Evolution of polysemous word senses from metaphorical mappings

Yang Xu (yang_xu_ch@berkeley.edu)
Department of Linguistics, Cognitive Science Program, UC Berkeley

Barbara C. Malt (barbara.malt@lehigh.edu)
Department of Psychology, Lehigh University

Mahesh Srinivasan (srinivasan@berkeley.edu)
Department of Psychology, UC Berkeley

Abstract

What forces have shaped the evolution of the lexicon? Languages evolve under the pressure of having to communicate an unbounded set of ideas using a finite set of linguistic structures. This suggests *why* the transmission of ideas should be compressed such that one word will develop multiple senses. Previous theory also suggests *how* a word might develop new senses: Abstract concepts may be construed in terms of more concrete concepts. Here, we bring these two perspectives together to examine metaphorical extensions of English word meanings over the past millennium, analyzing how senses from a source domain are extended to new ones in a target domain. Using empirical and computational methods, we found that metaphorical mappings are highly systematic and can be explained in terms of a compact set of variables. Our work shows how metaphor can provide a cognitive device for compressing emerging ideas into an existing lexicon.

Keywords: Word meaning; semantic change; polysemy; metaphorical mapping; systematicity

Words are fundamental components of language, but their meanings are not stable. For example, the English word *grasp* originally conveyed a physical action, as in "grasp a fruit," but was later extended to express an abstract sense of understanding, as in "grasping of an idea." The "physical action" sense first appeared around 1300, and the "understanding" sense emerged around AD 1600 (Christian, Roberts, Samuels, Wotherspoon, & Alexande, 2015a). The synchronic product of such historical sense extensions, known as polysemy - that a single word form can express multiple distinct but related senses - is widespread in natural language (Lakoff & Johnson, 1980; Brugman, 1988; Sweetser, 1991; Geeraerts, 1997). We investigate the extent to which historical polysemous patterns can be predicted by focusing on metaphorical mapping.

Metaphorical mapping is a key structuring force in sense extension and and semantic change (e.g., Lakoff & Johnson, 1980; Sweetser, 1991). It operates by mapping an existing sense of a word from its own *source* domain to another *target* domain based on structural similarities between the two domains. For instance, the "physical action" sense of *grasp* can be thought of as metaphorically extended to "understanding," namely holding onto an idea. The historical process through which polysemy develops raises the following question: Is the evolutionary path that metaphorical senses follow unpredictable, or is there a *systematicity* in how new senses develop from existing ones, driven by overarching evolutionary forces? We address this question by bringing together two influential theories that were not previously in contact.

On the one hand, a prominent theory of the nature of human cognition - Conceptual Metaphor Theory (Lakoff & Johnson, 1980; Reddy, 1979) - holds that thought is grounded in metaphor. By this account, metaphorical mappings in thought should occur in directions that are cognitively natural. For example, such a theory postulates that abstract thought is facilitated by linkage to concrete concepts (e.g., "understanding" construed in terms of concepts that are more concrete or directly related to human experience - a "physical action" of holding onto something). This perspective implies systematicity in the development of the lexicon, because it suggests that mappings will tend to be from concrete to abstract and not *vice versa*. However, it suffers in two important respects as an account of polysemy. First, variables proposed to be cognitively priviledged in metaphorical mapping (discussed below) have not been assessed against the historical record of lexical change, and therefore their explanatory powers and interrelations in accounting for empirical data with respect to the lexicon remain unknown. Second, the conceptual theory itself does not motivate why metaphor should serve as a key mechanism for the evolving lexicon, because its goal is to explain thought, not word meanings *per se*.

On the other hand, a growing line of research on principles of language evolution provides clues for addressing the *why* question. By this account, linguistic structures evolve under the dual pressures of communicative needs and cognitive constraints (e.g., learnability), such that languages should trade off between competing pressures of expressivity and compressibility (Kirby, Tamariz, Cornish, & Smith, 2015). Specifically, language evolution must allow for meaningful and informative communication of ideas while providing sufficient compression of ideas into existing linguistic structures (e.g., via compositional means) to ensure that languages do not grow without bound and are therefore learnable. Although such theories have been applied to explain synchronic features of language such as word length (Piantadosi, Tily, & Gibson, 2011) and the structure of semantic domains (e.g., Regier, Kemp, & Kay, 2015), they have not been used to help understand the role of metaphorical mapping (cf. Geeraerts, 1997; Blank & Koch, 1999) in the diachronic development of lexicons.

We offer such a perspective on metaphorical mapping by bringing together these perspectives. We propose that the

unique nature of metaphor in lexical evolution is precisely explained by functional needs for expressive communication and cognitive constraints for learnability. In one respect, metaphorical mapping facilitates meaningful communication of emerging ideas, such that when a speaker uses a word in a novel sense metaphorically, the listener should find it cognitively effortless to decode the intended meaning (cf., Traugott, 2003). In another respect, metaphorical mapping serves as a compression device that effectively folds emerging meanings into existing words without requiring construction of word forms *de novo*. As such, it serves as a strategic device that trades off between communicative and cognitive constraints. Our proposal predicts that metaphorical sense mappings should occur in systematic ways by conforming to both communicative and cognitive constraints.

We critically assess this idea by analyzing a large set of metaphorical mappings between source and target domains and spanning 1300 years, dating from Anglo-Saxon English to the present. We seek to predict the historical ordering of sense extension by testing whether some domains are more likely to be the starting point of an historical trajectory (the source domain) while others are more likely to be the endpoint (target domain). To our knowledge, this is the first large-scale study evaluating the predictability and directionality of polysemous metaphorical mappings against the record of historical change in a lexicon.

Candidate variables

We identify six candidate variables for metaphorical mapping based on communicative and cognitive considerations. Three of the variables we test are suggested by Conceptual Metaphor Theory (Lakoff & Johnson, 1980; Reddy, 1979). Applied to the development of word senses, meanings that are easier to understand and more richly experienced - because they are more concrete and tied to bodily experience - might develop earlier in historical time and might be especially accessible, well-structured, and conceptualized. As such, they may lend themselves to extension via metaphorical processes because such extensions would provide a high degree of efficiency and expressivity. These considerations provide the following predictions:

Concrete → Abstract. Word senses that refer to things perceived through the sensory systems should serve as a source of metaphorical mappings, relative to ones labeling less perceptible referents (Lakoff & Johnson, 1980).

Embodied → Disembodied. Word senses that refer to things that are more directly and viscerally experienced through our bodies should serve as a source of metaphorical mappings, relative to senses that label referents less directly experienced through our bodies (Lakoff & Johnson, 1980).

External → Internal. Word senses that refer to entities in the external world should serve as a source of metaphorical mappings, relative to senses that label internal, mental entities (e.g., emotions or feelings). This would be predicted if internal entities are understood in terms of external things (Sweetser, 1991), and/or if external entities are

more easily lexicalized than internal entities (because external things are easier to indicate ostensively).

An additional three variables can be motivated based on the proposal that metaphorical extensions tend to be communicatively efficient and expressive, which includes describing target domains in vivid and emotionally-valenced ways.

Animate → Inanimate. Insofar as animate entities hold a special status within our mental lives (Silverstein, 1976; Traugott, 2003), words referring to them may be more expressive than those referring to less animate entities, and thus may be recruited by speakers to vividly convey salient features of an intended referent to addressees. Thus, words that refer to animate entities could serve as a source of metaphorical mappings, relative to words that label inanimate entities.

Less valenced → More valenced. If metaphorically-derived senses arise in part because of their expressive power, we might expect derived senses to be more emotionally valenced than originating senses. This predicts that when word senses from source domains are extended, the resulting senses in the target domain will be more emotionally valenced (Ullmann, 1957).

More Intersubjective → Less intersubjective. Intersubjectivity refers to the degree to which people experience something the same way and agree about the nature of that experience (Traugott, 2003). For example, most people would agree about whether or not a chair is wooden (making it more intersubjective) but not necessarily about whether it is beautiful (making it less intersubjective). Word meanings that are more intersubjective can be easier to understand or establish labels for;[1] They would thus tend to serve as a source of metaphorical mappings, relative to senses with less intersubjective meanings.

In sum, at least six potential variables are worthy of evaluation. Some of these variables are likely to be correlated with one another—an important aspect that we address in our analysis. For example, many of the same word meanings are likely to be external, concrete, intersubjective, and embodied. Despite this fact, these variables are not identical and can in principle be teased apart from one another. For example, some highly concrete word meanings, like *table* and *arm* differ in how embodied they are: e.g., arms are more viscerally and directly experienced than tables.

Materials and methods

To explore metaphorical sense mappings in history, we drew on the corpus of metaphorical mappings of English provided by the Mapping Metaphor project (2015). We describe this database and the empirical methods for obtaining behavioral ratings of the six variables.

Historical database We obtained data from the Metaphor Map of English database (Christian, Roberts, Samuels, Wotherspoon, & Alexande, 2015b), which identifies metaphorical links among different semantic domains (as defined within the database) over more than a millennium. This

---

[1]Personal communication via Eve Sweetser.

database provides metaphorical sense changes identified and coded from 793,742 word forms and 225,131 semantic domains (e.g., some domains include "textiles," "digestive organs," "anger," "pride," etc.) in the Historical Thesaurus of English project (Christian et al., 2015a), based on the *Oxford English Dictionary* and A Thesaurus of Old English (Roberts, Christian, & Grundy, 2015). Each semantic domain defines a category of meaning that word senses can evolve from (source domain) or towards (target domain). We used all domains having full date and word information 1) source domain, 2) target domain, 3) earliest of period of extension between these domains, and 4) some sample words that participated in this extension. In sum, this data set contains records of historical metaphorical sense mappings among 400 semantic domains. The data set contains mappings spanning an 1100 year period, from the Old English period around AD 800 through to the present era, around AD 1950. Within the database, the metaphorical mappings were summarized separately for the Old English period (i.e., before AD 1100), and in 50-year steps for the subsequent 800 years, providing 18 unique historical time points for our analysis. In total, the database lists over 5,000 pairs of domains, indicating the historical direction of sense extension among each pair of domains. Three types of directions were recorded: 1) A→B, i.e. A is source and B is target; 2) A←B, i.e. A is target and B is source; 3) A↔B, i.e. bidirectional. Our analysis in the following section aims to predict these relationships among domains based on the ratings described.

Ratings We gathered ratings of the 400 semantic domains along the six target variables discussed above through an online survey. For each variable, participants rated each of the 400 domains on a 1-7 scale. Participants first read a definition of each domain (e.g., "Plant" - A living thing that grows in the ground, usually has leaves or flowers, and needs sun and water to survive). These definitions were assembled by consulting the *Merriam-Webster Dictionary*. After reading the definition of the domain, participants were asked to rate the domain on one of six variables. For example, for concreteness, participants selected a number between 1 and 7, where 1 represented "highly abstract," 7, "highly concrete," and 4, "intermediate." Similarly, for valence, 1 represented a "highly negative" emotional response , 7 "highly positive," and 4, "intermediate" or neutral. Because a large number (400 domains×6 variables) of ratings needed to be completed, each participant rated a block of 40 domains randomly sampled from the 400 for a single variable. Data were collected from 1448 participants using the Qualtrics survey software,[2] disseminated via Amazon's Mechanical Turk.[3] Data from participants whose native language was not English, and from those who did not respond correctly to three "catch" trials assessing attention to the task, were discarded. We obtained on average 18 (SD=2) ratings for each of the domain-variable questions with a standard deviation of 1.6 (SD=0.32)

in inter-subject agreement.

Computational analyses and results

To investigate systematicity in metaphorical mapping, we performed two analyses to account for 1) diachronic ordering and 2) synchronic asymmetry in metaphorical sense mappings recorded in the historical database. We describe each of these analyses.

Diachronic ordering of metaphorical mappings. We used a set of parameter-free models to predict directions of metaphorical mappings among semantic domains. Specifically, for each of the variables described above, we created a model that specified the predicted direction of mapping between a pair of semantic domains by calculating the difference in average empirical ratings of these two domains along the variable dimension. For example, for the animacy variable, we predicted the direction of mapping to be Animate→Inanimate, such that the domain that was rated as more animate on average should serve as the source, and the domain that was rated less animate on average should serve as the target. Our models also predicted bidirectional mappings (accounting for 6.76% of the available mappings) if the average ratings of two domains were equal. Since the valence predictor is polarized (i.e. a concept can be either positive or negative), we calculated the absolute value of its ratings, and for this model predicted the direction of mapping betwen domains to follow Less valenced→More valenced. Thus, a domain rated 7 (highly positive) would be treated as equally-valenced as a domain rated 1 (highly negative). Table 1 specifies the predictions of each model along with example mappings that the models successfully predicted in the database.

Each model made a prediction about the direction of extension (the source-target relation) at different historical epochs for a given pair of domains (e.g. *A* and *B*), out of three possibilities: *A→B*, *A←B*, or *A↔B*. To establish a baseline, we considered a random model that predicted the direction of mappings arbitrarily. The results appear in Figure 1a. Overall, all models predicted directional change above chance. In particular, externality (accuracy = 68.1%) and concreteness (67.6%) were roughly equivalent to one another, with each exceeding chance-level accuracies (33.3%) by twofold. These were followed by intersubjectivity, valence, and embodiment variables with accuracies all above 50%, while animacy (48.8%) explained the data the least. To verify that the specific predictive directions we proposed for these models were indeed more dominant than their opposite, e.g. Abstract→Concrete, we ran these models by reversing their predictive arrows. In each of these cases, we found that the predictive accuracy was substantially worse and closer to chance: animacy (44.2%), concreteness (25.4%), embodiment (40.5%), externality (25%), intersubjectivity (35.9%), valence (37%). These results support the view that historical metaphorical mappings are systematic rather than arbitrary, such that words are more likely to be extended from some domains toward other domains, compared to the re-

Table 1: Model predictions and examples of model-predicted metaphorical sense mappings in English.

Model prediction	Existing sense (rating on dimension: 1-7) → New sense (rating)	Attested word	Beginning of new sense
Animacy			
Animate→Inanimate	Birds(6.7) → Night(1.7)	owl	1400-1450
	Baby and young person(6.7) → Courage(2)	lad	1550-1600
Concreteness			
Concrete→Abstract	Reflection (of light)(6) → Virtue(1.6)	clear	1350-1400
	Plant(6.8) → Intellect(1.8)	vegetary	1550-1600
Embodiment			
Embodied→Disembodied	The human body(6.8) → Kinship and relationship(3.2)	fleshly/flæsclic	Old English
	Bodily tissue(6.1) → Individual colours(2.1)	incarnate	1500-1550
Externality			
External→Internal	Light(6.4) → Thought(1.1)	reflect	1550-1600
	Fireworks(6.7) → Esteem(1.4)	sky-rocket	1850-1900
Intersubjectivity			
Agreeable→Disagreeable	Place and position(5.2) → Supernatural(3.1)	presence	1650-1700
	Mathematics(5.9) → Wisdom(3.5)	calculative	1750-1800
Valence			
Neutral→Valenced	Relative position(4.1) → Excitement(6.4)	up	1300-1350
	Granular texture(3.6) → Moral evil(1.3)	dusty	1600-1650

verse. The fact that the variables we considered helped predict the direction of metaphorical mappings also suggests that the communicative and cognitive factors we described constrain metaphorical extension.

Next, we aimed to explore whether the variables we assessed act in complementary ways, and account for unique variance in explaining directions of metaphorical extension. Specifically, we performed a correlation analysis and a follow-up residual predictive analysis. Figure 1b shows the inter-correlations between all pairs of variables. Externality and concreteness were most strongly correlated (Pearson $r = 0.84$). We thus expected these variables to make similar predictions about mapping directions and for one to make little independent contribution to explaining the directionality of metaphorical mappings over the other. Intersubjectivity was most strongly correlated with both externality ($r = 0.44$) and concreteness ($r = 0.47$) among the remaining pairwise correlations. Animacy, embodiment and valence were generally less correlated with other variables.

To take into account these correlations, we applied a residual prediction procedure to analyze the total variance explained from all of these variables as follows. First, we found the variable that best predicted the directions of all available metaphorical mappings. We then iteratively searched for the variable that best predicted the remaining mappings until there existed no further variables. Figure 1c summarizes the total variance explained via this procedure. Strikingly, the variables together accounted for over 90% of all available mappings. Externality stood out as the single most accurate predictor, explaining 68% of mappings. Embodiment (14%) and valence (5%) explained the majority of the remaining data not predicted by externality, while animacy accounted for 2.3%. Finally, concreteness and intersubjectivity both ex-

plained no more than 1% of data. This is not to suggest that these latter variables play no role in explaining the data, but only that their power in explaining the residual variance was likely subsumed by the externality variable, with which they were highly inter-correlated.

Together, our findings suggest that externality, embodiment, and valence define a relatively orthogonal space that can help explain the directionality of the large majority of metaphorical mappings in the historical dataset. These findings provide the first large-scale evidence for the idea that metaphorical senses have been developed in systematic and constrained ways.

Synchronic asymmetry in metaphorical mappings. Our analysis so far has focused on metaphorical mappings at different historical points, but the outcome of these diachronic changes should also be reflected in more global, time-independent asymmetries as to which domains tend to be the sources of metaphorical extension, and which the targets. To explore this further, we examined the extent to which the variables we proposed would explain the asymmetry of semantic domains. Namely, can we predict which domains tend to be sources and which tend to be targets considering all available metaphorical mappings across time points?

To address this question, we first needed to specify a quantifiable indicator of source-target asymmetry for each semantic domain. To do so, we created an asymmetry index (*AI*), formally defined as the difference between two probabilities:

$$AI = p(source) - p(target) \qquad (1)$$

Here $p(source)$ is the probability of a domain serving as a source of metaphorical mappings across all time points in history (i.e. a synchronic summary of "source" strength), and
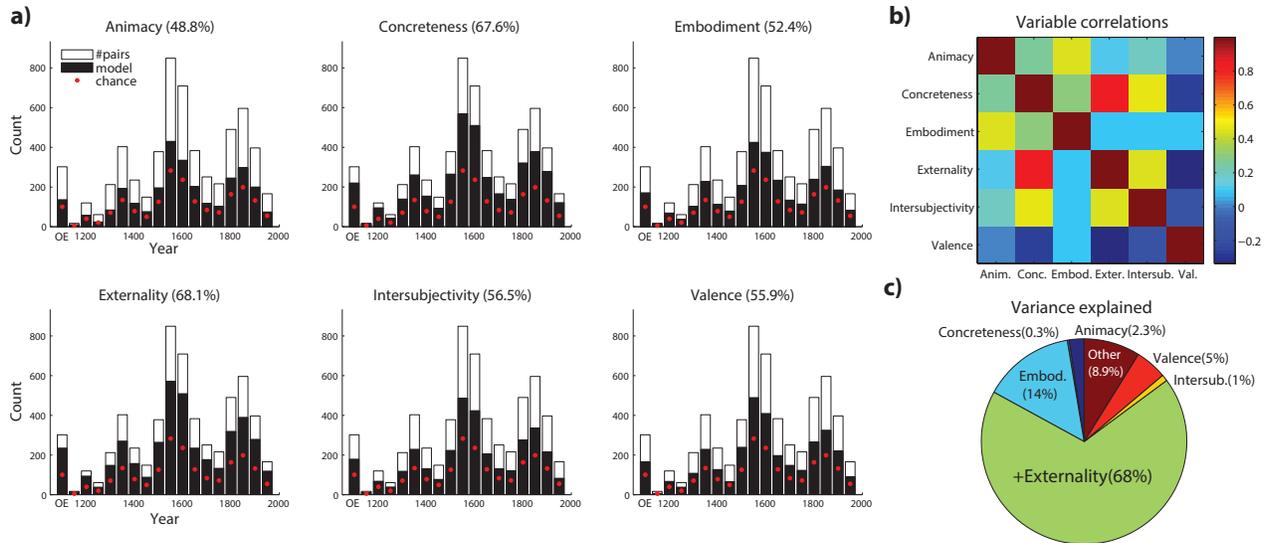
Figure 1: Summary of results on predicting historical order of metaphorical sense mappings. a) Prediction from individual variables. "OE" stands for Old English. b) Correlations among the variables. c) Variance explained by the variables.

$p(target)$ is the probability of a domain being a target. A high positive value of *AI* indicates a strong source domain, whereas a low negative value of *AI* indicates a weak source (or strong target) domain. We computed these probabilities by summing all cases of metaphorical mapping where a domain served as a source or target respectively in history and normalized them by the total number of mappings. Table 2 lists the strongest source and target domains in the dataset based on this measure.

Having obtained the asymmetry indices, we then correlated the mean empirical ratings of each variable with the degree of source-target asymmetry in these domains. Figure 2a summarizes the results. All variable ratings except for those of animacy correlated significantly with the asymmetry indices ($p < 0.01$ for the five predictive variables). Specifically, externality (Pearson $r = 0.46$) and concreteness ($r = 0.42$) showed the highest strengths of correlation, similar to their leading roles in predicting direction of metaphorical mappings. Intersubjectivity ($r = 0.22$) and embodiment ($r = 0.13$) were moderately correlated with the AI. These results suggest that domains that are external, concrete, embodied, and share a conceptualization across people, tend to serve as the source of metaphorical mappings across history. Finally, valence negatively correlated ($r = -0.19$) with the asymmetry index, suggesting that more valenced domains tend to be the target, also consistent with our previous findings.

Again taking into account the intercorrelated nature of the variables, we performed a multiple linear regression to examine the relative contributions of the six variables. We expected the most important (and orthogonal) predictor ratings to be weighted the highest when regressed against the asymmetry index. Figure 2b shows the relative contributions of these variables as reflected by their weights. The regression fit has a Pearson $r = 0.42$ ($p < 0.001$). The candidate with the

Table 2: Listing of the strongest source and target domains based on the source-target asymmetry index.

Strongest sources	**Strongest targets**
Textiles	Excitement
Supernatural	Pride
Digestive organs	Anger
Hardness	Hatred & hostility
Softness	Bad
Ruminants	Behaviour & conduct
Cultivated plants	Money
Wetness	Literature
Darkness	Fear
Solidity & density	Vigorous action & degrees of violence

largest absolute weight was externality ($p < 0.001$), followed by embodiment ($p < 0.09$) and valence ($p < 0.26$) although neither contributed significantly beyond externality. This result is largely consistent with the previous results where we found that the same set of variables accounted for most of the variance in in explaining metaphorical mapping directions during different historical epochs.

Taken together, our diachronic and synchronic analyses suggest that metaphorical sense extensions are highly systematic and support efficient compression of emerging senses along a compact set of variable dimensions.

Conclusion

The present study focused on one of the principal mechanisms of polysemy generation: metaphorical sense mapping. We found that mappings recorded in a large historical database of English can be explained by a compact set of variables. Our investigation lends convincing support that
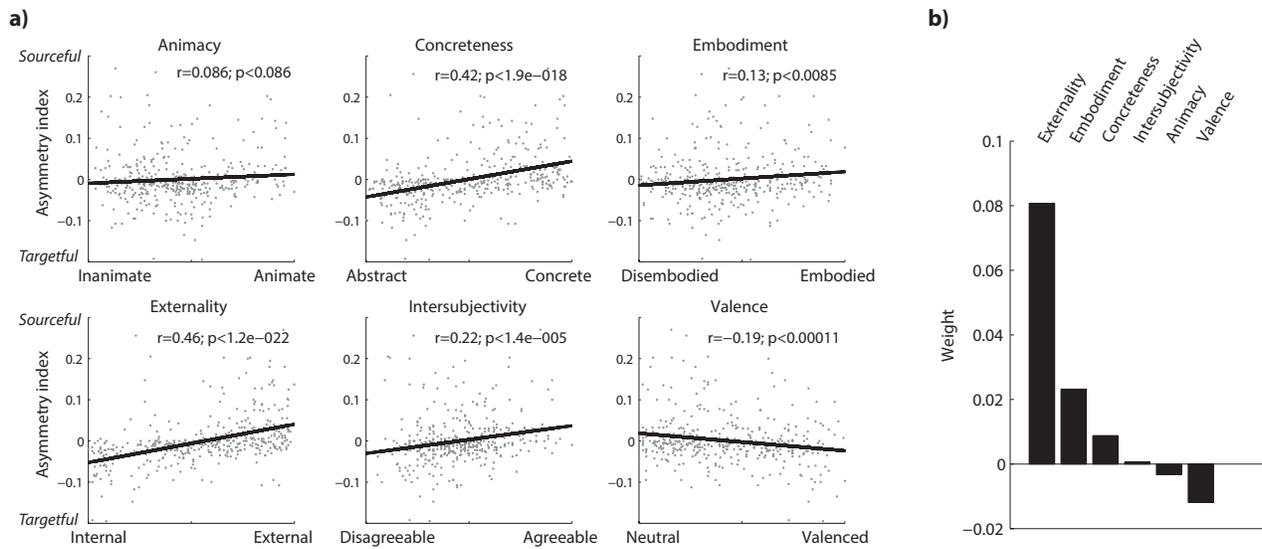
Figure 2: Summary of results on predicting source-target asymmetry of domains. a) Prediction from individual variables. b). Relative weighting of the variables in multiple linear regression.

the evolution of polysemy proceeds in systematic ways. It also offers a novel perspective on the nature of metaphor that bridges theories of conceptual metaphor and language evolution. That is, metaphor provides a cognitive device for compressing emerging concepts into existing words, hence facilitating a finite, learnable lexicon.

Our work also opens many questions for future research. For example, we assumed sense change at the domain level, whereas a full account should explain semantic change at the word level. Future work can also explore cognitive machineries other than metaphor that support polysemous sense extensions, the socio-cultural factors that explain the rise or fall in communicative needs for different semantic domains, and the degree to which these findings generalize beyond English. The current work has provided an empirical approach for in-depth explorations into the nature of meaning.

Acknowledgments

References

Blank, A., & Koch, P. (1999). *Historical semantics and cognition*. Walter de Gruyter.

Brugman, C. M. (1988). *The story of over: Polysemy, semantics, and the structure of the lexicon*. Taylor & Francis.

Christian, K., Roberts, J., Samuels, M., Wotherspoon, I., & Alexande, M. (2015a). *The Historical Thesaurus of English, version 4.2*. Glasgow: University of Glasgow. (http://historicalthesaurus.arts.gla.ac.uk/)

Christian, K., Roberts, J., Samuels, M., Wotherspoon, I., & Alexande, M. (2015b). Metaphor Map of English. In *Mapping Metaphor with the Historical Thesaurus*. Glasgow: University of Glasgow. (http://mappingmetaphor.arts.gla.ac.uk)

Geeraerts, D. (1997). *Diachronic prototype semantics*. Oxford University Press.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. *Metaphor and Thought*, *2*, 164–201.

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. *The Handbook of Language Emergence*, *87*, 237.

Roberts, J., Christian, K., & Grundy, L. (2015). *A Thesaurus of Old English*. Glasgow: University of Glasgow. (http://oldenglishthesaurus.arts.gla.ac.uk/)

Silverstein, M. (1976). Hierarchy of features and ergativity. In R. Dixon (Ed.), *Grammatical categories in Australian languages*.

Sweetser, E. (1991). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure* (Vol. 54). Cambridge University Press.

Traugott, E. C. (2003). From subjectification to intersubjectification. *Motives for Language Change*, *124*, 139.

Ullmann, S. (1957). *The principle of semantics*. Glasgow: Jackson, Son & Co.