Cognition 111 (2009) 329–344

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Language, thought, and real nouns

David Barner [a,*], Shunji Inagaki [b], Peggy Li [c]

[a] Department of Psychology, University of California, 9500 Gilman Drive, San Diego, La Jolla, CA 92093-0109, United States
[b] Department of Language and Culture, Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan
[c] Department of Psychology, Harvard University, 25 Francis Ave, Cambridge, MA 02138, United States

ARTICLE INFO

Article history:
Received 21 March 2008
Revised 19 January 2009
Accepted 19 February 2009

Keywords:
Syntax
Semantics
Individuation
Mass nouns
Count nouns
Japanese
Mandarin
Language and thought
Whorfian hypothesis
Lexical statistics

ABSTRACT

We test the claim that acquiring a mass-count language, like English, causes speakers to think differently about entities in the world, relative to speakers of classifier languages like Japanese. We use three tasks to assess this claim: object-substance rating, quantity judgment, and word extension. Using the first two tasks, we present evidence that learning mass-count syntax has little effect on the interpretation of familiar nouns between Japanese and English, and that speakers of these languages do not divide up referents differently along an individuation continuum, as claimed in some previous reports [Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman, & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge University Press]. Instead, we argue that previous cross-linguistic differences [Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*, 169–200] are attributable to "lexical statistics" [Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. Holyoak, & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge University Press]. Speakers of English are more likely to think that a novel ambiguous expression like "the blicket" refers to a kind of object (relative to speakers of Japanese) because speakers of English are likely to assume that "blicket" is a count noun rather than a mass noun, based on the relative frequency of each kind of word in English. This is confirmed by testing Mandarin–English bilinguals with a word extension task. We find that bilinguals tested in English with mass-count ambiguous syntax extend novel words like English monolinguals (and assume that a word like "blicket" refers to a kind of object). In contrast, bilinguals tested in Mandarin are significantly more likely to extend novel words by material. Thus, online lexical statistics, rather than non-linguistic thought, mediate cross-linguistic differences in word extension. We suggest that speakers of Mandarin, English, and Japanese draw on a universal set of lexical meanings, and that mass-count syntax allows speakers of English to select among these meanings.

1. Introduction

What does syntax add to the meaning of words? In psychology and linguistics, this question has spawned investigations of compositionality in language – how the meanings of complex expressions, like noun phrases, are determined by the meanings of their parts and the rules that combine them (Fodor & Lepore, 2002; Frege, 1892; Partee, 1995). It has also led to cross-cultural investigations, which have asked whether cross-linguistic syntactic variation leads to differences in the content of nouns. For example, English, but not Japanese, makes a distinction between count syntax (e.g., *a string*/*some strings*) and mass syntax (*some string*), which corresponds to a referential distinction between a kind of individual and a kind of stuff. According to some, these syntactic differences lead not only to differences in noun content, but also to differences

* Corresponding author. Tel.: +1 858 246 0874.
E-mail address: barner@ucsd.edu (D. Barner).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.02.008

in non-linguistic thought and the perception of things in the world.

The mass-count distinction provides an ideal test of how language affects thought because it is subject to systematic cross-linguistic variation, making it easier to determine the effect of syntax on interpretation. In English and many other languages, names for countable individuals often appear in count syntax. Count nouns occur in singular and plural forms (e.g., *a cat*/*some cats*), with quasi-cardinal quantifiers and determiners (e.g., *these cats*, *those ideas*, and *many blocks*), and can be directly modified by cardinal numbers (e.g., *one cat*, *two ideas*, and *five blocks*). Mass nouns (e.g., *sand* and *dirt*) can occur in none of these contexts. For example, mass nouns cannot be used in singular or plural contexts (e.g., **some sands* and **two dirts*). However, they can occur with quantifiers like *much* and *little* (e.g., *not much milk*). Most of the time, mass nouns do not denote kinds of individuals, though mass syntax does permit individuation (Barner & Snedeker, 2005, 2006; Barner, Wagner, & Snedeker, 2008; Chierchia, 1998; Gillon, 1992, 1999).

Nouns in classifier languages, like Japanese and Chinese, lack count syntax, and behave syntactically like English mass nouns (see Allan, 1980; Chierchia, 1998). For example, Japanese lacks obligatory singular–plural morphology, has few quasi-cardinal quantifiers, and does not permit nouns to be modified directly by numerals. Thus, to name a ball, a set of multiple balls, or even a portion of non-solid stuff like water, the same syntactic structure can be used in Japanese:

(1) Kenji-wa booru-o motteiru.
Kenji-Top ball-Acc have[1]
"Kenji has a/some balls"
(2) Kenji-wa mizu-o motteiru.
Kenji-Top mizu-Acc have
"Kenji has water"

Further, nouns in Japanese require a measure word or classifier when counting, even when counting discrete physical objects:

(3) Kenji wa ni-ko-no booru-o motteiru.
Kenji-Top two-CL-Gen ball-Acc have
"Kenji has two balls"

Syntactically, classifiers resemble English measure words, which are required when counting portions that are named by mass nouns – e.g., two *cups* of water; one *piece* of chocolate; three *sheets* of paper (see Allan, 1980; Borer, 2005; Cheng & Sybesma, 1998; Chierchia, 1998; Li, Barner, & Huang, 2008; for discussion of how these structures are related cross-linguistically).

The lack of count syntax in classifier languages has led some researchers to argue that nouns in these languages are fundamentally different from nouns in English. Meanings encoded directly by nouns in English may be derived in classifier languages, via the combination of nouns, which denote properties or substances, and classifiers, which specify units. For example, Lucy (1992) argues that in classifier languages like Yucatec Mayan, nouns fail to encode individuation. Instead, "Yucatec nouns, lacking such a specification of unit, simply refer to the substance or material composition of an object" (p. 89). Thus, the Mayan word for *banana* (*ha'as*) does not denote a kind of individual by default, but is equally consistent with meanings like "banana-fruit", "banana-leaf", "banana-tree", "banana-bunch", and "banana-stuff", among others. According to Lucy, speakers of Yucatec can distinguish these meanings via the use of classifiers: "all the lexical nouns of Yucatec are unspecified as to unit since they all require supplementary marking (i.e., numeral classifiers) in the context of numeral modification" (p. 73). Classifiers do not merely select from among multiple meanings provided by a particular noun, but actually supply units of individuation, just as English mass nouns require unitizers like *piece* and *bit* (see Section 4, for details).

Lucy argues that this hypothesized compositional difference between Yucatec and English causes speakers of each language to think differently about objects and stuff in the world: "Use of the English lexical items routinely draws attention to the shape of a referent insofar as its form is the basis for incorporating it under some lexical label. Use of Yucatec lexical items, by contrast, routinely draws attention to the material composition of a referent insofar as its substance is the basis for incorporating it under some lexical label". (p. 89) In support of this, he presents data from a similarity judgment task, in which Yucatec Mayan and English-speaking subjects were presented with an object, and then asked to judge which of two alternatives was most similar – one matching the original in shape, or one matching in substance. Lucy found that whereas English speakers preferred the shape-matched choice, Yucatec Mayans divided their choices between the two alternatives (see also Lucy & Gaskins, 2001, 2003). According to Lucy, shape was a more salient dimension than material for speakers of English, a mass-count language, than for speakers of Yucatec Mayan, a classifier language.

The idea that learning mass-count syntax draws attention to stimulus properties like solidity and shape has also gained support from studies of other classifier languages, like Japanese (Imai & Gentner, 1997; Imai & Mazuka, 2003, 2007). Following up on a study by Soja, Carey, and Spelke (1991) and Imai and Gentner (1997) tested English and Japanese-speaking 2-year-olds with a word extension task. Children were presented with novel labels for either a simple-shaped object, a complex-shaped object, or a portion of non-solid stuff. Novel labels were presented in mass-count ambiguous syntax for English children (e.g., *Look at the blicket*) to approximate the absence of mass-count syntax in Japanese. Results indicated that both groups of children were more likely to extend novel words by shape when they named solid objects than when they named portions of non-solid stuff (Imai, Gentner, & Uchida, 1994; Imai & Mazuka, 2007; Landau, Smith, & Jones, 1988, 1992, 1998, for more on the shape bias for solid objects). However, the two groups differed in the extent to which they extended words on the basis of shape. For simple objects and non-solid substances (but not complex objects),

[1] "Top" = topic marker; "Acc" = accusative case marking; "CL" = classifier; "Gen" = genitive case marking; in example (4), below, "Nom" = nominative case marking.