Extension, ontological type, and morphosyntactic class: Three ingredients of countability

David Clausen, Alex Djalali, Scott Grimm, Sven Lauer, Tania Rojas-Esponda and Beth Levin

Stanford University

September 22, 2010

Revised: December 1, 2010

Current approaches

Mass or *count* is a property of *nouns*, not *extensions*.

- **Inherency** ⇒ Choice is *predetermined* by the nature of the entity named
- **Arbitrariness** ⇒ Choice is *not-predetermined*, though there may be some regularities or tendencies in lexicalization as mass or count

Doublets as support for arbitrariness

"In fact, the same slice of reality can be classified as either count or as mass, as attested by the existence of near synonyms" (Chierchia 1998: 56)

Such **doublets** include:

- *foliage* is mass, but *leaves* is count
- *mail* is mass, but *letters* is count
- *luggage* is mass, but *suitcases* is count
- *change* is mass, but *coins* is count

Doublets as support for arbitrariness

The arbitrariness position: A noun's status is not *predetermined* by the nature of the entity named, though there may be some regularities or tendencies in lexicalization as mass or count.

"Nor can I see anything that would explain the count/mass difference between 'footwear' and 'shoe', 'clothing' and 'clothes', 'shit' and 'turd', or 'fuzz' and 'cop'. These are normally mass nouns and count nouns for basically the same thing."
(Ware 1979: 22)

Doublets as support for inherency

Wierzbicka (1985) claims *conceptual* and *cultural factors* influence a noun's classification as mass or count.

- *Mode of interaction* with the relevant entity
- *Distinguishability* of any constituent element, which is influenced by their *size* and *contiguity*

Doublets as support for inherency

Zwicky (1997) points out that plants that cover areas of ground in a garden are usually mass nouns, as *ice plant* is.

Yet *petunias*, which can be used as a ground cover, is count.

The reason is distinguishability.

Petunias

[Picture of petunias as ground cover]

Figure: Like a typical ground cover, it seems difficult to distinguish individual plants.

Petunias

[Picture of a petunia plant in a pot]

Figure: But petunias are actually easily divisible into individual plants

Ice plant

[Picture of ice plant]

Figure: Not only are individual plants difficult to distinguish, but it is difficult to actually separate them out.

Universal grinder

- **Universal grinder** ⇒ Every count noun, given the right context, can have a mass interpretation

(1) There is dog all over the highway.

Universal packager

- **Universal packager** ⇒ The 'inverse' operation, which results in count interpretations for typically mass nouns

(2) Three beers please. [= three servings of beer]

Universal grinder and packager

Universal grinder and packager data are often taken as evidence that a noun's status is not tied to the lexical item itself but is necessarily computed at the NP level (Allan 1980, Bunt 1985).

Non-universality of universal grinder/packager

If the effects of the grinder and packager were truly universal, they should apply uniformly across all nouns, but these operations are *restricted*.

Non-universality of universal grinder

- Grinding is restricted. In particular, it is difficult to grind highly individual objects, especially artifacts (Chierchia 2010).

(3) There is dog all over the highway.

(4) #There is mug/toaster on the table.

(5) #Would you care for some more pea? (Fillmore 1989: 49)

Non-universality of universal packager

- Packaging is largely restricted to those nouns whose referents are already associated with conventionalized units of packaging

(6) Three beers please. [= three servings of beer]

(7) #Rices adorn the altar.

Experiment: Acceptability Judgements

Question: To what extent is the Universal grinder truly *universal*?

Prediction: Grinding might be differentially available depending on the nature of the noun involved

Methods

Test the extent to which native speakers of English judge sentences involving an application of the Universal Grinder *acceptable* using a 1 (*unacceptable*) – 7 (*acceptable*) value *Lickert scale*.

Materials

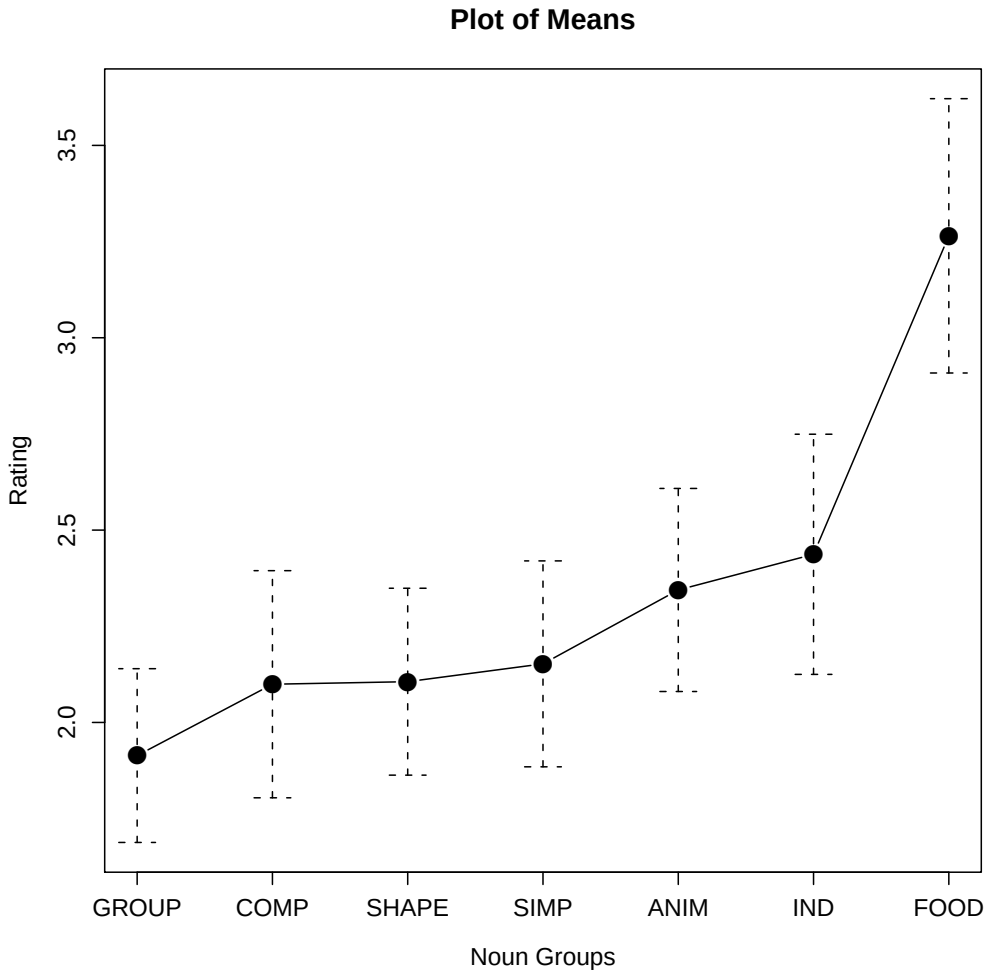
7 noun types; 5 tokens of each

- **shape**: *tube, cylinder, sphere, cone, cube*
- **group terms**: *forest, bouquet, fleet, swarm, committee*
- **members of group terms**: *tree, flower, ship, bee, person*
- **simplex artifacts**: *hammer, towel, shirt, bucket, pencil*
- **complex artifacts**: *toaster, car, computer, violin, forklift*
- **animals**: *squirrel, snake, robin, butterfly, pig*
- **food stuff**: *steak, apple, cracker, yam, pea*

Results

- Target sentences were given *low* acceptability ratings. On average, **2.3** out of 7 (SD 1.81)
- Filler sentences were given **5.7** out of 7 (SD 1.85)

Results


Plot of Means
Rating
2.0
2.5
3.0
3.5
GROUP
COMP
SHAPE
SIMP
ANIM
IND
FOOD
Noun Groups

Results (low to high acceptability)

$$\text{group terms} < \left\{ \begin{array}{l} \text{complex artifacts} \\ \text{shape} \\ \text{simplex artifacts} \\ \text{members of group terms} \end{array} \right\} < \text{animals} < \text{foodstuff}$$

Our proposal

- The members of the pairs used to argue that these phenomena illustrate arbitrariness in the semantics-syntax mapping at best represent near-synonyms.
- Even if the members may sometimes overlap in their extensions, they differ in meaning or, more precisely, in the construal of entities or events in the world they lexicalize.
- In each pair, the difference in meaning is critical to a difference in behavior.
- Thus, these grammatical phenomena do not illustrate arbitrariness in the semantics-syntax mapping.

Our proposal

There are three distinct levels at work, with two mapping relations:

$$[\![\text{entity}]\!]$$
$$\Downarrow$$
Ontological Type/Conceptual Level
$$\Downarrow$$
Morphosyntactic Class

For the conceptual level something intended akin to Bierwisch (1983), Lexical-Conceptual Structure (Rappaport and Levin 1988)

Our proposal

The rest of the talk will address:

- the mapping between entities in the world and ontological types
  - this portion will be motivated by the evidence that was just reviewed (doublets)
- the mapping between ontological types and morphosyntax
  - this portion will be motivated by cross-linguistic evidence

Re-Examining the Doublets

The arbitrariness said to be manifested by mass/count noun doublets is largely apparent.

Such claims are based on a consideration of **extensions**.

Considering such doublets purely extensionally breaks down:

- the members name distinct perspectives on the relevant entities;
- however, this is not always evident from an examination of their extensions;
- the different countability status of the member nouns arises from distinct conceptualizations/perspectives.

Mail vs. Letters

Mail: the set of objects that one receives via the post;

- may include letters, but also magazines, packages, postcards, and the like.

Letters: a far narrower class of entities, that need not actually have been mailed.

The key point: Not all letters are mail, nor is all mail letters.

Luggage vs. Suitcases

Luggage: the ensemble of items that one is travelling with;

- may include suitcases, hat boxes, duffle bags, make-up bags, and more.

Suitcases: the most prototypical and frequent form of luggage, though a suitcase could be used for storage rather than travel.

Change vs. coins

Change: the leftover money received after a sale;

- may include (but is not limited to) coins.

Coins: a narrower class of entities that need not have been received after a particular financial transaction.

Foliage vs. Leaves

When observing a tree, a speaker may freely choose to talk about its *leaves* or its *foliage*.

When discussing leaves raked into a pile, *leaves* is appropriate, whereas *foliage* is not.

Distinguishing the Doublets

In a doublet what sets the member with mass morphosyntax apart from the member with count morphosyntax?

The doublets involve nouns naming for sets of items.

Two factors favoring mass morphosyntax for such nouns:

- Functional similarity of set members
- Contiguity/connectedness of set members

Factor: Functional Similarity

Some nouns name sets of entities that

- participate together in an event:
  - *mail* names a set of entities that travel through the postal system
  - in fact, these nouns are often deverbal
- arise together as a result of an event:
  - *change* is a result of a monetary transaction

Factor: Functional similarity

- These nouns name sets whose members are identical with respect to their role in an associated event.
- Functional similarity can be seen as an analogue of the more familiar similarity among elements of granular aggregates, such as *gravel*, *rice*, *salt*.

  Compare the common need for unitizers:
  grain of rice/sand and *piece of luggage/mail*

Factor: Contiguity/Connectedness

Some nouns provide a holistic perspective on a co-occurring, contiguous and normally connected aggregate of things.

- *foliage* (compare *leaves*): the collectivity and the interconnectedness of leaves with one another rather than individual leaves
- *plumage* (compare *feathers*): the ensemble of feathers on a bird, but not the contents of a down pillow, which may be referred to as *feathers*.

Further reflected in allowable adjective combinations:

- *dense foliage* / *?dense leaves*

Beyond Extensionality

The doublets show that a noun conveys more than its extension.

Noun meanings may encompass:

- spatial and temporal contiguity and connectedness (*foliage*, *plumage*)
- similarity of form (*foliage*, *plumage*) and function (*mail*, *luggage*) of the members of the aggregate

These observations motivate the mapping:

$$[\![\text{entity}]\!]$$
$$\Downarrow$$
Ontological Type/Conceptual Level

Why do such doublets arise?

Meanings are construals of the world, so that even if in some instances *leaves* and *foliage* might have the same extension — that is, refer to the same entity — the basis for the synonymy claim — the two words lexicalize different perspectives on this entity.

In fact, this is precisely the key claim in Wierzbicka's well-known study of the mass/count distinction (1985): *conceptual* and *cultural factors* influence a noun's classification as mass or count:

- *Mode of interaction* with the relevant entity.
- *Distinguishability* of any constituent element, which is influenced by its *size* and *contiguity*.

The lesson from mass/count doublets

- Doublets are significant *not* because they illustrate supposed arbitrariness, but because they demonstrate the availability of multiple perspectives on certain entities in the world.
- Precisely those entities that are open to the appropriate multiple perspectives may show both mass and count names:
  - when these perspectives align with the factors that contribute to mass vs. count status.

(See Middleton et al. 2004, Wierzbicka 1985)

A caution

Functional aggregates, then, demonstrate there is more systematicity in mass/count classification than has sometimes been claimed.

Nevertheless, a residue of arbitrariness in the classification of nouns as mass or count is likely.

A caution

A prediction: Some of this residue should arise precisely where the criteria for mass/count classification do not make clear cuts.

Support:

- Wierzbicka notes that size and distinguishability play a part in mass/count classification.
- The differential status of *rice* and *lentils* might follow because the relevant unit size is on the boundary between what qualifies as mass vs. count (Cruse 2004).

Countability and Morphosyntax

We now explore the second half of the mapping:

Ontological Type/Conceptual Level

$\Downarrow$

Morphosyntactic Class

Need to determine which ontological 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 markedness with respect to countability

Cross-Linguistic Differences

We argue associations with different morphosyntactic classes are

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

Examine three languages:

- 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 marking
  - 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 Markedness

Class 1 has a markedness distinction:

- the singular interpretation has the unmarked form
- the plural interpretation has a marked form

English: Morphosyntactic Markedness

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

Welsh: Morphosyntactic Classes

Welsh has a three-way split (Stolz 2001):

- Class 1: Nouns allowing plural marking
  - includes primarily animates and other individuals
- Class 2: Nouns allowing singulative marking
  - includes granular aggregates (*turf*, *sand*) as well as collective aggregates such as small animals and insects, vegetables/grains/fruits, inherently plural body parts (*ribs*) (cf. Acquaviva's 2008 'inherent plurals')
- Class 3: Nouns having one form
  - includes liquids and substances

Welsh: Morphosyntactic Markedness

Class 1 and Class 2 differ in the direction of markedness:

- for Class 1 (singular/plural), the singular is morphologically unmarked
- for Class 2 (collective/singulative), the aggregate is unmarked and the singulative is marked

Welsh: Morphosyntactic Markedness

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 (Gur;Niger-Congo) has a four-way split (Grimm 2009):

- Class 1: Nouns with plural marked
  - individuals (*child, dog*)
- Class 2: Nouns with singular marked
  - 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

Dagaare: Morphosyntactic Markedness

The classes differ in the direction of markedness:

- Class 1: the singular is morphologically unmarked
- Class 2 and 3: the aggregate is unmarked and the singular/singulative is marked

Mapping the Terrain

Ordering classes from those most unmarked in the plural to those most unmarked in the singular imposes an order on the ontological 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 ontological types form a scale (Grimm 2010)

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

Individuation

Does a semantic property organize the scale?

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

Individuation

Individuation serves as a cover term for these factors which characterize the propensity for an entity to appear as an individual unit.

- ease of distinguishability of elements
- size of elements
- spatial and/or temporal contiguity among elements
- canonical mode of interaction

Identified in a variety of work in philosophy, linguistics and cognitive science (see Quine 1960, Mufwene 1980, Wierzbicka 1985, Bloom 1994, Zwicky 1997, Middleton et al. 2004)

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 small and not easily distinguishable; one does not canonically interact with them

Collective aggregates represent an intermediate category:

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

Relating the scale to the morphosyntax

For a given language, entities are realized in the manner that their location on the individuation scale is realized in that language.

A particular ontological type can be assigned

- a unique morphosyntactic class (Dagaare liquids)
- or the same class as the type to its left, right, or both

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*)	

Relating the scale to the morphosyntax

The morphosyntactic classes respect the structure of the scale:

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

Entities of a given ontological 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*)	

Mapping between extensions, ontological types and morphosyntactic classes

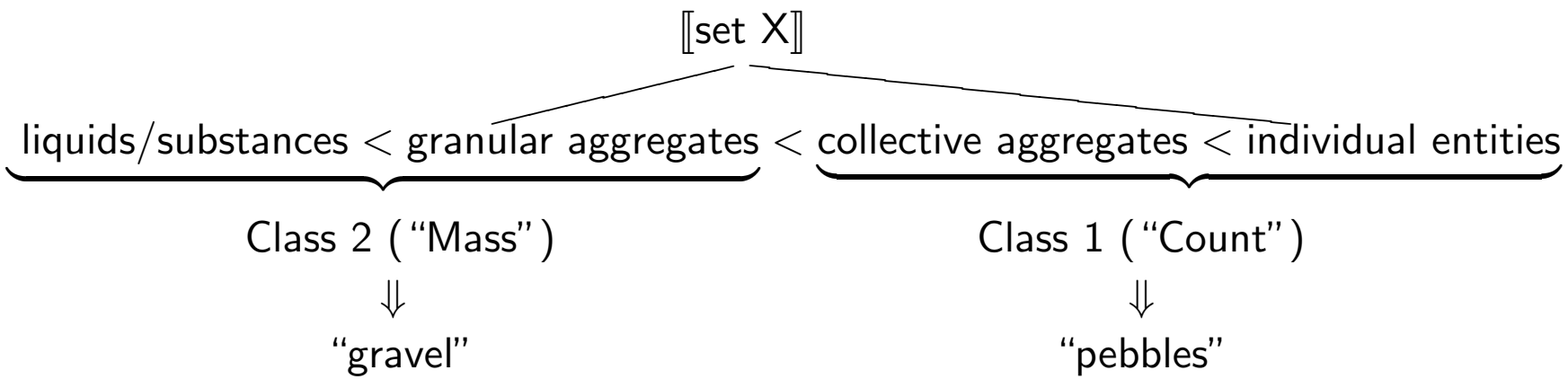
The mapping between extensions, ontological types and morphosyntactic classes in a given language conforms to a picture as below:

$$\underbrace{\overset{[\![\text{entity1}]\!]}{\text{ont. type 1}} < \text{ont. type 2}}_{\text{Morphosyntactic Class 1}} < \underbrace{\text{ont. type 3} < \text{ont. type 4} < \overset{[\![\text{entity 2}]\!]}{\text{ont. type 5}}}_{\text{Morphosyntactic Class 2}}$$

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



Application: Within Language Variation

Grinding can be seen as a function resulting in a shift in individuation type

$$\underbrace{\underset{\llbracket \text{GRIND(apple)} \rrbracket}{\text{liquids/substances} < \text{granular aggregates}}}_{\text{“Mass”}} < \underbrace{\underset{\llbracket \text{apple} \rrbracket}{\text{collective aggregates} < \text{individual objects}}}_{\text{“Count”}}$$

Application: Between Language Variation

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

Language 1:

$$\underbrace{[\![\text{entity X}]\!] \mid \text{ont. type 1} < \text{ont. type 2}}_{\text{Morphosyntactic Class 1}} < \underbrace{\text{ont. type 3}}_{\text{Morphosyntactic Class 2}}$$

Language 2:

$$\underbrace{\text{ont. type 1}}_{\text{Morphosyntactic Class 1}} < \underbrace{[\![\text{entity X}]\!] \mid \text{ont. type 2} < \text{ont. type 3}}_{\text{Morphosyntactic Class 2}}$$

Conclusion

The larger picture sketched here recognizes three levels:

- extensional
- conceptual
- morphosyntactic

The picture helps make sense of

- empirical challenges to the mass/count distinction
- the cross-linguistic diversity of mass/count-related morphosyntactic distinctions

Provide avenues for further investigation of the empirical phenomena

Thank you

Thank you!

We would like to thank a previous audience at Stanford University for their helpful feedback.

Selected Bibliography

Acquaviva, P. (2008) *Lexical Plurals: A Morphosemantic Approach*, Oxford University Press, Oxford.

Allan, K. (1980) "Nouns and Countability", *Language* 56, 541-567.

Bierwisch, M. (1983) "Semantische und konzeptuelle Representation lexikalischer Einheiten", W. Motsch and R. Ruzicka, eds., *Untersuchungen zur Semantik*, Akademie-Verlag, Berlin, 61-99.

Bloom, P. (1994) "Possible Names: The Role of Syntax-Semantics Mappings in the Acquisition of Nominals", *Lingua* 92, 297-329.

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

Chierchia, G. (1998) "Plurality of Mass Nouns and the Notion of "Semantic Parameter", in S. Rothstein, ed., *Events and Grammar*, Kluwer, Dordrecht, 53-103.

Selected Bibliography

Chierchia, G. (2010) "Mass Nouns, Vagueness, and Semantic Variation", *Synthese* 174, 99-149. Fillmore, C.J. (1989) "Grammatical Construction Theory and the Familiar Dichotomies", in R. Dietrich and C.F. Graumann, eds., *Language Processing in Social Context*, North-Holland, Amsterdam, 17-38.

Grimm, S. (2009) "Number Marking and Individuation: A View from Dagaare", unpublished ms., Stanford University.

Grimm, S. (2010) "Number and Individuation", dissertation proposal, Stanford University.

Selected Bibliography

Middleton, E.L., E.J. Wisniewski, K.A. Trindel, and M. Imai (2004) "Separating the Chaff from the Oats: Evidence for a Conceptual Distinction between Count Noun and Mass Noun Aggregates", *Journal of Memory and Language* 50, 371-394.

Pelletier, F.J. (1979) "Non-Singular Reference: Some Preliminaries", in F.J. Pelletier, ed., *Mass Terms*, Reidel, Dordrecht, 1-14.

Rappaport, M. and B. Levin (1988) "What to Do with Theta-Roles", in W. Wilkins, ed., Syntax and Semantics 21: Thematic Relations, Academic Press, New York, 7-36.

Soja, N.N., S. Carey, and E.S. Spelke (1991) "Ontological Categories Guide Young Children's Inductions of Word Meaning: Object Terms and Substance Terms", *Cognition* 38, 179-211.

Selected Bibliography

Stolz, T. (2001) "Singulative-Collective: Natural Morphology and Stable Classes in Welsh Number Inflexion on Nouns", *Sprachtypologie und Universalienforschung* 54, 52-76.

Ware, R.X. (1979) "Some Bits and Pieces", in F.J. Pelletier, ed., *Mass Terms*, Reidel, Dordrecht, 15-29.

Wierzbicka, A. (1988) "Oats and Wheat: The Fallacy of Arbitrariness", in A. Wierzbicka, *The Semantics of Grammar*, John Benjamins, Amsterdam, 499-560.

Zwicky, A.M. (1997) "Count versus Mass in English: How to Talk about Plants", unpublished ms., Stanford University and Ohio State University.