Grammatical Number and the Scale of Individuation

Scott Grimm

1 Introduction

The noun *dog* disposes of morphological and syntactic means to be counted, such as pluralization, but the noun *water* doesn't. Average speakers of English are hardly mystified by this contrast–for *dog* names a type of object in the world which is easily countable, namely discrete entities, while *water* refers to a (non-discrete) substance. Despite the strong intuition that grammatical form and ontological type are related, this putative relation is the central point of controversy in the count/mass literature.

Many researchers have proposed that semantic categorization does not underlie the count/mass distinction. One argument hinges on crosslinguistic incommensurability: If the count/mass distinction reflects the ontological nature of entities in the world, viz. discrete entities vs. substances, it is unexpected that a particular entity would be named by a countable noun in one language but a non-countable noun in another. Yet, such doublets are easy to find (see, inter alia, Chierchia 1998; Palmer 1971; Pelletier and Schubert 2004; Rothstein 2010; Ware 1975). For instance, *hair* in English permits non-countable syntax while its French equivalent, *cheveux*, is obligatorily plural (Palmer, 1971). A second argument concerns nouns, like *leaves* and *foliage*, which on certain occasions may describe the same stuff, yet nonetheless fall into different countability classes, countable and non-countable, respectively (Palmer, 1971). Again, so the argument goes, if the grammatical count/mass distinction is a reflection of an ontological distinction, such interchangeability is unexpected.

Based on such evidence, many have come to the same conclusion as found in Chierchia (1998, 57): "while the mass/count distinction is not altogether indifferent to how things are inherently structured, it appears to be independent of it, which is what makes such a distinction a strictly grammatical one." This critique targets 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 correct that a strong view upon which

———————————————
Address(es) of author(s) should be given

the count/mass distinction directly reflects a binary ontological distinction between objects and substances faces fatal challenges from crosslinguistic and intra-linguistic data. Yet, two important assumptions underlie this argument: (i) countability is a binary phenomenon (count vs. mass) and (ii) the *only* relevant extralinguistic distinction is the object/substance distinction. I argue that neither assumption holds.

First, many languages possess three or more categories of grammatical number, indicating that countability is richer than a binary count/mass contrast. Second, increasing evidence from the psycholinguistic literature has shown that, while the object/substance distinction is important, many other factors also play a role, such as ease of distinguishing individual units. Taken together, this indicates that grammatical countability is a much more heterogeneous domain than typically assumed, with multiple grammatical categories of number and multiple extralinguistic factors that influence into which of those categories a noun will fall.

This paper further rejects the supposition that countability must either be a grammatical phenomenon or a semantic one; it instead steers a middle course, recognizing that both semantic and grammatical categorization are active in the organization of grammatical number systems. I argue that while there is certainly some variation as to how grammatical number systems are organized, they also obey broad semantic constraints.

The paper examines the interplay between COUNTABILITY and INDIVIDUATION. I use the term countability to refer to morphosyntactic distinctions made in languages related to grammatical number, such as singular, plural, collective, etc., and the grammatical categories formed upon these distinctions, e.g. the class of nouns with a singular/plural contrast. Section 2 contributes data from several languages which make more than a binary countability contrast. I investigate in particular these languages' collective classes, which have an intermediate status between prototypical countable nouns and prototypical non-countable nouns.

Individuation refers to the propensity for an entity to be construed as an independent individual.[1] After reviewing relevant results from the psycholinguistic literature in section 3.1, I argue that individuation is not a binary distinction (individual vs. non-individual), but a scalar distinction, which tracks the degree to which the referents of nouns are construed as independent individuals. I then argue that several INDIVIDUATION TYPES can be identified, wherein noun meanings are organized into equivalence classes based on shared individuation properties, and the individuation types can be ordered based on the degree of individuation, giving rise to a SCALE OF INDIVIDUATION.

The core thesis of the paper is that the organization of grammatical number systems reflects the semantic organization present in the scale of individuation. This framework developed yields a range of predictions, such as what possible grammatical number systems are and what form they may take, i.e. patterns of markedness in grammatical number systems. In section 4, I explore the typological predictions

---

[1] 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.

of this scalar view on individuation and consider how this view connects with other factors known to be relevant, such as animacy.

Section 5 discusses how this view speaks both to the regularity and variation found in grammatical number systems. Grounding countability distinctions in entities in the world and our perception of and interaction with them provides an explanation for the regularity found across grammatical countability systems, since the physical properties as well as how speech communities perceive or interact with, say, water, are likely to be quite commensurable. Yet, the model does not over-determine how reality is mapped to vocabulary: There is often a choice to be made in how entities are to be described and which perceptual or interactional properties of an entity are highlighted. There is a further decision point concerning how many and what type of grammatical categories a language manifests, but this is constrained by the scale of individuation. This view then allows for principled variation, and further provides predictions as to where variability is more likely to be (not) found and what forms of variability one might expect. I show how this view extends to treat cases of lexical variation (*leaves* vs. *foliage*) as well as contextual variation due to "grinding" and "packaging" uses. I then conclude in section 6.

Before proceeding, it is useful to fix some terminology which I use throughout the rest of the article. While the subject of this article is normally described as a distinction between "count" and "mass" terms, the various ways these terms have been used has rendered them unfortunately confusing. For some "mass" is been taken to designate a grammatical category (Bloomfield 1933 and many following him), but for others, it has a narrower range and aligns only with certain types of entities in the world, namely substances (see discussion in Joosten 2003). I will instead describe nouns as COUNTABLE or NON-COUNTABLE, where these two terms are restricted to designating nominal behavior in terms of morphosyntactic characteristics, and make no reference to semantic characteristics. I will also draw a terminological distinction between ENTITY, referring to some portion or element of the external world, and INDIVIDUAL, referring to a construal of an entity as an individual object.

2 Beyond Binary Systems

A set of distributional characteristics have been taken as reflecting a distinction between countable and non-countable nouns in English. As summarized in table 1, countable nouns, like *dog*, pluralize (*dogs*) and combine with numerals (*two dogs*), determiners implicating cardinality (*several dogs*), and the indefinite article (*a dog*). Non-countable nouns do not permit pluralization nor these other combinations. Instead, non-countable nouns may appear bare, or combine with quantifiers such as *much* and *a little* or with measure terms such as *three kilos of*.

The picture of countability presented in table 1 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. First, many nouns may have several uses, some countable and some not, so these diagnostics must be used with great care (see discussion in Pelletier and Schubert 2004). Second, not all nouns which could be classified as countable or non-

Table 1 Core Distributional Properties of Nominal Countability

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*)	*	✓	✓

countable by *some* distributional properties behave identically with respect to *all* of the distributional properties. As discussed in Allan (1980), there are many more nuanced distinctions than a strictly binary division can represent, e.g. the noun *admiration* combines with the indefinite article (*a great admiration*), but does not combine with quantifiers implicating plurality (*several admirations*). Unsurprisingly, different grammatical number systems also make distinctions that do not completely align with a simple division into countable and non-countable nouns.

This section examines grammatical number systems of Welsh, Turkana, Maltese and Dagaare, which express a greater number of countability distinctions beyond the simple binary countable/non-countable contrast expressed in English. For each language, I will discuss how the countable/non-countable distinction is manifested and then other number distinctions that the language makes.

I have restricted this investigation in several ways. First, I focus 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 or elaborate gender systems, since a possible confound lurks in those systems' tendency to code number along with other categories of nominal meaning, such as shape or animacy (see Denny, 1976; Katamba, 2003; Maho, 1999, for Niger-Congo systems).

Second, I focus on what I term NATURAL CONCRETE ENTITIES (*dog*, *water*) to the exclusion of ARTIFACTUAL ENTITIES (*hammer*, *furniture*) or ABSTRACT ENTITIES (*arrival*, *happiness*). Natural concrete entities provide the firmest foundation for comparative study since the relation between nouns and their referents is reasonably clear. If two languages have nouns for describing the entity *apple*, it is reasonable to assume that these nouns refer to the same thing. Less clear is whether words glossed as 'sadness' across different languages 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, artifactual 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. I return to the issue of artifactual and abstract nouns in section 4.3.

Finally, while it is well known that a particular noun's countability status may be able to shift due to contextual influence, in this section and in section 3 I will focus on comparing canonical uses of nouns across languages. Thus, while *dog* may have a use in English which designates dog-stuff in "grinding" contexts such as *dog all over the street*, I will focus on its canonical use as designating a living being. Similarly for "packaging" uses of nouns such as *water* such as *the waiter brought 3 waters*. The

primary reason to set them aside is that there is a high degree of variation, both as to which nouns within a language permit multiple uses, and as to which languages allow different types of nominal flexibility. To illustrate the last point, languages considered here include those that are highly permissive of packaging (Yujda) and also those for which packaging is virtually impossible (Dagaare). I return to contextual variation in section 5.1, but for now, I focus on understanding the core organization of grammatical number systems and countability.

2.1 Countability in Welsh

Welsh resembles English in possessing certain canonical morphosyntactic diagnostics of countability, which distinguish, e.g., non-countable nouns. Yet, Welsh differs from English in possessing, in addition to nouns where the singular value is zero-coded and the plural is overtly coded, a class of nouns for which the plural value is zero-coded (collective) while the singular value is overtly coded (unit). Thus, Welsh grammatically recognizes three classes of nouns: singular/plural, collective/unit and non-countable. I discuss the different classes in turn.

Countable and non-countable nouns Grammars of Welsh typically distinguish countable and non-countable nouns based on 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 and Thomas, 1977, p. 161). Examples of countable and non-countable nouns are given in (1) and (2), respectively.

(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. The examples in (3)–(4), adapted from Jones and Thomas (1977, p. 175–176), show certain quantifiers in Welsh indicating cardinality are not acceptable with nouns designating substances.[2]

(3)   mwyafrif o 'r   llyfr-au/*eira
        majority of DET.DEF book-PL/snow
        'majority of the books/snow'

---

[2] The glossing here follows the Leipzig Glossing Rules. Additional glosses used are: DETPL = determinate plural, DISTPL = distributive plural, HUM = human, NHUM = non-human, SING = singulative, REDUPL = reduplicated plural, SCM = specific class marker, and 3DIM = 3-dimensional class.

Table 2 Grammatical Number Categories in Welsh

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

(4)   amryw o 'r   llyfr-au/*eira
          several of DET.DEF book-PL/snow
          'several of the books/snow'

So far, the grammatical contrast in English between countable and non-countable nouns finds a parallel in Welsh, further strengthened by a parallel notional contrast, at least for natural concrete entities: Countable nouns correspond to entities which are in some sense individuals, while non-countable nouns correspond to material stuff, namely substances and liquids.

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.

Examples (5)–(6) (Jones and Thomas, 1977, p. 157–158) demonstrate how Welsh nominals may differ in their default form: for *adar* ('birds'), 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 also be observed through pronominal agreement: the pronoun in the responses in (5)–(6) agrees in number with the noun in the question, both of which are in bold. (The singulative morpheme is glossed as SING as opposed to SG, which glosses the singular.)

(5)   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.

(6)   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.

Table 3 Sub-types of Collective Nouns in Welsh (adapted from Stolz 2001)

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'
vegetation/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 4 Borrowings from English into the Collective Class in Welsh

English Term	Collective	Unit
bricks	*brics*	*bricks-en*
figs	*ffigys*	*ffigys-en*
gooseberries	*gwsberys*	*gwsberys-en*
peas	*pys*	*pys-en*

The grammatical number system of Welsh provides a clear instance of a tripartite number system. As displayed in table 2, three different categories of opposition in terms of countability are manifested: singular/plural, collective/unit, non-countable.

Descriptions of Welsh's number system inevitably point to a notional distinction accompanying the formal distinction between the singular/plural and collective/unit categories, namely the distinction between those types of entities which habitually co-occur in the world and those which do not. According to this view, that Welsh has two types of countable nouns is not some morphological oddity, but a systematic distinction. Stolz (2001) analyzes the lexical items which comprise the collective class in Welsh and characterizes them as those nouns designating 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. His categorization is partially reproduced in table 3.

Stolz (2001) argues that this grammatical category and the notional category underlying it have been both historically stable, and furthermore, have gone through expansion via borrowing. Stolz (2001) and Jones and Thomas (1977) observe that 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 4.

The typical pattern of borrowing is that the morphologically uncoded form is borrowed (see Tiersma, 1982); however, the borrowings in table 4 are remarkable in that the plural, and morphologically complex, English form 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, designated in English in the plural, and

thus borrowed the plural form and treat it as uncoded, to which the singulative can then apply.

Altogether, the data from Welsh demonstrates that countability distinctions may be richer that a binary countable/non-countable contrast. The morphosyntactic patterns of the language isolate three classes of nouns: singular/plural, collective/unit, and non-countable.[3] Further, the morphological pattern of collective/singulative seems to reflect a categorization of entities in the world.

Implications for Countability One may question 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 aids 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 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 individual entities, 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), showing that certain collective nouns permit pluralization of the singulative form. 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. The collective form and the plural form have distinct meanings: The pluralized singulative form designates a set of individual pieces in contrast to, e.g. a heap of grain.

---

[3] Two more noun types, not relevant to the points being made here, can also be established through contrastive singular/plural marking and verbal agreement. First, Welsh distinguishes characteristic "group" nouns (see Landman (1989), Joosten et al (2007)), nouns designating social bodies or organizations. These nouns manifest a singular/plural distinction; however, the singular form requires the plural form of the verb (Jones and Thomas, 1977, p. 161). Welsh also distinguishes pluralia tantum, i.e. nouns which only appear in plural form. This class includes nouns such as *nefoedd* 'heaven' or *trigolion* 'inhabitants'. Unlike group nouns or collective/unit nouns, they do not manifest a singular/plural contrast.

Treating collectives as plurals faces a serious challenge in accounting for these three-way contrasts.[4]

(7)	Welsh (Stolz, 2001, p. 70)		(8)	Breton (Press, 2009, p. 445)
a.	grawn		a.	deil
	grain			leaf
	'grain'			'leaves' (collective)
b.	gron-yn		b.	deli-enn
	grain-SING			leaf-SING
	'a single grain'			'leaf'
c.	gron-ynn-au		c.	deli-enn-où
	grain-SG-PL			leaf-SING-PL
	'grains'			'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 nouns such as *leaves* and *ants*, which are not canonical examples of countable or non-countable nouns, provide much interest for understanding countability contrasts[5]. These types of nouns will reappear in the other languages that are examined in this section.

2.2 The Collective/Singulative in Nilo-Saharan Languages

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.[6] Number is coded in Turkana by both number suffixes and agreement prefixes, which also code gender. 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,

---

[4] Further notional arguments are given in King (2003, p. 48). Nurmio (2016) provides arguments from morphological patterning against conflating plural and collective values, as well as against treating the collective as simply possessing a null suffix with a plural value, as suggested by Awbery (2009). For instance, the *-en/-yn* suffix does not apply to regular plurals, and further, both the singular and the collective base forms, but not the plural, may feed derivation.

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

[6] The contrast between inherently plural and inherently singular non-countable nouns does not appear to be semantically significant, rather due to etymology. See Dimmendaal (2000, p. 230) for discussion.

Table 5 Countable and Non-Countable Nouns in Turkana

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 6 Types of Countable Nouns in Turkana

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'

the agreement prefix shows whether the noun has a singular or plural form. Table 5, adapted from Dimmendaal (1983, 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'

Countable nouns are further classified by their coding of singular and/or plural interpretations. Table 6 displays the three patterns in Turkana. The singular/plural class codes the plural, while the collective/singulative class codes the singular value via 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. The focus here will be on the contrast between the singular/plural class and the collective/singular class, setting aside the replacement pattern as it does not bear countability per se.

As for Welsh, there is a clear lexical semantic generalization underlying the collective/singulative class which stands in contrasts to the singular/plural class. Dimmendaal (2000, p. 229) observes that "entities congregated in large numbers or quantities tend to be morphologically unmarked in the plural in Nilo-Saharan languages" and for these nouns the singulative codes "an individuated item from a collective or group".

Despite their geographical and genetic separation, Nilo-Saharan and Celtic number systems show substantial overlap in the types of entities inhabiting their collective/singulative classes. Table 7 shows examples from the lexicon of Turkana which include nouns of the same entity types found in Welsh, along with two further entity types—types of people and paired/grouped body parts.

Table 7 Sub-types of Collective Nouns in Turkana (extracted from Ohta 1989)

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

paired/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'

vegetation/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/nga-cok* 'seed'/'seeds'; *e-tab-a/nga-tab* 'piece/pieces of tobacco'

Table 8 Countable Noun Paradigm in Maltese

	Singular	Plural	Indefinite Quantifiers	Cardinals 2-10
Form	ˈskola	ˈskeyyel	ˈhafna ˈskeyyel	ˈerbaˤ ˈskeyyel
Gloss	school.SG	school.PL	many school.PL	four school.PL
Form	ˈraǧel	irˈǧil	ˈhafna irˈǧil	ˈerbaˤ irˈǧil
Gloss	man.SG	man.PL	many man.PL	four man.PL

In sum, the Turkana grammatical number system, and more broadly those of Nilo-Saharan languages, makes a three-way distinction among 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 grammatically recognized categories of countability, namely singular/plural, collective/singulative, and non-countable. 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.

2.3 The Collective/Singulative in Maltese

Maltese, a Semitic language historically in intensive contact with European languages, 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 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. Cardinal modification from 2 to 10 requires the plural form of the noun, while cardinal modification from 11 onwards requires the singular form. Table 8 shows the paradigm for countable nouns. 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.

As in Welsh and Turkana, Maltese possesses a collective/singulative contrast; however, unlike in Welsh and Turkana where the collective is formally a plural, e.g. in terms of agreement, the Maltese collective is zero-coded in the collective interpreta-

Table 9 Collective Noun Paradigm in Maltese

	Collective	Unit	Indefinite Quantifiers	Cardinals 2-10
Form	hût	ˈhûta	ˈhafna ˈhût	ˈerbaˌhûˈtit
Gloss	fish	fish.SING	many fish	four fish.DETPL
Form	ˈnemel	ˈnemla	ˈhafna ˈnemel	ˈerbaˌnemˈlit
Gloss	ant	ant.SING	many ant	four ant.DETPL

tion and formally singular and masculine. The unit interpretation is achieved through the 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 9.

The types of entities which exhibit the collective/singulative paradigm in Maltese are given in table 10. There is substantial overlap with the types of entities seen in Welsh and Turkana—the core members of vegetation, insects and granular aggregate are present. Yet, there are some striking differences—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 and Turkana. 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.

The literature on Maltese has often noted that the collective/singulative class spans entity types that are both countable nouns (*shoes*) and logically non-countable nouns (*iron*, *wood*) (Mifsud 1996, p. 32, Sutcliffe 1936, Borg 1980). A closer look at the data shows that the singulative has two distinct uses, which depend on the type of entity. When the singulative combines with an aggregate noun, the singulative designates one member of the aggregate. When the singulative combines with foodstuff nouns, as well as for certain material nouns, the singulative designates one conventional portion, e.g. the unit interpretation of 'bread' results in 'a loaf'. These different interpretations are shown in table 11. The countable interpretations of the entities in the foodstuff and material categories are reminiscent of interpretations attributed to the "Universal Packager" in English (see Pelletier 1975), e.g. *a water* to designate a glass or bottle of water. Like the Universal Packager, the use of the singulative with non-countable nouns is restricted just to those which have a conventionalized portion interpretation, and may not co-occur with any arbitrary non-countable noun. [7]

---

[7] Since *soap* and other material nouns take singulative/plural coding in Maltese, it is tempting to analyze these nouns as countable rather than non-countable. Yet, other diagnostics demonstrate clearly that the class of nouns taking the singulative is formed from two types of citizens, viz. logically countable entities and logically non-countable entities. First, adjectival modifiers which presuppose individuals, such as *big*, *square*, combine with collectives comprised of naturally countable entities, such as *ut* 'fish', but not those comprised of material, such as *kartun* 'cardboard' unless in combination with the singulative or determinate plural (Gil, 1996). See Author #1 for further examples and discussion.

Table 10 Sub-types of Collective Nouns in Maltese (after Mifsud 1996)

paired entities: 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'

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

granular aggregate: 'ramel 'sand' ; trab 'dust'

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

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

Table 11 Two Interpretations of the Singulative in Maltese: Member or Conventional Portion

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

The grammatical number system in Maltese, as in Welsh, shows every sign of being productive. Evidence from borrowing shows that entities of the appropriate type are simply integrated into the collective category. Mifsud (1996, p. 44-45) discusses borrowings such as the Maltese term for sandals, based on the Italian plural *sandali*, which is integrated into the collective/singulative paradigm as *sandli/sandl-a*. Similarly, the term for brick, based on English *bricks*, appears as *briks/briks-a*. As in Welsh, here it is the plural form in the source language that is borrowed.

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. Further, the lexical semantic domains of each category overlap substantially with those found in Welsh and Turkana.

2.4 Dagaare: Beyond Tripartite Number Systems

Dagaare, a Gur language spoken in Northern Ghana, presents an even more complicated grammatical number system[8]. Dagaare, as spoken in the Central Dialect in and around the town of Jirapa, Ghana, not only grammatically recognizes countable and non-countable nouns, but makes finer distinctions within both classes.

---

[8] The data in this section were gathered over the course of 7 months of fieldwork conducted in 2008 and 2011.

Table 12 Nouns with and without a Singular/Plural Contrast in Dagaare

Noun	Plural Form	Gloss
tìɛ́	tììrí	'tree'
pìɛ́	pèrí	'basket'
dòɔ́	dòrí	'pig'
kòɔ́	-	'water'
kāā	-	'oil'
zīī	-	'blood'
sáálá	-	'charcoal', 'coal'

Countable/Non-Countable Contrast The countable/non-countable contrast in Dagaare can be established by using some of the core distributional properties. First, countable nouns display a singular/plural contrast while non-countable nouns do not have a plural form, as shown in table 12. Similarly, countable, but not non-countable, nouns combine with cardinal modifiers, as shown in (10).[9]

(10)   dò-rí   à-yí
         pig-PL NHUM.PL-two
         'two pigs'   (NHUM = Non-Human Prefix)

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*.[10] More broadly, contextual variability familiar from the discussion of English in the form of grinding, packaging and sorting context was not available in Dagaare. For instance, Dagaare speakers strongly reject the use of nouns such as *báá* 'dog' in "grinding" contexts.[11]

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*. Examples of nouns with the distributive plural are given in table 13. The distributive plural, however, is not able to combine with cardinal numbers, as shown in (11). 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 (10).

---

[9] 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.

[10] Only one non-countable noun, *dāā*, referring to the local alcoholic beverage, was accepted by speakers in combination with cardinal modifiers under a packaging reading. All other nouns were strongly rejected.

[11] Sentences constructed to elicit grinding interpretations, such as in (i-a), were judged unacceptable by Dagaare speakers, who instead required modification, as shown in (i-b).

(i)   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.'

Table 13 Distributive Plural in Dagaare

Noun	Dist. Pl.	Gloss
kòɔ́	kɔ̀nnɛ́ɛ́	'water; types of waters; water in different locations'
móɔ́	móɔ́nɛ́ɛ́	'grass; types of grasses; grass in different locations'

Table 14 Number Marking Possibilities for Non-Countable Nouns in Dagaare

Noun	Singulative	Dist. Pl.	Gloss
kòɔ́	—	kɔ̀nnɛ́ɛ́	'water/–/(types of) waters'
móɔ́	móɔ́rúú	móɔ́nɛ́ɛ́	'grass/blade of grass/grasses'

(11)   *kɔ̀nnɛ́ɛ́   à-yí
         water.DISTPL NHUM.PL-two
         'two waters'

Varieties of Non-Countable Nouns Non-countable nouns can be further classified as to whether they permit a singulative suffix -*ruu*, designating 'a piece of'. Unlike the singulative morphemes discussed so far, the distribution of the singulative in Dagaare is very limited: It combines with fewer than 100 nouns, designating primarily granular aggregate terms or foodstuffs (*sɛ́rímáání* 'pepper', *kpɛ́ɛ́* 'malt', *múó* 'grass', *kàmáánì* 'corn').[12]

The distribution of -*ruu*, shown in table 14, implicates that the non-count nouns 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 which is morphologically coded varies by noun, giving rise to a pattern known as "inverse number marking".[13] Table 15 contrast the nouns for 'child' and 'seed': They share the same stem, yet for 'child' the morpheme -*ri* codes the plural interpretation while for 'seed' it codes the singular interpretation. Additional examples in tables 16 and 17 demonstrate -*ri* coding plural and singular interpretations, respectively. Note -*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 *rɪ* and -*ri*.)

Agreement with demonstratives and number words establish that the singular and plural forms code bona fide singular and plural interpretations. The demonstrative agreement patterns for plural-coded and singular-coded nouns are shown in (12) and (13), respectively: The demonstrative has a singular (*ŋâ*) and plural (-*má*) form, the

---

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

[13] See AUTHOR #1 for further discussion of Dagaare's inverse number marking system and its implications for the understanding of plurality and markedness.

Table 15 The Inverse Number Marking Paradigm in Dagaare

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

Table 16 Marked Plural Nouns in Dagaare

-V Singular	-rI/-nI Plural	Gloss
tìέ	tììrí	'tree'
gbìé	gbèrí	'forehead'
pìέ	pèrí	'basket'
nàŋá	nànní	'scorpion'

Table 17 Marked Singular Patterns

-rI/-nI Singular	-V Plural	Gloss
lúgrí	lúgó	'prop, pillar'
nyágrí	nyágá	'root'
fìlí	fìlέ	'sore'
íílí	íílè	'horn'

plural further distinguishing via a prefix between human (*ba-*) and non-human (*a-*) referents.

(12)   a.   bíé   ŋâ
              child.SG DEM.PROX.SG
              'this child'
       b.   bíí-rí   bà-má
              child-PL HUM.PL-DEM.PROX.PL
              'these children'   (HUM = Human Prefix)

(13)   a.   bí-rì   ŋâ
              seed-SG DEM.PROX.SG
              'this seed'
       b.   bíè   à-má
              seed.PL NHUM.PL-DEM.PROX.PL
              'these seeds'

Number words designating 'two' or more select for plural nouns and again take plural agreement prefixes. The word designating 'one', *yénì*, does not show agreement but acts as a modifier of the noun, directly attached to the stem. The number word agreement patterns for plural-coded and singular-coded nouns are shown in (14) and (15), respectively.

(14)   a.   bì-yénì   (bì- + yénì)
              child-one
              'one child'
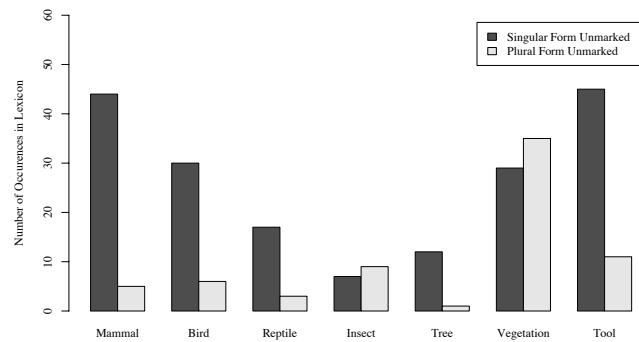
Fig. 1 Number Marking Across Semantic Domains

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

(15)   a.   bí-yénì   (bí- + yénì)
           seed-one
           'one seed'
    b.   bíè   à-yí
           seed.PL NHUM.PL-two
           'two seeds'

As one might suspect in light of the examinations of the Welsh, Turkana and Maltese number systems, a systematic examination of the lexicon developed in the field shows reliable asymmetries across different nominal domains. Figure 1 shows counts in the lexicon for the domains of animals, insects, trees, vegetation, and tools. The dark-shaded bars indicate nouns with the singular unmarked, while the light-shaded indicate those with the plural unmarked, e.g. for *mammal*, 43 entries are unmarked in the singular and 5 in the plural.[14]

Overall, 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, the graph includes data for nouns designating tools, which designate items that could be reasonably hypothesized as entities which naturally occur as self-standing individuals. These nouns showed a strong tendency towards being unmarked in the singular.

An asymmetry in the domain of body parts is displayed in figure 2. Nouns designating body parts which inherently come in pairs or groups, e.g. *eye* or *rib*, are

---

[14] In these counts, I exclude derived forms, since they follow their own patterns, which tends to obscure any generalization.
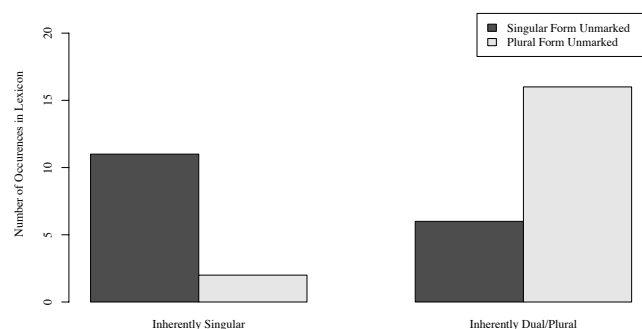
Fig. 2 Number Marking and Inherent Plurality in the Domain of Body Parts

Table 18 Sub-types of Unmarked Plural Nouns in Dagaare

paired body parts: *dumi/duno* 'knee'/'knees'; *gbɛrɪ/gbee* 'leg'/'legs'; *sɪŋgkpɪgrɪ/sɪŋgkpɪgɛ* 'hip'/'hips'; *nukumi/nukuno* 'fist'/'fists'; *yagrɪ/yaga* 'cheek'/'cheeks'
insects/reptiles/bugs: *dunduli/dundulo* 'maggot'/'maggots'; *soori/sobo* 'locust'/'locusts'; *yʊɔnɪ/yʊɔmɔ* (a type of small) 'frog'/'frogs'
vegetation/cereals/fruits: *kamaanɪ/kamaana* 'corn'; *lungguri/lungue* 'lemon'/'lemons'; *guori/goe* 'cola nut'/'cola nuts'

more likely to be unmarked in the plural; yet nouns designating body parts which are inherently singular, e.g. *head*, are more likely to be unmarked in the singular.

Thus, Dagaare makes a distinction similar to what has been shown for the other collective/singulative systems: Nouns which are likely to co-occur or form a natural pair or group, shown in table 18, 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.

Altogether, the distribution of the different morphological markers implicates five nominal types that receive grammatical recognition in Dagaare: substance, 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.

2.5 Summary: Comparing Grammatical Class, Coding, and Entity Type

The data presented so far in this section would appear to add arguments, alongside those discussed in section 1, against the hypothesis that semantic classification underlies the count/mass distinction. If English and French cognates (viz. *hair* and *cheveux*) do not cleanly fall into the same countability categories, expanding the cross-linguistic database only exacerbates the problem.

These mismatches however do not prove that semantic classification is completely absent—they simply indicate there is no one-to-one correspondence between entities and their countability classification valid across all languages. Yet, from what we know about other grammatical categories, such as grammatical relations or definiteness, it would be extremely surprising if there was such agreement among languages—total alignment of the content of grammatical categories across languages is rare.

Despite these mismatches, much overlap is present in these different grammatical number systems. The goal here is to illuminate where languages do converge and provide a framework which represents this information in a predictive manner. This clearly cannot be achieved at the level of particular entities in the world, as cognates with opposite countability classification prove, and I will argue in the following that this does not even hold at the level of types of entities. This in turn motivates couching the explanation of countability in more abstract properties related to individuation, which occupies section 3.

The different languages examined show substantial variation in the inventory of entity types appearing in their collective/singulative classes. A first question to ask is if any systematicity can be found amongst this variation. Table 19 plots the relation between entity types and their manner of coding in languages with a singulative marker. The entity types given here have no formal status but are simply descriptive categories based on the discussion in this section. The rows of the table give the entity types, while the columns give the languages. Each cell specifies the coded value, e.g. singular by *sing*; additionally, shaded cells signify entity types falling in collective/unit classes in the different languages. (Dagaare codes the singular value for two categories of nouns: those with the singular marked by *-ri* and those disposing of the optional singulative *-ruu*, which are shaded in dark and light grey, respectively.)

The relation between entity types and their coding in table 19 reveals that some entity types are coded via a collective category more often than others across the different languages. While insects and vegetation are uniformly treated as members of the collective category, higher animates are less often treated as such. Granular aggregates are always distinguished from liquids, although foodstuffs are coded variably. This table indicates that 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.

Further, organizing the entity types on the basis of their coding preferences across these different languages, from zero-coding to singular-coding to plural coding, establishes an ordering, the scale in (16). This scale is not strictly ordered, since some entity types are ranked equally (as indicated by $\leq$).

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

The ordering in (16) appears to reflect the accessibility of the unit interpretation. The higher in the scale a entity type is, the more salient the unit interpretation. The entity types lowest on the scale have zero-coded nouns, where units are either not

Table 19 Entity Types and Their Number Coding in Welsh, Turkana, Maltese and Dagaare

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
vegetation/cereals/fruits:	sing	sing	sing	sing
granular aggregates:	sing	sing	sing	0/sing
foodstuffs:	0	0	sing	0/sing
liquids:	0	0	0	0

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. 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, namely granular aggregates and foodstuffs, 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 (16) might be, it falls short in several ways. First, 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 2.4 showed, these lexicalization patterns aligning entity type and grammatical coding are strong correlations, but not absolutes. Further, the scale in (16) gives the impression that the size of an entity is a relevant factor to its countability classification, yet, as will be discussed, the influence of size is epiphenomenal. Table 19 indicates 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, elements and (ii) to the accessibility of units. In sum, while a noun's countability classification cannot be directly deduced from the type of entity, grammatical number does indeed seem to pattern with what nouns designate.

The next section ties accessibility of the unit interpretation to the more abstract notion of INDIVIDUATION. As discussed in section 1, 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 argues for a relation between speakers' use of grammatical number and how they interact with and view the world. I now turn to discussing some evidence which both supports such a relation, yet indicates it is not always a simple relation.

3 The Scale of Individuation

The ordering on the scale in the preceding section appears to be linked to the propensity for the entity described by the noun to occur as an individual—a property known

as individuation which 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 entities to the grammatical realization of count and mass nouns. The simplest hypothesis posits 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 and Snedeker 2005). I will argue that, given the data from section 2, a dichotomous distinction is too coarse and that a scalar approach to individuation is necessary.

This section first discusses evidence from the psycholinguistic literature that (i) grammatical countability is related to the object/substance distinction, and thereby an individual/non-individual distinction, and (ii) that individuation is related to factors beyond only the object/substance distinction. Considering these factors together implicates a scalar approach to individuation.

3.1 Individuation and the Object/Substance Distinction

Researchers in language acquisition have brought together a body of evidence that children are sensitive to a object/substance distinction prior to learning language. Soja et al (1991) conduct a "word extension task" which demonstrate that children distinguish between solid objects and shapeless substances. If children learn a name for a novel solid object, they extend that name to another object of the same shape, but not to one of the same material. If they learn a name for non-solid material, they extend the name to an item formed from the same material, not to one with a similar shape. For further results indicating the robustness with which children distinguish between solid objects and shapeless substances, see Landau et al (1988), Imai and Gentner (1997), and Samuelson and Smith (1999) and references therein.

Prasada et al (2002) provide evidence that multiple factors influence the countability classification of nouns beyond the object/substance distinction. Adult speakers confronted with novel entities were more likely to use countable nouns to describe them if possessing a regular rather than irregular shape. They further observed that if speakers saw a set of entities with the same irregular shape, then they were more likely to use a countable noun to describe the entity than if they saw just a single irregular-shaped entity. Additionally, if entities were shown to be designed for some function, then they were also more likely to be described by a countable noun. As all the stimuli were discrete objects, it is apparent that countability categorization involves more than a simple discrete vs. non-discrete distinction. Instead, Prasada et al (2002) provide evidence that speakers categorize entities as (non-)countable according to whether they *construe* an entity, viz. some portion of the world, as an individual worth counting on the basis of multiple factors.

Further studies demonstrate that multiple factors bear on countability classification. Middleton et al (2004) examine two hypotheses from the work of Wierzbicka (1988). The first hypothesis, DISTINGUISHABILITY, proposes that nouns designating entities for which the constituents are more easily distinguishable are more likely to be used as countable nouns, while those nouns designating entities for which the constituents are not easily distinguishable will be used as non-countable nouns. Middleton et al (2004) present subjects with two displays of novel aggregates, each containing 40 elements, which vary in the spatial proximity among the elements (*Close* versus *Apart*) and the size of elements (*Large* versus *Small*). For example, a subject would see two sets of an element: For one set, each element was large and spatially separated from the others, for the other set, each element was small and spatially contiguous with other elements. Participants then match a nonce countable phrase ('These are worgels') or non-countable phrase ('This is worgel') with one of the displays. Subjects' choices of countable or non-countable phrase were significantly influenced by spatial proximity, but not by the size, of the elements.

Middleton et al (2004) also investigate whether the canonical MANNER OF INTERACTION with an entity influences countability. Wierzbicka argues that entities with which one interacts one-by-one will be described by countable nouns, while entities with which one interacts in quantities, rather than individually, will be described by non-countable nouns. Middleton et al (2004) present subjects with a novel aggregate, yellow decorative coarse-grained sugar, in a cardboard box. As hypothesized, when subjects had to interact with individual grains, under the guise of engaging in a game, they were significantly more likely to describe the novel aggregate with a countable phrase rather than a non-countable phrase.

Real-world cases provide supporting evidence for the role of interaction. 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. For instance, the word *chad*, which designates the tiny pieces of paper left over from punch-cards ballots, underwent a shift in countability status during the tight 2000 U.S. presidential election. It became critical after the election to count each vote by determining whether the ballot had been completely 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, and punch-cards in general, typically used the word as a non-countable noun. 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, representing a vote. 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.

The various experimental results imply a complex relation between entities in the world and their countability classification mediated by individuation factors. Countability classification cannot reduce to only objective ontological properties of entities in the world: Many of the effects observed concern not properties intrinsic to entities, but relations those entities are perceived to stand in, such as spatial proximity

amongst similar entities or relations to functions or to common interactions. Accordingly, individuation is not only related to intrinsic features of objects, such as shape, but also to whether elements are recognized as sufficiently independent from one another, either spatially or functionally, to be construed as individuals. Yet, countability classification cannot be completely independent of the external world, for how things are construed depend on features of the external world and our interaction with them. The various pieces of experimental evidence indicate that when features of an entity in the world are changed, this may affect its likelihood of being classified as countable or not.

As we move on to consider how individuation factors relate to nominal meaning, it is important to recognize a difference between the experimental contexts in this section and how individuation is encoded in lexicalized meaning. Within the experiments in, e.g., Prasada et al (2002), participants must reason from their observations about a novel object to a countabilty classification. I do not assume such a process underlies daily language use. Rather a noun, i.e. an entity description, encodes a construal of the entity it describes. This construal is not calculated on-line, but is rather a generalization over speakers' experiences with entities in the world, which in turn have provided the very properties which lead to the construal of those entities as individuated or not. As the example of *chad* indicates, this relation between the entity and how it has been encoded is not static, but is open for renegotiation when speakers' experiences change sufficiently.[15]

3.2 A scalar view of individuation

The experimental evidence reviewed implicates that countability classification involves several distinct ingredients. I now argue that these individuation factors relate to countability classification in a structured way. 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, while still restricting the investigation to natural concrete entities and their canonical interpretations. First, a core distinction is the presence of perceptible[16] minimal units. Second, whether these minimal units are sufficiently individuated to merit being described by a countable noun depends (at least) on various factors experimentally validated, regularity and repetition of shape (Prasada et al, 2002) as well as contiguity between elements (Middleton et al, 2004). Unlike the presence or absence of perceptible minimal units, which makes a categorical distinction between two classes of entities, the contiguity between elements is a matter of degree.

Given the various distinct factors, a view on individuation which divides entities into individuated and non-individuated would appear to be an over-simplification. Also from the perspective of countability classes, the grammatical classification that

---

[15] Thus, the position here, or invoking individuation generally, should not be equated with an "internalist" cognitive account with no connection to the world.

[16] 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).

individuation is invoked to explain, a simple contrast between individuated and non-individuated will be too coarse-grained. Even for English, often considered to have a binary countable/non-countable distinction, the differential behavior of nouns across a variety of diagnostics demonstrates that there are "degrees of countability" that are grammatically detectable (Allan, 1980).[17] Thus, from different perspectives, it is plausible that individuation is better viewed as a scalar phenomenon rather than one that can be reduced to a simple binary feature $+/-$ *individuated*.

A scalar view on individuation in terms of the factors discussed so far would provide a straightforward interpretation for the different entity types discussed in section 2.5 in terms of different levels of individuation. Least individuated are entities without any perceptible elements—the standard description for substances. Entities with perceptible units, but not typically 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. Entities with perceptible units that are separated from one another but still connected in some fashion, whether spatially near or functionally united, as for collective aggregates such as *ants* or *cherries*, are then again more individuated. 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. More precisely, individuation types are equivalence classes of nominal descriptions based in individuation properties, i.e. *granular aggregates* contains all the nouns (or noun senses) which encode a construal of entities that includes having perceptible units while being connected to one another or clumped together.[18] I will refer to the scale constructed out of the individuation types, given in (17), as the SCALE OF INDIVIDUATION.[19]

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

This view of grammatical countability, where a set of properties bear on lexical classification, is reminiscent of the view of argument realization and thematic roles put forth by Dowty (1991). If, in English, a noun encodes a lexical meaning which rates highly on the various individuation properties, it will be lexicalized as a countable noun, similarly to how if a verbal argument has a large number of Proto-Agent entailments and few Proto-Patient entailments, it will be lexicalized as the verb's subject. There are critical differences, however. Dowty's (1991) Proto-Properties are entailments, and as such are active in the truth conditional semantics of the verbs. In contrast, the individuation properties, such as likelihood of contiguous elements, are active in the concept expressed by the noun, but are not active in the online calculation of truth conditions. For instance, the knowledge that *rice* refers to a type of

---

[17] Although Allan (1980) was the first to recognize different degrees of countability, this was strictly limited to a characterization of the syntactic distribution of nouns in English. His approach purposefully refrained from making any connection to the semantics of countability (p. 547, fn 3), and thus differs fundamentally from the approach developed here.

[18] For a formal semantic analysis of these different individuation types, along with number morphology in Welsh and Dagaare, see Author #2.

[19] The scale presented (17) is built from four individuation types, although we will see that the scale of individuation is potentially more fine-grained. See Author #1 for discussion of prior formulations of animacy (Lucy, 1992) and individuation continuums (Gentner and Boroditsky, 2001).

entity for which the individual elements typically come in contiguous bunches, is an instance of what I term an individuation property, but it is not entailed or presupposed to hold of the referents designated by the use of *rice*—the statement *rice is on the table* may be true even if the elements of rice are scattered and not contiguous. Instead, the individuation properties inhere in the conceptualization of a noun, which then has grammatical effects, such as determining lexicalization patterns. As such, individuation properties, inasmuch as they delineate information about the conceptualization underpinning a noun, apply at the level of nominal kinds, in contrast to the more familiar mereology-based properties, such as divisiveness or cumulativity, which apply at the level of referents—instantiations of a kind.

I now turn to applying the scale of individuation in (17) to the grammatical number systems discussed so far. I argue that the scale of individuation just developed underlies the structure of grammatical number systems. Laying out the interaction between individuation 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 individuation types.

3.3 Relating Entity Types and the Scale of Individuation

I first relate the scale of entity types developed in section 2.5, repeated in (18), to the scale of individuation, repeated in (19), to clarify the relation between entity type and individuation type.

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

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

The scale of entity types, while based on the morphological coding across the different languages examined, fails to be predictive of the structure of grammatical number systems. This was seen for Welsh, where, for instance, not all nouns designating small-sized animals fall under the collective class, but only those whose "characteristic way of living together [is] in swarms, flocks, herds and shoals" (Stolz, 2001, p. 65). In contrast, the individuation properties, e.g. habitually appearing in groups, cross-cut entities types. Thus, it is the behavior of entities with respect to individuation properties which provide generalizations about the relationship between meaning and grammatical number categories.

While there is no direct relation between the scale of entity types and the scale of individuation, the scale of entity types does contain information about 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 available comparative data, it appears that languages are comparable in how entity types are classified under individuation types, such as large animals

Table 20 The Scale of Individuation: Dagaare, Welsh and English

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*)	

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 an entity should be described in one fashion or another. Future typological work may be able to determine whether 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.

3.4 Relating Individuation Types, Grammatical Categories of Number, and Morphological Coding

Having proposed that individuation types are the proper level from which to examine the underpinnings of grammatical number systems, I examine the relation among individuation type, grammatical categories of number, and the manner of morphological coding across the languages under discussion. The view developed here implies a relation among these different elements: A grammatical number category is established based on patterns of morphosyntactic coding, yet, the morphosyntactic coding pattern of a noun, e.g. how it codes singular or plural, relies on its degree of individuation, i.e. its individuation type.

Table 20 compares the systems of Dagaare, Welsh and English in terms of the scale of individuation, where Welsh represents the tripartite systems discussed in section 2. Table 20 displays the individuation types, ordered according to the scale of individuation, along the top. The language is named in the leftmost column. The table displays each of the grammatical categories in the different languages as a shaded region spanning the relevant individuation types. Each of the shaded regions also indicates the grammatical category's manner of coding, e.g. for Welsh, (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*). For Dagaare, the four individuation types are each manifested by a distinct grammatical category of number.

The general picture implied by the table is that morphosyntactic classes are based in the scale of individuation. Each of these systems differs substantially from the others, yet they are still all comparable: They all partition the scale of individuation into a discrete set of grammatical categories. On this view, morphosyntactic categories of number are grammatical phenomena—different languages have established different categories. Yet, 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.

Recalling that the individuation types are a partitioning of nominal descriptions into their equivalence classes defined by individuation properties, and since morphosyntactic categories of number are also a partitioning of nominal descriptions, one can consider the relation between the two partitions. Although the relation between the cells of the two partitionings could be arbitrary, I argue that morphosyntactic classes adhere to the ordering of the scale of individuation, which I implement by requiring an order-preserving mapping between the cells of two different partitionings.

To put this problem in more precise, and more general, terms, let the set of nominal descriptions in a language be partitioned into a set of individuation types, $\Pi_I = I_1, I_2 \ldots I_n$ and also into a set of (language-particular) grammatical number categories, $\Pi_G = G_1, G_2 \ldots G_n$. An order-preserving mapping between two arbitrary ordered sets, e.g. $(P, \leq)$ and $(Q, \leq)$, is a function f such that for all $p_1, p_2 \in P$, if $p_1 \leq p_2$, then $f(p_1) \leq f(p_2)$. In our case, let $Ind = (I, \leq_I)$, the set of individuation types with the ordering given by the scale, and let $Gram = (G, \leq_G)$, the set of grammatical countability classes, also with an ordering on them which will be subject to the order-preserving mapping. Then, any mapping $f : Ind \rightarrow Gram$ is such that if $I_i \leq_I I_j$, then $f(I_i) \leq_G f(I_j)$. In prose, if two individuation types, say *granular aggregates* and *collective aggregates* stand in a less-than-or-equal-to relationship with respect to the scale of individuation, they will map to categories in the the language's grammatical class inventory which preserve the ordering.[20]

The requirement that grammatical categories of countability respect the ordering of the scale of individuation enforces that grammatical number systems should manifest certain properties, which also has typological implications. First, since grammatical number categories respect the the scale of individuation, up to equivalence, grammatical number categories will correspond to either individuation types, or coherent combinations of individuation types. Second, since the order-preserving map is a function, each cell in Π_I will map to exactly one and only cell in Π_G. Accordingly, no individuation type will map to two grammatical categories, from which it follows grammatical number partitioning will be either equivalent to or more coarse-grained than the individuation type partitioning. Finally, since the ordering is preserved, there should not be systems where a category of grammatical number (such as DUAL or

---

[20] To see how this works, consider Welsh. Here f(LIQUIDS/SUBSTANCES) = NON-COUNTABLE, f(INDIVIDUAL ENTITIES) = SINGULAR/PLURAL and f(GRANULAR AGGREGATES) = f(COLLECTIVE AGGREGATES) = COLLECTIVE/SINGULATIVE. The following implications hold: LIQUIDS/SUBSTANCES $\leq$ GRANULAR AGGREGATES $\Rightarrow$ f(LIQUIDS/SUBSTANCES) $\leq$ f(GRANULAR AGGREGATES) (since NON-COUNTABLE $<$ COLLECTIVE/SINGULATIVE), while GRANULAR AGGREGATES $\leq$ COLLECTIVE AGGREGATES $\Rightarrow$ f(GRANULAR AGGREGATES) $\leq$ f(COLLECTIVE AGGREGATES) (since COLLECTIVE/SINGULATIVE = COLLECTIVE/SINGULATIVE). It can be checked that the other ordering relations are also preserved.
Note that different cells of the Π_I partition could map to a single cells in the Π_G partition, as is the case with the mapping in Welsh from GRANULAR AGGREGATES and COLLECTIVE AGGREGATES to COLLECTIVE/SINGULATIVE, therefore the mapping is not injective. Yet, from the construction of the problem, it follows that it will be surjective, i.e. every cell in the Π_G partition will be mapped to from the Π_I partition.

Table 21 Hypothetical Grammatical Number System Violating the Prediction of the Scale of Individuation

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

COLLECTIVE or COUNT) spans two disconnected segments of the scale. An example of a system which would violate this condition is shown in table 21 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.[21]

The theory here does not provide direct predictions concerning, say, which categories of grammatical number a language will develop, but rather provides broad constraints governing how grammatical number systems behave. These predictions will be refined in section 4.2 when the influence of animacy is also considered.

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. I suggest that understanding grammatical number categories in light of the scale of individuation provides a basis for the coding choices made in different languages.

Following Jakobson (1957/1971) and Horn (1989), I take morphological markedness to be rooted in semantic markedness (see Koontz-Garboden 2007 and Kiparsky and Tonhauser 2012 for further discussion). In particular, I assume that a grammatical number category's coding preferences are linked to the semantic markedness patterns of its associated individuation type(s), and that the default coding value of the grammatical category aligns with the default semantic value of the individuation type. For instance, the individuation type *individual entities* is comprised of things in the world which have a high propensity to appear as free-standing individuals, thus a grammatical class which covers *individual entities* would likely designate singular entities by default as opposed to multiple entities, i.e. the plural would be the marked value, as has been seen in the grammatical number systems so far. Conversely, the individuation type *collective aggregates* is comprised of entities where the default is co-occurrence. For a grammatical class which covers the individuation type *collective aggregates* reference to multiple entities, the plural value, would be default, while the singular value would be marked. The limiting case is the portion of the scale for which there is no countability contrast. In general terms, the empirical prediction is that the higher the level of individuation of a grammatical class, the more likely the designation of single entities will be the default, while the lower the level of individ-

---

[21] Here, f(GRANULAR AGGREGATES) = f(INDIVIDUAL ENTITIES) = SINGULAR/PLURAL and f(COLLECTIVE AGGREGATES) = COLLECTIVE/SINGULATIVE. Although GRANULAR AGGREGATES < COLLECTIVE AGGREGATES < INDIVIDUAL ENTITIES, the grammatical categories COLLECTIVE/SINGULATIVE and SINGULAR/PLURAL will not preserve the ordering, since either f(GRANULAR AGGREGATES) $\not\leq$ f(COLLECTIVE AGGREGATES) or f(COLLECTIVE AGGREGATES) $\not\leq$ f(INDIVIDUAL ENTITIES).

Table 22 Countable and Non-Countable Nouns in Miraña

Countability Category	Base Form	Plural Inflected Form	Gloss
Countable	ka:ni	ká:ni-mɯ	'father'
	i:βa	i:βa-mɯ	'macaw'
Non-Countable	ka:	*ka-mɯ	'ants'
	í:nɯ	*í:nɯ-:nɛ	'earth'

uation of a grammatical class, the more likely the designation of multiple entities will be the default.[22] I return to these issues in 4.4.

4 Applications of the Scale of Individuation

The scale of individuation as presented accounts for the number systems in section 2 which motivated it. This section demonstrates its wide empirical reach, discussing its application to other grammatical number systems, its interaction with animacy, its application beyond natural concrete entities to artifactual and abstract nouns, and its relation to frequency.

4.1 Extension to Other Types of Number Systems

I now show how the scale of individuation aligns with several other non-binary grammatical number systems beyond those discussed in section 2. While these systems have distinct means of manifesting grammatical number, e.g. classifiers, the categories of nouns that can be distinguished accord with the scale of individuation.[23]

Miraña Miraña, a highly endangered Amazonian language spoken in southern Colombia, possesses both nominal classifiers and inflectional number. Seifart (2009) discusses distributional properties which distinguish the different nouns types. First, whether a noun combines with inflectional number markers distinguishes countable from non-countable nouns: Countable nouns, when designating multiple entities, obligatorily inflect, while non-countable nouns do not, shown in table 22.

Non-countable nouns instead combine with class markers to refer to single objects and these forms then permit inflectional number (as well as cardinal modification), shown in (20)–(21).

(20)   a.   í:nɯ
              earth
              'earth'

---

[22] This formulation generalizes the notion of "local markedness" due to Tiersma (1982). See Stebbins (2007) for relevant discussion.

[23] See further, for instance, Mathieu (2012) for an application of the scale of individuation to Ojibwe. Other languages for which the predictions of the scale appear to be borne out include Arabic (Acquaviva, 2008; Mathieu, 2012), Breton (Press, 2009), Czech, Hebrew (Doron and Muller, 2014), Lingala (Mufwene, 1980), Swahili (Contini-Morava, 2000), Uncunwee (Jade Comfort, p.c.), and Walman (Matthew Dryer, p.c.).

Table 23 The Grammatical Number Categories of Miraña along the Scale of Individuation

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

    b. íːnɯ-ba
       earth-SCM.3DIM
       'piece of earth'   SCM = Specific Class Marker
    c. íːnɯ-báː-kɯ
       earth-SCM.3DIM-PL
       'pieces of earth'

(21)   a. kaː
        ant
        'ants'
     b. ka-ʔba
        ant.SCM.3DIM
        'an ant'
     c. ka-ʔbá-mɯ
        ant.SCM.3DIM-PL
        'some ants'

There is a distinct class of of grammatically non-countable 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 (21), 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, refers to smaller numbers of entities, glossed by "some ants". Thus, this class of nouns in Miraña with default reference to collections aligns with the collective/singulative classes previously explored. Further, the entity types in Miraña's number categories overlap with what would be expected: The aggregate class include insects, along with lower animates, such as rats, whereas the singular/plural class includes humans and big or culturally salient animals, e.g. the macaw.

In sum, despite the different grammatical means employed, Miraña's classes of non-countable, aggregate and countable nouns align with the predictions of the scale of individuation, shown in table 23.

Yudja Another Amazonian language, Yudja (Lima, 2014), (Juruna family, Tupi stock, spoken in Brazil), serves as a limiting case: The countable/non-countable distinction is only very weakly present. Nouns in Yudja are allowed to appear bare, unspecified for number or definiteness, as shown in (22), from Lima (2014, p. 33).

(22)   ali   ba'ï ixu
        child paca eat

Table 24 The Grammatical Number Categories of Yudja along the Scale of Individuation

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

> '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 (2014) reports that the distribution of numeral modification does not distinguish between countable and non-countable nouns: Numerals combine freely with nouns designating substances or individuals. Examples in (23) and (24) (Lima's (31c) and (28b)) 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 (2014, (p. x)) proposes that in Yudja "all nouns can be used as count nouns".

(23)   Yauda ali   eta   apapa
          two   child sand drop.REDUPL
          'Children dropped two (portions of) sand(s) (in different events)'

(24)   Txabïu apeta ipide   pepepe
          three   blood on the floor to drip.REDUPL
          'Three (drops of) blood dripped on the floor (in different events)'

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 where 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.

Kiowa The application of the scale of individuation extends to much more complicated systems, such as 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 show how the different classes of grammatical number correspond to natural lexical semantic classes in a fashion consistent with 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 the major classes align 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 4.2, 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 avoids these difficulties and, instead, speaks to the richness of different number systems.[24] Further, since the notion of individual is a graded concept, and it is expected that some languages, like Yudja, might restrict their expression of grammatical number to only the types of individuated entities highest in the scale. I now turn to exploring the relation between the scale of individuation and animacy, followed by its relation with frequency.

4.2 Countability and Animacy

I have been arguing that a scalar structure based on individuation underlies countability phenomena; however, other factors, especially animacy, have been shown to influence number marking. While animacy has been reliably tied to the manifestation of plural marking in a large number of languages, 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 that viewing grammatical number as grounded in individuation leads to understanding how animacy influences

---

[24] An anonymous reviewer remarks that distinguishing singular/plural contrast from the mass/count contrast is sufficient to explain phenomena like plural mass nouns (*clothes*, *dregs*) and collective nouns, citing Gillon (1992) as an example for which a feature system is constructed from two equipollent syntactic features, +/-Count and +/-PL, resulting in three categories: count nouns (+Count) (which may be singular or plural), singular mass nouns (-Count,-PL), and plural mass nouns which are -Count but lexically specified as +PL. Despite the fact such systems are richer than a strictly binary +/-Count, they do not provide a rich enough framework to treat the range of phenomena discussed here. Even the variation in collective/singulative systems discussed in section 2 poses difficulties. Collectives would be relegated to the third category, -Count,+PL, which could, at a first pass, function for Welsh and Turkana, whose collective forms impose plural agreement, but not for Maltese where the collective agrees in the singular. Such a treatment would further remain silent on a range of issues, e.g. on interpretational differences between the collective classes discussed here and plural mass terms ('brains') or on grammatical number systems whose complexity exceeds three categories, as for Dagaare or Kiowa.

Table 25 The Smith-Stark Hierarchy

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

Table 26 Tamil's Grammatical Number System along the Smith-Stark Hierarchy

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

Table 27 Smith-Stark Hierarchy Violated by Maltese Collective

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

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 25, which he considers 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 express plurality for any noun type higher on the hierarchy.

An instantiation of this prediction is provided by Tamil, as shown in table 26. The shaded area of the animacy hierarchy in table 26 indicates the portion which manifests a singular/plural contrast. According to the discussion in Smith-Stark (1974, p. 662), nouns in Tamil designating rational beings typically manifest plurality while those designating living beings not considered rational, including *child* and *infant*, are rarely pluralized. Smith-Stark's claim is that any noun type higher on the hierarchy would also have obligatory number marking, which is true—e.g., all pronouns manifest a singular/plural distinction.

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—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. In each case, the number categories, such as the collective, are apparently not sensitive to the animacy level of the noun.

For instance, Corbett points out that the collective in Maltese contains a mixture of animate ('flies') and inanimate ('corn','shoes') nouns, but cannot be expressed on, for instance, 1st and 2nd person pronouns, violating the predictions of the Smith-Stark hierarchy. This is shown in table 27, where the shaded region indicates for which part of the hierarchy the collective/singulative is expressed.

The question arises, then, what precisely the relation is between grammatical 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 grammatical 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 Cor-

bett 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 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 section 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 possesses 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. Here 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. This distinction was already seen in Miraña, which categorizes 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 Aissen lattice in figure 3 exhausts the combinatoric space of the individuation scale and the animacy hierarchy. For the segment of the individuation scale from *collective aggregate* and higher, the entirety of the (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 4, 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 lat-
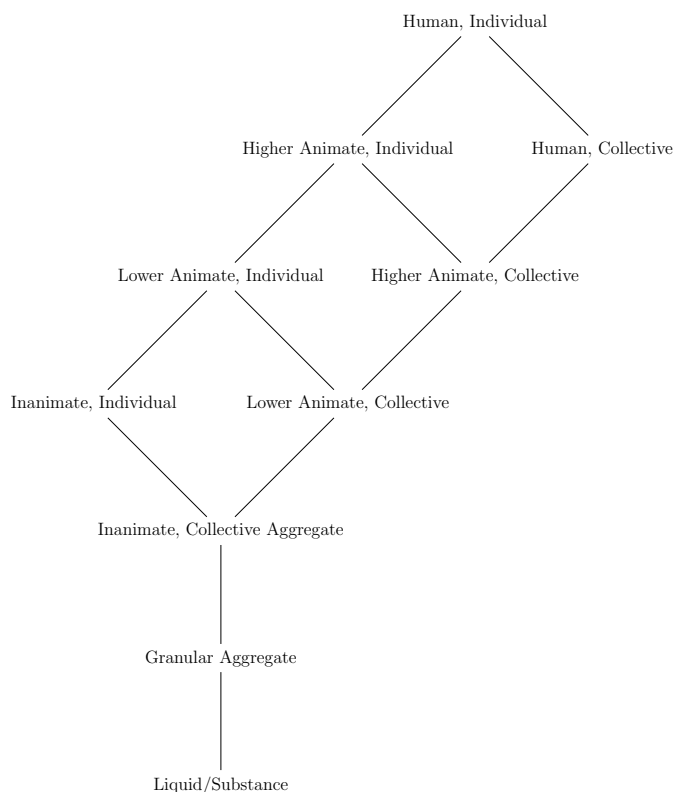
Human, Individual

Higher Animate, Individual   Human, Collective

Lower Animate, Individual   Higher Animate, Collective

Inanimate, Individual   Lower Animate, Collective

Inanimate, Collective Aggregate

Granular Aggregate

Liquid/Substance

Fig. 3 The Lattice of Animacy and Individuation

tice.[25] 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

---

[25] I simplify by leaving aside the packaging uses of Maltese as this is not relevant for the moment.
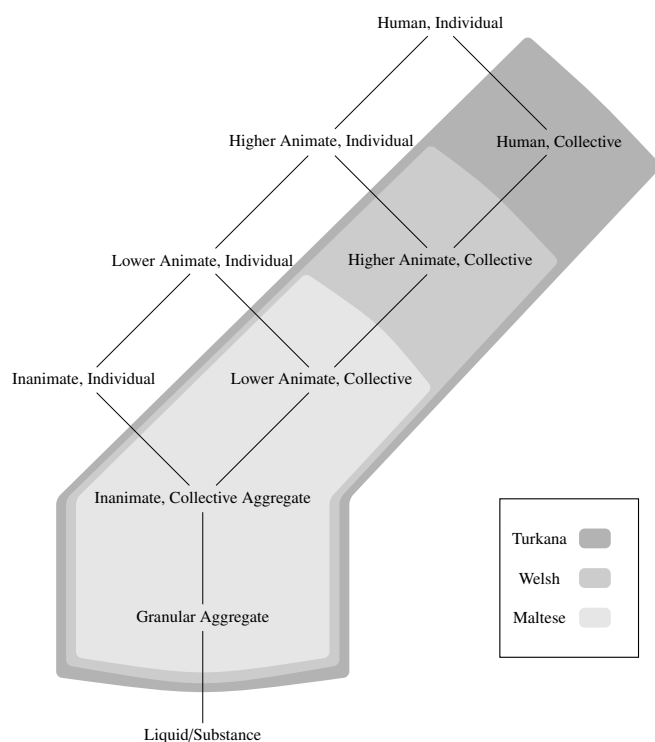
Fig. 4 The Collective/Singulative Classes in Turkana, Welsh and Maltese on the Lattice of Animacy and Individuation

as cows, and collections of humans, e.g. 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 other 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.

4.3 Beyond Natural Concrete Entities

The proposed view extends to other types of nouns, such as artifactual and abstract nouns. On the view here, such nouns are also categorized into grammatical countability classes based on individuation properties, yet the nature and kind of individuation properties are different. While perceptual properties (shape, spatial proximity) are relevant for individuating natural concrete entities, they will not be relevant for individuating, say, eventive nouns, where properties such as temporal discreteness or continuity of the event described are likely to be relevant (Bloom, 1990; Mourelatos, 1978). In this section, I sketch how artifactual nouns, such as *chair* or *furniture*, can be treated in the view developed here.[26]

Furniture has long been regarded as paradoxical: It is non-countable, yet, unlike non-countable substance nouns, *furniture* designates individual objects, such as chairs or tables, which are otherwise countable. Certain linguistic contexts demonstrate that the individual objects are grammatically accessible, as with comparatives (Barner and Snedeker, 2005) or adjectives like *big*, as in *big furniture* (Schwarzschild, 2011). If countability only relates to a binary $+/-individual$ distinction, then the behavior of *furniture* is unexpected as it is grammatically non-countable despite being capable of referring to individuals.

The behavior of *furniture* would be similarly unexpected if individuation could only be achieved through perceptual properties. Yet, artifactual nouns differ from natural concrete nouns in that they do not describe merely objects, but objects with respect to a function.[27] I argue that individuation for artifact nouns is bound up with the function described by the noun.

The functions associated with artifactual nouns further provides a contrast between individuals and aggregates of individuals, from which an explanation of *furniture*'s countability classification follows. The functions associated with *chair* or *hammer*, i.e., to sit on or to hammer with, are canonically achieved through using just one entity. In contrast, the function associated with *furniture* is to furnish a space, as indicated by its derivational source, *to furnish* ($>$ *fournir* Fr.). Furnishing a space is canonically achieved through using multiple entities. Thus, an individual/aggregate distinction emerges for artifactual nouns as well: Nouns such as *chair* describe entities which individually are capable of performing the associated function while nouns such as *furniture* describe entities which typically perform the associated function as an aggregate. Accordingly, *furniture* is classified in English in the same individuation

---

[26] See Author #3 for a full lexical and formal semantic analysis of artifactual nouns.

[27] This has been widely noted since Aristotle. See Brown (1999) and Nichols (2008), among others, for linguistic effects of this distinction.

type as other aggregates where access to individuals is not relevant to their description, such as *foliage* or other granular aggregates. Since aggregate nouns such as *foliage* are treated as non-countable in English, *furniture* is also classified as non-countable. An immediate prediction is that cognates of *furniture* may appear in a collective/singulative class should the language have one. This prediction is borne out by Welsh, where the cognate for *furniture* appears in the collective/unit class, *dodrefn* 'furniture'/ *dodrefn-yn* 'a piece of furniture', as do other functional aggregates such as *offer* 'tools' /*offer-yn* 'a tool' or *dillad* 'clothes'/*dilled-yn* 'a piece of clothing'. Thus, once distinct principles of individuation are recognized for natural concrete and artifactual nouns, then nouns such as *furniture* need not be seen as an exception to the principles of countability classification, but as aligned with them.

Placing the explanatory burden on domain-specific modes of individuation yields a more nuanced understanding of countability preferences: Although *oil* and *furniture* may both be non-countable in English, they need not be so for the same reason. Rather, their respective grammatical classifications as non-countable nouns derive from distinct modes of individuation, related to their differing lexical semantics. This differs from accounts in which the non-countable status of *oil* and *furniture* in English are given a uniform explanation (Chierchia, 1998), or where *furniture* is treated as a lexically-specified exception (Bale and Barner, 2009). The countability behavior of artifactual nouns, such as those lexicalizing a description of furniture, will systematically differ across languages according to which type of function is described by the noun, its related individuation properties, and the grammatical system of the language at hand. The resultant countability status may coincide with that of substance nouns, e.g., *oil*, but often will be distinct. This view anticipates cross-linguistic differences exactly when the nominal description differs. Cognates of *furniture*, e.g., *meuble* in French, are countable, which has been taken as problematic. Yet, this is consistent with the approach advocated here since *furniture* and *meuble* differ at the level of entity description: *meuble* is derived from a root meaning *to move*, thus designates 'movable entity', which in turn evokes different individuation properties than entities related to the root *to furnish*.

Applying this discussion to the scale of individuation indicates that ultimately the units of the scale must be generalized, composite types. Clearly, there will be multiple, domain-dependent modes of individuating entities, including spatial separation, temporal separation, or individuation via function. The individuation type *individual* or *collective aggregate* will contain entities whose individuation properties differ in their domain-specific characteristics, such as being spatially-separated versus temporally-separated, but which share in abstract individuation properties, namely being separated. It is incumbent on this view to understand the principles of individuation that underlie nominal categorization for each type of entity and to articulate in detail the relation between the lexical semantics of a noun and the individuation properties relevant to the noun type. The expectation then is that, modulo the specific facets of the domain-dependent individuation properties, the categorization of natural concrete, artifactual, eventive and abstract nouns will maintain the higher-level individuation contrasts which ground grammatical classification of countability.

4.4 Individuation, Morphological Coding, and Frequency

I will now consider the relationship between the proposed individuation account and frequency. 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. Typological studies have sometimes attributed such coding asymmetries to frequency asymmetries (Haspelmath, 2008). Two notions of frequency are relevant here: First I discuss non-linguistic frequency, i.e. frequency in relation to elements in the external world, and then linguistic, or textual, frequency, e.g. frequency in relation to occurrences of words or morphemes. 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 reviewed here.

Reference and Frequency It is tempting to think that nominal countability categorization could be explained purely in terms of how frequently entities co-occur in the world, as suggested by Tiersma (1982) or Haspelmath (2008). For instance, it is reasonable to suppose 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 examined here, 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.

Stolz (2001) argues explicitly that for Welsh 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. Rather, it is whether the entities are *perceived* as habitually coming together which is determinate, according to Stolz.[28] Similarly for Turkana, Dimmendaal observes "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, i.e. a plural value. 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 precon-

---

[28] 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 aligns with the results of the first experiment of Middleton et al (2004), presented in section 3.1, where size is not a significant factor.

dition, rather than the ultimate explanation of why a noun may be classified in the collective/singulative class.

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. Although textual frequency has been causally related to the structuring of morphological coding (Greenberg, 1966), the different strains of evidence brought together here, e.g. the psycholingustic evidence in section 3.1, indicate that countability classification cannot be reduced to textual frequency. However, I demonstrate that there is a correlation between morphological coding asymmetries and textual frequency asymmetries: Nouns which are typical members of a morphological coded collective/singulative class correspond, for languages which only manifest a singular/plural contrast, to nouns which have greater textual frequency in the plural than in the singular.

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 English 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 5, 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.[29]

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.[30] In English, for many words such as the insects in figure 5, 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 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 En-

---

[29] 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.

[30] I would like to thank Paul Kiparsky for leading me to the connections in this section.
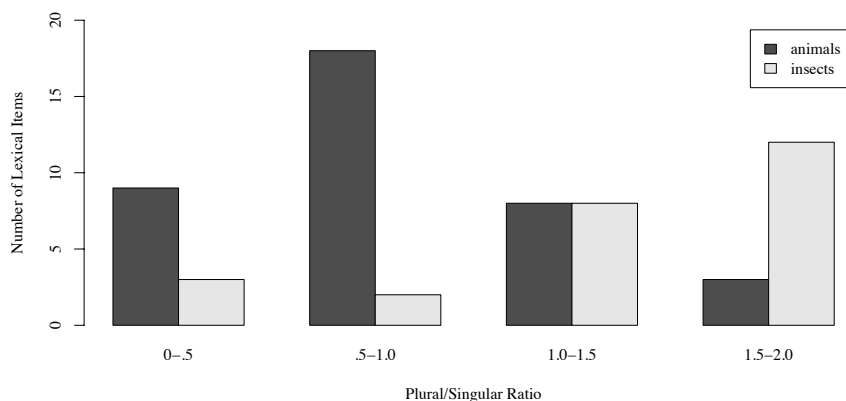
Fig. 5 Number Marking Frequency Patterns in English for nouns in the semantic domains *animals* and *insects*

glish, 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 and 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.

This concludes my discussion of the scale of individuation. I now turn to examining its implications for the theoretical understanding of countability.

5 A Meaning-Based Account of Countability

The previous section 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. The proposal is represented in figure 6, which unpacks the four different levels: things-in-the-world, lexical nouns and their properties, i.e. entity descriptions as discussed at the end of section 3.1, individuation types

and grammatical classes.[31] The figure displays the full generality of arbitrarily many individuation types and grammatical classes, and, to exemplify, mappings are given of two entities which belong to the highest and lowest individuation classes.

The level *entity*—the thing in the world picked out by the noun—contrasts with *entity description*, the noun itself as a *description* of a type of entity. Importantly, the entity description encodes a particular construal of the entity, not just the extensional content. Although this distinction is not typically made in the countability literature, a distinction between *events* and *event descriptions* has long been recognized as crucial in the aspect/telicity literature (Krifka, 1992). The same event can be described by two different expressions which differ in telicity, viz. *running* (atelic) and *running a mile* (telic). If one assumes that these expressions refer directly to events, then this situation is paradoxical, for the same event is at once atelic and telic. Recognizing that these expression are simply two different descriptions of the same event immediately resolves this conflict. As I will discuss in section 5.2, recognizing the distinction between *entities* and *entity descriptions* similarly resolves paradoxical examples such as *leaves* and *foliage*—both words may refer to the same stuff in the world on a given speech occasion, but they are different descriptions of that stuff.

The proposal in figure 6 is a shift in perspective from previous theories of countability. First, there is a higher degree of complexity both in the relation between the things in the world and the grammatical class, represented by the vertical dimension, and also among the individuation classes and the grammatical classes, represented by the horizontal dimension. Second, on this view countability involves a *principled relation* between things in the external world and grammatical classification which is mediated by individuation.
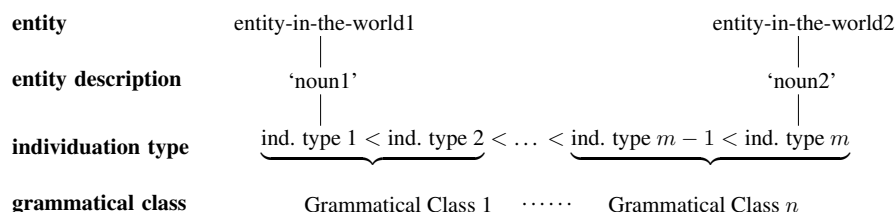
entity	entity-in-the-world1	entity-in-the-world2
entity description	'noun1'	'noun2'
individuation type	$\text{ind. type } 1 < \text{ind. type } 2 < \ldots < \text{ind. type } m-1 < \text{ind. type } m$	
grammatical class	Grammatical Class 1 $\cdots\cdots$ Grammatical Class n	

Fig. 6 Mapping between Entities-in-the-World, Lexical Nouns, Individuation Types and Grammatical Classes

Both directions of this relation are relevant for different facets of countability. One direction represents the world-to-word fit, guiding lexicalization patterns. Things in the external world and how we interact with them set the possible ways the entity may be construed, or individuated. A noun, or in more complicated cases one sense of a noun, lexicalizes one of these construals. The entity description encoded by the noun falls under an individuation type, which determines the noun's countability status with respect to the grammatical number system of the language. The other direction

---

[31] 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), although only a binary countability distinction is put forth there.

represents the word-to-world fit, guiding how interpretation is achieved. There is a relation between grammatical classification and types of entities in the world and, accordingly, interpreting a noun in tandem with a grammatical classification that specifies particular individuation properties will aid in picking out the intended referent in the world.

Countability does not arise from any single level, but is a product of the interrelation among these different elements. The additional complexity invoked here has a pay-off. Other theories of countability are primarily motivated either by the regularity of the grammatical classification of things in the world into countabilty classes or by the variability of this classification, or yet again, by the variability possible due to different contexts. Accordingly, while these theories have the advantage of simplicity, they explain one facet of countability classification at the expense of the others. The complex of relations in figure 6 addresses these different aspects of countability simultaneously. I now relate a range of other theories in reference to the different levels given in figure 6.

The ontological and grammatical views, introduced in section 1, only address part of the picture given in figure 6. The ontological view on countability places the explanatory burden of countability on the properties of entities in the external world, the *entity* level, and presents countability as a function from entities to grammatical classification. Accordingly, for the ontological view, countability only involves two of the levels in figure 6 and one direction, as it is a function from the level of *entity* to *grammatical class*. The grammatical view, on the other hand, places the explanatory burden on the relation between nouns and grammatical class, an ultimately arbitrary relation independent of "how things are inherently structured". Accordingly, the grammatical view also involves a relation only between two of the levels in figure 6, namely *entity description* and *grammatical class*.

While the ontological and grammatical views speak, respectively, to the regularity and variability observed in countability classification, these simpler models of countability run aground on the complexity of the data reviewed here. The exclusion of either the level of *entity* or of *entity description* renders many of the phenomena reported mysterious. For instance, meaning shifts, as exemplified by the word *chad* discussed in section 3.1, do not accord with either view. For the ontological view, the inherent properties of the entities falling under the noun *chad* have not changed, so no change in countability class should arise. For the grammatical view, the relation to the external world is irrelevant, so no change in countability class should arise either.

A very different set of theories take countability to be a function of (morpho-)syntax. Borer (2005), building on Sharvy (1978), argues that nominal roots, the level of *entity description* in figure 6, designate only "stuff" and contain no information related to countability—only through morphosyntactic structure, i.e. plural marking or classifiers, can a countable interpretation result. This is the mirror-image of the ontological view. There to determine if a noun is countable, one looks to the world, which then determines the grammatical class. Here to determine what a noun refers to, one must look at the grammatical class information, which determines whether the noun refers to discrete entities or stuff. Thus, these structure-based theories provide a radically different view on the locus of countability: It arises solely, or primarily, depending on the variant of the theory, via grammatical structure and little or no in-

formation about countability is contained at the lexical level, i.e. at the level of *entity description*. This is one step further than the grammatical view: Now countability is not just independent of "how things are inherently structured", but truly indifferent to it. Since these theories are motivated by contextual variability of countability, that *chad* has both a countable and non-countable variant poses no difficulty. Yet, these theories make strong predictions about the pervasiveness of contextual variability, which I will return to in section 5.1.

In the remainder of this section, I address how the theory proposed here differs in its predictions from the grammatical, ontological, and syntactic views on countability. I first discuss the challenges from the syntactic approach, since it poses the most dramatic challenge by denying lexical categorization altogether. I discuss its empirical short-comings and then how contextual variability works within the account here. After arguing that lexical categorization cannot be dispensed with, I turn to examining the different predictions which these accounts make for regularity and variability in countablity classification both within and across languages. I demonstrate that lexical and cross-linguistic variation is not incompatible with a meaning-based theory of countability, but is actually a prediction of this more articulated account.

5.1 Reconciling Lexical Meaning and Contextual Variability

The model in figure 6 speaks to the regularity of the relation between nominal reference and countability classification: A noun encodes an entity description, a particular construal of the entity, which then uniquely determines its countability classification. Yet, this would seem to be obviously contradicted by the many nouns that occur in more than one countability category. One set of cases, termed "flexible nouns" by Barner and Snedeker (2005), concern nouns such as *string* that are regularly used with both countable and non-countable syntax (*three strings/a lot of string*). The other set of cases include very many nouns that regularly fall into one countability category but can nevertheless be used in a different countability category when the context is appropriate, as in the well-known cases of "grinding", "packaging" and the like. I will consider these two types of cases in turn under the heading of CONTEXTUAL VARIABILITY.

The existence of contextual variablity has been taken by some researchers to vitiate, either partially or totally, the possibility that the lexical meaning of a noun incorporates a predictable relation to a countability categorization. Considering flexible nouns, Barner and Snedeker (2005) assert that there is no lexical specification, rather, for those nouns, countability only arises through the syntactic environment. That many nouns, given an appropriate context, can have both countable and non-countable uses, has led to even stronger views, such as that put forth by Borer (2005), where the lexical level plays *no* role, and all information about countability is contributed by the syntax. Although providing a highly economical theory, this strong view has suffered a number of empirical setbacks. Bale and Barner (2009) point out that the English noun *furniture* and other "object-mass" nouns such as *mail* are counter-examples: They must have access to individuals in their denotations despite strongly resisting pluralization or other count morphology. Thus, there must be lexical

classification at minimum for such nouns. Other researchers point out that in several languages bare nouns designate individuals rather than "stuff", despite lacking the syntactic structure which, by hypothesis, would be required, as in Dëne Sųłıné (Wilhelm, 2006) and Karitania (Doron and Muller, 2014). Dagaare, as has been discussed, is also such a language—the bare use of Dagaare's term for *dog* is only acceptable as referring to individual dogs, not to dog meat or other interpretations that may be found with non-countable terms in English. In sum, the strong view that there is no lexical classification and that nouns can be interpreted as individuals only through the contribution of syntax appears untenable in light of recent empirical work.

One could consider a mixed view, as in Bale and Barner (2009), where a portion of the lexicon, viz. their object mass nouns (*furniture*), is specified to have only individuals in the denotation, but the rest of the lexicon is under-determined and only differentiated by syntactic context. This view clearly speaks to the observed *variability* of nominal reference for this larger portion of the lexicon, for there are no constraints put on it. Yet, there is a high price attached to this analysis since it does not make predictions for *regularity* of nominal reference, and so does not connect to the typological generalizations.

It turns out that both sorts of contextual variability are straightforward to accommodate in the model advocated here. Flexible nouns, such as *string* or *chicken*, are simply polysemous between two different entity descriptions. Each of these entity descriptions follows the predictions of the mapping given in figure 6 when applied to English: The sense of *chicken* which describes living animals falls into the grammatical class of countable nouns, while the sense which describes meat falls into the grammatical class of non-countable nouns. From this perspective, flexible nouns provide little of interest for theories of countability aside from possessing a type of polysemy that straddles two countability classes, while other types of nominal polysemy, for instance between the institution and building senses for *bank*, do not. Some cases which have been treated as grinding (*chicken*) or packaging (*a beer*) fall within this explanation. Thus, these nouns present no particular obstacle to the view that nouns lexicalize countability information—one must only grant that the phenomenon of polysemy exists, which I take to be uncontroversial. This would have an impact on our theories of countability only if flexibility was a general property of nouns—a consideration to which I now turn.

The claim that it is possible to use nouns which to all appearances are countable nouns in non-countable contexts and vice versa has a long history, going as far back as Gleason (1965) (see also Allan 1980; Borer 2005; Bunt 1985; Pelletier 1979; Pelletier and Schubert 2004 among others): "Every noun, given the right context can occur in either type of usage, count or mass" (p. 136-7). Should this statement be empirically correct, then a clear solution is to assimilate all nouns to the status of flexible nouns: All nouns will permit some meaning consistent with a non-countable use and some meaning consistent with a countable use. A lexical entry for nouns is then at once impoverished in terms of possessing any sort of countability information, but extremely well-off in terms of access to meanings which are compatible with different countability contexts. A risk for this position of course is that it generalizes to the worst case.

I argue that one can accommodate contextual variability while still maintaining a contentful lexical semantics of nouns, namely through pragmatic reasoning underpinned by the model in figure 6. Take the utterance *There is armadillo all over the road*. By hypothesis, there is no single lexical item which corresponds to the meaning of splattered armadillo, and accordingly, there is a lexical gap. Using a bare form of *armadillo* serves to fill the lexical gap in this context. Even though neither a speaker nor a hearer has access to a lexicalized bare form of *armadillo* which indicates "armadillo stuff", the use of *armadillo* given the context is enough for the hearer to work out the intended meaning, of which the speaker is aware, based on analogy with bare nouns which do designate stuff and in the standard way pragmatic models have led us to expect. I now elaborate predictions that this pragmatic-based makes but that a theory for which all or most nouns are flexible nouns does not.

First, the lexicalized meaning of a noun provides the *sine qua non* use of the noun, e.g. *hat* as clothing item worn on the head, and extralexical uses arise through pragmatic reasoning, e.g. *hat* as referring to shredded material coming from a hat in utterances like *There is a hat all over the floor*. Accordingly, the extralexical uses of a given noun should be pragmatically marked. Indeed, in contrast to bona fide flexible nouns where both senses occur frequently and unremarkably in countable and non-countable contexts, many of the grinding examples in the literature have a novelty and/or a jocular nature, as noted by Bloom (1990), and may also involve re-imaginations of reality, such as the early example from Gleason (1965).

(25)   Mother termite complains about her son Johnny: "Johnny is very choosy about his food. He will eat book, but he won't touch shelf."

These qualities are hallmarks of nonce uses, as opposed to standardized uses, of a word form. Such nonce uses are well-known from the literature on language change: Novel uses of a word form often serve an immediate communicative need for the speaker, which may include "economizing, filling in a conceptual/lexical gap, or creating a stylistic effect" and which furthermore "must be understood from its context" (Brinton and Traugott, 2005, p. 45). The vast majority of examples of grinding fit this description.

This explanation of contextual variability of countability connects to other widely discussed uses of nouns which are extralexical and whose interpretation is guided by contextually-based pragmatic reasoning, such as in *The ham sandwich wants his check* (Nunberg, 1979). For these cases as well, there is a stylistic effect achieved (see Nunberg 2004 on "noteworthiness" as a condition on predicate transfers). It is telling that for these cases, the type of referent achieved by the noun in the particular context is not one that would be contained in a standard lexical entry of the noun: No one thinks that the lexical entry of *ham sandwich* refers disjunctively to the food item or to customers who order it because of these uses, yet that the lexical entry of *hat* refers disjunctively to the clothing item or its stuff has been frequently advanced in the countability literature.

A second set of predictions concern when contextual variability of countability should *not* occur. There should be two sorts of gaps: (i) principled gaps, where a secondary use of the noun is unavailable since the resultant shift in meaning would

conflict with the intrinsic meaning of the noun, and (ii) idiosyncratic gaps, where a secondary use of the noun could conceivably occur, but doesn't. Both of these gaps are found.

Despite various claims that all nouns have a countable and non-countable use, there are many types of nouns for which this appears impossible, which I term IN-FLEXIBLE NOUNS. For instance, nouns which designate a unit do not admit grinding: Galmiche (1989, 68) observes the unacceptability in French of *du kilo*, *de la catégorie*, and *du chapitre*, which also holds for the English counterparts, *much kilo*, *much category*, or *much chapter*. Similarly, nouns designating two-dimensional shapes such as *triangle* only possess a countable use and are incompatible with grinding contexts.[32] While such restrictions would be puzzling under a view upon which the grammatical nature of nouns ensures equal access to countable and non-countable uses, they follow naturally from a meaning-based approach to countability: The descriptive content of nouns such as *kilometer* restricts their interpretation to conventionalized units of measurement, which is simply incompatible with referring to non-discrete stuff.

Idiosyncratic gaps in contextual variability are well-attested, even for English which is quite permissive. For instance, grinding from an animal to its foodstuff is well-attested, but fails from a fruit to its juice, as in (26). Many other such exceptions are given in Ostler and Atkins (1991) and Nunberg and Zaenen (1992).

(26)   ??I enjoyed a glass of orange with my breakfast. Nunberg (2004, p. 352)

The availability of contextual variability also varies cross-linguistically. While some languages may have certain contextually licensed uses of bare nouns, these may differ from those in English. Nunberg (2004), citing personal communication with Jerrold Saddock, discusses the impossibility in Greenlandic Eskimo of canonical grinding from animals to their meat, but the presence of grinding from trees to their wood. (See also Cheng et al 2008 for the failure of grinding in Mandarin.) Similar variation exists in the availability of packaging: Wiese and Maling (2005) demonstrate that even in closely related Germanic languages such as English, German, and Icelandic, the availability and grammatical strategies for packaging may differ substantively.

The view that emerges of contextual variability is that on one extreme fall cases of true polysemy and at the other extreme fall extralexical cases which succeed due to pragmatic reasoning. Between these two poles, there are many semi-productive relations between countable and non-countable uses that may hold, such as in English that a word referring to an animal may also refer to its meat, and less often fur. In addition, there is a large amount of conventionalization (see discussion in Ostler and Atkins 1991; Nunberg and Zaenen 1992). Yet, here too, the broad view on countability couched in individuation makes predictions that more structure-based accounts will not. The way in which speakers interact with things in the world is a predictor of their nominal descriptions' contextual variability statuses. When we only interact with a given type of thing as discrete individual objects, such as for *hat*, then access to a meaning of, e.g., stuff is far less likely and is accordingly less likely to be stored

---

[32] See AUTHOR #1 and references therein for further discussion of inflexible nouns.

in a lexical entry. Yet, when we interact with some type of thing in the world both as discrete individual pieces and non-discrete portions, as in the case of string, then the noun describing this type of thing is more likely to refer to both discrete and non-discrete quantities. Such a correlation is obvious, yet not tractable in theories of countability unless they incorporate how speakers interact with objects in the world.

5.2 Regularity and Variation in Countability Classification: Lexical and Cross-Linguistic

I now return to considering the puzzle, brought up in section 1, of lexical variation (*foliage* vs. *leaves*) and cross-linguistic variation (*hair* vs. French *cheveux*) in countability classification. These doublets lead to an apparently paradoxical situation: How could two words which refer to the same things in the world belong to two different countability categories?

As observed by Joosten (2003), considering the relation between things and the world and their countability classification as based in objective ontological properties leaves unexplained these types of variation; however, considering the relation as arbitrary leaves unexplained the regularity witnessed in countability systems. Basing the relation between things and the world and their countability classification in individuation speaks to both the regularity and variability found. Broad typological tendencies, such as substances being typically non-countable or animate entities being typically countable, are grounded in typical manners in which speech communities individuate and interact with such things. Another typological tendency, as seen in section 2, is that entities in the world that come as aggregates are much more variable in how they are described within and across languages. This variation is, however, regular and also finds an explanation in an individuation-based account: It is precisely for these sorts of things that there is regularly a choice of how to individuate them and, consequently, how to describe them, as a single entity or as collections of entities.

As pointed out in section 1, the puzzle of variability of countability classification only arises when under the assumption that a semantic theory of countablity must be a direct, one-to-one relation between entities in the world and countability categories. It is clear that those making these arguments have such a theory in mind. For instance, as Chierchia (2010, p. 151) puts it while discussing the cross-linguistic variation of the countability of hair in English and Italian (*capelli*, pl.) as "*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."[33] This argument, then, assumes that if countability classification is semantic, it should follow from the nature of the entity in the world that is referred to. In other words, this argument presupposes that objective ontological properties,

---

[33] Although clearly the physical hair would not change if the language changes, the description of hair may change as we change language, and this may have linguistic effects. Barner and Snedeker (2005, p. 57, fn. 7) found that French speakers "judged six tiny pieces of hair to be 'plus de cheveux' (i.e. more hairs) than two long ones, while English speakers picked the two long pieces as 'more hair'."

such as the individual entity/stuff distinction, are the only relevant candidates which could influence nominal countability.

These arguments are effective against such a view of countability based in objective ontological properites, but do not engage the explanation of countability based in individuation properties. Most work on countability which invokes the notion of individuation clearly distinguishes entities in the world from *construals* of those entities (see McCawley 1975; Mufwene 1984; Bloom 1990; Wisniewski et al 2003, inter alia). Thus, an individuation account entirely agrees with those who think that the explanatory burden should not be placed on the objective ontological properties of entities in the external world. Instead, the burden is placed on construals of entities in the world.

In a given situation, such as the sentences in (27), *leaves* and *foliage* may be interchangeable.

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

Yet, this apparent synonymy does not hold in general—the different nouns have different descriptive content, which lexicographers are careful to note. The dictionary entries for *leaf* and *foliage* from the Oxford English Dictionary (OED) are given in (28) and (29), respectively.

(28)   *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)

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

These different descriptions are in turn associated with different individuation properties. For *leaves*, the focus is on the individual leaves, while for *foliage*, the focus is on the collectivity of leaves and other associated vegetation.

If within a language some entity may be described differently, it should be unsurprising that across different languages a given entity may be described differently, resulting in different countability classifications. Thus, mismatches that Chierchia brings up between *hair* and *capelli* (It.) should not be disturbing—although hair doesn't change as we change language, how hair is described may well change as we change language. This change, however, does not force one to conclude that no semantic categorization is in effect in countability classification.

Even when the descriptive content of cognates is nearly identical, the words in the different languages may have different countability classifications simply because the grammatical number systems in which they are found are structured differently. For instance, while English resorts to two different lexicalizations to convey the distinction between a collection of leaves and individual leaves, in languages with different morphological possibilities, these interpretations may be formally related. Thus, in discussing the Welsh grammatical number system, King (2003) states, "The true relationship between collective and unit nouns is particularly clear in the English translations for some of the above pairs. Foliage, for example, is a very close approximation to the actual sense of *dail* [the collective form of 'leaf'], conveying as it does

the idea of 'leaves as one homogeneous body." (p. 67) Thus, while English and Welsh presumably give voice to essentially the same meanings of *collection of leaves*, *leaf*, and *plurality of leaves*, because of the different structure of their grammatical number systems, these terms come out with different countability classifications in the different languages.

The scale of individuation makes clear predictions as to which portions of the lexicon will have a high degree of cross-linguistic agreement, namely the poles of the scale, and which sections of the lexicon will display less agreement, namely the middle region. This is exactly what one sees: Although closely related languages like English and French may agree on nouns which correspond to the two poles of the scale, they differ on lexical items that fall somewhere in the middle zone. Thus, the fact that there are cross-linguistic mismatches does not invalidate the view of semantic categorization underlying countability classification, rather these mismatches indicate for which entity types we must achieve a more nuanced understanding of the principles of semantic categorization underlying countability classification.

6 Conclusion

Based on a wide range of data, this paper has argued that one of the main suppositions of the count/mass literature stands in need of revision: Grammatical number systems are not simply coding a binary countability contrast. Instead, grammatical number classes are based in a scale of individuation. Once this step is taken, the putative paradoxes confronting a semantic grounding of countability resolve themselves. Variation among grammatical number systems, rather than being an obstacle for semantic categorization, points to regions of the lexicon that can add to our understanding of countability and sharpen our theories of word meaning and semantic categorization. The framework presented here provides a broad perspective on countability, and leaves open many paths for future investigation. At the level of lexical semantics, there are many entity types such as "paired entities" or "types of people" which have (understandably) received scant attention so far, but provide interesting cases for further study. At the typological level, a balanced and thorough examination of the relation between grammatical number coding and entity types would clearly be a valuable contribution.

References

Acquaviva P (2008) Lexical Plurals: A morphosemantic approach. Oxford University Press, Oxford

Aissen J (2003) Differential object marking: Iconicity vs. economy. Natural Language and Linguistic Theory 21:435–483

Allan K (1980) Nouns and countability. Language 56(3):41–67

Awbery GM (2009) Pluralization in Welsh, and the need for phonologically null suffixes. Journal of Celtic Linguistics 13:1–20

Baayen H, Burani C, Schreuder R (1997) Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In: Booij G, van Marle J (eds) Yearbook of Morphology 1996, Kluwer Academic Publishers, pp 13–33

Baayen R, Piepenbrock R, Gulikers L (1996) CELEX2. Linguistic Data Consortium, Philadelphia, PA

Bale A, Barner D (2009) The interpretation of functional heads: Using comparatives to explore the mass/count distinction. Journal of Semantics 26:217–252

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

Bloom P (1990) Semantic structure and language development. PhD thesis, Massachusetts Institute of Technology

Bloom P (1994) Possible names: The role of syntax-semantics mappings in the acquisition of nominals. Lingua 92:297–329

Bloomfield L (1933) Language. Henry Holt and Co., New York

Borer H (2005) Structuring Sense I: In Name Only. Oxford University Press, Oxford

Borg AJ (1980) An aspectual distinction in verbs and nouns in Maltese. Archivum Linguisticum 11(2):107–117

Brennan A (1988) Conditions of identity. Clarendon Press, Oxford

Brinton LJ, Traugott EC (2005) Lexicalization and Language Change. Research Surveys in Linguistics, Cambridge University Press, Cambridge

Brown CH (1999) Lexical acculturation in Native American languages, Oxford studies in anthropological linguistics, vol 20. Oxford University Press, NewYork and Oxford

Bunt HC (1985) Mass Terms and Model-Theoretic Semantics. Cambridge University Press, Cambridge

Cheng LL, Doetjes J, Sybesma R (2008) How universal is the universal grinder? In: van Koppen M, Botma B (eds) Linguistics in the Netherlands 2008, John Benjamins, Amsterdam/Philadelphia, pp 50–62

Chierchia G (1998) Plurality of mass nouns and the notion of "semantic parameter". In: Rothstein S (ed) Events and Grammar, Kluwer Academic Publishers, Dordrecht, pp 53–104

Chierchia G (2010) Language, thought, and reality after Chomsky. In: Franck J, Bricmont J (eds) The Chomsky Notebook, Columbia University Press, New York, pp 142–169

Contini-Morava E (2000) Noun class as number in Swahili. In: Contini-Morava E, Tobin Y (eds) Between Grammar and Lexicon, John Benjamins, Amsterdam/Philadelphia, pp 3–30

Corbett GG (1996) Minor number and the plurality split. Rivista di Linguistica 8(1):101–122

Davidson D (1973) On the very idea of a conceptual scheme. In: Proceedings and Addresses of the American Philosophical Association, vol 47

Denny JP (1976) What are noun classifiers good for? In: Proceedings of the 12th Chicago Linguistic Society, Chicago Linguistics Society, Chicago, pp 122–132

Dimmendaal GJ (1983) The Turkana Language. Foris Publications, Dordrecht

Dimmendaal GJ (2000) Number marking and noun categorization in Nilo-Saharan languages. Anthropological Linguistics 42(2):214–261

Doron E, Muller A (2014) The cognitive basis of the mass-count distinction : evidence from bare nouns. In: Hofher PC, Zribi-Hertz A (eds) Crosslinguistic studies on noun phrase structure and reference, Syntax and Semantics, vol 29, Brill, Leiden and Boston, pp 73–101

Dowty D (1991) Thematic proto-roles and argument selection. Language 67(3):547–619

Fargetti CM (2001) Estudo fonológico e morfossintático da língua Juruna. PhD thesis, Unicamp University

Galmiche M (1989) Massif/comptable: de l'un à l'autre et inversement. In: David J, Kleiber G (eds) Termes massifs et termes comptables (Recherches Linguistiques 13), Klincksieck, Paris, pp 63–77

Gentner D, Boroditsky L (2001) Individuation, relational relativity and early word learning. In: Bowerman M, Levinson S (eds) Language acquisition and conceptual development, Cambridge University Press, Cambridge, pp 215–256

Gil D (1996) Maltese 'collective' nouns: A typological perspective. Rivista di Linguistica 8(1):53–87

Gillon B (1992) Towards a common semantics for English count and mass nouns. Linguistics and Philosophy 15:597–639

Gleason HA (1965) Linguistics and English Grammar. Holt Rinehart and Winston, New York

Greenberg J (1966) Language Universals. Mouton, The Hague

Harbour D (2008) Morphosemantic Number: From Kiowa Noun Classes to UG Number Features. Springer, Dordrecht

Haspelmath M (2005) Occurrence of nominal plurality. In: Haspelmath M, Dryer MS, Gil D, Comrie B (eds) The World Atlas of Language Structures, Oxford University Press, Oxford, pp 142–145

Haspelmath M (2008) Creating economical morphosyntactic patterns in language change. In: Linguistic Universals and Language Change, Oxford University Press, Oxford, pp 185–214

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

Imai M, Gentner D (1997) A cross-linguistic study of early word learning: universal ontology and linguistic influence. Cognition 62:169–200

Jakobson R (1957/1971) Shifters, verbal categories, and the russian verb. In: Jakobson R (ed) Selected Writings, Vol. 2: Word and Language, Mouton, The Hague, pp 386–392

Jones M, Thomas AR (1977) The Welsh Language: Studies in its syntax and semantics. University of Wales Press for the Schools Council, Cardiff

Joosten F (2003) Accounts of the count-mass distinction: A critical survey. Linguisticae Investigationes 26:159–173

Joosten F, de Sutter G, Drieghe D, Grondelaers S, Hartsuiker RJ, Speelman D (2007) Dutch collective nouns and conceptual profiling. Linguistics 45:85–132

Katamba F (2003) Bantu nominal morphology. In: Nurse D, Philippson G (eds) The Bantu Languages, Routledge, New York, pp 103–120

King G (2003) Modern Welsh: A Comprehensive Grammar. Routledge

Kiparsky P, Tonhauser J (2012) Semantics of inflection. In: Maienborn C, von Heusinger K, Portner P (eds) Semantics: An International Handbook of Natural Language Meaning, Mouton de Gruyter, Berlin, pp 2070–2097

Koontz-Garboden A (2007) States, Change of State, and the Monotonicity Hypothesis. PhD thesis, Stanford University, Stanford, CA.

Krifka M (1992) Thematic relations as links between nominal reference and temporal constitution. In: Sag I, Szabolcsi A (eds) Lexical Matters, CSLI, Stanford, pp 29–53

Landau B, Smith LB, Jones SS (1988) The importance of shape in early lexical learning. Cognitive Development 3:299–321

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

Landman F (2011) Count nouns — mass nouns neat nouns — mess nouns. In: Partee B, Glanzberg M, Šķilters J (eds) The Baltic International Yearbook of Cognition, Logic and Communication, New Prairie Press, Manhattan, KS, vol 6 Formal Semantics and Pragmatics: Discourse, Context, and Models

Lima S (2014) The grammar of individuation and counting. PhD thesis, University of Massachusettes - Amherst

Lucy JA (1992) Grammatical categories and cognition: A case study of the linguistic relativity hypothesis. Cambridge University Press, Cambridge, MA

Maho JF (1999) A Comparative Study of Bantu Noun Classes. Acta Universitatis Gothoburgensis, Göteborg

Mathieu E (2012) Flavors of division. Linguistic Inquiry 43(4):650–679

McCawley JD (1975) Lexicography and the count-mass distinction. In: Proceedings of the 1st Annual Meeting of the Berkeley Linguistics Society, Berkeley Lingustics Society, Berkeley, pp 314–321

Middleton EL, Wisniewski EJ, Trindel KA, Imai M (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

Mifsud M (1996) The collective in Maltese. Rivista di Linguistica 8(1):29–51

Mourelatos A (1978) Events, processes and states. Linguistics and Philosophy 2:415–434

Mufwene SS (1980) Number, countability and markedness in Lingala LI- MA- noun class. Linguistics 18:1019–1052

Mufwene SS (1984) The count/mass distinction and the English lexicon. In: Parasession on lexical semantics, Chicago Linguistic Society, Chicago, pp 200–221

Nichols L (2008) Lexical semantic constraints on noun roots and noun borrowability. Studies in Language 32(3):683–700

Nunberg G (1979) The non-uniqueness of semantic solutions: Polysemy. Linguistics and Philosophy 3:143184

Nunberg G (2004) Pragmatics of deferred intrepretation pp 343–364

Nunberg G, Zaenen A (1992) Systematic polysemy in lexicology and lexicography. In: Euralex II, pp 387–396

Nurmio S (2016) Collective nouns in Welsh: A noun category or a plural allomorph? Transactions of the Philological Society pp 1–21

Ohta I (1989) A Classified Vocabulary of the Turkana in Northwestern Kenya. African Study Monographs, The Center for African Area Studies at Kyoto Uni-

versity, Kyoto, Japan

Ostler N, Atkins B (1991) Predictable meaning shift: Some linguistic properties of lexical implication rules. In: Pustejovsky J, Bergler S (eds) Proceedings of the First SIGLEX Workshop on Lexical Semantics and Knowledge Representation, Springer-Verlag, pp 87–100

Palmer FR (1971) Grammar. Penguin, Harmondsworth

Pelletier FJ (1975) Non-singular reference: Some preliminaries. Philosophia 5:451–465

Pelletier FJ (1979) Non-singular reference: Some preliminaries. In: Pelletier F (ed) Mass Terms: Some Philosophical Problems, Reidel, Dordrecht, pp 1–14

Pelletier FJ, Schubert LK (2004) Mass expressions. In: Gabbay D, Guenthner F (eds) Handbook of philosophical logic, vol 10, Kluwer, Dordrecht, pp 249–335

Prasada S, Ferenz K, Haskell T (2002) Conceiving of entities as objects and stuff. Cognition 83:141–165

Press I (2009) Breton. In: Ball MJ, Müller N (eds) The Celtic Languages, Routledge, New York, pp 427–487

Quine WVO (1960) Word and Object. MIT Press, Cambridge, MA

Ravid D, Hayek L (2003) Learning about different ways of expressing number in the development of Palestinian Arabic. First Language 23:41–64

Rothstein S (2010) Counting and the mass/count distinction. Journal of Semantics 27:343–397

Samuelson LK, Smith LB (1999) Early noun vocabularies: do ontology, category structure and syntax correspond? Cognition 73:1–33

Schwarzschild R (2011) Stubborn distributivity, multiparticipant nouns and the count/mass distinction. In: Lima S, Mullin K, Smith B (eds) Proceedings of the 39th meeting of the North East Linguistics Society, GLSA, Amherst, MA, vol 2, pp 661–678

Seifart F (2009) Towards a typology of unitization: Miraña noun classes compared to numeral classifiers and singulatives Unpublished Ms.

Sharvy R (1978) Maybe English has no count nouns: Notes on Chinese semantics. Studies in Language 2:345–365

Smith-Stark T (1974) The plurality split. In: Galy ML, Fox R, Bruck A (eds) Papers from the 10th Chicago Linguistic Society, Chicago Linguistic Society, Chicago, pp 657–671

Soja NN, Carey S, Spelke ES (1991) Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. Cognition 38:179–211

Stebbins T (2007) Asymmetrical nominal number marking: a functional account. STUF - Language Typology and Universals 50(1):5–47

Stolz T (2001) Singulative-collective: Natural morphology and stable cases in Welsh number inflexion of nouns. Sprachtypologie und Universalienforschung 54(1):52–76

Sutcliffe EF (1936) A Grammar of the Maltese Languages. Progress Press, Valletta

Tiersma P (1982) Local and general markedness. Language 58:832–849

Ware RX (1975) Some bits and pieces. Synthese 31:379–393

Watkins L (1984) A Grammar of Kiowa. University of Nebraska Press, Lincoln

Wierzbicka A (1988) The semantics of grammar. John Benjamins, Amsterdam/Philadelphia

Wiese H, Maling J (2005) Beers, kaffi, and schnaps: Different grammatical options for restaurant talk coercions in three Germanic languages. Journal of Germanic Linguistics 17:1–38

Wiggins D (1980) Sameness and Substance. Blackwell, Oxford

Wilhelm A (2006) Count and mass nouns in Dëne Sųłıné. In: Baumer D, Montero D, Scanlon M (eds) Proceedings of the 25th West Coast Conference on Formal Linguistics, Cascadilla Proceedings Project, Somerville, MA, pp 435–443

Wisniewski EJ, Imai M, Casey L (1996) On the equivalence of superordinate concepts. Cognition 60:269–298

Wisniewski EJ, Lamb CA, Middleton EL (2003) On the conceptual basis for the count and mass noun distinction. Language and Cognitive Processes 18:583–624

Zwicky A (2001) Counting chad, handout from Stanford Semantics Fest, March 16, 2001