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

Scott Grimm

Universitat Pompeu Fabra

February 15, 2013

# Grammatical Number

The relation between language and number has been of interest in philosophy, psychology and linguistics

The stage for these discussions is often set by English

Two primary oppositions:

- Singular-Plural contrast (*dog* vs. *dogs*)
- Count-Mass contrast (*dog* vs. *sand*)

# Grammatical Number

**Goal**: gain an understanding of how grammatical number systems encode entities in the world

- What types of entities can be counted?
- What morphological forms are available, e.g. singular, plural, dual, paucal, etc.
- Are there principles governing the structure of grammatical number systems?

# Morphosyntactic Traits of Countability

What does it mean for a noun to be count or non-count?

- morphosyntactically defined

Count nouns (*dog*, *chair*):

- plural coding (*dogs*, *chairs*)
- modification by cardinal quantifiers (*two dogs/chairs*)
- modification by determiners implicating plurality (*several dogs*, *several chairs*)

Non-count ("Mass") nouns (*water, sand*)

- do not permit plural coding (**waters, *sands*)
- nor cardinal quantifiers or those implicating plurality (**two waters*, **several sands*)
- may allow modification by *much* or *a lot of*

# Count-Mass Distinction

The count-mass distinction has been a hotbed of controversy for decades

Wide variety of positions

- grammatical account (Bloomfield 1933; Chierchia 1998)
- semantic account (Bunt 1985; Wierzbicka 1985; Jackendoff 1992)
- contextual-pragmatic accounts (Pelletier 1979)

# Count-Mass Distinction: Grammatical Account

The count-mass distinction is grammatical

- no connection with extra-linguistic factors

"The mass/count distinction is specific to human language and to its grammar" (Chierchia 2010)

Challenges:

- leaves unexplained clear intuitive (and cross-linguistic) correlations between individuals and countable nouns (*dog*) and substances and uncountable nouns (*water*)

# Count-Mass Distinction: Semantic Account

The count-mass distinction reduces to a semantic/conceptual distinction:

- individuated entities (those appearing as an individual unit) vs. unindividuated entities (Wierzbicka 1985)
- bounded vs. unbounded (Jackendoff 1992)

Explains correlations between individuals and countable nouns (*dog*) and substances and uncountable nouns (*water*)

Challenges:

- Many intermediate cases are not given a principled explanation (*peas* vs. *rice*)
- Leaves open many questions about cross-linguistic differences—do they correspond to differences in conceptualization?

# Grammatical Number: Proposal

**Proposal**: Morphosyntactic organization of grammatical number systems works in tandem with the semantic organization

- Number systems obey broad semantic constraints, yet at the same time allow for some variation as to how they are organized

# Outline

1. Examine some lesser known number systems (Welsh, Dagaare)
2. Develop a typological picture of how countability is encoded by grammatical systems
3. Provide a formal treatment of countability, which requires an extension of the standard tools

# Welsh

Welsh provides a simple demonstration that grammatical number systems can be more complex than a binary count/mass contrast

- Welsh has a tripartite system

| Countability Category | Singular | Plural | Gloss |
|---|---|---|---|
| Singular/Plural | cadair | cadair-**iau** | 'chair' |
| Plural/Singulative | cacyn-**en** | cacwn | 'hornet' |
| Non-Count | llefrith | | 'milk' |

# Dagaare

Dagaare (Gur; Niger-Congo, spoken in NW Ghana) has a unique system that

- grammatically recognizes different types of count nouns
- grammatically recognizes different types of non-count nouns

Based on work in 2008 and 2011

- with Mark Ali (College of Education, Winneba, Ghana)
- currently completing a Dagaare-English dictionary

# Count-Mass Contrast in Dagaare

There is a clear opposition between objects and substances in Dagaare:

- substance terms do not fall into the standard singular/plural patterns
- they have special number suffixes: a distinct distributive plural suffix *-ree*

| **Mass** | **2nd Pl.** | **Gloss** |
|---|---|---|
| kùó | kònnéé | 'water/ (types of) waters' |
| múó | múónéé | 'grass/grasses' |

# Count-Mass Contrast in Dagaare

Substance terms can be further divided as to whether they permit the singulative suffix *-ruu*:

| **Singulative** | **Mass** | **2nd Pl.** | **Gloss** |
|---|---|---|---|
| — | kùó | kònnéé | 'water/ (types of) waters' |
| múórúú | múó | múónéé | 'blade of grass/grass/grasses' |

Implicates two types of mass nouns in Dagaare

# Singular/Plural Opposition in Dagaare

Singular and plural distinctions are coded by the same suffix

- synchronically resembles an "inverse number marking" system

Demonstrated by near-minimal pair below

- same stem, yet *-ri* codes the plural interpretation for 'child' and the singular interpretation for 'seed'

| **Singular** | **Plural** | **Stem** | **Gloss** |
|---|---|---|---|
| bíé | bíírí | bi- | 'child' |
| bìrí | bíè | bi- | 'seed' |

# Inverse Number Pattern in Dagaare

This is a rare system but attested in at least North America (Kiowa) and the Pacific (New Ireland) (see Corbett 2000)

*Mock English Inverse Pattern*

| **Singular** | **Plural** | **Gloss** |
|---|---|---|
| cat | cat-s | 'cat' |
| dog-s | dog | 'dog' |

# Inverse Number: *-ri* plural

| -V Singular | -rI/-nI Plural | Gloss |
|---|---|---|
| bíé | bíírí | ‘child’ |
| tìɛ́ | tììrí | ‘tree’ |
| gbìé | gbèrí | ‘forehead’ |
| pìɛ́ | pɛ̀rí | ‘basket’ |
| dùó | dòŕi | ‘pig’ |
| nàŋá | nànní | ‘scorpion’ |

# Inverse Number: *-ri* singular

| rI/-nI Singular | -V Plural | Gloss |
|---|---|---|
| bìrí | bìé | 'seed' |
| kùùrí | kùé | 'hoe, metal instrument' |
| lúgrí | lúgó | 'prop, pillar' |
| nyágrí | nyágá | 'root' |
| fìlí | fìlɛ́ | 'sores' |
| dólì | dólò | 'dry spot' |
| íílí | íílɛ̀ | 'horn' |

# Markedness

- This pattern poses a serious potential problem for theories of markedness:
  - Usually, singular is unmarked and plural is marked (Jakobson, Greenberg's Universal 35)
  - The inverse number marking pattern clearly contradicts this

# Understanding the distribution

Dagaare appears to have a greatly reduced noun class system, where, aside from the human class, almost all the singular/plural number coding is performed by -ri

- No viable phonological explanation

Does the distribution adhere to any principles?

*Individuation*: the propensity for an entity to appear as an individual unit

- explored in the philosophical, linguistic and psycholinguistic literatures (Quine 1960; Mufwene 1980; Bloom 1994)

# Understanding the distribution: Individuation

*Individuation Hypothesis*:

- Cognitive or perceptual qualities influence the grammatical realization of count and mass nouns
  - count nouns (*dog*) correlate with individual entities
  - mass nouns (*water*) correlate with non-individuated substances
- Open Question: Is individuation relevant for other grammatical number distinctions, such as those found in Dagaare?

# Understanding the distribution: Individuation

Individuation has been used as a cover term for many distinct properties (cf. agentivity)

Grimm (2011) considered the potential influence of four individuating factors on the realization of nominals in Dagaare:

- animacy (Smith-Stark 1974; Corbett 1996, 2000)
- ease of distinguishability (Wierzbicka 1988; Middleton et al. 2004)
- manner of interaction (Wierzbicka 1988; Middleton et al. 2004)
- "inherent plurality" (Acquaviva 2008)

# Individuation: Animacy

Animacy is known to influence number marking cross-linguistically (Smith-Stark 1974; Corbett 1996, 2000)

- The higher the entity referred to by a noun rates on an animacy hierarchy, the greater likelihood that the noun is capable of expressing a singular/plural contrast.

# Individuation: Distinguishability

"Distinguishability" (Wierzbicka 1988)

- entities whose constituents are easily distinguishable are more likely to be used as a count nouns
- entities whose constituents are not easily distinguishable are used as mass nouns.
- *beans* is more likely to be a count term than *rice* since individual beans are in principle easier to distinguish than individual grains of rice.

# Individuation: Distinguishability

**Experimental Evidence for Distinguishability**
(Middleton et al. 2004)

Subjects matched a nonce count or mass term with one of two graphical displays of novel aggregates which varied in terms of distinguishability along two dimensions:

(i) spatial proximity to other elements (Close versus Apart)

(ii) size of elements (Large versus Small)

# Individuation: Distinguishability

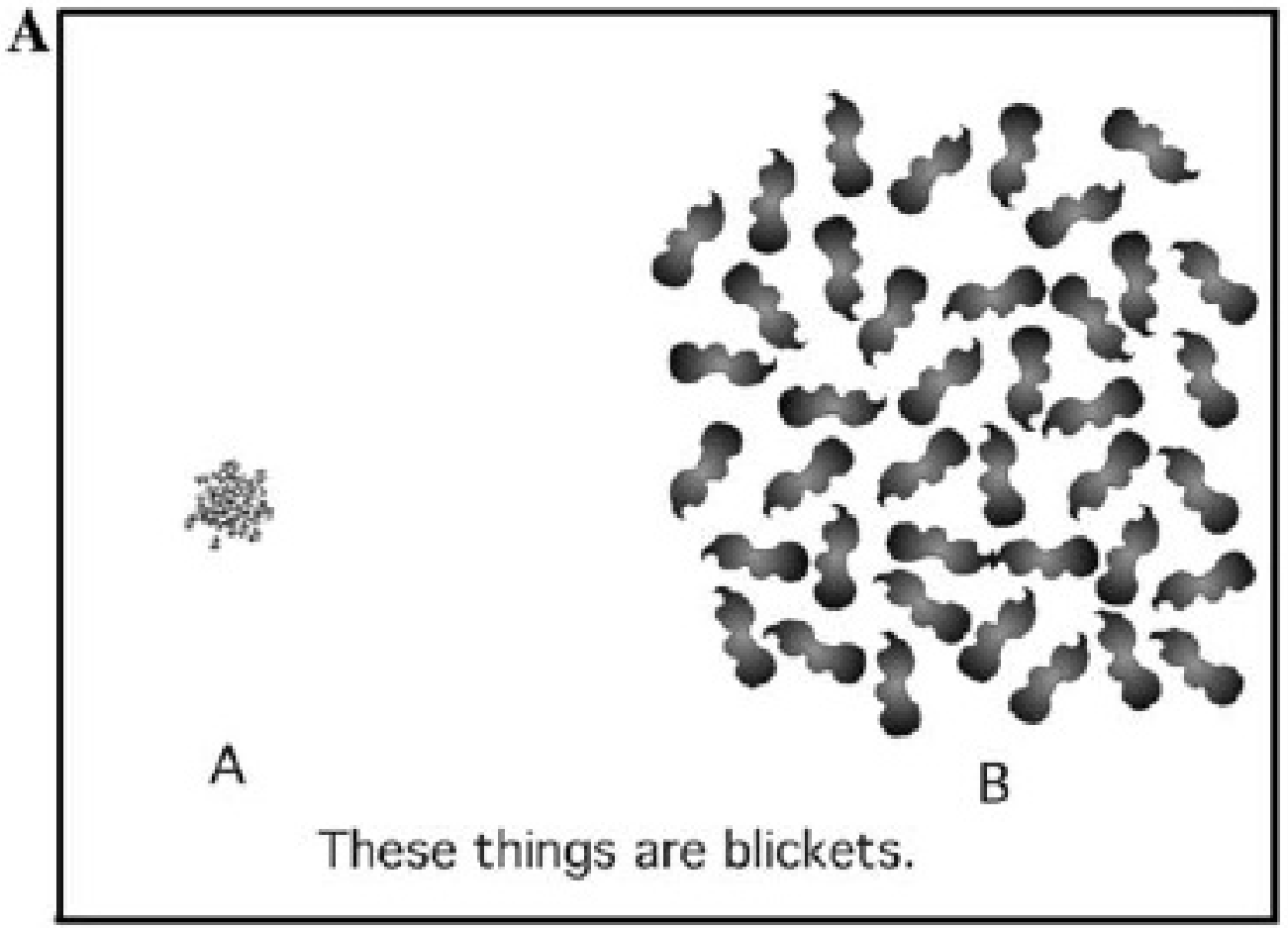
A subject would see two pictures of a set of elements:

- each element was spatially separated from the others
- each element was spatially contiguous with others.

Subjects are asked which picture aligns with phrases such as "This is blicket" / "These are blickets"

# Individuation: Distinguishability

Novel Aggregates Used in Middleton et al. (2004)



# Individuation: Distinguishability

Results:

Subjects' choice of count or mass terms was very significantly influenced ($p < .001$) by the spatial proximity, but not the size, of the elements.

# Individuation: Manner of Interaction

Whether one interacts with individual elements has an effect on categorization of the nouns as mass or count

Middleton et al. (2004) provide experimental evidence with a forced choice design (mass vs. count syntax)

When subjects were presented with a novel aggregate—"yellow decorative coarse-grained sugar" in a cardboard box—the majority described it with mass syntax ("This is worgle").

When subjects interacted with the sugar by scooping up individual grains, the majority described with count syntax ("These are worgles").

# Individuation: Inherent Plurality

Certain entity types (*eyes*) are predisposed to come, not as individual units, but as multiple units

- often coded by duals/collectives
- qualitatively different from entities which canonically appear as individuals (Acquaviva 2008)

# Back to Dagaare: The Hypothesis

Nouns possess lexical information, i.e. nouns come with a 'basic' number determined by the noun's semantic properties

- the application of *-ri* gives the inverse value

The more likely the entity is to be viewed as individuated, the more likely the singular will be the default (unsuffixed) form and *-ri* will code the plural

- Highly Individuated N + *-ri* $\Rightarrow$ plural

The more likely the entity is to be viewed as coming in groups or non-individuated, the more likely the plural will be the default (unsuffixed) form and *-ri* will code the singular

- Less Individuated/Inherently Plural N + *-ri* $\Rightarrow$ singular

# Validation Across Semantic Domains

**Prediction:** there should be asymmetries in the distribution of *-ri* in semantic domains relevant to the different individuation factors:

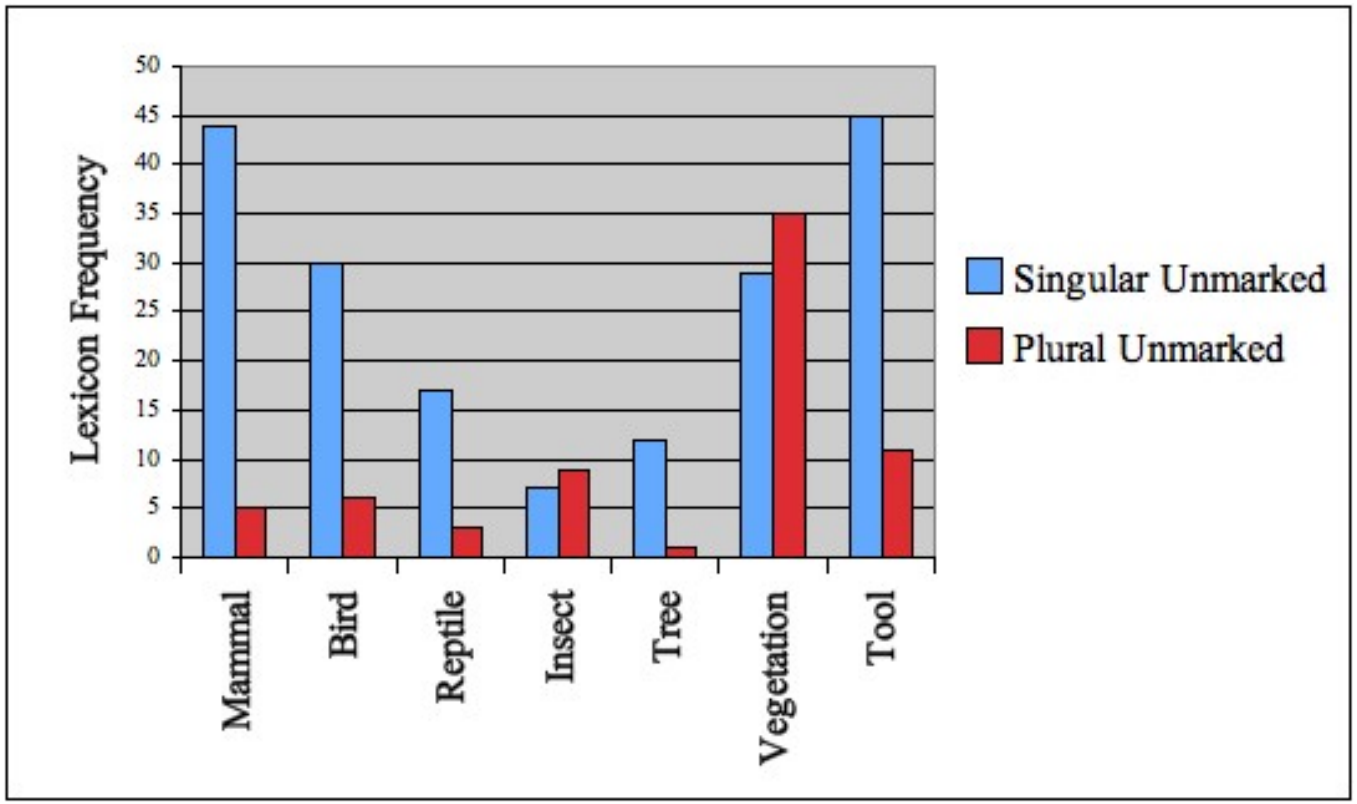
(i) Nouns naming larger/more salient animals should be more likely to refer to a single animal in their unsuffixed form in contrast to nouns naming insects/bugs (*animacy*)

(ii) Nouns naming trees should be more likely to have the singular as the unsuffixed form in comparison to vegetation (*distinguishability*)

(iii) Nouns naming tools should be more likely to be unsuffixed in the singular (*canonically interact with them individually*)

(iv) Nouns naming body parts in pairs/groups are more likely to be unsuffixed in the plural while nouns naming non-paired/grouped body parts are more likely to be unsuffixed in the singular (*inherent plurality*)

# -ri across semantic domains



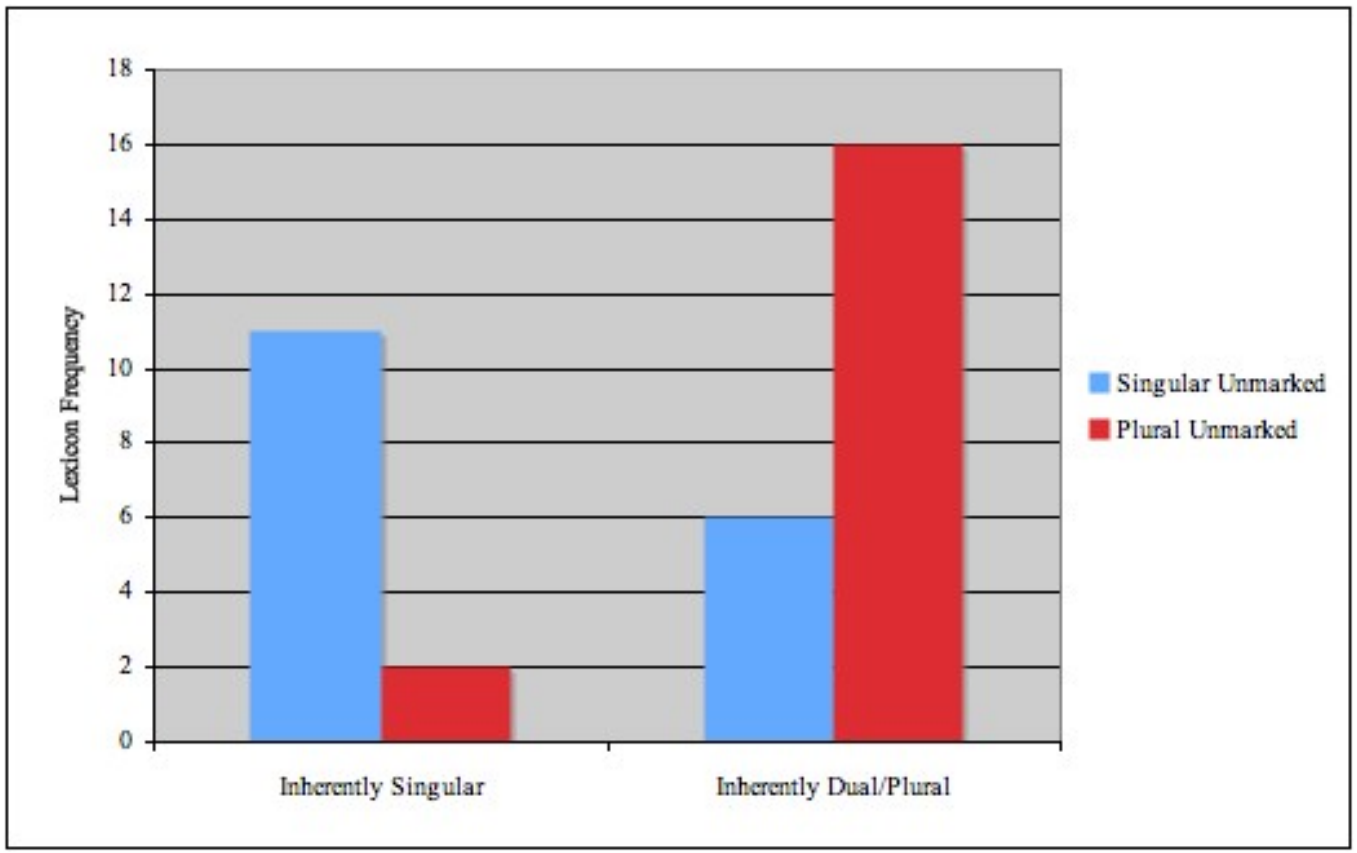
# Validation Across Semantic Domains

Reliable asymmetries are visible across the semantic domains:

- Nouns naming higher-level animates, trees and tools typically have the singular as the default (unsuffixed) form
- The majority of nouns naming bugs and vegetation have the plural as the default (unsuffixed) form

This study controls for derived forms, since they follow their own patterns which tends to obscure the generalizations

# Validation for Body Parts



# Validation Across Semantic Domains

Nouns in a given domain that do not conform to the expected pattern also show semantic subregularities:

- most of the bugs for which the singular is the unsuffixed form are those capable of causing harm (e.g. scorpion, wasp, spider)

# Varia

There are two words in Dagaare which are glossed almost identically:

| Singular | Plural | Gloss |
|---|---|---|
| wɛ́gɛ̀ | wɛ́grì | 'log' |
| lúgrí | lúgó | 'log, pillar' |

# Varia

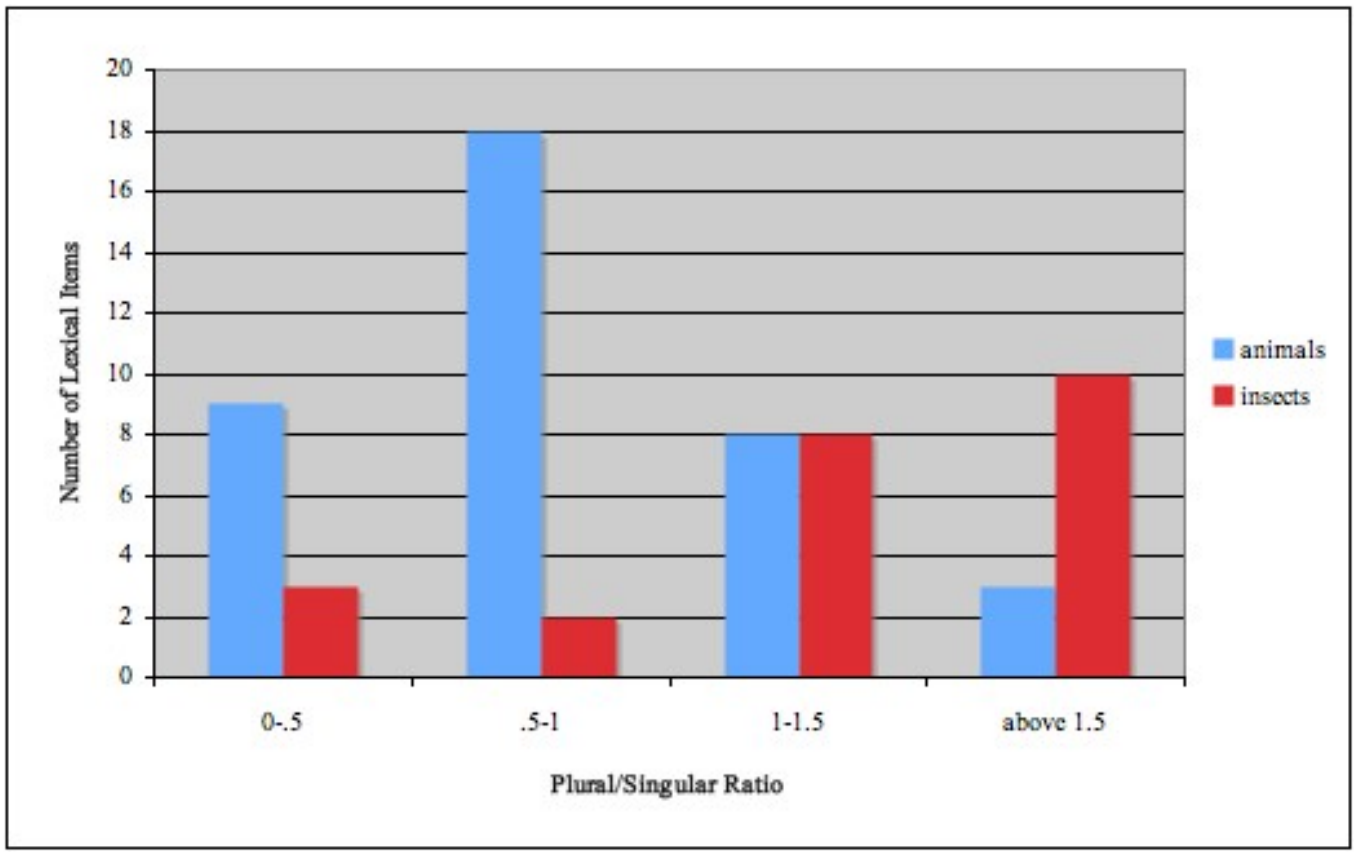
# Evidence from English Frequency Patterns

**Prediction**: "unmarked plurals" should be detectable even in languages which do not code them morphologically

Specifically, such nouns should be used more frequently in the plural, making the plural the "locally unmarked" form (Tiersma 1982)

- Examined frequencies of English words using the COBUILD corpus (18 million words)
- Calculated plural-to-singular ratio for animals and bugs (basic terms and those consistent with the vocabulary of Dagaare)

# Evidence from English Frequency Patterns



# Markedness: Economy and Learnability

Languages with plural defaults are economical:

- omit more plural morphology than undifferentiated languages like English

Yet, learnability is presumably much more difficult

# Summary

Bottom line:

- Dagaare singular/plural morphology is sensitive to the degree of individuation of a noun's referent
- *-ri* codes singular when a noun is considered to be less individuated, otherwise codes the plural

More generally

- individuation factors are relevant beyond just the mass/count distinction

# Countability and Morphosyntax

Four elements relevant to understanding countability:

- things in the world (extensions)
- lexical items (nouns)
- individuation types (e.g. granular aggregate)
- morphosyntactic classes (count, mass)

Is there an organizing principle determining which individuation types are associated with which morphosyntactic classes?

This is not trivial as

- languages dispose of different numbers of morphosyntactic classes related to countability
- additionally characterized by differences in which member of an opposition is coded

# Cross-Linguistic Differences

I argue the relation between individuation types and morphological classes are

- systematic rather than arbitrary
- cohere to a scale of individuation

Examine three languages and the relation between individuation type and morphological class:

- English [2 classes]
- Welsh [3 classes]
- Dagaare [4 classes]

# English: Morphosyntactic Classes

English makes a two-way split in terms of morphosyntactic type:

- Class 1: Nouns allow plural coding
  - individuated things (*apple*, *pencil*)
  - collective aggregates (*bees*, *grapes*)
- Class 2: Nouns have one form
  - liquids (*water*, *oil*)
  - substances (*granite*, *wood*)
  - granular aggregates (*flour*, *rice*, *sand*, *sugar*)

# English: Morphosyntactic Classes

| Language | liquids/ substances | granular aggregates | collective aggregates | individual entities |
|---|---|---|---|---|
| English | 0 | | 0/Plural (*–s*) | |

# Welsh: Morphosyntactic Classes

Welsh has a three-way split:

- Class 1: Nouns allowing plural coding
  - includes primarily animates and other individuals
- Class 2: Nouns allowing singulative coding
  - includes granular aggregates (*turf*, *sand*) as well as collective aggregates (small animals and insects, vegetables/grains/fruits, inherently plural body parts)
- Class 3: Nouns having one form
  - includes liquids and substances

# Welsh: Morphosyntactic Classes

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

# Dagaare: Morphosyntactic Classes

Dagaare has a four-way split :

- Class 1: Nouns with plural coded
  - individuals (*child, dog*)
- Class 2: Nouns with singular coded
  - collective aggregates such as vegetation, insects, or inherently plural body parts
- Class 3: Nouns with optional singulative
  - granular aggregates such as *pepper*, *straw*, *grass*
- Class 4: Nouns with one form
  - liquids, materials

# Mapping the Terrain

Ordering classes from those least likely to be coded in the plural to those least likely to be coded in the singular imposes an order on the individuation types

| **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*) | |

# The Scale of Individuation

The picture that emerges from the table suggests that the individuation types form a scale

liquids/substances $<$ granular aggregates $<$ collective aggregates $<$ individual entities

This scale can be viewed as organized under the principle of individuation

# Understanding the ordering of the scale

The poles of the scale are liquids/substances vs. individual entities

This opposition in turn corresponds to minimally and maximally individuated entities:

- **Liquids/substances**: minimal elements are continuous and not distinguishable: one does not interact with individual elements at all
- **Individual entities**: the inverse holds

This fundamental opposition appears early in child development (Soja et al. 1991).

# Understanding the ordering of the scale

**Granular aggregates** have individuation properties similar to liquids and tend to pattern with them morphosyntactically:

- often have minimal elements (*a grain of sand*), which are not easily distinguishable; one does not canonically interact with them

**Collective aggregates**:

- the minimal elements are more accessible than for granular aggregates; interaction with their minimal elements is also more frequent.

# Connecting Number and Individuation

Can this scale help us understand morphosyntactic number classes?

This approach constrains the possible inventory of morphosyntactic systems of number

- they must respect the scale
- although different languages can lump together different portions of the scale in different ways

# Establishing morphosyntactic classes

Morphosyntactic classes are carved out of the scale of individuation

- categories of number are grammatical categories
- yet based in individuation types, or coherent combinations of individuation types

$$\underbrace{\text{ind. type 1} < \text{ind. type 2}}_{\text{Morphosyntactic Class 1}} < \underbrace{\text{ind. type 3} < \text{ind. type 4} < \text{ind. type 5}}_{\text{Morphosyntactic Class 2}}$$

# How the scale relates to the morphosyntactic categories

Entities of a given individuation type may receive distinct treatments in different languages

| **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*) | |

A generalization across morphosyntactic inventories

- no morphosyntactic class spans two individuation types that are not contiguous on the scale

# Revisiting the Count-Mass distinction

From the viewpoint developed, which recognizes distinct roles for morphosyntax and semantics, many of the traditional problems which have plagued the mass-count distinction vanish

# Revisiting the Count-Mass distinction

The mass-count distinction has often served as a touchstone for those who would deny that meaning and grammar are related

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

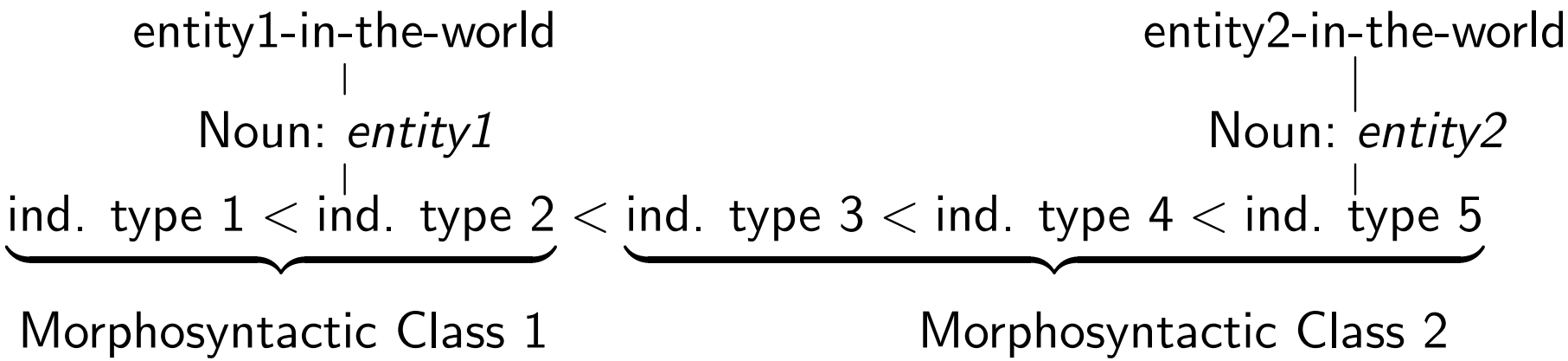
# Things-in-the-world, individuation types and morphosyntactic classes

Challenges for a semantic account:

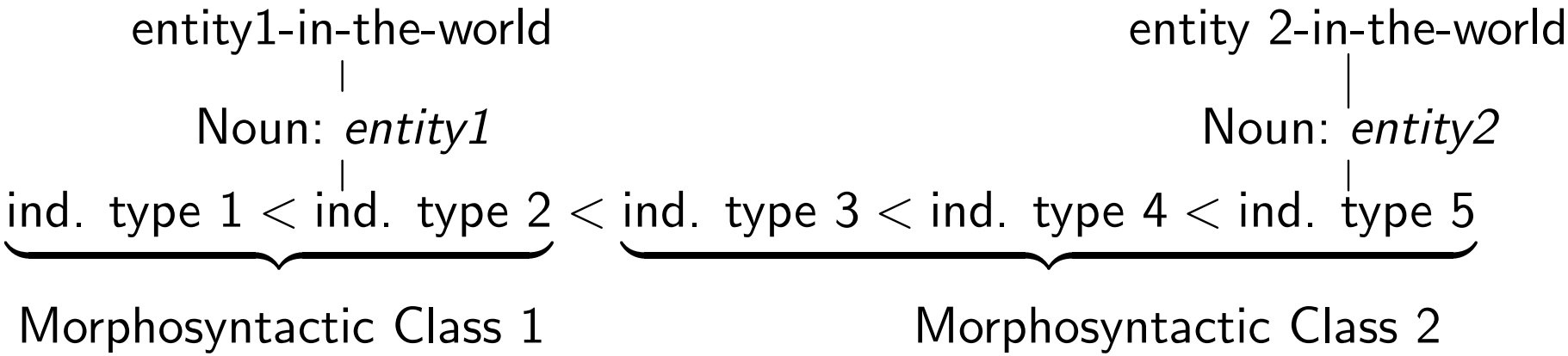
- Across-language variation:
  - *hair* vs. *cheveux*
- Within-language variation:
  - *foliage* vs. *leaves*

# Mapping between things-in-the-world, individuation types and morphosyntactic classes

The figure displays four different levels that have been discussed: things-in-the-world, nouns and their properties, individuation types and grammatical classes.



# Revisiting the Count-Mass distinction




Two types of indeterminacy are possible in the mappings between levels:

- how grammatical number systems put together different individuation types
- construal of nouns, and consequently the associated individuation properties

# Application: Across Language Variation

An entity mapped to a given individuation type may have a different morphosyntactic realization

Language 1:

$$\underbrace{\text{ind. type 1} < \overset{\text{name of entity X}}{\overset{|}{\text{ind. type 2}}}}_{\text{Morphosyntactic Class 1}} < \underbrace{\text{ind. type 3}}_{\text{Morphosyntactic Class 2}}$$

Language 2:

$$\underbrace{\text{ind. type 1}}_{\text{Morphosyntactic Class 1}} < \underbrace{\overset{\text{name of entity X}}{\overset{|}{\text{ind. type 2}}} < \text{ind. type 3}}_{\text{Morphosyntactic Class 2}}$$

# Application: Within Language Variation

*Key insight*: a single slice of reality can be described in different ways in a given moment, but that doesn't mean that the descriptions are in all manners equivalent

Some nouns provide a holistic perspective on a set of entities as opposed to providing a description of the individual entities

- *foliage* (compare *leaves*): the collectivity and the interconnectedness of leaves with one another rather than individual leaves

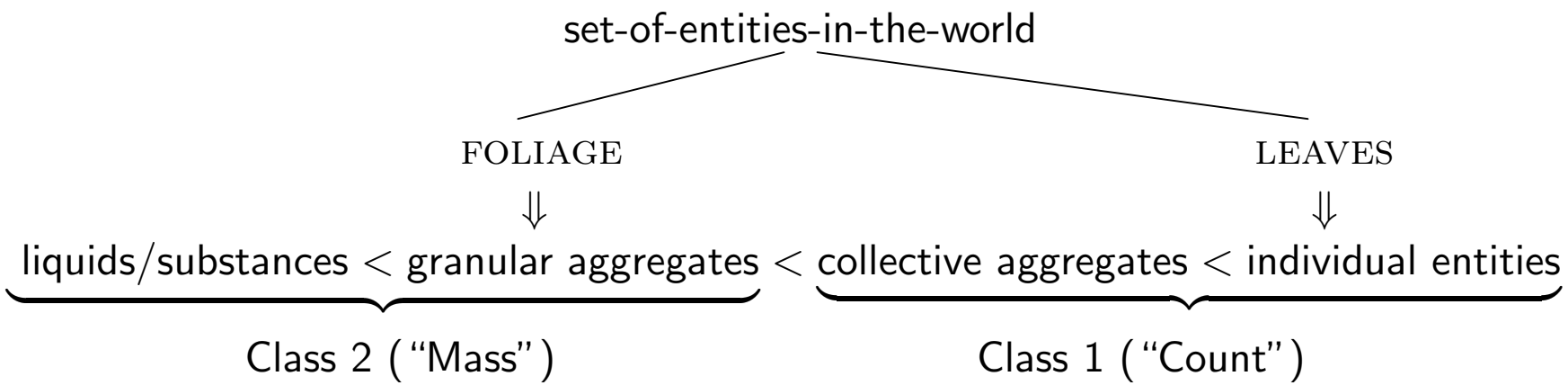
Further reflected in allowable adjective combinations:

- *dense foliage* / *?dense leaves*

# Application: Within Language Variation

A set of entities which are referentially interchangeable in certain situations may be construed differently

- corresponding to distinct individuation types
- in turn, having distinct morphological classes



# Formal Account

Desiderata for a formal account:

- be able represent different individuation types, in particular aggregates
- give an account of grammatical number morphology
- connect to the typological generalizations just made

# Formal approaches

Long tradition of using mereology (part-structures) to model the nominal domain (Quine 1960, Link 1983, Landman 1989)

Proposal:

- using mereotopology—mereology enriched with connectedness relations—allows us to better model the nominal domain

# Mereology

Mereology has been used as an alternative to set theory as it is able to represent plural and mass terms

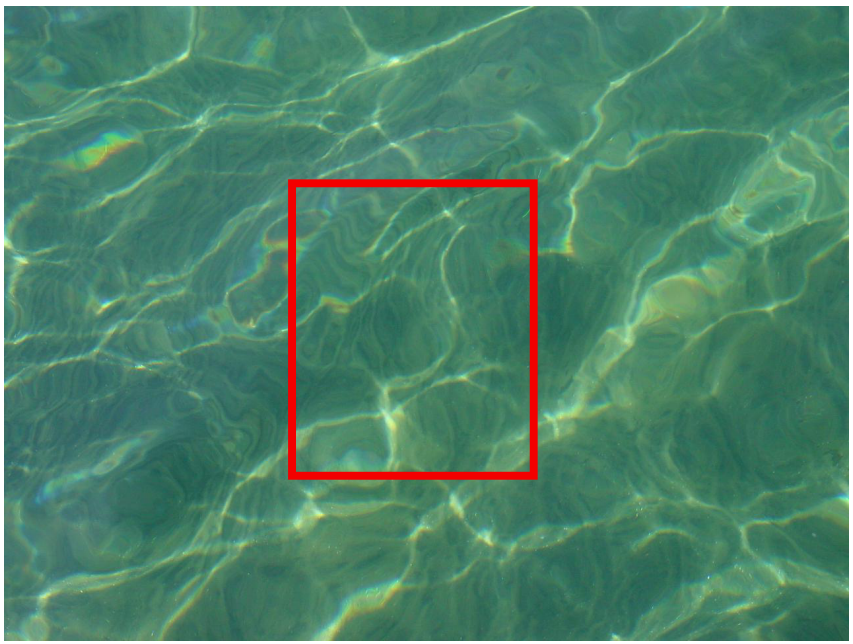
- can define useful properties such as *divisibility* or *atomicity*, which have been argued to correspond to mass and count terms, respectively

# Divisibility

$$\text{Divisive(P)} = [P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$$

Let *water* be divisive and let *water* be true of *a*

*a*

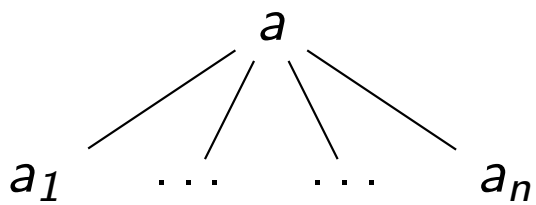


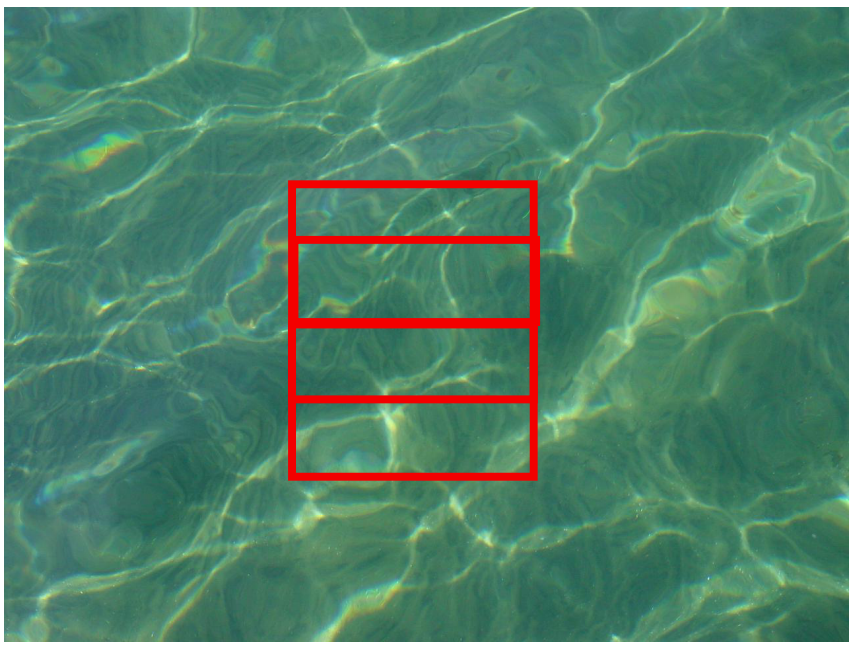
# Divisibility

$\text{Divisive(P)} = [P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$

Let *water* be divisive and let *water* be true of $a$

Then any part of $a$, say $a_1$, is also water



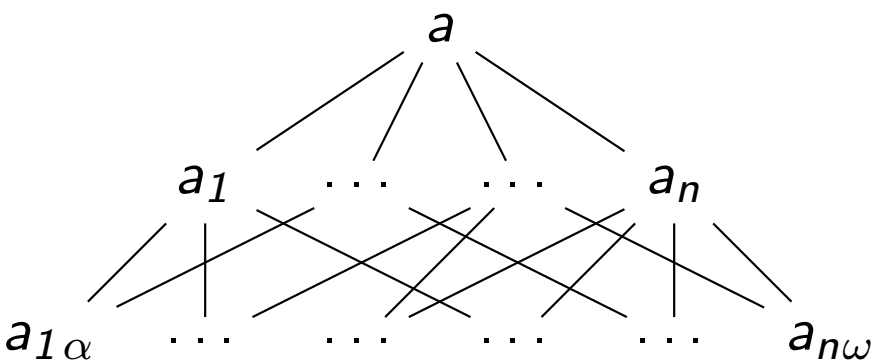
# Divisibility

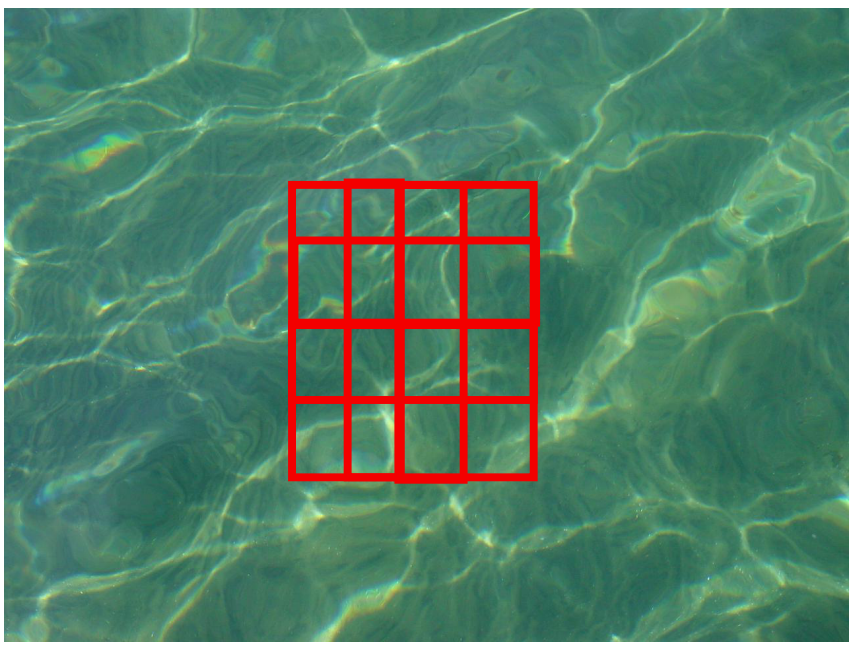
$$\text{Divisive(P)} = [P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$$

Let *water* be divisive and let *water* be true of $a$

Then any part of $a$, say $a_1$, is also water

. . . and so are the parts of the parts of $a$, such as $a_{1\alpha}$



# Divisibility

While 'divisibility' has been widely used, it has often been critiqued due to the "minimal parts" problem

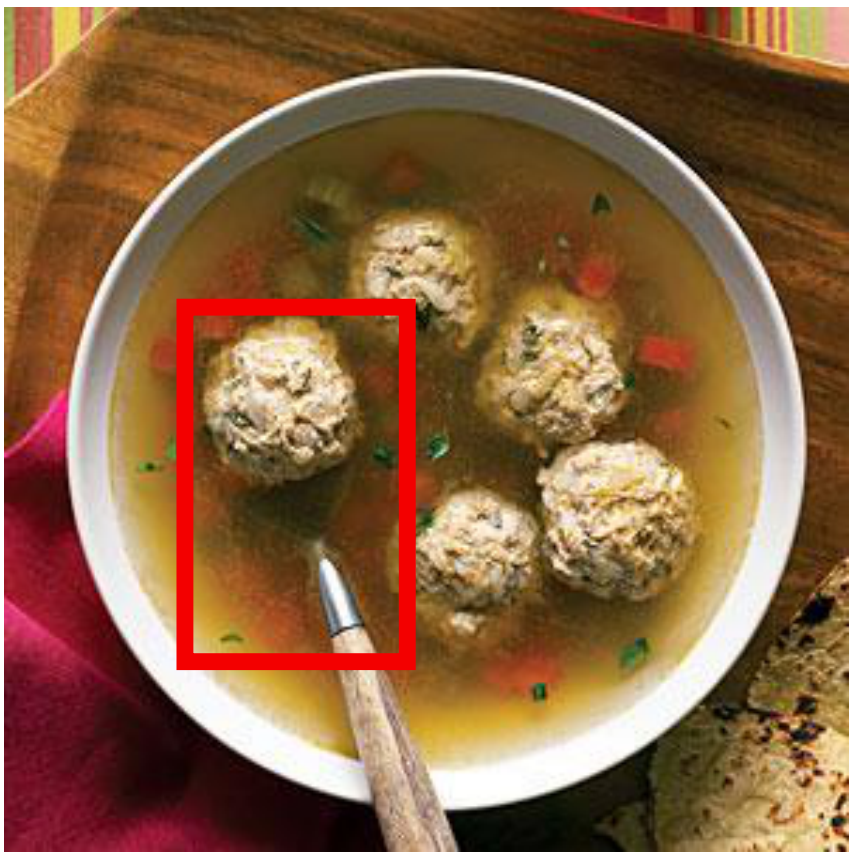
- many non-countable nouns (*furniture*) are not divisive
- even many substance nouns do not lend themselves to be infinitely divisible into the same type of stuff (Taylor 1977)

# Divisibility

$$\text{Divisive(P)} = [P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$$

Let *soup* be divisive and let *soup* be true of *a*

*a*

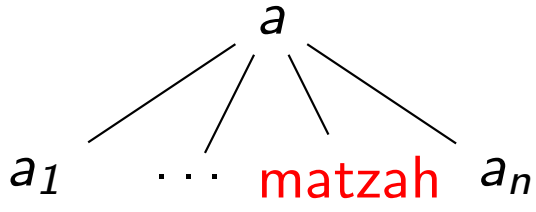


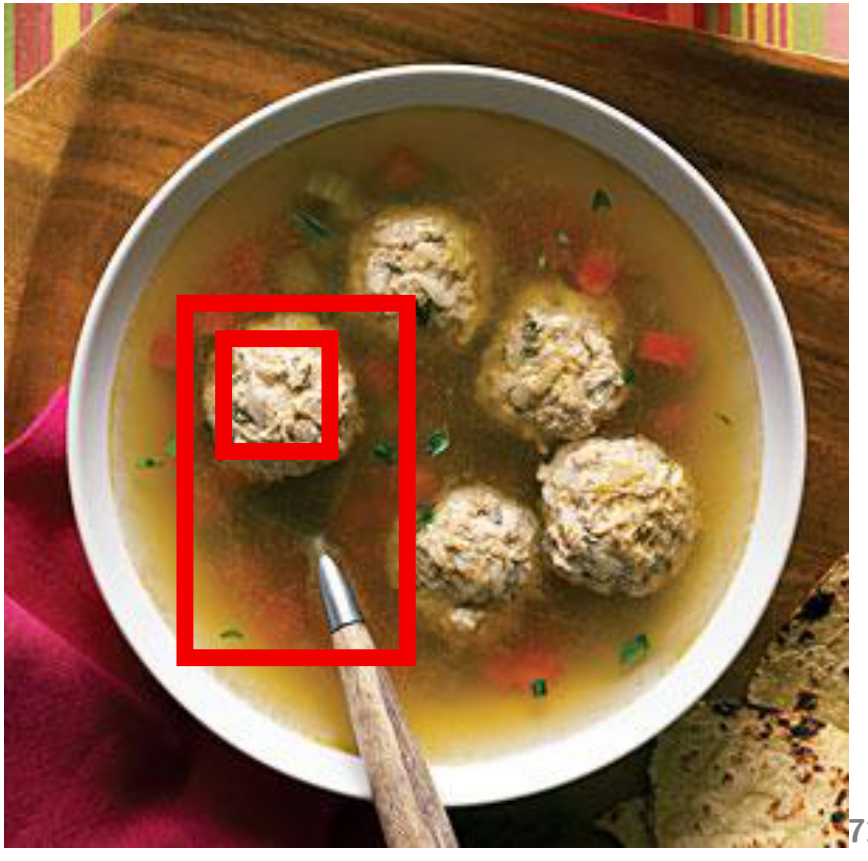
# Divisibility

$\text{Divisive(P)} = [P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$

Let *soup* be divisive and let *soup* be true of *a*

Then any part of *a*, say $a_1$, is also soup



# Divisibility

While the matzah ball is clearly a part of the soup, it does not satisfy the predicate SOUP

# Atomicity

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

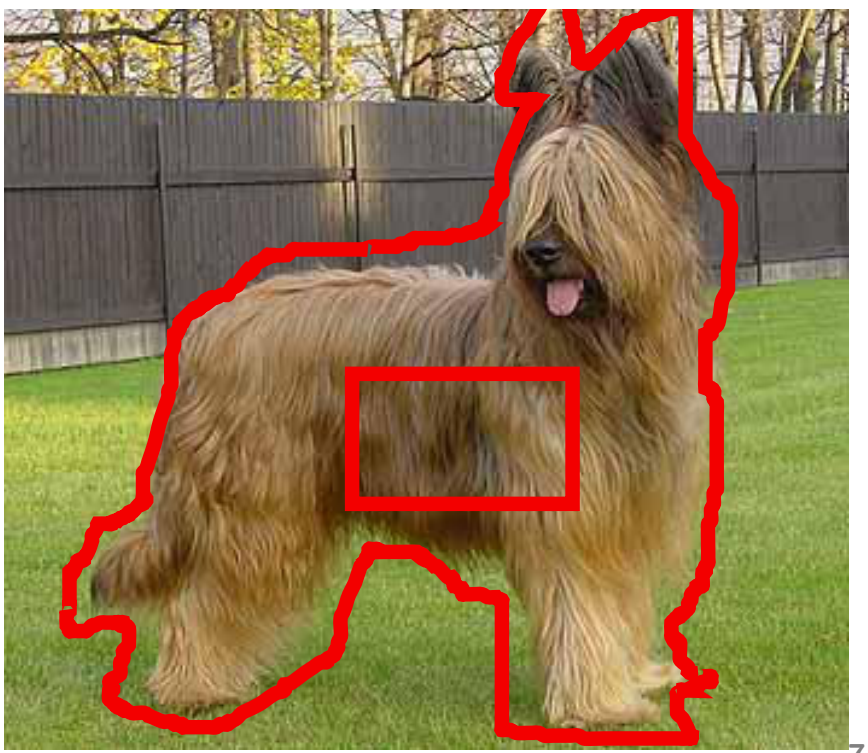
*a*

# Atomicity

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

Let *dog* be atomic and let *dog* be true of *a*

Then any part of *a*, say $a_1$, is not an instance of *dog*

# Atomicity

Classic challenge to atomicity (Zucchi and White 2001; Rothstein 2010):

- There are count predicates for which parts of the extension again (arguably) satisfy the count predicate
  - *fence*, *wall*, *sequence*, *line*

# Atomicity

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

*a*

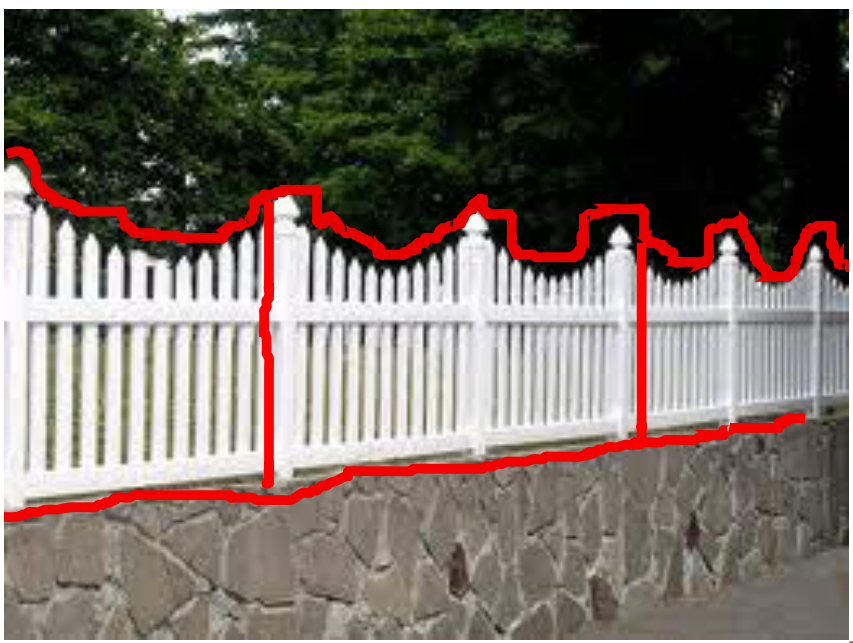
# Atomicity

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

Let *fence* be atomic and let *fence* be true of *a*

Then any part of *a*, say $a_1$, is not an instance of *fence*

$\rightarrow$ leads to the wrong result

# Further issues

In addition to the classical challenges, the data here does not sit well with analyzing countability as resulting from divisibility or atomicity:

- Aggregate nouns cannot be assimilated to being divisive or atomic predicates
- Part-structures are not sensitive to *contiguity between elements*
  - Yet Middleton et al. found *contiguity* to be predictive of whether a term is classified as count/non-count

# Extending the mereological approach

Recent work in philosophy/ontological modeling (Smith 1996, Casati and Varzi 1999, *inter alia*) has explored extending standard mereology with connectedness relations: "mereotopology"

Two fundamental motivations:

- represent the connections *between* m(ereological)-individuals
- distinguish between m-individuals forming true wholes and those forming only arbitrary collections
  - The sum of two halves of a sphere (a true whole) differs from the sum of my left shoe and the Eiffel Tower (an arbitrary collection)

# Extending the mereological approach

Basic connectedness relation $\mathrm{C}$

- holds when two m-individuals touch at least on their boundaries

Interaction with the pure mereological relations *overlap*, $\mathrm{O}$, and *part*, $\leq$:

- if two m-individuals overlap, it implies that they are connected.
- if one m-individual is part of the other, they are connected

# A mereotopological approach to countability

Connection relations come in a variety of strengths

The two primary types:

- Strongly Connected: two m-individuals are connected via overlapping

  $\mathrm{StrongC}(x, y) = \mathrm{O}(x, y)$

- Externally Connected: two m-individuals are not connected by overlapping but by touching

  $\mathrm{ExtC}(x, y) = \mathrm{C}(x, y) \wedge \neg \mathrm{O}(x, y)$

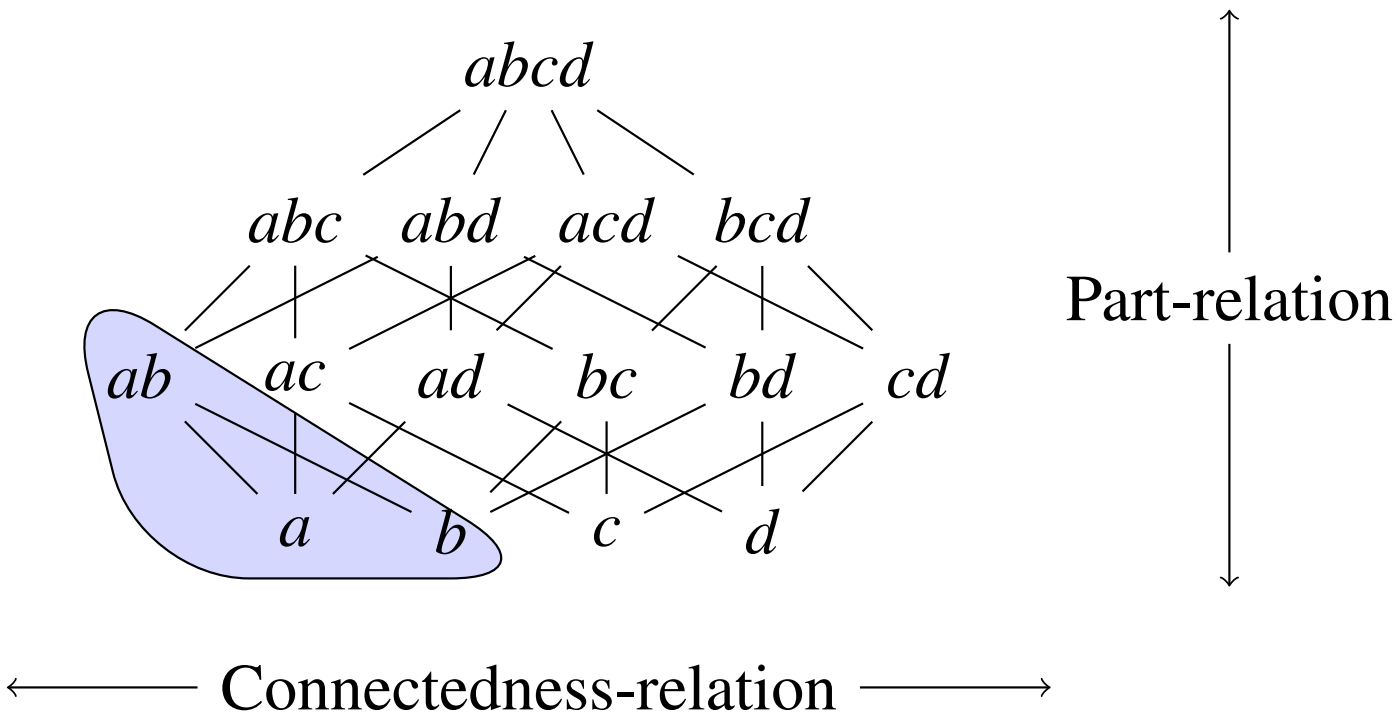
# A mereotopological approach to countability

Can still use a lattice representation, but decorate it with regions representing which elements are connected

- Suppose $a$ and $b$ are two halves of a sphere, and $c$ and $d$ are my left shoe and the Eiffel Tower
- $a + b$ is a connected sum, while the other sums are not connected

# A mereotopological approach to countability

*a* and *b* are two halves of a sphere, and *c* and *d* are my left shoe and the Eiffel Tower



# Denotation Types

Modelling four denotation types within a mereotopology:

- individual (*dog*)
- substance (*water*)
- collective aggregate (*ant*)
- granular aggregate (*rice*)

# Individuals

## Individuals as Self-Connected Wholes

Individual entities correspond to integrated wholes (Moltmann 1997), but using a stricter notion of Maximally Strongly Self-Connected (MSSC) relative to a property (Casati and Varzi 1999)

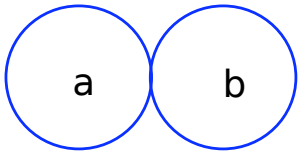
An m-individual is Maximally Strongly Self-Connected (MSSC) relative to a property if

- every (interior) part of the individual is connected to (overlaps) the whole (*Strongly Self-Connected*)
- and anything else which has the same property and overlaps it is once again part of it (*Maximal*)

# Individuals

Connectedness of interior parts rules out a sum of entities which just "touch"

- so $a + b$ below is not strongly self-connected




Maximality ensures that the biggest entity satisfying a property is taken:

- A sphere may contain many, many other m-individuals which also satisfy being a sphere, but the only one that satisfies being a MSSC individual is the largest one that contains the others
- corresponds to the cohesive objects that research on infant object perception has shown that infants expect (Spelke 1990)

# Individuals

(1) INDIVIDUAL(P)$\rightarrow \forall x[R(x, P) \rightarrow \exists y[y \leq x \wedge MSSC(y)]]$

A predicate is INDIVIDUAL iff any m-individual that realizes the predicate has a (not necessarily proper) part which is a MSSC m-individual.

For instance, the predicate *dog* will be restricted to MSSC-individuals

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

# Individuals

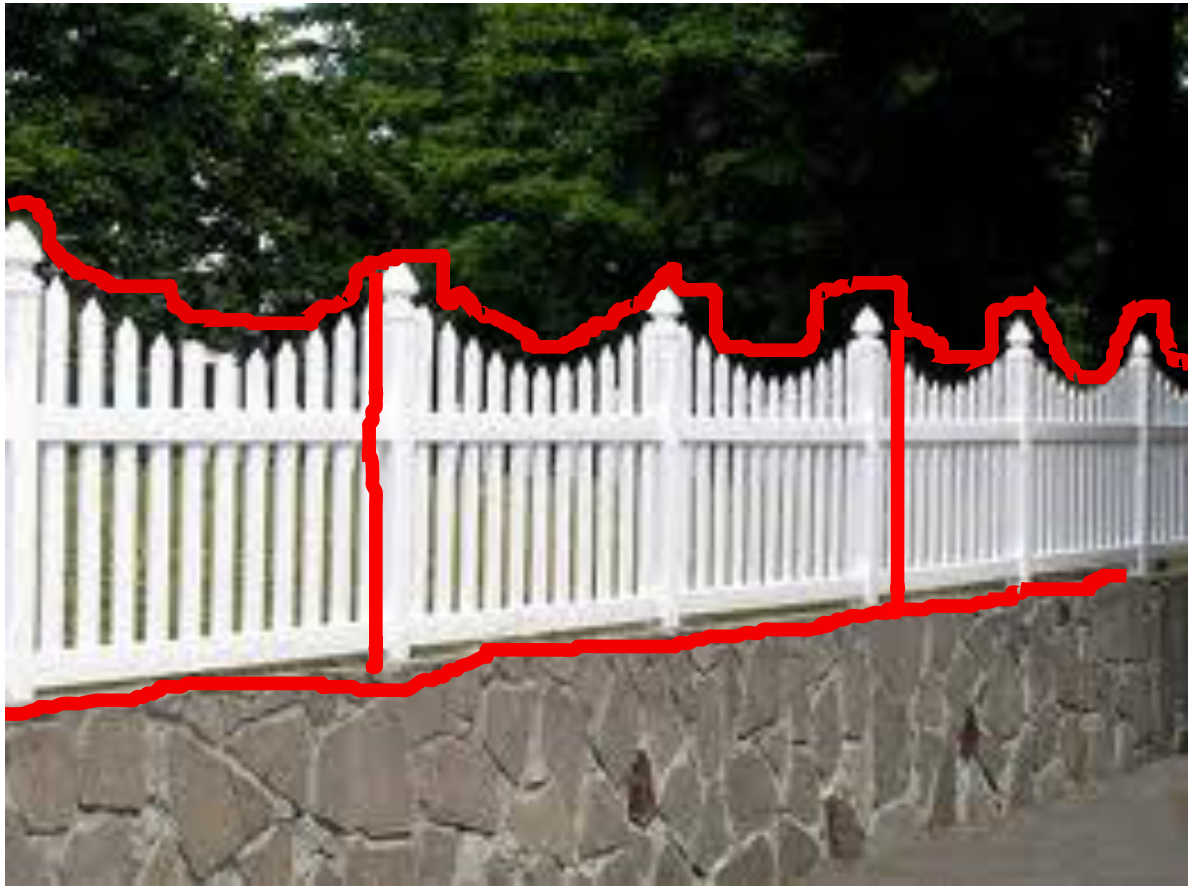
This will pick out, for example, individual whole dogs

# Individuals

**Advantage:** *fence*-predicates are not a problem because of the maximality condition

- a part of a fence may satisfy the property *fence*, but will not be a MSSC-individual
- Individual-predicates only pick out the largest m-individual

# Individuals



# Substances

Substances are characterized as entities that always come strongly connected to other entities of the same type

**Intuition:**

- If *a* is a *water* m-individual, it overlaps with a different *water* m-individual *b*
- In fact, every time one has a liquid/substance m-individual *a*, that m-individual will be strongly connected to (overlap) another m-individual of the same type

# Substances

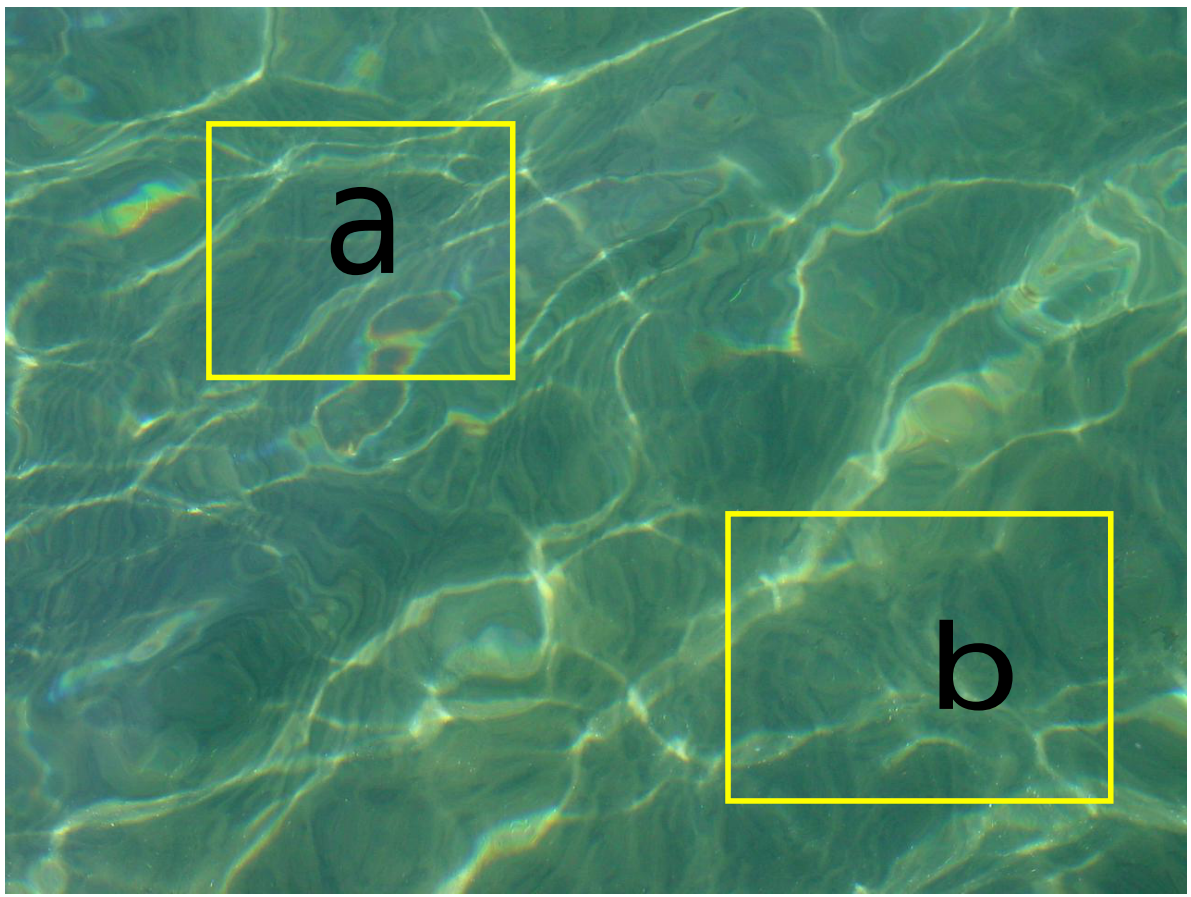
(3) $\text{SUBSTANCE(P)} \rightarrow [\forall x[R(x, P) \rightarrow \exists x'[R(x', P) \wedge x' \neq x \wedge \text{StrongC}(x, x')]]$

If P is a substance predicate then all m-individuals that realize P are strongly connected to a distinct m-individual of the same substance.

A portion of water can be thought of as covered by many overlapping sub-portions of water

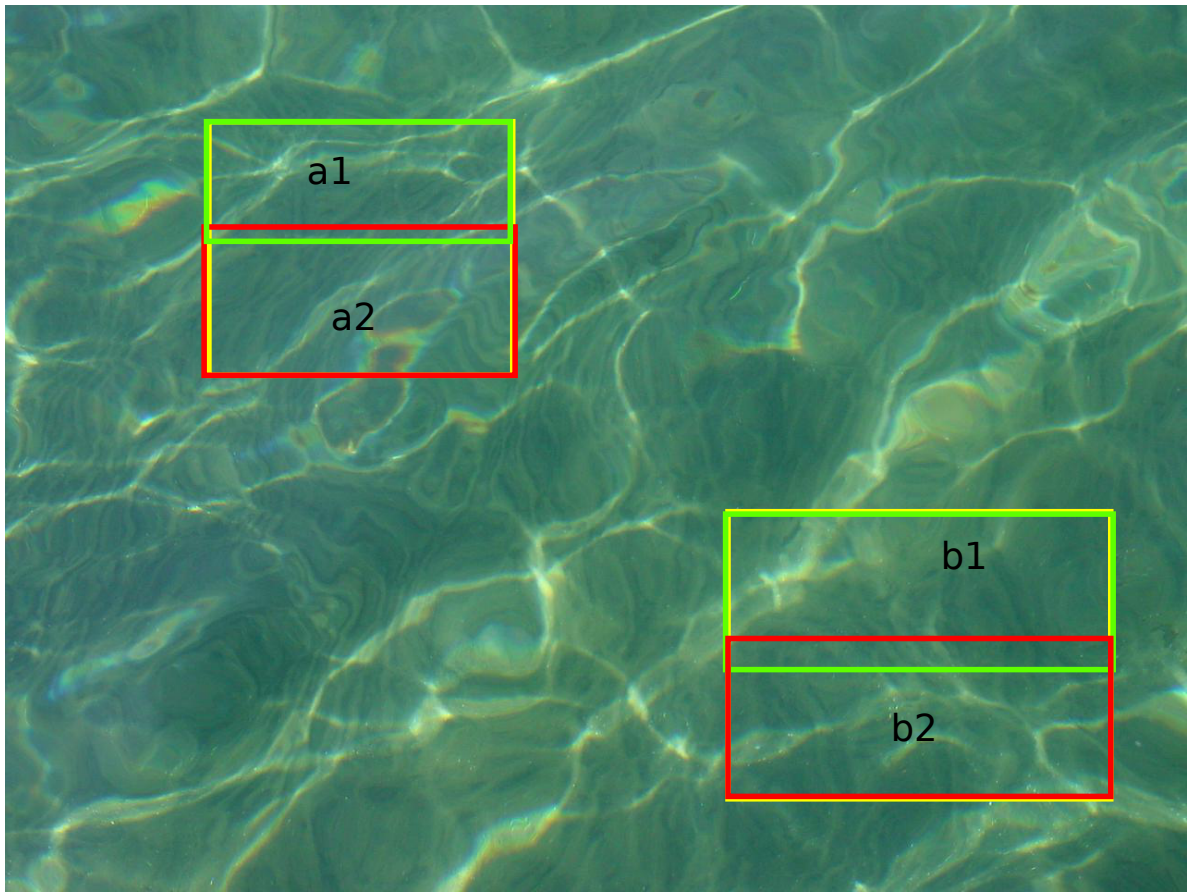
# Substances

Suppose an instance of water was the sum of two portions of water $a + b$



# Substances

Then the condition on substances says we can find distinct m-individuals which are strongly connected—any of the overlapping parts will do: $a_1, a_2, b_1, b_2$



# Substances

**Advantage:** *soup*-predicates are not a problem since there is no infinite part-taking

# Aggregates

Aggregates, such as *sand* or *rice*, are a hybrid of the first two categories:

- have minimal parts
- designate groups which come together in some manner

*sand* is true of single grains (MSSC individuals), clumps of sand or combinations of the two

# Aggregates

Can use connectedness relations to define clusters of m-individuals

- A clustered m-individual will correspond to the sum of a set of connected entities
- A clump of sand consists of a set of sand individuals which are (indirectly) connected via touching

# Aggregates

Need a notion of *transitive connection* relative to a property and a connection relation:

TransitiveC(x,y,P,C,Z) $=_{def} \forall z \in Z[R(z, P) \wedge (x = z_1 \wedge y = z_n) \wedge Cz_1z_2 \wedge Cz_2z_3 \ \ldots \ \wedge Cz_{n-1}z_n]$
where $Z = \{z_1, z_2, \ldots, z_n\}$ (Transitively Connected)

($x$ and $y$ are transitively connected relative to a property $P$, a connection relation $C$, and a set of entities $Z$, when all members of $Z$ satisfy $P$ and $x$ and $y$ are connected through the sequence of $z_i$s in Z.)

# Aggregates

Cluster(x,P,C)$=_{def}$
$\exists Z[x = \bigoplus Z \wedge \forall z, z' \in Z[TransitiveC(z, z', P, C)]]$
(Clustered Individual)

($x$ is a cluster relative to a property $P$ and a connection relation $C$ iff $x$ is a sum entities falling under the same property which are all transitively connected relative to the same property and connection relation.)

# Aggregates

Varying the connectedness relations gives more tightly or more loosely connected clusters:

$\text{CLUSTER}_{ExtC}$ represents clusters which are related via external connectedness (granular aggregates: *rice*, *sand*)

To represent colletive aggregates such as insects, need a looser form of connection:

- Proximately Connected: two m-individuals are proximately connected if they are with in a distance $d$ of one another

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

$\text{CLUSTER}_{PC}$ represents clusters which are related via proximate connectedness (collective aggregates: *ants*, *berries*)

# Return to the collective/unit systems

Armed with these new distinctions, can appropriately treat nouns in collective classes and singulative morphemes

- Collective nouns in Welsh have their denotational space limited to clustered individuals:

(4) $[\![\textit{cacwn}]\!] :=$
$\lambda x_O[\mathrm{R}(x_O, \mathrm{Hornet}) \wedge x_O \in \mathrm{CLUSTER}_{PC}]$

- The singulative morpheme in Welsh is treated as an operator which, presupposing a clustered individual, returns MSSC individuals:

(5) $[\![-\textit{en}/-\textit{yn}]\!] := \lambda Q \lambda x.Q_{\textit{cluster}}[x \leq Q \wedge x \in \mathrm{MSSC}]$

Analogous to the standard analysis of the English plural morpheme (Link 1983)

# Mereotopological approach

Relation between degrees of connectedness and degrees of individuation:

- ordering the denotation types by the strength of the connection among elements corresponds to ordering of the scale of individuation

| | | | | | | |
|---|---|---|---|---|---|---|
| strongly connected | $<$ | externally connected | $<$ | proximately connected | $<$ | MSSC-whole |
| liquids | $<$ | granular aggregates | $<$ | collective aggregates | $<$ | individuals |

# Conclusion

Examining rich grammatical number systems such as that of Dagaare is not only intrinsically interesting, but helps to make generalizations about which types of entities are relevant for countability across languages

The larger typological picture sketched here recognizes four levels:

- things-in-the-world
- lexical nouns
- individuation types
- morphosyntactic classes

# Conclusion

Understanding the relation between the different levels:

- provides some answers to empirical challenges to the mass/count distinction
- makes predictions for the cross-linguistic diversity of mass/count-related morphosyntactic distinctions

Part-structures can be extended with connectedness relations to represent the semantic distinctions observed cross-linguistically

Much more to do!

# Thanks!

Thanks!