Cross-linguistic Evidence for Cognitive Foundations of Polysemy

Huichun Zhu (huichun.zhu@gmail.com)
Department of Psychology, 17 Memorial Drive East
Lehigh University, Bethlehem, PA 18015, USA

Barbara C. Malt (barbara.malt@lehigh.edu)
Department of Psychology, 17 Memorial Drive East
Lehigh University, Bethlehem, PA 18015, USA

Abstract

Existing discussions of polysemy describe the relations that extended senses may have to the most central sense of a word, but they do not explain in more detail how particular senses are generated for a given word. We propose that extended senses are initially built on the salient features of referents of core senses (and further senses may be generated from those). We provide evidence for the role of salient features of core senses in generating extended senses through three studies. These studies use speakers of English and Chinese, historically unrelated languages.

Keywords: polysemy; word meaning; word senses; cross-linguistic comparison

Introduction

Polysemous words are those that have multiple different but related meanings (e.g., *foot* as in "my left *foot*", "the foot of the *chair*," "at the *foot* of the mountain"). Polysemy is the rule more than the exception for words of moderate to high frequency (e.g., Berlich, Todd, Herman & Clarke, 2003). This one-to-many mapping of form to meaning has raised questions about the representation and processing of polysemous words (e.g., Bohrn, Altmann & Jacobs, 2012; Caramazza & Grober, 1976; Klein & Murphy, 2001; Simpson, 1994). But studies addressing such questions leave unanswered a logically prior question: How do polysemous senses of words arise?

At one extreme, it could be that the generation of new senses from an initial word meaning is entirely unpredictable. The first extended senses that arise for any individual word might be highly influenced by the particular lexical gaps that exist and other characteristics of the language involved (such as where it falls on the synthetic-to-isolating continuum, which affects how morphology can be used to help build sets of related senses). They may also be influenced by the cultural conditions present (e.g., introduction of new products or practices) that drive innovation for the linguistic community. The result of these influences could be that the senses tend to be idiosyncratic to that word and that language. Furthermore, because extended senses may build on one another to create chains moving farther from the original (Lakoff, 1987), once the first extended sense emerges for a word, the additional senses it spawns may set the chain for that word off in directions not followed by other languages for their most closely related word.

The possibility of complete unpredictability is defeated by the observation that some relationships between central sense and extended senses of words tend to recur. For instance, at least within English, the same word can be used to refer to both the object and the substance (e.g., *chicken; fish*); both the text and the object (e.g., *newspaper; book*); both the act and the instrument (e.g., *drill; brush*); and both the actor and the act (e.g., *cook; scout*) (e.g., Blank, 1999; Norrick, 1981; Nunberg, 1979). More broadly, metaphor and metonymy frequently characterize the relations between the most literal sense of a word and its extended senses (Stern, 1931; Ullmann, 1962; Blank, 1999). For instance, the foot of a mountain may be metaphorically similar to a human foot, and the tongue spoken by a linguistic community is the language that emerges by means (partially) of the tongue in their mouth.

Still, such observations are largely descriptive and post-hoc. They do not indicate, for any given word, what senses are likely to arise, and they provide no explanation of why those senses may come about. At a broad level, they do suggest that language users engage in metaphorical and metonymic thinking and also that they are sensitive to more specific relations such as that of an object to its substance. But that does not reveal what particular senses are likely to be created or why. Indeed, Lakoff (1987) and Langacker (1988) have suggested that although the array of senses conventionally associated with a word is not arbitrary, neither is it readily predictable.

These analyses of recurring relations have generally been carried out without closer consideration of the cognitive processes that might be involved. We suggest that further consideration of the cognitive processes contributing to generation of extended senses can help provide predictive power. In particular, we propose that salient characteristics of the default referents of the core senses (the most central, literal senses) provide the basis for sense extensions. Consider the word *head*. Suppose that the core sense of the word refers to a certain body part of animals. Speakers of a language know many things about typical referents of this sense of the word. But those things they know are not all equally salient or important. Although people know that heads are made of substances such as flesh, bones, blood, and so on, these features are not likely to be the ones they bring to mind when thinking about heads. Instead, properties such as *located at top of body* and *organ of decision making* are more uniquely and saliently associated with heads. In turn, these salient features will be more

uniquely and saliently associated with the core sense of the word *head*. These salient characteristics of the word *head* are the ones that will be most available to build upon in creating extended senses, as exemplified in expressions such as "*head* of the table" and "*head* of the department". On the other hand, for other body parts, there may be different salient properties. For noses, for instance, the location with respect to the rest of the body may be less important but the role in olfaction may be critical, making smell-related extended senses of the word *nose* more likely to emerge.

Much of the past analysis of polysemy has been based on English and, to a lesser extent, on other Indo-European languages. The generalizability of the analyses to other languages is not well-established. Even if sense relations such as object-substance, text-object, and act-instrument do appear across languages, the particular words of a language for which specific polysemous senses arise may vary unpredictably due to the sorts of influences mentioned earlier: The earliest extended senses generated for any particular word might be strongly influenced by elements of the linguistic and cultural context of the original sense, and the diverse senses generated in turn may send languages off in different directions as sense chains develop for a word.

In contrast, our proposal argues for the likelihood of shared senses across languages. Although some aspects of what is salient about an entity are bound to be culture-specific, many features of the world will be salient across cultures due to shared sensory, perceptual, attentional, and other cognitive processes and shared human needs and goals. This assertion is supported by data from similarity-sorting on diverse sets of real-world entities including both objects and actions. Speakers of different languages produce convergent similarity judgments despite the featural richness of the stimuli, which would have allowed different sorting patterns to emerge (e.g., Malt et al., 1999; Malt, Ameel, Imai, Gennari, Saji, & Majid, in press). Given shared appreciation of some feature of entities, then where two languages have words with similar core senses, our proposal suggests that the specific senses they will spawn will tend to overlap, even for historically independent languages. This should be most true for extended senses closest to the core, since they will be generated most directly from those salient features.

In three studies, we tested the proposal that salient features associated with core senses of words provide the foundation upon which extended senses are built. Evidence is provided for this proposed language-independent mechanism using data from speakers of English and Chinese, two historically unrelated languages. Pre-tests first established core senses and salient features for a set of words. The studies then asked (1) to what extent extended senses are shared between English and Chinese; (2) whether the shared ones tend to be closer to the core sense than non-shared ones; (3a) whether the salient features are more applicable to shared than non-shared senses; and (3b) whether the degree of applicability of salient features predicts distance of senses from the core.

Participants and Materials

Participants for all studies were English and (Mandarin) Chinese native speakers. None of the English participants knew any Chinese. All the Chinese participants knew some English but identified Chinese as their dominant language. The English-speaking participants were undergraduates at Lehigh University in the U.S. The Chinese-speaking participants for the two pre-tests and for Study 1 were recruited from the Chinese community at Lehigh. The Chinese participants for Study 2 and 3 were recruited from universities in Kaifeng and Chengdu, China. Participants did only one task and only in their native language.

We selected 36 pairs of Chinese and English words for which the members of each pair were considered likely to share the same core sense (verified in Pre-test 1, below). Because higher frequency words tend to have more senses than lower frequency ones (Zipf, 1949), words were selected such that the two members of each pair fell into the same frequency rank. Frequencies were determined using Subtlex-US for English (Brysbaert & New, 2009) and Subtlex-CH for Chinese (Cai & Brysbaert, 2010). Twelve nouns, 12 verbs, and 12 adjectives were included, as follows. Nouns: head /头,world/世界, hand/手, doctor/医生, foot/脚, tree/树, face/脸, door/门, heart/心, flower/花, mouth/嘴, table/桌. Verbs: push/推, follow/跟, run/跑, win/赢, touch/触, hang/挂, listen/听, leave/离开, smell/闻, lead/领, eat/吃, die/死. Adjectives: sweet/甜, safe/安全, short/短, rich/热, high/高, true/真, hot/热, broken/破, tight/紧, empty/空, heavy/重, simple/简单.

Pre-test 1: Verifying Core Senses

Pre-test 1 examined whether the words of each chosen pair share a core sense. For each word in a language, 30 sentences were selected to represent its varied senses. English sentences were chosen from The Corpus of Contemporary American English (COCA, Davies, 2011). Due to lack of a parallel corpus for Chinese, Chinese sentences were chosen from www.baidu.com, the most powerful search engine for modern Chinese. Sentences were chosen to represent all distinct senses found for a given word. The determination of "distinct' senses was based on the dictionary numbering of senses for English and on experimenter intuition for Chinese; verification of psychologically real distinctions among senses is provided in Study 1. The 30 sentences were printed on paper slips, one per slip. Because the number of senses per word varied, some words had more sentence examples per sense than others.

For each target word, 10 participants were asked to select all and only the sentences representing the most fundamental and basis sense of the word. Sentences selected by 70% or more of participants were considered to represent the core sense. Each word had at least 3 sentences that met this criterion, confirming that native speakers agree on the core sense for words.

Two Chinese-English bilingual judges who were blind to the experimental hypotheses then judged the cross-linguistic agreement of these data independently. Each was given the sentences representing the core sense (based on the 70% criterion) and the non-core senses (all remaining sentences) for each word of a pair. One judge started from English and the other from Chinese. They were asked to imagine that they were translating the core sense of the starting language into the other language. They were to judge whether the sense represented by the core sense set of sentences would fall into the core or non-core set of sentences in the other language. For all 36 word pairs, both judges selected the core sense sentences of the other language as the set containing the core sense of the starting language. This outcome confirms the similarity of core senses between the languages for these 36 word pairs.

Pre-test 2: Establishing Salient Features

The second pre-test established the salient features of the core sense of each word and determined whether the same features emerged for paired words of the two languages. Twenty participants per language were given 75 seconds to list all the features that came to mind after reading a given word (Rosch & Mervis, 1975). This procedure was repeated for each of the 36 words. They were given an example for *dog* of properties such as has fur, four legs, barks, etc. Properties listed by at least 1/3 of participants per word were retained for further consideration. For both languages, about 6 features per word met this criterion (for English, $\overline{x}$ = 6.12, s.d. = 2.14; for Chinese, $\overline{x}$ = 5.89, s.d. = 2.3).

The features of the two languages that met the criterion were then compared for each word pair by a Chinese-English bilingual judge who was blind to the experimental hypotheses. The judge determined, for each word of a language, whether she considered a feature listed in one language to have an appropriate translation among those listed for its paired word in the other language. A mean of 4.47 (s.d. = 1.89) per word were judged to be shared across the languages. Thus, most features having consensus within a language were also agreed on across languages.

This result provides a set of features to use in further tests. Furthermore, it provides supporting evidence that speakers from two distinct cultures, speaking historically independent languages, perceive many of the same features of a diverse set of entities to be salient.

Study 1: Do Senses Correspond between English and Chinese?

The first study asked to what extent the extended senses of our 36 word pairs correspond across the two languages. At one extreme, it could be that few or no senses of the words correspond beyond the core sense. This outcome would suggest that the linguistic and cultural forces idiosyncratic to individual languages at particular moments in time dominate the outcomes of the sense generation process. At the other extreme, the languages could correspond largely or fully. This would come about if language- and culture-specific forces make minimal contributions to the sense generation process and the process is instead heavily dominated by shared perception of salient features of entities along with shared processes of generating senses (including tendencies toward metaphorical and metonymic thinking, as well as appreciation of specific relations such as object-substance and act-instrument). Given the likelihood of some contribution of language- and culture-specific forces, our perspective does not predict exactly what percent of senses might be shared between Chinese and English. It does predict that it should not be close to the lower extreme.

Because the senses used in Pre-test 1 to establish core meanings were obtained from a dictionary and the internet, not directly from language users, we first sought evidence of the psychologically distinct senses for each word. Twenty-eight English and 20 Chinese speakers participated in a sorting task. Participants sorted the 30 sentences for each word of their native language (as described above) into piles. They were asked to put all sentences that represented the same sense in a single pile, and to put sentences representing different senses into other piles.

A 30 x 30 matrix of the similarity values for each possible pair of sentences was then created from the sort. The number of participants who sorted a pair into the same pile was taken as the similarity value for that pair (e.g., Kruskal & Wish, 1978). To create a visual representation of the perceived similarities among the senses represented by the sentences, multi-dimensional scaling (MDS) (IBM SPSS, 2010) was then carried out on the similarity matrix for each word. The dimensionality of the solution that was retained for further use for each word was determined by looking for the lowest dimensionality with acceptable fit (following the rule-of-thumb for acceptable Young's Stress of <.2 and that for R^2 of >.9), along with applying the "elbow" rule (looking for where a further increase in dimensionality results in declining additional benefit to fit).

To further help identify discrete senses, hierarchical clustering (IBM SPSS, 2010) was then applied to the coordinates of the chosen MDS solution for each word. Selection of the optimal or most meaningful clustering solution for each word (that is, the number of different clusters to allow) was made based on two considerations. First, we again applied the elbow rule, which limited candidate solutions to the region of 5-15 clusters. Second, we sought a solution for each word that provided a level of granularity showing agreement with the results of Pre-test 1's core sense task. We looked for the solution in which the sentences deemed to represent the word's core sense in Pre-test 1 were maximally present within one cluster and minimally present within any others. The outcome of these steps was a clustering solution for each word of each language, where we take each cluster to represent a psychologically distinct sense.

To evaluate to what extent senses were shared between the two languages, two new Chinese-English bilingual judges who were blind to the experimental hypotheses judged agreement of the obtained senses independently. For

each word pair, each judge began with the senses within one language (as represented by the sets of sentences constituting the clusters just determined) and looked for a match among the sentence clusters of the other language. Two senses were considered a match between the languages if and only if both judges proposed the same clusters as matching. Based on their judgments, the average number of shared senses across word pairs of the two languages was 56%. (Because the English and Chinese paired words sometimes differed in the number of sense clusters obtained, in principle, calculating the percent of English senses shared with Chinese and the percent of Chinese senses shared with English could produce somewhat different values. In practice, however, both values were within rounding range of 56%.) There was no effect of word type in either language; outcomes were similar for nouns, verbs, and adjectives.

In short, slightly over half of all senses identified in English and in Chinese were shared between the two languages. This value indicates that generation of extended senses from core senses of words is influenced to some extent by language- and culture-specific factors. At the same time, though, it shows a level of agreement that is impressive given the historical independence of the English and Chinese languages, as well as the markedly different cultural histories. This level of agreement argues for a substantial influence of shared cognitive representations and processes on the generation of extended senses.

Study 2: Are Shared Extended Senses Closer to the Core than Non-shared Ones?

The results of Study 1 are compatible with the proposal that salient features associated with core senses of words provide the foundation upon which extended senses are built. However, it does not provide a direct test of this notion. As the next step in evaluating our proposal, we examined whether extended senses that are shared between the two languages are closer to the core senses of the words than those that are not shared. If senses are generated from cores by building on salient features of the cores (and further senses may be generating by building upon those in ways that create chains of senses increasingly distant from the core; Lakoff, 1987), then those closest to the core should also be the most likely to be shared. More distant senses are more likely to be influenced by language- and culture-specific factors and therefore to be more idiosyncratic.

Twenty-one English and 20 Chinese participants participated in a task to judge the distance of extended senses from core senses. For each target word of their language, they were given the 30 sentences presorted into clusters representing the word's psychologically distinct senses as determined in Study 1. The cluster representing the core sense was marked for participants. Participants judged the distance between each extended sense and the core using a 1-7 scale, where 1 was labeled "very close", 4 was "neither far nor close", and 7 was "very far".

Mean judgments for each extended sense were calculated. Collapsing across the two languages, shared senses were rated as closer to the core senses (M = 3.91, SD = .71) than non-shared senses (M = 4.30, SD = .57). The main effect of sharing status was significant, $F(1, 39) = 40.78$, $p < .001$, $\eta_p^2 = .51$.

We further analyzed the two languages separately. For each language separately, shared senses were rated as closer to core senses than non-shared senses. The main effect of sharing status was significant for both languages. For English, the mean for shared senses was 3.73 (SD = .67), and for Non-Shared senses, 4.20 (SD = .59), $F(1, 20) = 38.50$, $p < .001$, $\eta_p^2 = .66$. For Chinese, the mean for shared senses was 4.12 (SD = .69), and for non-shared senses, 4.42 (SD = .52), $F(1, 19) = 9.97$, $p = .005$, $\eta_p^2 = .34$.

This outcome supports the proposal that salient features of core senses provide the foundation for extended senses by showing that the specific senses appearing in two unrelated languages are those closest to parallel core senses of the languages. In doing so, it also provides evidence that similar cognitive processes drive sense generation for English and Chinese.

Study 3, Part A: Are Salient Features More Applicable to Shared Senses than Non-shared Ones?

To further evaluate whether salient features associated with core senses provide the foundation on which extended senses are built, we asked directly whether, within each language, salient characteristics of core senses are more applicable to shared senses than to non-shared ones.

Nineteen English and 20 Chinese participants participated. The 30 sentences were again presented presorted into clusters representing the word's psychologically distinct senses as determined in Study 1. In addition, the features that passed the 1/3 criterion in Pre-test 2 were presented. The participants judged, for each sense, whether each of the features for that word was applicable to that sense. Judgments were made using a 0-7 scale, where 0 was labeled "completely inapplicable", 4 was "moderately applicable", and 7 was "perfectly applicable".

Mean ratings for the feature set for each sense were calculated. The feature sets of the core senses were judged as more applicable to shared senses than to non-shared senses in both languages. For English, the mean feature rating for shared senses was 2.42 (SD = .53); the mean rating for non-shared senses was 1.85 (SD = .48), $p < .001$ (planned comparisons LSD). For Chinese, the mean rating for shared senses was 2.06 (SD = .65); the mean rating for non-shared senses was 1.71 (SD = .70), $p < .001$ (planned comparisons LSD).

This result provides evidence that features saliently associated with the core senses of words for speakers of both languages are projected into extended senses. It also confirms that senses shared between the two languages are more closely tied to the features jointly perceived as salient than are non-shared senses.

Study 3, Part B: Does the Applicability of Salient Features Predict the Perceived Distance of Senses from the Core?

As a final step in testing our proposal, we asked whether ratings of feature applicability to senses (as derived from the data of Study 3, Part A) predict the perceived distance of senses from the core (as derived from the data of Study 2). If extended senses are generated from cores via salient features, then the more applicable the features are to a sense, the closer the sense should be to the core.

For each language, we calculated the Pearson correlation between applicability ratings to each given sense of each word from Study 3 Part A and distance ratings of the same senses from Study 2. For both languages, the more applicable the salient features, the less distant the extended senses to the core senses. The correlations were significant for both languages: for English, $r = -.613$, $p < .001$; for Chinese, $r = -.647$, $p < .001$. These data provide evidence for the proposed mechanism of polysemy. Salient features of core senses provide the foundation for generating initial extended senses (and additional senses may then build upon those).

Discussion

The current findings indicate that despite language-and culture-specific influences, psychological forces common to speakers of different languages play a prominent role in polysemy generation. English and Chinese speakers can and do agree on what features associated with paired words are salient. Furthermore, this agreement is reflected in a shared set of extended senses of those words. This provides evidence that there exists a common psychological mechanism driving generation of polysemy in different languages. By combining a cross-linguistic perspective with a cognitive one, this project disentangles psychological forces from non-psychological ones.

Although there have been past proposals about the existence of general principles of polysemy (e.g., Langacker, 1988; Nunberg, 1979; Norrick, 1981), none of them has been detailed and concrete enough to provide predictive power about what senses might be generated for specific words or what senses might be shared across languages. The current data demonstrate that a more predictive account is possible by giving closer consideration to the psychological processes involved. Our proposal is consistent with past ideas about the importance of metaphorical and metonymic thinking. However, it builds on them to help create a more complete account of how specific senses of words come about.

There are at least two major directions to consider in further developing accounts of when and where polysemous senses of words will arise. One is to better understand how language characteristics do influence generation of senses in individual languages. We alluded earlier to the possibility that language morphology (the synthetic-to-isolating continuum) may matter. A heavily agglutinating language such as Turkish, for instance, can form many new words by piling additional morphemes onto root words. This feature opens possibilities for the creation of word senses that would typically be captured in multi-word phrases in many other languages. (It also complicates determination of what should count as the senses of a single root word.) For Chinese, a related issue arises. Chinese often creates compound words by combining two characters, which serve as conventional (not created-on-the-fly) units of meaning. Generating new compounds and generating new senses of a single character are two different ways for sense extensions to come about in Chinese, while English (among other languages) has only something similar to the latter. Further issues arise with other languages. In languages with grammatical gender, the gender marker can suggest different senses of the same word. In French, *le pendule* is equivalent to English *pendulum* while *la pendule* is equivalent to English *clock*, and in German, *der See* means *lake* while *die See* means *sea* (Ullman, 1962). In light of these considerations, it is particularly impressive that English and Chinese do, in fact, share a sizeable set of extended senses that can be predicted on purely psychological grounds, apart from linguistic considerations. The more languages that can be studied, and the more different samples of words that are used, the clearer it will be what portion of variance the psychological processes account for. Ultimately, however, a full account will need to be able to understand the language-specific factors as well.

The second additional consideration is about the role of bottom-up forces. Our focus has been more on the top-down than the bottom-up. That is, the current discussion has been about how cognitive processes drive new senses, without addressing how external stimuli help inspire them. However, more often than not, new senses are not created just because they can be. Rather, novel usages of existing words arise when there is a need. For example, think about a computer mouse. When this device was newly invented, a name for it was needed. Its body and tail shape presumably activated mental representations of the mouse mammal, and so *mouse* became a possible name for it. A full account of how extended senses arise will need to incorporate such bottom-up influences. In doing so, it will help articulate how cultural needs of the moment and lexical gaps with respect to those needs play into the extension of senses. Note, however, that this bottom-up triggering of *mouse* as a possibility does not operate in isolation from the top-down. If the body shape and tail were not salient features of the mouse mammal and salient associates of the pre-existing sense of the word *mouse*, then *mouse* as the name for the computer gadget would be unmotivated and puzzling to users. Both directions of influence must be considered.

Acknowledgments

This work was supported by a grant from the National Science Foundation (#BCS-1057855) to Barbara Malt and Ping Li. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the

author(s) and do not necessarily reflect the views of the National Science Foundation. We thank Kathleen Costello, Yishuang Li, Shenchi Tian, Andrew DeLena, Nikolas Rivera, Lauren O'Brien, Xinyue Geng, Yihan Lin, Nikolas Rivera, and Aislinn Rowan for assistance with data collection and analysis, and Padraig O'Seaghdha, Amanda Brandone, Mark Bickhard, and Kiri Lee for helpful discussion.

References

Nerlich, B., Todd, Z., Herman, V. & Clarke, D. (2003). *Polysemy: Flexible Patterns of Meaning in Mind and Language*. Berlin/New York: Mouton de Gruyter.

Blank, A. (1999). Why do new meanings occur? A cognitive typology of the motivations for lexical semantic change. In A. Blank & P. Koch (Eds.), *Historical Semantics and Cognition* (pp. 61–90). Berlin/New York: Mouton de Gruyter.

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language — A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia, 50(11),* 2669-2683.

Brysbaert, M. & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 2009, 41 (4),* 977-990.

Cai, Q. & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *Plos ONE, 5(6), e10729.*

Caramazza, A., & Grober, E. (1976). Polysemy and the structure of the subjective lexicon. In C. Rameh (Ed.), *Georgetown University roundtable on languages and linguistics. Semantics: Theory and application* (pp. 181-206). Washington, DC: Georgetown Univ. Press.

Davies, M. (2011). The corpus of contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing, 25,* 447-465.

IBM Corp. Released 2010. *IBM SPSS Statistics for Windows, Version 19.0*. Armonk, NY: IBM Corp.

Klein, D. & Murphy, G.L. (2001). The representation of polysemous words. *Journal of Memory and Language*, *45*, 259-282.

Kruskal, J. & Wish, M. (1978), Multidimensional scaling. *Sage University Paper series on Quantitative Application in the Social Sciences* (pp.7-11). Beverly Hills and London: Sage Publications.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Langacker, R. (1988). Review of George Lakoff: *Women, fire and dangerous things*. *Language, 64(2),* 384-395.

Malt, B. C., Ameel, E., Imai, M., Gennari, S., Saji, N., & Majid, A. (in press). Human locomotion in languages: Constraints on moving and meaning. *Journal of Memory and Language*.

Malt, B., Sloman, S.A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language, 40*, 230-262.

Norrick, N. (1981). *Semiotic Principles in Semantic Theory*. Amsterdam: John Benjamins.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy, 3 (1),* 143-184.

Rosch, E. & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7(4),* 573-605.

Simpson, G. B. (1994). Context and the processing of ambiguous words. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 359-374). San Diego, CA: Academic Press.

Stern, G. (1931). *Meaning and change of meaning with special reference to the English language.* Göteborg: Elander.

Ullmann, S. (1962). *Semantics: An Introduction to the Science of Meaning.* Oxford: Blankwell.

Zipf, G. (1949). *Human Behavior and the Principle of Least Effort.* Addison-Wesley Press, Inc.