Synonymy and Arbitrariness in Linguistic Argumentation

Beth Levin

Stanford University

Revisiting Synonymy

Helsinki, October 2010

Setting the stage

The mass/count distinction revisited

The unaccusative/unergative distinction revisited

The causative alternation revisited

Overview

- A question posed in the Call for Papers:

  "What does synonymy (at any level/in any form) reveal about language?"

- A recurring answer:

  Synonymy — or even near-synonymy — can be used to argue for arbitrariness in the semantics-syntax mapping.

- My goal:

  To show this argument is often flawed: many examples previously cited in support of arbitrariness actually support systematicity in the semantics-syntax mapping.

Outline

Setting the stage

The mass/count distinction revisited

The unaccusative/unergative distinction revisited

The causative alternation revisited

Examples of purported arbitrariness

- The mass count distinction: the domain of entities
- The causative alternation: the domain of events
- The unaccusative/unergative distinction: the domain of events

An example of purported arbitrariness

The domain of entities: The mass/count distinction
(e.g., Chierchia 1998, Rothstein 2010, Ware 1979)

"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)

An example of purported arbitrariness

In English, certain entities in the world can be referred to by two nouns, one mass and the other count.

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

An example of purported 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)

An example of purported arbitrariness

An extension of the argument for arbitrariness:

Just as **near-synonyms** within a language have been appealed to in arguments for arbitrariness, so too have **translation equivalents** across languages.

- English *hair* is typically mass, but French *cheveux* is count.

A second example of purported arbitrariness

The domain of events: The causative alternation

- Verbs with related transitive and intransitive uses:

  The plate **broke**.
  Tracy **broke** the plate.
  (i.e. Tracy caused the plate to break.)

- Claimed to be a hallmark of "change" verbs (Smith 1970)

  Examples: *break*, *close*, *cool*, *dry*, *open*, *shatter*, ...

A second example of purported arbitrariness

How to test the arbitrariness hypothesis:

- "collect a fairly large number of verbs that are members of the change class, and then for each member find a verb that is semantically related, i.e., a verb that is a (near-)synonym or a (near-)antonym." (Farsi 1974: 21)
- Check whether (near-)synonyms show the alternation.
- If they do not, there is evidence for arbitrariness.

A second example of purported arbitrariness

Confirmation for the argument for arbitrariness:

- *shake* shows the alternation, but *shudder* is only intransitive.

  I shook the tree./The tree shook.
  I shuddered./*The news shuddered me.

- *heal* shows the alternation, but *cure* is only transitive.

  The wound healed./The doctor healed the wound.
  *The patient cured./The doctor cured the patient.

(Farsi 1974: 22)

Conclusion: "the capacity of a verb to be used both transitively and intransitively is a matter of lexical idiosyncrasy" (Farsi 1974: 21)

A third example of purported arbitrariness

The domain of events: The unaccusative/unergative distinction

Proponents of arbitrariness have appealed to translation equivalents to argue that unaccusativity is not semantically determined (e.g., Rosen 1984).

- English *blush* and its Dutch translation equivalent *blozen* are unergative.
- Their Italian counterpart *arrossire* is unaccusative.

The argument against arbitrariness encapsulated

- 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.

The argument against arbitrariness encapsulated

There are three distinct linguistic levels at work:

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

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

The mass/count distinction revisited

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 members of 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 adjectival modification:

- *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).

The unaccusative/unergative distinction revisited

The challenge: 'blushing' across languages.

Italian *arrossire* and Dutch *blozen* are translation equivalents, but:

They are said to differ in their classification:

- *arrossire* is said to be unaccusative:
  it takes the auxiliary *essere* 'be'.
- *blozen* is said to be unergative:
  it takes the auxiliary *hebben* 'have'.

'Blushing' across languages

- Italian conceptualizes 'blushing' as a change of state:

  arrossire is *a–* + *rosso* + *–ire*, literally 'become red'.

- Dutch conceptualizes 'blushing' as a process.

  Evidence from temporal modifiers:

  - J heeft een uur lang gebloosd
    'J has one hour long blushed'
  - *J heeft in een uur gebloosd
    'J has in one hour blushed' (McClure 1990: 314, Table 4)

Why do *blozen* and *arrossire* differ?

- Blushing involves an internally caused involuntary process, accompanied by reddening of the face.
- *blozen* lexicalizes this process.
- *arrossire* lexicalizes the change of state which happens to accompany it, and is conventionally used to refer to the process.

The lesson from *blozen* and *arrossire*

blozen and *arrossire* show that a verb conveys more than its extension.

These observations motivate the mapping:

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

The lesson from *blozen* and *arrossire*

- This pair is significant *not* because it illustrates supposed arbitrariness, but because it demonstrates the availability of multiple construals for certain events in the world.
- When a single event in the world is open to multiple construals, there may be distinct verbs available to describe it, one for each construal.
- Depending on the nature of the construals, these verbs may differ with respect to whether they are unaccusative or unergative.

The causative alternation revisited

The challenge: "the capacity of a verb to be used both transitively and intransitively is a matter of lexical idiosyncrasy" (Farsi 1974: 21)

Shake and *shudder* are near-synonyms, but

- although *shake* shows the causative alternation:

  I shook the tree./The tree shook.

- *shudder* is only intransitive:

  I shuddered./*The news shuddered me.

Externally vs. internally caused eventualities

The relevant lexical semantic distinction and its relevance to the causative alternation:

- Verbs denoting internally caused eventualities:
  these verbs do not show the causative alternation.
- Verbs denoting externally caused eventualities:
  these verbs show the causative alternation.

(Levin & Rappaport Hovav 1995)

Externally vs. internally caused eventualities

An **externally caused eventuality** inherently involves an external cause with immediate control over the event: agent, natural force, or circumstance.

The core verbs are Jespersen's "Move" and "Change" Verbs:

- Change of position: bounce, move, roll, rotate, spin, ...
- Change of state: bake, blacken, break, close, cook, cool, dry, freeze, melt, open, shatter, thaw, thicken, whiten, widen, ...

Externally vs. internally caused eventualities

An **internally caused eventuality** "cannot be externally controlled" (Smith 1970: 107), but is conceived of as arising from inherent properties of its argument. An inherent property of the argument is "responsible" for the eventuality denoted by an internally caused verb.

Externally vs. internally caused eventualities

- Internal causation subsumes agentivity: Monadic agentive verbs are internally caused.
- The prototypical internally caused eventuality involves an agentive argument with a self-controlled body acting volitionally.
- Examples: *play*, *run*, *shout*, *speak*, *swim*, *whistle*, *work*, ...

Externally vs. internally caused eventualities

- However, not all internally caused verbs are agentive.
- Such verbs tend to exert strong selectional restrictions on their subject since the eventuality they denote must result from inherent properties of the verb's argument, and thus the argument must have the requisite properties.
- Examples:
  - Verbs of bodily process (animate nonagentive argument): *blush, cry, cough, laugh, sigh, sneeze*, ...
  - Verbs of emission (inanimate nonagentive argument): *burble, creak, glow, ooze, shine, smell, sparkle*, ...

shake and *shudder* revisited

shake and *shudder* are near-synonyms whose meanings differ in a significant respect:

- *shudder* denotes an internally caused eventuality.
- *shake* denotes an externally caused eventuality.

Given this, *shake*, but not *shudder*, should and does show the causative alternation.

shake and *shudder* revisited

Evidence for the classification of the two verbs:

- Things that *shudder*: people, animals, earth, machines/engines
  $\Rightarrow$ have "self-controlled bodies"
- Things that *shake*: the above and leaves, furniture, dishes, ...

shake and *shudder* revisited

- *shudder* describes an involuntary emotional or physical reaction.
- The reaction is internally initiated, even if it may be in response to an external stimulus.
- The indirect, external stimulus may be expressed, but in an *at* phrase.
  - shuddered **at** the thought of drowning in such loneliness . . . (Oxford Corpus; 5Gates)
  - . . . shuddering **at** tales of the Big Bad Wolf . . . (Oxford Corpus; OnEdge)

Why may near-synonyms differ as to the causative alternation?

- Verb meanings involve construals of events in the world.
  - For the causative alternation, what matters is whether an event is construable as externally caused.
- Sometimes an event may be open to multiple construals, and there may be distinct verbs naming each construal.
- Depending on the nature of the construals, these verbs may differ with respect to the causative alternation.
  - *shudder* and *shake* illustrate this possibility.

The lesson from the causative alternation

Near-synonyms are significant *not* because they illustrate supposed arbitrariness, but because they demonstrate the availability of multiple construals of certain events in the world.

Specific construals align with the availability of the causative alternation.

Conclusions

- Near-synonyms constitute a rich domain for investigating the semantics-syntax mapping.
- Near-synonyms have figured in arguments for arbitrariness because of a focus on potential overlaps in their extensions.
- Their study contributes to developing finer-grained lexical semantic analyses, revealing more systematicity in the semantics-syntax mapping than is sometimes claimed.
- The larger picture sketched here recognizes three linguistic levels: extensional, conceptual, and morphosyntactic.

Thank you!

Selected bibliography

Atkins, B.T.S. and B. Levin (1995) "Building on a Corpus: A Linguistic and Lexicographical Look at Some Near-synonyms", *International Journal of Lexicography* 8, 85-114.

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

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

Clausen, D., A. Djalali, S. Grimm, S. Lauer, T. Rojas-Esponda, and B. Levin (2010) "Extension, Ontological Type, and Morphosyntactic Class: Three Ingredients of Countability", slides, Conference on Empirical, Theoretical and Computational Approaches to Countability in Natural Language, Ruhr-Universität Bochum.

Cruse, D.A. (2004) *Meaning in Language*, Oxford University Press, Oxford. (Second edition.)

Farsi, A.A. (1974) "Change Verbs", *Language Sciences* 31, 21-23.

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

Selected bibliography

Levin, B. and M. Rappaport Hovav (1995) *Unaccusativity*, MIT Press, Cambridge, MA.

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.

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.

Rothstein, S. (2010) "Counting and the Mass Count Distinction", *Journal of Semantics* 27, 343-397.

Selected bibliography

Smith, C.S. (1970) "Jespersen's 'Move and Change' Class and Causative Verbs in English", in M.A. Jazayery, E.C. Palome and W. Winter, eds., *Linguistic and Literary Studies in Honor of Archibald A. Hill*, Vol. 2,, Mouton, The Hague, 101-109.

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.