Do words reveal concepts?

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

Eef Ameel (eef.ameel@psy.kuleuven.be)
Department of Psychology, University of Leuven, 3000 Leuven, Belgium,

Silvia Gennari (s.gennari@psych.york.ac.uk)
Department of Psychology, University of York, York, England

Mutsumi Imai (imai@sfc.keio.ac.jp) and Noburo Saji (nons@sfc.keio.ac.jp)
Department of Environmental Information, Keio University at Shonan-Fujisawa, Kanagawa, Japan

Asifa Majid (asifa.majid@mpi.nl)
Max Planck Institute for Psycholinguistics, 6500AH Nijmegen, The Netherlands

Abstract

To study concepts, cognitive scientists must first identify some. The prevailing assumption is that they are revealed by words such as *triangle*, *table*, and *robin*. But languages vary dramatically in how they carve up the world by name. Either ordinary concepts must be heavily language-dependent or names cannot be a direct route to concepts. We asked English, Dutch, Spanish, and Japanese speakers to name videos of human locomotion and judge their similarities. We investigated what name inventories and scaling solutions on name similarity and on physical similarity for the groups individually and together suggest about the underlying concepts. Aggregated naming and similarity solutions converged on results distinct from the answers suggested by any single language. Words such as *triangle*, *table*, and *robin* help identify the conceptual space of a domain, but they do not directly reveal units of knowledge usefully considered "concepts."

Keywords: concepts; naming; cross-linguistic diversity; universality; locomotion

Introduction

Concepts have been said to give human experience stability (Smith & Medin, 1981), to hold our mental world together (Murphy, 2002), and to provide the foundation of human learning (Bloom, 2004). Fodor (1998) considered concepts so fundamental to cognition that he declared that the heart of a cognitive science is its theory of concepts. If concepts are so important, then cognitive scientists need to be able to identify concepts to study. We ask here how concepts are to be found, and in particular what role words can play in identifying them.

The prevailing assumption has been that many important concepts can be easily identified because they are revealed by words – in fact, for many researchers, the words of English. English nouns such as *hat*, *fish*, *triangle*, *table*, and *robin* are used to identify concepts in work encompassing not only the adult concepts literature but developmental work (e.g., Carey, 2009), computational models (e.g., Rogers & McClelland, 2004), conceptual combination (e.g., Hampton, 1997), and neuroscience (.e.g., Mahon & Caramazza, 2007).

But from a cross-linguistic perspective, this approach is startling. There are many possible ways to map between words and the world, and languages vary dramatically in how they carve up the world by name. Substantial cross-linguistic variation has been documented in domains including color, causality, mental states, number, body parts, containers, motion, direction, spatial relations, and terms for acts of cutting and breaking and of carrying and holding (see chapters in Malt & Wolff, 2010, for illustrations).This variation occurs even in concrete domains labeled by nouns, where structure in the world might seem most likely to provide universally recognized groupings captured by words. *Hand* vs. *arm*, *bottle* vs. *jar*, or *dish* vs. *plate* may seem to English-speakers to be self-evident distinctions based on obvious discontinuities in the distribution of properties in the world, but not all languages observe these same distinctions (e.g., Malt et al., 1999; Majid, Enfield, & Van Staden, 2006).

In light of the documented diversity, there are three possibilities for the relationship between words and concepts. The first is that the words of a language do effectively reveal much of the stock of general-purpose concepts that a person holds. Given cross-linguistic variability in naming patterns, this possibility implies that word-learning creates much of the language user's non-linguistic representations of the world. Under this scenario, it is not possible to hold that any substantial stock of basic concepts is shared across speakers of different languages, since the language-specific sets will be substantially different from one another.

The second possibility is that concepts are dissociated to some notable extent from the large differences in naming patterns, and it is therefore impossible to use words to identify concepts. After all, much learning about the world comes from direct interaction, rather than through language. Non-linguistic representations may be substantially shared (while still allowing that language could have some influence on them). Crucially, if linguistic and non-linguistic representations are distinct and only loosely linked, then the words of a language cannot routinely and straightforwardly be used to identify a person's concepts.

The third possibility is that the relation of words to concepts is not as straightforward as current practice assumes (also suggested by the second possibility), but still, if examined in the right way, words may reveal something useful about conceptual representations shared across speakers of different languages. By applying more sophisticated techniques to extract structure from language data, it may be possible to discern shared elements of meaning that indicate constraints on cross-language variability and reflect common underlying aspects of how knowledge is represented.

Here, we assess the relation of words - specifically, verbs for human locomotion - to conceptual representations. The data discriminate among the three possibilities just described and address what an appropriate use of words is for researchers whose interest is in underlying non-linguistic representations rather than in knowledge about the word meanings of a particular language.

The Studies

Speakers of four languages named instances of human locomotion such as walking, running, shuffling, and jumping, and made similarity judgments about them. Human locomotion provides a useful case study because there are reasons to expect both commonalities and differences in how speakers of different languages understand and name the domain. Using film clips of biomechanically defined instances of walking and running, we (Malt et al., 2008) found that English, Dutch, Spanish, and Japanese speakers all sorted the walking and running clips into separate piles on the basis of similarity and also drew a lexical distinction between them. We also found that speakers of English and Dutch named more distinctions within the two biomechanical gaits than speakers of Spanish and Japanese. Together, these considerations mean that it is useful to ask whether there is some shared understanding of more varied instances of locomotion, and whether there is any analysis of the potentially diverse naming patterns across languages that can reveal this shared understanding. The current studies again used speakers of English, Dutch, Spanish, and Japanese. Although the first three all belong to the Indo-European family, their histories are different enough that the languages have shown substantial variation in naming patterns in other domains (e.g., containers: Malt et al., 1999) as well as in naming of locomotion for the more restricted set of exemplars in Malt et al. (2008).

Study 1: Using Names to Look for Concepts

We presented video clips of varied human locomotion and asked native speakers of English, Dutch, Spanish, and Japanese to name the motion in each. We first determined the name inventories for the four languages to see what they suggest as the basic concepts for human locomotion. We then performed scaling analyses on the naming data of individual groups to look further for common patterns underlying the names produced. Last, we combined the naming data of the four groups to see if greater coherence arises from the aggregated information.

Method

Thirty English speakers were recruited at Lehigh University, U.S.; 22 Spanish speakers at Comahue National University, Argentina; 26 Dutch speakers at the University of Leuven, Belgium; and 25 Japanese speakers at Keio University, Japan. To generate the stimulus videos, we selected all the verbs that named gaits of an individual moving forward, upright, on a trajectory from a list of over 250 English verbs of manner of movement provided by D. Slobin. We added familiar gaits done in place (e.g., walking and running in place), walking in high heels, and eight variants of locomotion suggested by Japanese and Argentinean informants that were not covered by the English terms. An American college student was filmed portraying each gait for 3-4 seconds on outdoor walkway at Lehigh University. The stimulus set was then reduced by selecting the movements that most clearly contrasted with each other. For instance, if the clips filmed in response to *amble* and *saunter*, or *strut* and *swagger*, looked very similar to us, we kept only one of each. The final set contained 36 clips. Figure 1 shows sample frames from four clips.

Figure 1. Sample frames from clips.

The clips were embedded in a web page. Participants read instructions in the relevant language telling them that they would see a series of video clips, and that they should watch each one carefully and type into the response box the word or phrase that best described what they saw in the clip. Following the instructions were the 36 clips, each with a response box preceded by the words, "What is the woman doing? She is…." or their translation.

Results and Discussion

Name Inventories We determined the name inventories for each language to see what concepts are implied if names are taken to directly identify concepts, as is the common practice in the literature. We counted as an instance of a name all surface forms containing the same root word(s) labeling a manner of movement and then determined the dominant (most commonly produced) name for each clip in each language. All names dominant for at least one clip in a language are given in Table 1. This tally makes clear that if there are universally shared locomotion concepts, the name inventories do not directly reveal what they are. The different languages provide different answers about what that set would be.

Table 1: Inventory of mono- and multi-morpheme names

Language			
English	Dutch	Spanish	Japanese
creep	hinkelen	caminar	aruku
gallop	huppelen	correr	hashiru
hop	joggen	marchar	**sukippu-suru**
jog	lopen	saltar	**ashibumi-suru**
jump	marcheren	trotar	**kenken-suru**
leap	rennen		**koushin-suru**
march	slenteren		**janpu-suru**
run	sluipen		
skip	springen		
stomp	stappen		
walk	wandelen		
shuffle			
tiptoe			
power walk			

Note. **Boldface** indicates multi-morphemic terms.

Individual Languages' Naming as Similarity Data We next asked whether commonalities emerge from the naming data of the four languages if we make use of the full set of names produced by all participants. We created name similarity matrices that reflect the extent to which each pair of objects received the same name across speakers of a language. We assigned, for each participant, a 0 or a 1 to each possible pair of clips according to whether the person gave the two clips a different name or the same name. We constructed the similarity matrix for each language group by summing the distance values for each of the 630 possible pairs across the participants in that language group. This use of the data is similar to using confusion matrices as similarity data (e.g., Shepard & Chang, 1963).

We first correlated the name similarity matrices for each pair of languages to give an overall sense of the correspondence in the naming patterns, using the Mantel test for correlation of matrices. Table 2 shows that these correlations are all substantial and significant. The full patterns of name use, while diverse, still share some substantial commonalities.

Table 2: Between-group correlations of naming matrices

	English	Dutch	Spanish
English			
Dutch	0.82		
Spanish	0.69	0.65	
Japanese	0.76	0.76	0.79

Note. All correlations are significant at $p < .0009$.

We then carried out multi-dimensional scaling (MDS) on the matrix for each language with Additive Tree clusters drawn on the solutions to help interpret the results (Sattath & Tversky, 1977). Due to space constraints we do not present the solutions here but note that they showed both similarities and distinct differences. For all four languages, the X-axis was interpretable in terms of the basic biomechanical distinction between elastic, bounce-and-recoil gaits (running, hopping, jumping, etc.) and the pendulum gaits where one foot is on the ground at all times (walking, striding, etc.). The Y-axis for the most part seemed to reflect a dimension of speed and/or aggressiveness, but the Japanese solution less clearly conformed to this possibility. The Addtree clusters reinforced the idea that the biomechanical distinction is salient for two of the languages – Dutch and Japanese – which had similar top-level clusters separating essentially the same clips. English and Spanish clusters were less like the Dutch and Japanese results: for English, running actions clustered with pendulum motions at the top level, and for Spanish, walking backwards and several forms of marching combined with bounce-and-recoil motions, as well as walking in place. In the next level of clusters within these top-level clusters, each language more or less separated the faster/more aggressive pendulum actions from slower, more cautious pendulum actions, but the exact composition of the clusters was variable. The Dutch solution, in particular, did not honor this separation as much as the others.

These solutions indicate that the naming patterns of the four languages reflect a shared orientation to the same dimensions of the movements. This outcome supports the idea that speakers of the four languages may have more in common in their perception of the domain than their name inventories indicate. In light of the variability of the Addtree clusters across the four solutions, though, it remains difficult to specify exactly what could be identified as shared discrete concepts in the traditional sense.

Aggregated Naming as Similarity Data Last, we created an MDS solution combining the naming data of all four language groups. MDS looks for commonalities in the data, and to the extent that it finds them, produces a coherent solution. If a coherent solution emerges, this result would support the idea of a shared conceptualization of the domain while underscoring the inadequacy of individual words of a single language to reveal it.

We carried out the MDS using a stacked name similarity matrix consisting of 36 columns (the clips) and 4 x 36 (language x clips) rows. Stacking allows the program to compute different weights for each language for the dimensions extracted. Figure 2 shows the two-dimensional solution, with clusters provided by Addtree. This solution shows a neat horizontal separation of the bounce-and-recoil motions (toward the right) from the pendulum motions (toward the left). The vertical dimension appears to reflect speed and aggressiveness of the actions, with slower/less aggressive actions toward the top and faster/more aggressive ones toward the bottom.

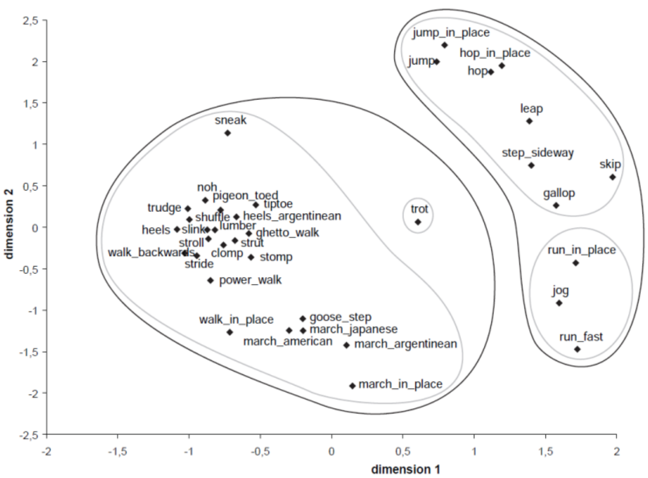

Figure 2. MDS solution based on the four languages' stacked naming data. Clip names refer to names bestowed by the experimenters, not names produced by participants.

At the top level of clustering, the bounce-and-recoil motions are separated from the pendulum-based ones, with the exception of the TROT clip falling into the pendulum cluster. Within these clusters, sub-clusters separate the running clips from the other bounce-and-recoil actions, and separate the true pendulum motion clips from the intermediate TROT clip. These clusters are thus readily interpretable, although they do not seem to map directly onto the words of any of the languages.

Thus, the naming data when aggregated across the four languages provides more indication of a systematic conceptual space than looking at scaling solutions of the four languages individually. This shared space emerges out of the noisiness of the individual name inventories, which make different distinctions and numbers of distinctions. Because MDS can discover commonalities in data but it cannot invent them, the simplicity of the solution is evidence in favor of a shared underlying understanding of the domain. At the same time, though, if the clusters identified by Addtree are taken to indicate discrete concepts within this space, they do not seem to be picked out by words of the languages.

Conclusions from Using Names to Look for Concepts

These analyses demonstrate that languages differ in what their name inventories would tell us the concepts for the domain are. If there are shared basic concepts, then the words of any one language do not directly reveal what they are. Despite the diversity evident in the name inventories, other ways of analyzing the data provide more evidence of commonalities underlying the naming patterns. In particular, scaling of the combined naming data of four languages produces a coherent and interpretable solution, suggesting a shared orientation to certain dimensions of the space. Still, the clusters within the scaling solution indicated by Addtree do not neatly correspond to those labeled by the names of the languages, raising questions about what, if anything, can be identified as discrete concepts in the traditional sense.

Study 2: Using Judgments of Physical Similarity to Look for Concepts

We collected judgments of the physical similarity of the actions in the video clips. We can evaluate to what extent those perceived similarities are shared across speakers of different languages, and to what extent they are related to the naming patterns of the participant's languages. Although physical attributes most likely do not exhaust conceptual knowledge in this domain, they are a large component of it.

Method

Twenty English speakers were recruited at Lehigh University, U.S.; 20 Dutch speakers at the University of Leuven, Belgium; 15 Spanish speakers at the Bariloche Atomic Centre and Balseiro Institute, Argentina; and 24 Japanese speakers at Keio University, Japan. Stimuli were the same 36 video clips of human locomotion used in Study 1. A computer program presented the clips in a 6 x 6 array on a computer screen, with each clip running in a continuous loop. Participants sorted the clips according to the physical similarity of the actions by dragging and dropping clips (still running) into boxes on the right side of the screen, creating as many boxes as they wished. Following Boster (1994), to reduce variability in the number of piles created, if they had created five or fewer boxes in the first sort, they were then asked to divide the boxes further. If they had created more than five, they were asked to combine boxes.

Results and Discussion

For each participant, a distance of zero reflects the case where two clips are not in the same pile in either sort; a distance of one reflects the case where they are in the same pile in one sort but apart in the other, and two reflects the case where they are in the same pile in both (Boster, 1994). We constructed a similarity matrix for each language group by summing the distance values for each of the 630 possible pairs of clips across its members. Table 3 shows that the similarity judgments of the four groups are strongly correlated. The correlations among the sorting matrices are significantly higher than those for naming were with the exception of Dutch to Japanese, *p*s $< .0001$. Speakers of the four languages are more alike in how they sort the gaits based on similarity than in how they name them.

Table 3: Between-group correlations of sorting matrices

	English	Dutch	Spanish
English			
Dutch	0.87		
Spanish	0.84	0.83	
Japanese	0.89	0.75	0.89

Note. All correlations are significant at $p < .0009$.

An aggregated MDS solution was constructed as for the aggregated name similarity by stacking the matrices of the four groups. The two-dimensional solution is shown in Figure 3, overlaid with an Additive tree cluster analysis.

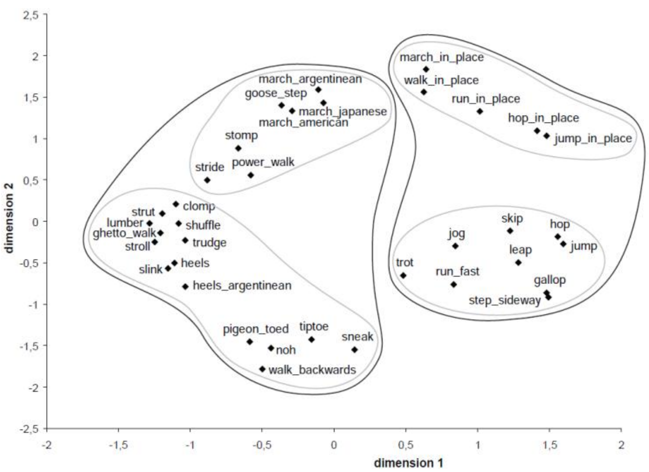

Figure 3. MDS solution based on four language groups' stacked sorting data. Clip names refer to names bestowed by the experimenters, not names produced by participants.

As for naming, the horizontal dimension of the MDS solution is readily interpreted in terms of the biomechanical gait distinction described earlier. Toward the left are gaits with the pendulum motion. To the right are the ones with the elastic bounce-and-recoil motion. Toward the middle are some stimuli having knees lifted high but lacking real bouncing off the ground (sneaking, walking in place, various forms of marching, etc.). Also as for naming, the vertical dimension seems to reflect something about the speed/aggressiveness of the actions, although on the left it seems to be that faster, more aggressive motions are at the top and on the right this pattern reverses. Allowing MDS to use three dimensions (not pictured here) this reversal between left and right goes away, with the extra dimension allowing a speed dimension to separate itself from an aggressiveness dimension. Consistent with the importance of the biomechanical distinction and speed/aggressivness dimensions, the two top-level clusters that Addtree identifies largely separate the gaits with the bounce-and-recoil motion from those with a pendulum motion. Within the left-hand large cluster, the sub-clusters may be characterized as distinguishing groups of pendulum motions that are more slow and cautious (lower cluster) from those that are more fast and aggressive (top cluster). On the right, one sub-cluster encompasses all the bounce-and-recoil actions that involve forward (or sideways) motion. The other encompasses those that involve bounce-and-recoil movements in place (hopping, jumping, and running in place), plus marching in place and walking in place.

The sorting solution points to largely the same sets of dimensions as the combined naming data did, despite the somewhat different spatial layout and clustering. There was one notable exception to the close correspondence with the naming solution. In the sorting data, the five clips that entail motion in place (hopping, jumping, running, marching, and walking in place) do form a sub-cluster, but in the naming they did not. It seems that being carried out in place is a salient physical property of the actions, even though none of the four languages honors it with a basic level name that sets the actions apart on that basis.

Conclusions from Using Physical Similarity Judgments to Look for Concepts Speakers of our four languages make closely corresponding judgments of the similarity among the 36 actions. Furthermore, there is higher correspondence among language groups on this sorting task than on naming, consistent with other domains (e.g., Malt et al., 1999). The data indicate that the understanding of locomotion is more shared, and more tied to the perceptual and biomechanical experience of the domain, than the varied word inventories would imply. Again, the data argue against assuming a close alignment of words with concepts.

Although the sorting data clearly indicate the inadequacy of the inventories of individual names to reveal shared understanding of the locomotion domain, the scaling solutions do show a marked resemblance to the combined naming data. This outcome suggests that both reflect some deeper commonality in the dimensions of locomotion space that people find salient. At the same time, and despite some proposals in the literature that concepts fall directly out of similarity space (e.g., Rogers & McCelland, 2004), the scaling solution still does not directly reveal concepts in the sense of bounded units of knowledge stored in long-term memory. Would they be the top-level distinctions in the cluster solution that seem to be grounded in the biomechanical contrast between pendulum and bounce-and-recoil motions? Would they be at the level of the next set of clusters, which corresponds better, but far from perfectly, to the words of the languages? Or might they be at some other level? And what about the discrepancies that do exist between the sorting and naming solutions? In other words, how can we identify exactly what the basic concepts are that should be the subject of investigation by those researchers who want to study them? We consider these points below.

General Discussion

The data indicate that the relation of words to concepts is not straightforward. Pervasive linguistic diversity, amply documented in other research but rarely taken into consideration within "concepts" research, is by itself cause for serious concern. It would still be possible to preserve a commitment to word-concept alignment by subscribing to a strong version of the idea that language shapes thought. However, the evidence here from aggregated naming and from sorting suggests that any conceptual differences between speakers of different languages are less than implied by the large differences in word inventories. This outcome is consistent with findings in other domains and with the observation that attention to aspects of the world is likely shaped by multiple forces that include but are not limited to language. It seems unavoidable to conclude that researchers need to stop relying on the word inventories of English, or any other single language, to know what constitutes the concepts of a domain.

Despite the complex nature of the relationship between language and concepts, our data suggest that it is still possible to use linguistic data to gain insight into something more fundamental about the nature of conceptual space. Combining naming information across languages does seem to provide useful information, because the aggregate allows the commonalities to emerge over the "noise" of individual language idiosyncracies. However, for researchers whose usual methodologies entail only members of one language group, obtaining such cross-linguistic data may not be feasible. Developing other methods of avoiding reliance on a faulty word-concept equivalence will be crucial.

The scaling solutions still leave it unclear exactly what units of knowledge should count as the most fundamental, basic concepts of the domain. Different data sets (sorting vs. naming) and different levels of the cluster analysis produce somewhat different potential answers. The difficulty of specifying exactly what constitute the basic concepts suggests that it may be time for psychologists to more radically rethink conceptual understanding of a domain. An alternative approach to understanding where shared and possibly innate elements of mental representation are to be found is represented by the search for smaller units of knowledge such as EVENT, STATE, THING, PATH, PLACE, GOAL, MEANS, and END, etc. (see, e.g., Pinker, 2007). In this type of approach, the goal is not to identify discrete, bounded, and stable units of knowledge stored in long-term memory. Instead, what is identified are the dimensions of experience to which people attend under various circumstances. Such an approach may prove more fruitful than looking for traditional concepts.

Conclusion If the heart of a cognitive science is its theory of concepts, then cognitive scientists need to re-think how to find the concepts. Words can help identify the conceptual space of a domain, but they do not directly reveal bounded units of knowledge that can be labeled "concepts."

Acknowledgments

This work was supported in part by the Max Planck Gesellschaft, the Japanese Ministry of Education grant-in-aid for Scientific Research (#15300088), and research grants from Keio University. Eef Ameel is supported on a postdoctoral fellowship by the Fund for Scientific Research-Flanders. We thank Adam Darlow, Kiri Lee, Yo Matsumoto, Kyoko Ohara, Dan Slobin, Steven Sloman, and Phillip Wolff for useful discussion. Kristine Schuster and Hiroyuki Okada developed the web experiment for Study 1, Naoaki Tsuda created the sorting program for Study 2, and Stephanie Sterrett served as walkway actor. Ludy Cilissen, Lisa Guest, Celina Hayes, and Erin Howard assisted with data collection and analysis.

References

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Boster, J. S. (1994). The successive pile sort. *Cultural Anthropology Methods*, *6*, 7-8.

Carey, S. (2009). The origin of concepts. Oxford: Oxford University Press.

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Oxford University Press.

Hampton, J. (1997). Conceptual combination: Conjunction and negation of natural concepts. *Memory & Cognition, 25,* 888-909.

Mahon, B. Z & Caramazza, A. (2007). The organization and representation of conceptual knowledge in the brain: Living kinds and artifacts. In E. Margolis and S. Laurence (Eds.). *Creations of the mind: Theories of artifacts and their representation*. New York: Oxford University Press.

Majid, A., Enfield, N. J., & van Staden, M. (2006). Parts of the body: Cross-linguistic categorisation. Special issue of *Language Sciences*, *28*(2-3).

Malt, B.C., Gennari, S., Imai, M., Ameel, E., Tsuda, N., & Majid, A. (2008). Talking about walking: Biomechanics and the language of locomotion. *Psychological Science*, *19*, 232-240.

Malt, B. C., 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.

Malt, B. C. and Wolff, P. (Eds.) (2010). *Words and the mind: How words encode human experience*. New York: Oxford University Press.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: The MIT Press.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking.

Rogers, T. T. & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Sattath, S. & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, *42*, 319-345.

Shepard, R. N. & Chang, J.-J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology*, *65*, 94-102.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.