggregates, such as *sand* or *rice*, are a hybrid of the first two categories. The discussions in chapters 2 and 3 brought out that aggregates have the distinguishing characteristics of possessing at once minimal parts and also designating groups which come together in some manner. The various quotes from grammatical descriptions of the languages in chapter 2 showed that the propensity of the referents of these terms to appear together was central to the meaning of the various "collective" categories across the languages. The contiguity of elements again was relevant for the countability of English nouns in the experiments

of Middleton et al. (2004). The mereotopological framework developed in the preceding sections provides sufficient means to represent aggregates as a category of entities that differ from individuals or sums.

I will discuss the two most prominent types of aggregates, what I have termed granular and collective aggregates, which have been most salient in the discussion of the countable/non-countable distinction and in the discussion of collective/singulative classes in chapter 2. The core idea that I will pursue is that while aggregate terms designate minimal parts, and accordingly will designate sums of these minimal parts, the different connection relations can be exploited to designate "connected clusters" of the entities in question. In other words, the aim is to capture that, for instance, sand occurs as individual grains, as sums thereof, as well as in clusters. Similarly, hornets may come individually, in sums thereof, or in clusters. Different aggregate types can be distinguished according to the different connection types.[23] The property of coming in connected clusters has not been recognized in treatments of nominals semantics, but this property has ramifications for these terms, I will argue, not only with respect to their membership in particular morphological countability classes, but also with respect to different types of modification than nouns which only have plural manifestations.

Individual and substance predicates differ as to what type of m-individuals may serve as instantiations of the concept. Individual predicates require MSSC individuals, while substance predicates require strongly connected individuals, thus disallowing MSSC individuals, and therefore the two predicate types are clearly distinct. This does not appear to be the correct strategy for aggregate terms—for there is no evidence that the difference between *foliage* and *leaves* is about different types of m-indiviudals. The relevant characteristic of aggregate predicates is rather the canonical configuration of their referents, i.e. typically coming in clusters. The strategy that I will pursue is then to specify the denotational space of aggregates to include, in addition to MSSC and sum individuals, clustered individuals. The denotational space of an aggregate will then contrast with that of an individual predicate, which will contain only MSSC and sum individuals.

[23]This discussion will not necessarily exhaust the different denotation types which can be defined within this system and are useful for natural language semantics. As mentioned in section 4.2.3, items such as pair body parts may be best represented through an indirect connection relationship, although much more investigation of such terms is necessary.

To arrive at this condition on aggregates, two auxiliary notions must be defined. First, I give a generalized version of the relation of *by-connection*, Transitively Connected (TransitiveC) relative to a property, in D32. This relation allows us to state whether two entities are connected through a series of like entities. For instance, two grains of sand on opposite sides of a pile of sand are not externally connected to one another, but they are transitively (externally) connected through a series of grains of sand within the pile.

D 32. *TransitiveC(x,y,P,C,Z)* $=_{def} \forall z \in Z[R(z, P) \wedge (x = z_1 \wedge y = z_n) \wedge Cz_1z_2 \wedge Cz_2z_3 \ \ldots \ \wedge Cz_{n-1}z_n]$
where $Z = \{z_1, z_2, \ldots, z_n\}$ *(Transitively Connected)*
(x and y are transitively connected relative to a property P, a connection relation C, and a set of entities Z, when all members of Z satisfy P and x and y are connected through the sequence of z_i*s in Z.)*

With the definition of transitive connection in hand, the notion of a connected cluster can be given in D33. This definition provides a way to refer to collections of m-individuals that are related by a particular connection relation, such as a pile of sand or a swarm of locusts. At the same time, this does not enforce taking maximal clusters—a desirable property as, for instance, a pile of sand, which is a clustered individual, includes many other clustered individuals.

D 33. *Cluster(x,P,C)* $=_{def} \exists Z[x = \bigoplus Z \wedge \forall z, z' \in Z \ \exists Y[TransitiveC(z, z', P, C, Y)]]$ *(Clustered Individual)*
(x is a cluster relative to a property P and a connection relation C iff x is a sum of entities falling under the same property which are all transitively connected relative to some set Y under the same property and connection relation.)

A general condition on aggregates is given in (13) below, which states the denotational space of aggregate nouns includes MSSC individuals, clusters and sums. In (13) and subsequent examples, I use the notation C as a variable over connectedness relations, e.g. External Connectedness and Proximately Connected. I will also use CLUSTER_C to represent the set of clustered individuals under the relevant connection relation C, SUM to represent the set of sum individuals and MSSC for the set of MSSC individuals.

(13) AGG(P,C) $\rightarrow \forall x[P(x) \rightarrow x \in \text{CLUSTER}_C \cup \text{SUM} \cup \text{MSSC}]$
(*If P is an aggregate predicate relative to a connection relation then all m-individuals that satisfy P have cluster, MSSC or sum reference.*)

The type of connection relation has been underspecified in the definition, but instantiating it with different connection types results in the different aggregate types discussed. A clustered individual under the External Connectedness relation designates a cluster of individuals which are connected by touching, which is an appropriate designation for a canonical instance of *sand* or *rice*. The Proximately Connected relation, which, as discussed in section 4.2.3, holds when two entities are co-located and near one another. A clustered individual under the Proximately Connected relation then will specify a group of individuals all within a particular distance of one another, for instance, berries which are all within one inch of each other. This serves as a representation of collective aggregates such as insects or berries, which do not appear in groups where each individual touches another, but where they are all at some distance from one another.

This definition is very permissive, but is useful as a general definition of aggregates. *Sand*, for example, will be true of single grains (MSSC individuals) or sums thereof, as well as clusters of sand (e.g. piles), which are clusters of externally connected individuals, or a combination of these types.

I have argued that it is necessary to add a different sort of type of theoretical entity, a clustered individual, in order to adequately account for aggregate terms—but would there not be a more parsimonious way to treat these nouns simply as sums or groups, which are already well-established in semantic theory? For instance, clustered individuals are after all just a special type of sum, viz. one where the parts are connected in the relevant manner. In fact, the categorization of a noun as an individual or aggregate noun would not be different quantitatively, as long as the minimal parts are the same. Taking the example of a set of leaves which can be described as either *leaves* (a sum) or *foliage*. Whether one treated *foliage* as a clustered individual or a sum, there would be the same amount of stuff and the same number of component parts. In either case, the number of entities in the model would be the same in terms of the cardinality of the two sets. But assimilating *foliage* to a plural individual on a par with *leaves* would leave out the connectedness element, and just this is

what seems to be crucial. An arbitrary assortment of leaves does not constitute *foliage*, as was discussed in chapter 3 and demonstrated by the example repeated in (14). Rather, it is necessary that the leaves, and other elements, designated by *foliage* are locally connected through the tree or stem from which they originate. *Foliage* then refers to these clustered individuals (or sums of clustered individuals).

(14) I raked the leaves/#foliage into a pile.

The distinction being made here is a qualitative one concerning the types of entities falling under the concept—aggregate nouns are specifying more information than individual nouns do. This is not only relevant for accounting for how these nouns are treated in different grammatical number systems, as will be made explicit in section 4.4, but also for understanding certain types of modification which differ in interpretation depending on whether an aggregate or plural (sum) individual is being modified. A set of adjectives including *dense*, *spare*, *thick* and *thin* have distinct interpretations when used with aggregates that are not generally available with plural interpretations. I provide examples in (15) and (16) demonstrating this variability of interpretation.

(15) a. *dense foliage* (= the components of the foliage are in close proximity to one another)
b. *dense leaves* (=the material constitution of the individual leaves is dense)

(16) a. *thin hair* (=the hair is thinly distributed over the head)
b. *thin hairs* (=each hair is thin, the diameter is small)

While a full analysis of these sorts of adjectives is not the immediate concern here, the contrast is clear: when the adjectives are used with plural nouns, the modification proceeds distributively, when the adjectives are used with aggregates, they modify the proximity of elements to one another distributed over the entire aggregate. In other words, when these adjectives modify nouns such as *foliage*, the degree of connection between the elements is being modified. Distinguishing nouns which designate clustered individuals from sum individuals provides a way into analyzing these types of modifiers.

Aggregate terms are also distinct from canonical “group nouns”, such as *committee*.

Nouns such as *committee* designate something distinct from their constituent elements, since, for instance, a committee may remain the same committee even if all of its members have changed. A cluster individual is not a different type of thing from its constituent elements, i.e. a cluster of sand is simply sand. A clustered individual is only particular type of configuration of that type of individual.

These two types of nouns, group nouns and aggregates, are also handled in distinct ways by languages. For instance, in Welsh as discussed in chapter 2, canonical group nouns such as *committee* or *government* fall into a distinct class from the aggregate terms. The group nouns do not dispose of the singulative marker, but allow pluralization. This is in line with the standard treatment of groups as atomic entities that can feed pluralization, and so distinct from aggregate nouns which cannot.

In sum, this analysis brings out the similarities and differences between aggregates and both substance-predicates and individual-predicates. Similar to substance-predicates, their referential domains include connected m-individuals, although the type of connection is categorically different. Yet, like individual-predicates, the referential domain also includes natural minimal elements which are MSSC individuals, although the referential domain includes more than these individuals. These distinctions are summarized in table 4.2, where the nominal types are ordered according to the strength of the connection relation.

Nominal type	contains MSSC individuals	contains clustered individuals	connection type
individual (*dog*)	Yes	No	None
collective aggregate (*ant*)	Yes	Yes	Proximate
granular aggregate (*rice*)	Yes	Yes	External
substance (*water*)	No	Yes	Strong

Table 4.2: Nominal Types and Their Properties

This analysis further directly yields the scale of individuation. The degree of connectedness can be seen as related to the degree of individuation for a given predicate type: the most individuated predicate types will be those where their denotation space consists of separated m-individuals, while the more connected the m-individuals falling under a predicate are, the less their degree of individuation. Given this correspondence, the typological

findings of chapter 3, namely the scale of individuation, can be derived via the different strengths of the connection-relation. Table 4.3 show the correspondence between the different entity types and the different degrees of connectedness.

strong	<	external	<	proximate	<	MSSC (separated)
substance	<	granular aggregates	<	collective aggregates	<	individuals

Table 4.3: Relation between Degrees of Connectedness and Degrees of Individuation

This concludes the presentation of the different referential types. I now turn to examining the implications of this analysis.

4.4 Number Morphology

The last sections have developed formal machinery to treat different entities types. This section puts the various formal distinctions to work in giving an account of some of the different grammatical number systems seen in chapter 2. Against the backdrop of the richer framework developed, the actual account of the semantics of number morphology is, as one would hope, rather simple. Three different types of elements are in play. First, nouns which allow number distinctions may vary in different languages as to what is designated by the basic form of a noun, e.g. that which is not morphologically marked by a number morpheme. For instance, an unmarked countable noun in English refers to single entities, while an unmarked noun in Welsh may refer to clustered individuals. As for the number morphemes, they vary both as to (i) the nature of the operation, viz. pluralization or unitization, and (ii) the presuppositions in effect about the domains over which they operate. I begin by discussing English, the most familiar case, and then proceed to Welsh and finally Dagaare.

4.4.1 English

The core result that one would like from a compositional account of the number morphology of English is an adequate treatment of where number morphology, viz. pluralization via *-s*, is permitted to occur and where it is not. For one part of this problem, namely

treating countable nouns, I am able to follow traditional accounts of plurality with only minor modifications. The additional distinctions made in section 4.3 come into play for accounting for where plurality is disallowed—both nouns such as *water* and *sand* violate the prerequisites of plurality, but in different fashions.

Countable nouns in English, for which I will consider *dog* to be the parade example, do not require any radically different account from those used in the past, and I am able to maintain the standard view, with the exception that the property used is of course no longer atomicity but being a MSSC individual. Thus, we can state with, for instance, Link (1983) that the denotation of the singular form of *dog* is the property satisfied by singular entities, as given in (17), although MSSC is now the relevant property determining what a singular entity is.

(17) $[\![dog]\!] := \lambda x_O[\mathrm{R}(x_O, \mathrm{Dog}) \wedge \mathrm{MSSC}(x_O)]$

The plural morpheme *-s* can be treated as an operator, which presupposes a set of, in our case, MSSC individuals and then returns the set of sums.[24] The presupposition associated with the operation is expressed following a period after the lambda expression, as in "λP.Presupposition[$-\ -\ -$]". I also make use of the * operator from Link (1983), which given a set, returns the sums of things in that set, as defined in D15 in section 4.1.3. This analysis of *-s* is given in (18), which parallels Link's analysis of the plural.

(18) $[\![-s]\!] := \lambda P \lambda x.P_{MSSC}[P^*(x) - \mathrm{MSSC}(x)] = \lambda P.P_{MSSC}[P^*(x) \wedge x \in \mathrm{SUM}]$

Given this treatment of the plural, its incompatibility with substance predicates such as *water* is immediate: the plural morpheme requires MSSC individuals, which are not provided by substance predicates. There may be polysemy through which a second representation for, e.g. *water*, is available that presupposes a set of MSSC individuals, such as servings of water. But the availability of this meaning is here taken to be a fact about

[24] I am here assuming the "exclusive" analysis of the plural, where singular entities are *not* included. While the question of whether singular entities are included in the denotation of plural nouns is quite controversial, nothing crucial hinges on this for the points being made here about English or Welsh. The exclusive analysis of the plural is, however, critical for the analysis of Dagaare, as discussed in Grimm (to appear (a)). For further discussion of the exclusive and inclusive analyses of the plural, see Krifka (1989); Sauerland et al. (2005); Spector (2007); Farkas & de Swart (2010); Grimm (to appear (b)).

Countability Category	Singular	Plural	Gloss
Singular/Plural	cadair	cadairiau	'chair'
Collective/Unit	cacyn-en	cacwn	'hornet'
Non-Count	llefrith		'milk'

Table 4.4: Grammatical Number Categories in Welsh

lexicalization or, in some cases, coercion of noun meaning.

Similarly, granular aggregate nouns in English such as *sand*, *rice* or *gravel*, like substance terms, designate the entire semilattice, as in (19). Again, pluralization fails, since there is not a set of MSSC individuals from which sum individuals can be formed, rather the denotation of the noun includes much more.

(19) $[\![sand]\!] := \lambda x_O[R(x_O, \text{Sand}) \wedge x_O \in \text{CLUSTER}_{EC} \cup \text{SUM} \cup \text{MSSC}]$

I now turn to Welsh, which exploits the denotational space of such nouns in a different way.

4.4.2 Welsh

The countability categories of nouns in Welsh can be treated in an analogous manner to the analysis of English, despite the fact that the inventory differs. The basic data from Welsh discussed in chapter 2 is repeated in table 4.4.

The class of singular/plural nouns in Welsh admit of the exact same treatment as for English nouns with a singular/plural distinction. The denotation an individual predicate such as *cadiar* 'chair' and the plural morpheme *-ou*/*-au* parallel the entries for their English counterparts and are given in (20) and (21), respectively.

(20) $[\![cadair]\!] := \lambda x_O[\text{R}(x_O, \text{Chair}) \wedge x_O \in \text{MSSC}]$

(21) $[\![-ou/-au]\!] := \lambda P \lambda x.P_{MSSC}[P^*(x) - \text{MSSC}(x)] = \lambda P \lambda x.P_{MSSC}[P^*(x) \wedge x \in \text{SUM}]$

Similarly, nothing special needs to be said about substance terms in Welsh—they behave in the same manner as substance terms in English do.

The collective/unit class is the interesting case, but it too can be treated a manner similar

to the analysis of English. First, in the same way that unmarked count nouns in English are restricted to being interpreted as singular, collective/unit nouns in Welsh are in their basic form restricted to clustered individuals. The lexical entry for *cacwn* 'hornet' is given in (22), stating that *cacwn* designates entities which satisfy the property *hornet* and are clustered individuals by virtue of being proximately connected. The denotational space of the lexical entry does not include arbitrary sums. This is consistent with the discussion of Welsh in chapter 2, where it was noted that these nouns had a strong sense of designating a collection rather than an arbitrary set of individuals. Additionally, if arbitrary sums were included, it would be expected that these nouns could be directly modified by cardinal numbers—but nouns in this class can only take cardinal modification once the singulative has been applied.

(22) $[\![cacwn]\!] := \lambda x_O[\mathrm{R}(x_O, \mathrm{Hornet}) \wedge x_O \in \mathrm{CLUSTER}_{PC}]$

The singulative marker can then be treated as an operator as in (23). In parallel to the restriction on the plural morpheme which presupposes MSSC individuals, the singulative morpheme presupposes a set of clustered individuals. Given such a set of clustered individuals, it returns the parts which are MSSC, i.e. the units. This analysis respects the intuitions of the Welsh grammarians discussed in chapter 2 who noted that the collective form designated groups from which individual units were specified. Since the singulative presupposes clustered individuals, its application to count nouns such as *cadiar* 'chair' is correctly ruled out. Similarly, the application of the singulative would also be ruled out for substance nouns—although substance nouns possess clustered individuals, they do not possess MSSC-individuals which would serve as the output of the operation.

(23) $[\![-en/-yn]\!] := \lambda Q \lambda x.Q_{cluster}[x \leq Q \wedge x \in \mathrm{MSSC}]$

As mentioned in chapter 2, collective nouns in Welsh may also allow the application of the plural following the singulative, as shown in (24). The definitions of the singulative and plural given here allows for this possibility since the application of the singulative provides the MSSC individuals which in turn feeds the application of the plural.

(24) a. grawn
grain
'grain'
b. gron-yn
grain.SING
'a single grain'
c. gron-ynn-au
grain.SING.PL
'grains'

4.4.3 Dagaare

The number morphology of Dagaare poses a special set of challenges. The table in 4.5 recapitulates the different countability categories that could be distinguished in Dagaare. I will proceed by establishing the content of lexical entries for nouns of different categories and then treat the inverse morpheme *-ri* and the singulative *-ruu*.

The nouns which are individual predicates simply designate singular entities, and so can be treated in the same manner as countable nouns such as *dog* in English. The nouns of the collective aggregate class, however, typically contain clustered individuals, related by either proximate connection, as is plausible for insects or small fruits, or indirect connection, which is plausible for grouped body parts. Yet, according to my Dagaare informants, while the intuition that these entities come in clusters seems strong, these nouns also designate arbitrary sum individuals, whether they come as a clustered individual or not. This is consistent with the fact that, unlike in Welsh, nouns of this aggregate class accept cardinal modification, as shown in section 2.2. Thus, the denotational space of collective aggregate nouns in Dagaare includes clustered individuals and sum individuals, but not MSSC individuals. I give representative lexical entries for individual and collective aggregate nouns in (25) and (26), respectively.

(25) $[\![bie]\!] := \lambda x_O[\mathrm{R}(x_O, \mathrm{Child}) \wedge x_O \in \mathrm{MSSC}]$

(26) $[\![biri]\!] := \lambda x_O[\mathrm{R}(x_O, \mathrm{Seed}) \wedge x_O \in [\mathrm{CLUSTER}_{PC} \vee \mathrm{SUM}]]$

Countability Category	Singular	Plural	Gloss
Individual	bíé	bíírí	'child'
Collective Aggregate	bírì	bíè	'seed'
Granular Aggregate	múórúú	múó	'grass'
Substance	kùó		'water'

Table 4.5: Grammatical Number Categories in Dagaare

As discussed in chapter 2, the inverse number marking system differs significantly from the standard plural marking strategy observed for individual nouns in English or Welsh. Yet, there is a simple way to analyze *-ri* as a form of negation of the unmarked number value for the noun in question. This is an intuitive version of the function of inverse number marking, and is in essence a formal semantic update of the analysis of Kiowa in Wonderly (1954).[25]

-ri can be modeled as the operation of relative complementation (C), where the complementation operation is relativized to the domain. *-ri* applied to an individual noun will yield a plural denotation, while *-ri* applied to a collective aggregate noun will yield a singular denotation. Representative derivations are given in table 4.6, demonstrating that this analysis clearly secures the desired interpretations. In prose, for lexically singular nouns, the application of *-ri* gives the complement of the denotation of a singular noun, viz. the complement of the relevant set of MSSC individuals. The value returned is the sums formed from the atoms, less the atoms themselves, which is in turn exactly the value of the noun's plural denotation. For lexically plural nouns, the application of *-ri* gives the complement of the denotation of a plural noun, viz. the complement of the relevant set of sums. The value returned is the atoms which form the sums, which is in turn exactly the value of the noun's singular denotation.

Turning to the singulative *-ruu*, this morpheme can be treated in the same way as the singulative in Welsh, except that it selects for predicates which are granular aggregates, i.e. those predicates where the clustered individuals stand in an external connection relation to one another. Given a granular aggregate predicate, *-ruu* returns the MSSC individuals. In

[25]This line has also been developed independently in Bach (2007) and Bach (2008) for Kiowa.

Individual Noun	Collective Aggregate Noun
$[\![\text{bi-}]\!] := \lambda x_O[\text{R}(x_O, \text{Child}) \wedge x \in \text{MSSC}]$	$[\![\text{bi-}]\!] := \lambda x_O[\text{R}(x_O, \text{Seed}) \wedge x \in [\text{CLUSTER} \vee \text{SUM}(x)]$
$[\![\text{bi-}]\!]$ + ri	$[\![\text{bi-}]\!]$ + ri
$([\![\text{ bi- }]\!])^C$	$([\![\text{ bi- }]\!])^C$
$(\lambda x_O[\text{R}(x_O, \text{Child}) \wedge x \in \text{MSSC}])^C$	$(\lambda x_O[\text{R}(x_O, \text{Seed}) \wedge x \in [\text{CLUSTER} \vee \text{SUM}(x)])^C$
$\lambda x_O[\text{R}(x_O, \text{Child}) \wedge x \in \text{SUM}]$	$\lambda x_O[\text{R}(x_O, \text{Seed}) \wedge x \in \text{MSSC}$
= PL(bi-)	= SG(bi-)

Table 4.6: Derivations of Individual and Collective Aggregate Nouns

this way, singulatives across the different languages can be modeled in the same fashion—what differs is simply which type of aggregate is presupposed.

(27) $[\![-ruu]\!] := \lambda Q \lambda x. Q_{Gran.Agg}[x \leq Q \wedge x_O \in \text{MSSC}]$

Having treated the various morphological number markers, I now turn to a more general review of how the mereotopological system has fared according to the diagnostics laid out in section 4.1.4.

4.5 Taking Stock

The discussion in section 4.1.4 laid out some goals for what a formal system should represent to treat countability. In this section, I briefly recapitulate how those goals are met.

First, the system developed here formally distinguishes whole objects and substances. Whole objects are modeled as MSSC individuals, which are bounded entities and which do not overlap with any other individuals. In contrast, substances are modeled as entities which must be strongly connected to like entities, i.e. a substance m-individual always overlaps with another m-individual of the same substance. I argue that this treatment aligns with the intuitive distinction between objects and substances.

These are not distinctions that are immutably fixed by the world. Rather the distinction is whether speakers see certain predicates as referring to, for instance, MSSC individuals. Thus, in keeping with the quote from Krifka (1998) given in the opening of 4.3, predicates

specified in this system and their properties are intended to represent how speakers construe the world.

The system also is able to formally distinguish aggregates from plurals, which is not possible in standard mereology. Adding connection relations permits distinguishing clustered individuals from ordinary sum individuals. Further, through different degrees of connectedness, the system is capable of representing gradations of individuation and provides a basis for the scale of individuation. Having made these distinctions at an abstract level, section 4.4 shows how they can be implemented in different languages.

Chapter 5

Summary and Future Work

The major claim of this dissertation is that grammatical number categorization reflects different degrees of individuation associated with nominal descriptions. This claim contains two parts. First, countability is not a binary distinction, although some languages may only have two primary grammatical number categories, viz. countable and non-countable. Rather, countability is a scalar phenomenon. Second, a noun's countability status is not purely a grammatical fact, but is based in individuation properties associated with the entity being described. A subsequent claim of this dissertation is that to properly model different countability types, it is necessary to enrich standard mereological frameworks with topological relations.

The empirical core of this dissertation consists of the languages discussed in chapter 2, whose grammatical number systems possess three or more grammatical number categories. The examination of languages with a collective/singulative class provides a different perspective on what underlies countability than typical data from an investigation of, say, English does. These languages recognize aggregates as qualitatively different from singular entities or non-countable nouns, which in turn indicates that a binary view of countability based on a contrast between discrete objects and substances is insufficient. Further, there is a substantive lexical semantic generalization underlying the collective/singulative class. In contrast to nouns of the singular/plural class, which designate entities which habitually appear singly, the collective/singulative class is comprised of entities which habitually come

together. Accounts which relate grammatical number to a contrast between discrete objects and substances, I argue, oversimplify the typological space of grammatical number systems.

The scale of individuation elaborated in chapter 3 provides a view of grammatical number which accommodates the distinctions found in languages such as Welsh or Dagaare. Countability is no longer limited to a single distinction between countable and non-countable nouns, but may include a range of distinctions, including, of course, a collective/singulative class. I argue that grammatical number categories must cohere to a scale of individuation, but this does not mean that they must do so uniformly: languages may vary as to the number of divisions made along the scale and as to the different cut-off points. This predicts that there will be cross-linguistic variation in how languages realize their grammatical number categories, but this variation should be constrained by the scale of individuation: grammatical number categories must be based in coherent combinations of individuation types.

This view of countability also provides answers to many of the standard criticisms of meaning-based accounts. While I argue that countability is based in nominal meaning, this does not indicate that there must be a one-to-one relation between particular entity types and particular grammatical number categories—the empirical landscape shown here indicates that this would be far too restrictive. The relation between grammatical number categories and entities in the world is not direct, but mediated by how entities in the world are construed in terms of individuation properties, such as having a regular shape or being spatially contiguous with other entities of the same type. The fact that there may be two nouns which on certain occasions may describe the same set of objects, e.g. *leaves* and *foliage*, does not serve as an indication of arbitrariness in countability classification, as is often claimed. Instead, these two nouns provide different perspectives on this type of entity, and they in turn differ in their individuation properties and therefore their countability status.

The investigation of different grammatical number systems and the elaboration of the scale of individuation points to several short-comings of modeling countability in a standard mereological framework. Two properties which emerged from the discussion of individuation, being a whole object and habitually coming together, are absent in standard

formulations of mereology. I address this by enriching mereology with topological relations which can express these distinctions. Whole objects may be characterized in terms of Maximally Strongly Self-Connected individuals, while the notion of coming together may be characterized through various connectedness relations. I further showed how this shift in perspective from a mereological view to a mereotopological view on countability reconciled several recalcitrant problems with the standard mereological properties, such as the "minimal parts" problem and predicates such as *fence*.

There are several areas which I have set aside for future work. While this dissertation has extended the empirical base of grammatical number systems in investigating countability, this is a very tiny portion of the different grammatical number systems which exist. More typological work will undoubtably lead to refinement of the scale of individuation, and determine how other factors such as referentiality interact with the realization of grammatical number.

Additionally, many grammatical number systems make distinctions which are related, but not identical, to the collective/singulative classes examined here. For instance, Italian disposes of an irregular plural *-a* which applies to a lexically restricted set of nouns. The variety of different cases where this plural arises is too complex to set out here, but to illustrate, these nouns include paired body parts, such as *braccia* 'arms'. This irregular plural contrasts in meaning with the regularly inflected plural *bracci* which designates, for instance, arms of objects. Coming from the perspective of the collective/singulative class of, say, Welsh, it would be plausible to think that these special plurals might designate pairs of arms; however, this is not the case. The detailed discussion of Acquaviva (2008) shows that, for instance, *braccia* is used to talk about arms belonging to many different people, citing an advertising slogan for doughmixers: *la forza di venti braccia* 'the strength of 20 arms' (Acquaviva, 2008, p. 150). As Acquaviva (2008) indicates, the generalization appears to be that these plurals are distinguishing entities that are parts of "natural aggregates" (p. 150). While this differs from the typical collective, many of the components explored in this dissertation would still appear to have a role to play in specifying, for instance, the underlying structure of the natural aggregates from which the parts are taken. Other grammatical number systems no doubt hold many other surprising variants in store for researchers on countability.

Even within the English lexicon, much remains to be explored. Throughout this dissertation, I have primarily considered natural concrete entities. For the reasons given in the opening of chapter 2, this facilitated cross-linguistic comparison. Yet, a fuller account must investigate other types of nouns, such as artifactual nouns (*hammer*, *furniture*) or abstract nouns (*arrival*, *happiness*). It is currently unclear how many of the categorization principles related to individuation discussed for concrete nouns are applicable to artifactual or abstract nouns. Some studies indicates that for abstract nouns there are other factors in play (Payne & Huddleston, 2002; Grimm, 2012). For instance, qualities such as *kindness* are typically non-countable, but when referring to specific events may be countable, as in *his many kindnesses towards those in need*. A systematic study of the countability behavior of abstract nouns will surely reveal more.

Finally, there is much room to explore how the framework proposed here for the domain of objects may connect with the domain of events. Theories of event structure also make use of mereology or part structures and therefore it would be interesting to explore how the mereotopological system proposed could be extended to treat events. Many researchers, such as Bach (1986) and Krifka (1989), argue for a tight link between objects and events. It remains to be seen if the novel denotation types defined here for aggregates have natural analogues in the domain of events.

Bibliography

Acquaviva, Paolo. 2008. *Lexical plurals: A morphosemantic approach.* Oxford: Oxford University Press.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21.435–483.

Allan, Keith. 1977. Classifiers. *Language* 53.285–311.

Allan, Keith. 1980. Nouns and countability. *Language* 56.41–67.

Anderson, P. W. 1972. More is different. *Science* 177.393–396.

Anttila, Arto, and Adams Bodomo. 2009. Prosodic morphology in Dagaare. *Selected proceedings of the 38th Annual Conference on African Linguistics (ACAL 38)*, ed. by Fiona McLaughlin, Eric Potsdam, and Masangu Matondo. Cascadilla Proceedings Project, Somerville, Massaschusetts.

Baayen, Harald; Cristina Burani; and Robert Schreuder. 1997. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. *Yearbook of morphology 1996*, ed. by Geert Booij and Jaap van Marle, 13–33. Kluwer Academic Publishers.

Baayen, R.H.; R. Piepenbrock; and L. Gulikers. 1996. *Celex2.* Philadelphia, PA: Linguistic Data Consortium.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 15.5–16.

BACH, EMMON. 2007. Parochial and universal semantics: Semantic typology and little studied languages. Handout for pre-ALT Workshop on Language Documentation and Linguistic Typology, Paris, France.

BACH, EMMON. 2008. On morphosemantics: The internal meanings of words. Paper presented at the Copenhagen Linguistic Circle.

BALE, ALAN, and DAVID BARNER. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26.217–252.

BARNER, DAVE, and ALAN BALE. 2002. No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecication. *Lingua* 112.771–791.

BARNER, DAVID, and JESSE SNEDEKER. 2005. Quantity judgments and individuation: evidence that mass nouns count. *Cognition* 97.41–66.

BEHRENS, LEILA. 1995. Categorizing between lexicon and grammar: The MASS/COUNT distinction in a cross-linguistic perspective. *Lexicology* 1.1–112.

BLOOM, PAUL. 1990. *Semantic structure and language development*. Massachusetts Institute of Technology dissertation.

BLOOM, PAUL. 1994. Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua* 92.297–329.

BLOOM, PAUL. 1996. Intention, history, and artifact concepts. *Cognition* 60.1–29.

BLOOMFIELD, LEONARD. 1933. *Language*. New York: Henry Holt and Co.

BODOMO, ADAMS. 1997. *The structure of Dagaare*. Stanford: CSLI.

BODOMO, ADAMS. 2006. The second plural in Dagaare. Unpublished ms.

BORER, HAGIT. 2005. *Structuring sense I: In name only*. Oxford: Oxford University Press.

BORG, ALBERT J. 1980. An aspectual distinction in verbs and nouns in Maltese. *Archivum Linguisticum* 11.107–117.

Brennan, Andrew. 1988. *Conditions of identity*. Oxford: Clarendon Press.

Brinton, Laurel J. 1998. Aspectuality and countability: A cross-categorial analogy. *English Language and Linguistics* 2.37–63.

Bunt, Harry C. 1979. Ensembles and the formal semantic properties of mass terms. *Mass terms: Some philosophical problems*, ed. by Francis J. Pelletier, 249–277. Dordrecht: Reidel.

Bunt, Harry C. 1985. *Mass terms and model-theoretic semantics*. Cambridge: Cambridge University Press.

Carlson, Gregory N. 1980. *Reference to Kinds in English*. New York and London: Garland Publishing.

Cartwright, Helen. 1975. Amounts and measures of amount. *Noûs* 9.143–164.

Casati, R., and A. C. Varzi. 1999. *Parts and places: The structures of spatial representation*. Cambridge, MA: MIT Press.

Champollion, Lucas. 2010. *Parts of a whole: Distributivity as a bridge between aspect and measurement*. University of Pennsylvania dissertation.

Cheng, Chung-Ying. 1973. Comments on Moravcsik's paper. *Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics*, ed. by Jaakko Hintikka, Patrick Suppes, and J.M.E. Moravcsik, 215–220. Dordrecht: Reidel.

Cheng, Lisa L.S.; Jenny Doetjes; and Rint Sybesma. 2008. How universal is the universal grinder? *Linguistics in the Netherlands 2008*, ed. by Marjo van Koppen and Bert Botma, 50–62. Amsterdam/Philadelphia: John Benjamins.

Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of "semantic parameter". *Events and grammar*, ed. by Susan Rothstein, 53–104. Dordrecht: Kluwer Academic Publishers.

Chierchia, Gennaro. 1998b. Reference to kinds across languages. *Natural Language Semantics* 6.339–405.

CHIERCHIA, GENNARO. 2010a. Language, thought, and reality after Chomsky. *The Chomsky notebook*, ed. by Julie Franck and Jean Bricmont, 142–169. New York: Columbia University Press.

CHIERCHIA, GENNARO. 2010b. Mass nouns, vagueness and semantic variation. *Synthese* 174.99–149.

CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

CLARKE, B. L. 1981. A calculus of individuals based on "connection". *Notre Dame Journal of Formal Logic* 22.204–218.

CLAUSEN, DAVID; ALEX DJALALI; SCOTT GRIMM; SVEN LAUER; TANIA ROJAS-ESPONDA; and BETH LEVIN. 2010. Extension, ontological type, and morphosyntactic class: Three ingredients of countability. Extended abstract. Stanford University.

CONTINI-MORAVA, ELLEN. 2000. Noun class as number in Swahili. *Between grammar and lexicon*, ed. by Ellen Contini-Morava and Yishai Tobin, 3–30. Amsterdam/Philadelphia: John Benjamins.

CORBETT, GREVILLE G. 1991. *Gender*. Cambridge: Cambridge University Press.

CORBETT, GREVILLE G. 1996. Minor number and the plurality split. *Rivista di Linguistica* 8.101–122.

CORBETT, GREVILLE G. 2000. *Number*. Cambridge: Cambridge University Press.

DAGAARE LANGUAGE COMMITTEE. 1982. *A guide to Dagaare spelling*. Wa: Catholic Press.

DAKUBU, M.E. KROPP. 2005. *Collected language notes on Dagaare grammar*, *Collected Language Notes*, vol. 26. Legon, Ghana: Institute of Language Studies.

DAVIDSON, DONALD. 1973. On the very idea of a conceptual scheme. *Proceedings and addresses of the American Philosophical Association*, vol. 47.

DENNY, J. PETER. 1976. What are noun classifiers good for? *Proceedings of the 12th Chicago Linguistic Society*, 122–132. Chicago: Chicago Linguistics Society.

DIMMENDAAL, GERRIT J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42.214–261.

DIMMENDAAL, GERRIT JAN. 1983. *The Turkana language.* Dordrecht: Foris Publications.

FARGETTI, C. M. 2001. *Estudo fonológico e morfossintático da língua Juruna.* Unicamp University dissertation.

FARKAS, DONKA F., and HENRIËTTE DE SWART. 2010. The semantics and pragmatics of plurals. *Semantics and Pragmantics* 3.1–54.

GENTNER, DEDRE, and LERA BORODITSKY. 2001. Individuation, relational relativity and early word learning. *Language acquisition and conceptual development*, ed. by M. Bowerman and S. Levinson, 215–256. Cambridge: Cambridge University Press.

GENTNER, DEDRE, and MELISSA BOWERMAN. 2009. Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, ed. by J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, S. Özçalişkan, and K. Nakamura, 465–480. New York: Psychology Press.

GIL, DAVID. 1987. Definiteness, noun-phrase configurationality, and the count-mass distinction. *The representation of (in)definiteness*, ed. by E.J. Reuland and A.G.B. ter Meulen, 254–269. Cambridge: MIT Press.

GIL, DAVID. 1996. Maltese 'collective' nouns: A typological perspective. *Rivista di Linguistica* 8.53–87.

GIL, DAVID. 2003. English goes Asian: Number and (in)definiteness in the Singlish noun-phrase. *Noun phrase structure in the languages of Europe, empirical approaches to language typology*, ed. by Frans Plank, 467–514. Berlin/New York: Mouton.

GILLON, BRENDAN. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15.597–639.

GOODMAN, NELSON. 1951. *The structure of appearance.* Cambridge, MA: Harvard University Press.

GRANDY, RICHARD E. 1973. Reply to Moravcsik. *Approaches to natural language*, ed. by Jaakko Hintikka, Patrick Suppes, and J.M.E. Moravcsik, 295–300. Dordrecht: Reidel.

GRANDY, RICHARD E. 2007. Sortals. *Stanford encyclopedia of philosophy (fall 2008 edition)*, ed. by Edward N. Zalta. Online: `http://plato.stanford.edu/archives/fall2008/entries/sortals/`

GREENBERG, JOSEPH. 1966. *Language universals.* The Hague: Mouton.

GRIMM, SCOTT. 2012. Abstract nouns and countability. Talk delivered at Linguistic Society of America 2012 Annual Meeting.

GRIMM, SCOTT. to appear (a). Individuation and inverse number marking in Dagaare. *Count and mass across languages*, ed. by Diane Massam. Oxford: Oxford University Press.

GRIMM, SCOTT. to appear (b). Plurality is distinct from number neutrality. *Proceedings from the 41st meeting of the Northeast Linguistic Society*, ed. by Lena Fainleib, Nicholas LaCara, and Yangsook Park. Amherst, MA: GLSA.

GRIMM, SCOTT, and BETH LEVIN. 2011. *Furniture* and other functional aggregates: More and less countable than mass nouns. Talk delivered at Sinn und Bedeutung 16, Utrecht, The Netherlands.

GÜLDEMANN, TOM. 2000. Noun categorization systems in Non-Khoe lineages of Khosian. *Afrikanistische Arbeitspapiere* 63.5–33.

HARBOUR, DANIEL. 2008. *Morphosemantic number: From Kiowa noun classes to UG number features.* Dordrecht: Springer.

HASPELMATH, MARTIN. 2005. Occurrence of nominal plurality. *The world atlas of language structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 142–145. Oxford: Oxford University Press.

HASPELMATH, MARTIN. 2008. Creating economical morphosyntactic patterns in language change. *Linguistic universals and language change*, 185–214. Oxford: Oxford University Press.

IMAI, MUTSUMI. 1999. Constraint on word-learning constraints. *Japanese Psychological Research* 41.5–20.

IMAI, MUTSUMI, and DEDRE GENTNER. 1997. A cross-linguistic study of early word learning: universal ontology and linguistic influence. *Cognition* 62.169–200.

JESPERSEN, OTTO. 1913. *A modern English grammar on historical principles, part II: Syntax*, vol. 1. London: Allen & Unwin.

JESPERSEN, OTTO. 1924. *The philosophy of gammar.* London: Allen & Unwin.

JONES, MORRIS, and ALAN R. THOMAS. 1977. *The Welsh language: Studies in its syntax and semantics.* Cardiff: University of Wales Press for the Schools Council.

JOOSTEN, FRANK. 2003. Accounts of the count-mass distinction: A critical survey. *Linguisticae Investigationes* 26.159–173.

JOOSTEN, FRANK; GERT DE SUTTER; DENIS DRIEGHE; STEF GRONDELAERS; ROBERT J. HARTSUIKER; and DIRK SPEELMAN. 2007. Dutch collective nouns and conceptual profiling. *Linguistics* 45.85–132.

KATAMBA, FRANCIS. 2003. Bantu nominal morphology. *The Bantu languages*, ed. by Derek Nurse and Gérard Philippson, 103–120. New York: Routledge.

KENNEDY, JACK. 1966. *Collected field reports on the phonology of Dagaari.* Collected Language Notes 6. University of Ghana, Accra: Institute of African Studies.

KING, GARETH. 2003. *Modern Welsh: A comprehensive grammar.* Routledge.

KOSLICKI, KATHRIN. 1999. The semantics of mass-predicates. *Noûs* 33.46–91.

KRACHT, MARCUS. 2002. On the semantics of locatives. *Linguistics and Philosophy* 25.157–232.

KRIFKA, MANFRED. 1989. Nominal reference, temporal constitution and quantification in event semantics. *Semantics and contextual expressions*, ed. by Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, 75–115. Dordrecht: Foris Publications.

KRIFKA, MANFRED. 1992. Thematic relations as links between nominal reference and temporal constitution. *Lexical matters*, ed. by Ivan Sag and Anna Szabolcsi, 29–53. Stanford: CSLI.

KRIFKA, MANFRED. 1995. Common nouns: A contrastive analysis of English and Chinese. *The generic book*, ed. by Gregory N. Carlson and Francis J. Pelletier, 398–411. Chicago University Press.

KRIFKA, MANFRED. 1998. The origins of telicity. *Events and grammar*, ed. by Susan Rothstein, 197–235. Dordrect: Kluwer.

LANDAU, BARBARA; LINDA B. SMITH; and SUSAN S. JONES. 1988. The importance of shape in early lexical learning. *Cognitive Development* 3.299–321.

LANDMAN, FRED. 1989. Groups I. *Linguistics and Philosophy* 12.559–605.

LANDMAN, FRED. 2011. Count nouns — mass nouns neat nouns — mess nouns. *The Baltic international yearbook of cognition, logic and communication*, ed. by Barbara Partee, Michael Glanzberg, and Jurgis Šķilters, vol. 6 Formal Semantics and Pragmatics: Discourse, Context, and Models. Manhattan, KS: New Prairie Press.

LANGACKER, RONALD W. 1987. Nouns and verbs. *Language* 63.53–94.

LASERSOHN, PETER. 2011. Mass nouns and plurals. *Semantics: An international handbook of natural language meaning*, ed. by Klaus von Heusinger, Claudia Maienborn, and Paul Portner, vol. 2, 1131–1153. Berlin: De Gruyter.

LEONARD, HENRY S., and NELSON GOODMAN. 1940. The calculus of individuals and its uses. *Journal of Symbolic Logic* 5.45–55.

LEŚNIEWSKI, STANISLAW. 1916. *Podstawy ogólnej teoryi mnogosc*. Moskow: Prace Polskiego Kola Naukowego w Moskwie, Sekcya matematycznoprzyrodnicza. (Eng. trans.

by D. I. Barnett: 'Foundations of the General Theory of Sets. I', in S. Leśniewski, Collected Works, ed. S. J. Surma, J. Srzednicki, D. I. Barnett, and F. V. Rickey, Dordrecht: Kluwer, 1992, Vol. 1, pp. 129–173).

LEWIS, DAVID. 1991. *Parts of classes.* Oxford: Basil Blackwell.

LIMA, SUZI. 2010. About the count-mass distinction in Yudja: A description. *University of British Columbia working papers in linguistics: Papers for WSCLA 15, the fifteenth workshop on structure and constituency in languages of the Americas*, ed. by Beth Rogers and Anita Szakay, 157–164. Vancouver: University of British Columbia.

LINK, GODEHARD. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Meaning, use, and interpretation of language*, ed. by Rainer Bauerle, Christoph Schwarze, and Arnim von Stechow. Berlin: de Gruyter.

LØNNING, JAN TORE. 1987. Mass terms and quantification. *Linguistics and Philosophy* 10.1–52.

LUCY, JOHN A. 1992. *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis.* Cambridge, MA: Cambridge University Press.

MAHO, JOUNI FILIP. 1999. *A comparative study of Bantu noun classes.* Göteborg: Acta Universitatis Gothoburgensis.

MCCAWLEY, JAMES D. 1975. Lexicography and the count-mass distinction. *Proceedings of the 1st Annual Meeting of the Berkeley Linguistics Society*, 314–321. Berkeley: Berkeley Lingustics Society.

MIDDLETON, ERICA L.; EDWARD J. WISNIEWSKI; KELLY A. TRINDEL; and MUTSUMI IMAI. 2004. Separating the chaff from the oats: Evidence for a conceptual distinction between count noun and mass noun aggregates. *Journal of Memory and Language* 50.371–394.

MIEHE, GUDRUN, and KERSTIN WINKELMANN (eds.) 2007. *Noun class systems in Gur languages, vol. I Southwestern Gur languages (without Gurunsi).* Cologne: Rüdiger Köppe Verlag.

MIFSUD, MANWEL. 1996. The collective in Maltese. *Rivista di Linguistica* 8.29–51.

MIHATSCH, WILTRUD. 2007. Taxonomic and meronomic superordinates with nominal coding. *Ontolinguistics*, ed. by Andrea C. Schalley and Dietmar Zaefferer. Berlin: Mouton de Gruyter.

MOLTMANN, FREDERIEKE. 1997. *Parts and wholes in semantics.* New York: Oxford University Press.

MORAVCSIK, J. M. E. 1973. Mass terms in English. *Approaches to natural language*, ed. by Jaakko Hintikka, Patrick Suppes, and J.M.E. Moravcsik, 263–285. Dordrecht: Reidel.

MUFWENE, SALIKOKO S. 1984. The count/mass distinction and the English lexicon. *Parasession on lexical semantics*, 200–221. Chicago: Chicago Linguistic Society.

MÜLLER-REICHAU, OLAV. 2006. *Sorting the world: On the relevance of the kind-level/object-level distinction to referential semantics.* Universität Leipzig dissertation.

NSOH, AVEA E. 2002. Classifying the nominal in the Gurenɛ dialect of Farefare of Northern Ghana. *Journal of Dagaare Studies* 2.83–95.

NUNBERG, GEOFFREY, and ANNIE ZAENEN. 1992. Systematic polysemy in lexicology and lexicography. *Euralex II*, 387–396.

OHTA, ITARU. 1989. *A classified vocabulary of the Turkana in Northwestern Kenya.* African Study Monographs. Kyoto, Japan: The Center for African Area Studies at Kyoto University.

OXFORD ENGLISH DICTIONARY. 2012. Oxford University Press. Online Edition.

PALMER, FRANK R. 1971. *Grammar.* Harmondsworth: Penguin.

PAYNE, JOHN, and RODNEY HUDDLESTON. 2002. Nouns and noun phrases. *Cambridge grammar of the English language*, ed. by Rodney Huddleston and Geoff Pullum, 323–524. Cambridge: Cambridge University Press.

PELLETIER, FRANCIS J. 1979. Non-singular reference: Some preliminaries. *Mass terms: Some philosophical problems*, ed. by Francis.J. Pelletier, 1–14. Dordrecht: Reidel.

PELLETIER, FRANCIS J. 1991. Mass terms. *Handbook of metaphysics and ontology*, ed. by Hans Burkhardt and Barry Smith, 495–499. Munich: Philosophia Press.

PELLETIER, FRANCIS J. to appear. Lexical Nouns as Both +mass and +count , and Neither +mass nor +count. *Count and mass across languages*, ed. by Diane Massam. Oxford: Oxford University Press.

PELLETIER, FRANCIS J., and LENHART K. SCHUBERT. 2004. Mass expressions. *Handbook of philosophical logic*, ed. by Dov Gabbay and Franz Guenthner, vol. 10, 249–335. Dordrecht: Kluwer.

PIÑÓN, CHRISTOPHER. 1995. *An ontology for event semantics.* Stanford dissertation.

PRASADA, SANDEEP; KRAG FERENZ; and TODD HASKELL. 2002. Conceiving of entities as objects and stuff. *Cognition* 83.141–165.

PRESS, IAN. 2009. Breton. *The Celtic languages*, ed. by Martin J. Ball and Nicole Müller, 427–487. New York: Routledge.

QUINE, W. V. O. 1960. *Word and object.* Cambridge, MA: MIT Press.

RAVID, DORIT, and LUBNA HAYEK. 2003. Learning about different ways of expressing number in the development of Palestinian Arabic. *First Language* 23.41–64.

ROEPER, PETER. 1997. Region-based topology. *Journal of Philosophical Logic* 26.251–309.

ROTHSTEIN, SUSAN. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27.343–397.

SAMUELSON, LARISSA K., and LINDA B. SMITH. 1999. Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition* 73.1–33.

SAUERLAND, ULI; JAN ANDERSSEN; and KAZUKO YATSUSHIRO. 2005. The plural is semantically unmarked. *Linguistic evidence*, ed. by Stephan Kepser and Marga Reis, 413–434. Berlin: Mouton de Gruyter.

SCHWARZ, ANNE. 2010. Specificity and number: What really counts in the noun class systems of some Gur languages. Handout from presentation at Humboldt University, Institut für Asien- und Afrikawissenschaften, Linguistisches Kolloquium, November 23, 2010.

SCHWARZSCHILD, ROGER. to appear. Stubborn distributivity, multiparticipant nouns and the count/mass distinction. *Proceedings of the 39th meeting of the North East Linguistics Society*. Amherst, MA: GLSA.

SEIFART, FRANK. 2009. Towards a typology of unitization: Miraña noun classes compared to numeral classifiers and singulatives. Unpublished Ms.

SHARVY, RICHARD. 1978. Maybe English has no count nouns: Notes on Chinese semantics. *Studies in Language* 2.345–365.

SIMONS, P. M. 1987. *Parts. A study in ontology*. Clarendon: Oxford.

SMITH, BARRY. 1996. Mereotopology: A theory of parts and boundaries. *Data and Knowledge Engineering* 20.287–303.

SMITH-STARK, T.C. 1974. The plurality split. *Papers from the 10th Chicago Linguistic Society*, ed. by M.W. La Galy, R.A. Fox, and A. Bruck, 657–671. Chicago: Chicago Linguistic Society.

SOJA, NANCY N.; SUSAN CAREY; and ELIZABETH S. SPELKE. 1991. Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition* 38.179–211.

SPECTOR, BENJAMIN. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. *Presuppositions and implicatures in compositional semantics*, ed. by Uli Sauerland and Penka Stateva, 243–281. New York: Palgrave/ MacMillan.

STOLZ, THOMAS. 2001. Singulative-collective: Natural morphology and stable cases in Welsh number inflexion of nouns. *Sprachtypologie und Universalienforschung* 54.52–76.

SUTCLIFFE, EDMUND F. 1936. *A grammar of the Maltese languages*. Valletta: Progress Press.

TARSKI, ALFRED. 1935. Zur grundlegung der Booleschen Algebra. I. *Fundamenta Mathematicae* 24.177–198, Eng. trans. by J. H. Woodger: On the foundations of the Boolean Algebra, in A. Tarski, Logics, Semantics, Metamathematics, Papers from 1923 to 1938, Oxford: Clarendon Press, 1956, pp. 320-341.

TAYLOR, BARRY. 1977. Tense and continuity. *Linguistics and Philosophy* 1.199–220.

TIERSMA, PETER. 1982. Local and general markedness. *Language* 58.832–849.

TSOULAS, GEORGE. 2008. On the grammar of number and mass terms in Greek. *Proceedings of the 2007 workshop in Greek syntax*, *MIT Working Papers in Linguistics*, vol. 56, Cambridge, MA, 131–146.

VARZI, ACHILLE C. 2007. Spatial reasoning and ontology: Parts, wholes, and locations. *Handbook of spatial logics*, ed. by Marco Aiello, Ian E. Pratt-Hartmann, and Johan van Benthem, 945–1038. Berlin: Springer.

VARZI, ACHILLE C. 2011. Mereology. *The Stanford encyclopedia of philosophy (spring 2011 edition)*, ed. by Edward N. Zalta. Stanford, CA. Online: `http://plato.stanford.edu/archives/spr2011/entries/mereology/`

WARE, ROBERT X. 1975. Some bits and pieces. *Synthese* 31.379–393.

WATKINS, LAUREL (WITH THE ASSISTANCE OF PARKER MCKENZIE). 1984. *A grammar of Kiowa.* Lincoln: University of Nebraska Press.

WHITEHEAD, ALFRED N. 1920. *The concept of nature.* Cambridge: Cambridge University Press.

WHITEHEAD, ALFRED N. 1929. *Process and reality.* New York: The MacMillan Company.

WIERZBICKA, ANNA. 1988. *The semantics of grammar.* Amsterdam/Philadelphia: John Benjamins.

WIESE, HEIKE, and JOAN MALING. 2005. Beers, kaffi, and schnaps: Different grammatical options for restaurant talk coercions in three Germanic languages. *Journal of Germanic Linguistics* 17.1–38.

WIGGINS, DAVID. 1980. *Sameness and substance.* Oxford: Blackwell.

WILSON, WILLIAM A. 1962. *The vowel system of Dagaari and Lobiri.* University of Ghana, Accra: Institute of African Studies.

WISNIEWSKI, EDWARD J.; MUTSUMI IMAI; and LYMAN CASEY. 1996. On the equivalence of superordinate concepts. *Cognition* 60.269–298.

WISNIEWSKI, EDWARD J.; CHRISTOPHER A. LAMB; and ERICA L. MIDDLETON. 2003. On the conceptual basis for the count and mass noun distinction. *Language and Cognitive Processes* 18.583–624.

WONDERLY, WILLIAM J. 1954. Number in Kiowa: Nouns, demonstratives and adjectives. *International Journal of American Linguistics* 20.1–7.

YI, BYEONG-UK. 2009. Chinese classifiers and count nouns. *Journal of Cognitive Science* 10.209–225.

ZAMPARELLI, ROBERTO. 1999. *Layers in the determiner phrase.* New York, NY: Garland.

ZUCCHI, SANDO, and MICHAEL WHITE. 2001. Twigs, sequences and the temporal constitution of predicates. *Linguistics and Philosophy* 24.223–270.

ZWARTS, JOOST, and YOAD WINTER. 1997. A semantic characterization of Locative PPs. *SALT VII*, ed. by Aaron Lawson, 294–311. Ithaca, NY: Cornell University.

ZWICKY, ARNOLD. 2001. Counting chad. Handout from Stanford Semantics Fest, March 16, 2001.