



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at ScienceDirect

Journal of Memory and Language

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

Human locomotion in languages: Constraints on moving and meaning



Barbara C. Malt^{a,*}, Eef Ameel^b, Mutsumi Imai^c, Silvia P. Gennari^d, Noburo Saji^c, Asifa Majid^{e,f}

^a Department of Psychology, Lehigh University, United States

^b Department of Psychology, University of Leuven, Belgium

^c Department of Environmental Information, Keio University at Shonan-Fujisawa, Japan

^d Department of Psychology, University of York, United Kingdom

^e Centre for Language Studies, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, The Netherlands

^f Language and Cognition Department, Max Planck Institute for Psycholinguistics, The Netherlands

ARTICLE INFO

Article history:

Received 5 October 2012

revision received 13 August 2013

Available online 18 October 2013

Keywords:

Word meaning

Semantics

Lexical categorization

Locomotion verb

Cross-linguistic comparison

ABSTRACT

The distinctions between *red* and *yellow* or *arm* and *hand* may seem self-evident to English speakers, but they are not: Languages differ in the named distinctions they make. To help understand what constrains word meaning and how variation arises, we examined name choices in English, Dutch, Spanish, and Japanese for 36 instances of human locomotion. Naming patterns showed commonalities largely interpretable in terms of perceived physical similarities among the instances. There was no evidence for languages jointly ignoring salient physical distinctions to build meaning on other bases, nor for a shift in the basis of word meanings between parts of the domain of more vs. less importance to everyday life. Overall, the languages differed most notably in how many named distinctions they made, a form of variation that may be linked to linguistic typology. These findings, considered along with naming patterns from other domains, suggest recurring principles of constraint and variation across domains.

© 2013 Elsevier Inc. All rights reserved.

Introduction

To English speakers, the contrasts captured by the words *red* and *yellow*, *bowl* and *plate*, *arm* and *hand*, or *in* and *on* seem self-evident. This intuition of inevitability is compatible with the idea that the world contains “a series of discontinuities whose structure and content are seen by all human beings in essentially the same ways...” (Berlin, 1992, p. 9), and that category names label “intrinsically separate things” (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976, p. 383) (see also Anderson, 1991; Hunn, 1977; Rogers & McClelland, 2004).

But cross-linguistic work shows that these contrasts are not nearly as inevitable as they may feel. Languages often differ in the contrasts they lexicalize, a phenomenon documented for properties (such as colors and tastes), objects (such as plants, animals, and human-made objects), and relations (such as those captured by spatial terms and many verbs) (for review and discussion, see, e.g., Bowerman & Levinson, 2001; Malt & Majid, 2013; Malt & Wolff, 2010). Given the pervasive diversity, one could counter that the world must present few pre-existing distinctions to the observer, and differences between languages may be arbitrary and unpredictable.

Most likely, the truth lies in a middle ground. Many studies documenting variability across languages also uncover commonalities. For instance, languages tend to lexically distinguish walking from running gaits (Malt, Gennari, Imai, Ameel, Tsuda, & Majid, 2008; Vulchanova, Martinez, & Vulchanov, 2012), arms from torsos (Majid,

* Corresponding author. Address: Department of Psychology, 17 Memorial Drive East, Lehigh University, Bethlehem, PA 18015, United States. Fax: +1 610 758 6277.

E-mail address: barbara.malt@lehigh.edu (B.C. Malt).

Enfield, & van Staden, 2006), and red from black (e.g., Kay, Berlin, Maffi, & Merrifield, 1997). An account of how human experience in the world is mapped onto word meanings must explain both where shared tendencies come from and how variation arises. The current study contributes to such an account using the domain of locomotion, a domain useful in several ways described later.

Constraints on naming patterns: Evidence to date

Contributions of structure in the world

Most work on concepts and categorization in psychology assumes that structure in the world is the key driving force behind how word meanings develop in languages, and that similar named categories would emerge across languages (e.g., Anderson, 1991; Rogers & McClelland, 2004; Rosch et al., 1976). However, these assumptions are rarely tested. In other disciplines, however, researchers have tackled such issues directly. Anthropological research supports the idea that there are cross-cultural commonalities in named distinctions for plants and animals reflecting discontinuities in property distributions between biological genera (see Malt, 1995 for a review; see also Medin & Atran, 1999). Likewise, naming patterns for exemplars of walking and running show shared tendencies across languages that can be linked to the biomechanical gait distinction (Malt et al., 2008), and naming patterns for human body parts (Majid et al., 2006) have shared tendencies linked to segmentation points of the body. In contrast, for color, given the continuous nature of wavelengths, it seems that the input itself lacks discontinuities in structure or content that would explain observed shared tendencies in color naming. In a middle ground, Malt, Sloman, Gennari, Shi, and Wang (1999) found that for household containers, objects that clustered together in an overall similarity space tended to share a name in each language. Some objects did not fall into clusters, though, and their grouping by name differed more across languages. Together, these studies suggest that naming patterns reflect structure in input where it is present, but different domains may present different degrees of structure.

Contributions of the observer

Of course, cognitive processes intervene between world input and language output. There may be a tendency to partition domains in a way that maximizes similarity within categories and minimizes it across categories (Garner, 1974), as seems to be true for color (Jameson, 2005; Regier, Kay, & Khetarpal, 2007) and spatial terms (Khetarpal, Majid, & Regier, 2009). Within this constraint, different partitionings can be formed as long as they carve out contiguous areas of space. This observation raises the question of whether there are shared biases for using certain features to form partitionings, or whether some naming patterns reflect culture- or language-specific choices in the features used.

If there are shared biases toward certain features, one possibility is that they would be linked to how people interact with the domain. Functional or goal-related features are salient in many task contexts (e.g., Kemler

Nelson, Egan, & Holt, 2004; Lynch, Coley, & Medin, 2000; Ratneshwar, Barsalou, Pechmann, & Moore, 2001). It is not clear to what extent naming patterns reflect such features. Although named categories for plants or animals in principle could be based on utility (e.g., distinguishing toxic from edible), in fact they tend to be general-purpose (e.g., distinguishing birds from fish and pines from oaks based largely on external morphological features), even in traditional non-industrialized cultures (Berlin, 1992; Hunn, 1977). Behavior and relationships to humans do sometimes come into play but infrequently (e.g., in giving a ritually significant animal a special name; Bulmer, 1967, and in the English garden distinction between *weed* and *flower* which has no botanical basis). Similarly, the work on common household objects (Kronenfeld, Armstrong, & Wilmoth, 1985; Malt et al., 1999; see also Pavlenko & Malt, 2011) suggests that languages mainly draw on combinations of size, shape, material, and other physical attributes such as presence of a handle in naming. Ameel, Malt, and Storms (2008) did, however, find some functional features, along with physical ones, in the feature sets that best predicted use of some Dutch container names. Thus, features related to human interaction, including the functions entities serve and the goals they fulfill, may have a role in naming, but when and to what extent remains unknown.

An alternative possibility is that language focuses mainly on the most temporally enduring and externally visible attributes in lexicalized distinctions, since these may be the most readily perceived. This may not serve well as a general conclusion, either, given evidence from action domains implicating more dynamic properties. Presence or absence of movement along a trajectory over time is basic to the distinction between *hold* and *carry* observed across many languages (Saji et al., 2011). Predictability of the locus of separation was the most important dimension in accounting for named distinctions in acts of cutting and breaking across languages (Majid, Boster, & Bowerman, 2008). Our past study of locomotion on a treadmill (Malt et al., 2008) began with the observation that people switch abruptly from walking to running on a treadmill without transitional states (Diedrich & Warren, 1995). Each gait is characterized by a cluster of co-occurring properties that include its characteristic energy requirements, relative phase of the feet, fraction of the stride for which a given foot is on the ground, stride frequency, and stride length (Alexander, 2002; Bennett, 1992). These property clusters create a pendulum-like motion of the legs with one foot on the ground at all times for walking, and an impact-and-recoil motion of the legs with both feet off the ground at the same time at a point in each stride for running. We found that speakers of English, Dutch, Spanish, and Japanese naming video clips of a person on a treadmill all drew a strict lexical distinction between instances of walking and running. Other overtly manifest dimensions such as speed and direction, as well as inferred dimensions such as effort and goals of the agent (e.g., relaxing, exercising, hurrying, acting purposefully), are more constant across the movement than the biomechanical properties (where, e.g., impact is followed by recoil within a single stride). These dimensions could have served as the basis for named distinctions, but there was no evidence that they did. For

several sets of actions studied to date, then, properties and property clusters that are abstracted across time and space seem to be prominent bases for lexical distinctions across languages.

Moving forward in understanding naming patterns across languages

The preceding observations suggest that (a) structure in the world may ground meaning where it exists, but it is not necessarily equally present in all domains or parts of a single domain; and (b) observers are selective in the properties they attend to in exploiting this structure to create named categories. In some domains, static, external morphological features are prominent. In others, the salient features may include ones abstracted over time and space, including both individual properties and property clusters. Less directly visible properties such as object functions or an agent's efforts or goals may play a role in naming, but when and to what extent is not yet well understood. This state of knowledge indicates a need for more domains to be examined in greater detail, with an eye to three issues in particular.

First is the role of structure in the world. Are there only a handful of cases (including biological genera, body parts, and the walk-run distinction) where universally perceived discontinuities in property distribution exist and feed into patterns of lexicalization? Or could strong world structure be pervasive, with the limited structure observed for household objects only possible because artifact properties can be combined by their human creators in many ways? Either way, when present, does perceived structure inevitably serve to ground meaning, or do languages sometimes ignore it and build meanings on other bases?

Second, how do cognitive processes shape the mappings from input to naming patterns? What types of properties are used to create named groupings for languages? Are they the same properties as those salient non-linguistically? And is a single set for naming used throughout a given domain, or are different types of properties used to create different distinctions within a single domain?

The third key issue is how variation arises despite constraints. Languages may differ in which properties they use to divide up a domain by name. Alternatively, cross-linguistic variation may derive from different ways of using the same properties (by making cuts on the same dimensions at different points and/or making different numbers of cuts). If languages do vary in the properties they draw on, are some property types subject to more variation than others? If languages use the same properties, what gives rise to the variation?

Our previous study of locomotion naming (Malt et al., 2008) examined only walking and running on a treadmill to see whether this major biomechanical distinction would be reflected in naming across languages. Here, we investigate naming patterns for a wider range of locomotion exemplars that have more varied properties, produced under more natural conditions. This range allows us to more fully assess perceived structure across the domain, how this structure contributes to naming patterns for locomotion beyond the walk-run distinction, and how the

properties used in naming instances of locomotion are similar and different across languages and across different parts of this domain. Below, we describe how the three issues identified above are addressed to help build a general account of how human experience is mapped onto words.

Locomotion as a domain for addressing these issues

The term “human locomotion” receives definitions ranging from broad (all self-propelled motion by means of limbs) to narrow (walking). Our interest here is in forms of upright human movement across a solid substrate by characteristic movements of the limbs. Locomotion as so defined has several general properties that make it useful for addressing the three key issues outlined above. It plays a crucial role in human lives, since humans locomote many times a day in the course of essential activities. It also provides a domain where the exemplars are broadly, if not absolutely, shared across cultures. Artifact and natural kind exemplars often vary regionally, creating complications in pinning down sources of variation in naming patterns since different languages name different input. Finally, results are relevant to the argument (Gentner, 1982) that verb meanings are likely to vary more across languages than noun meanings because actions are less intrinsically separate from their context than objects are. Cross-linguistic labeling of action has been relatively little studied to date, and findings contribute evidence on this point.

Locomotion also has three characteristics of special interest. First, as already described, there are different potential bases for dividing instances by name. These include clusters of biomechanical properties, readily visible individual properties such as speed and direction, and inferred ones such as effort or goals. The latter, for a broad set of locomotion exemplars, may include not only relaxing, exercising, hurrying, and acting purposefully but also playing, displaying attitudes, and fulfilling roles associated with special occasions or purposes (e.g., in religious or military rituals). As such, this domain provides a chance to see how these different sorts of properties play out across languages. Second, the importance of different forms of locomotion to daily life varies, and in a way that is consistent across cultures. Walking is universally the most common, followed by running, with both used often across many contexts. Other forms such as hopping, jumping, and skipping are used far less and under restricted circumstances, such as in games or getting over an obstacle. One can ask whether the properties used to create named distinctions shift across parts of the domain. For instance, biomechanical properties might dominate in the more important and familiar part, whereas speed, direction, effort, or goals might become more prominent elsewhere within the same domain. This outcome would be expected if words in the more frequently experienced parts are general-purpose and those in more peripheral parts are linked to special purposes. We can also ask if naming patterns vary more across languages in the more peripheral parts, and whether any such variation is linked to lesser perceived domain structure and/or to greater variation in the features used as the basis for named groupings.

Finally, locomotion provides a chance to examine whether structural characteristics of languages contribute to the development of word meanings. Some languages (e.g., Germanic ones) most often encode manner of motion in their verbs (as in English *walk*, *run*, *slide*). Others (e.g., Romance languages, Greek) more often encode path or direction of movement (as in Spanish *subir* [ascend] and *salir* [exit]), although they do use manner verbs in some constructions (e.g., Slobin, 1996; Talmy, 1985). Previous research has focused on whether these linguistic differences have an influence on how scenes are remembered (e.g., Gennari, Sloman, Malt, & Fitch, 2002; Papafragou, Hulbert, & Trueswell, 2008). Of more interest here is whether the difference impacts locomotion verbs themselves. Malt et al. (2008) found that although both types of languages lexically distinguished between walking-type and running-type actions, the manner verb languages made more named distinctions within each biomechanical category (e.g., English *run*, *jog*, *sprint*). One can ask whether languages commonly expressing path in the main verb will have fewer locomotion verbs across the domain (as Slobin, 2004, suggests is true for motion verbs more broadly), while still respecting structural constraints in the terms they do have. A second noteworthy characteristic is that one of our languages, Japanese, has an especially rich set of grammatical devices beyond conventional verbs (described later) for expressing motion. We ask whether this feature influences the named distinctions for locomotion.

Method

Languages

The studies reported here examined naming patterns of English, Dutch, Spanish, and Japanese. The first three languages are in the Indo-European family, with Dutch and English being Germanic and Spanish, Romance. Japanese is most often considered to be an isolate, or perhaps Altaic (Crystal, 1987). Although the first three belong to the same family, their histories are different enough that they show variation in naming patterns in many domains (e.g., Bowerman, 1996, 2005; Majid, Gullberg, Van Staden, & Bowerman, 2007; Majid et al., 2008; Malt et al., 1999), including locomotion for a more restricted set of exemplars (Malt et al., 2008). The typology cross-cuts the family classification, with English and Dutch tending to express manner in the main verb and Spanish and Japanese, path (or path and ground).

Stimuli

We used a larger and more varied set of instances of locomotion than in Malt et al. (2008). To generate the stimuli, we began with a list of over 250 English verbs of manner of movement compiled by D. Slobin.¹ We selected all the verbs naming gaits of an individual moving forward upright on a trajectory via characteristic motor movement of

the legs (eliminating many such as *barge*, *bolt*, *bump*, and *burst* that do not identify a characteristic movement of the legs, and others such as *crawl* and *dive* that do not refer to upright locomotion). For further breadth, we added walking in high heels and some gaits done in place referred to by familiar English expressions such as *walking in place* and *running in place*. We also added five variants suggested by Japanese informants and three by Argentinean informants that the informants felt were familiar in their respective countries but were not covered by the English terms. These actions were conveyed to us via video clips.

An American college student trained in dance acted out each English gait word (e.g., *trot*, *hop*) or phrase (e.g., *walking in place*) along with the additional Japanese and Argentinean actions. For the latter, she studied their video clips and reproduced the actions to the best of her ability. She was filmed on an outdoor walkway, moving from a fixed starting point to a fixed end point (about 3–4 s for each action). Fifty-seven videos were made, including several variants for some verbs. The stimulus set was then reduced by selecting the movements that most clearly contrasted with each other (as judged by the first author and research assistants). For instance, if the clips for *amble* and *saunter*, or *strut* and *swagger*, looked very similar, we kept only one of each. The final set contained 36 clips of varied forms of upright locomotion that represented 63% of those originally filmed and did not correspond directly to the set of terms of any language. Fig. 1 shows sample frames from three clips (see Supplementary Video Part 1, 2 and 3).

Participants

Participants for all tasks were native speakers of the four languages. Some had familiarity with one or more additional languages, but none regularly used a second one. Most were undergraduates or graduate students; a few Spanish speakers in the physical similarity sorting task were instructors. English speakers were recruited at Lehigh University, United States. Spanish speakers were mainly recruited at Comahue National University, Argentina, with a minority from the Bariloche Atomic Centre and Balseiro Institute. Dutch speakers were from the University of Leuven, Belgium. Japanese speakers were from Keio University, Japan. None had participated in the treadmill study (Malt et al., 2008) and each participated in only one task.

Physical similarity judgment task

There is no readily available biomechanical analysis for most human locomotion actions beyond walking and running. Jumping and hopping have been referred to as separate gaits (Alexander, 2002), but there is little available commentary on whether biomechanical property clusters separate those from skipping or from running, or separate walking from marching, and so on. To determine what physical distinctions are salient that may feed into naming patterns, we collected judgments of the physical similarities among the exemplars and performed multidimensional scaling on the judgments.

Due to the number of stimuli, a sorting task was used to derive pairwise similarity values based on the number of

¹ The list was compiled by searching dictionaries and texts along with brain-storming. We thank Dan Slobin for sharing this list.



Fig. 1. Sample frames from three video clips. Row 1: hop; Row 2: stroll; Row 3: march_japanese (names are the experimenter-given identifiers used in Figs. 2 and 3).

participants who sorted each pair of stimuli into the same pile (e.g., Rosenberg & Kim, 1975). Similar to Malt et al. (2008), a computer program presented the clips in a 6×6 array on a computer screen, with each clip running in a continuous loop. Twenty English, 20 Dutch, 15 Spanish, and 24 Japanese speakers sorted the actions according to their physical similarity by dragging and dropping clips (still running) into boxes on the right side of the screen, creating as many boxes as they wished. When all the clips were sorted into boxes, participants performed a second sort. Following Boster (1987, 1994), to reduce variability in the number of piles created, if they had created five or fewer boxes in the first sort, they were asked to divide the boxes further; if they had created more than five, they were asked to combine boxes. No mention was made at any time of names for the actions. Participants were asked to focus on the physical, mechanical qualities of the movements such as how the limbs or parts of the limbs were moving and how the body as a whole was moving. They were further told that they could use any or all aspects of the movements that seemed important to them, but they should be sure to focus just on how the person is moving. Instructions were provided in the relevant language, translated from English by a native speaker.

Naming task

To determine naming patterns for the four languages, and ultimately to look for commonalities and differences and assess what drives each of these, we asked speakers of each language to name the video clips. Thirty English, 26 Dutch, 22 Spanish, and 25 Japanese speakers viewed the 36 video clips embedded in a web page. At the beginning of the web page, participants entered their native language, where they grew up, and their level of proficiency in any additional languages. They then read instructions telling them that they would see a series of video clips, and that they should type into the response box the word or phrase that best described what they saw in the clip. They were given examples of responses they might type if the clip were of a bird or inanimate object moving; the examples included both single-word answers and two-word answers (e.g., *soaring* and *flying fast* for the bird, for the English language version). Instructions also specified that there was no right or wrong answer to each clip and participants should give whatever word or phrase they felt was the best or most natural way to describe the action. Instructions were provided in the relevant language, translated from English by a native speaker. Each clip was

accompanied by a response box preceded by the words, “What is the woman doing? She is...” or their translation. Two random orders of stimuli were used. Because each clip varied little except in the manner of movement portrayed (stated in the instructions as well as apparent upon viewing the first few clips), participants should have been motivated to make their name choices contrastive rather than using broad descriptors such as *going somewhere*.

Results

Tabulation and transformation of raw data preliminary to the theoretical analyses

Similarity matrices from physical similarity sorting

Pairwise physical similarity matrices for each language group were created to use as input to multidimensional scaling analyses. For each participant, the sorting data were tabulated such that zero reflects the case where two clips are not in the same pile in either sort; one is where they are in the same pile in one sort but split apart in the other, and two is where they are in the same pile in both sorts (Boster, 1987, 1994). A similarity matrix for each language group was made by summing the values for each of the 630 possible pairs of clips across individuals.

The physical similarity matrices for the groups were internally consistent, shown by split-half reliabilities (followed by the Spearman–Brown formula) of .92 for Americans, .90 for Dutch, .91 for Spanish, and .92 for Japanese. Table 1 shows that the similarity judgments of the four groups also correlated strongly with one another (calculated using Mantel tests for the correlation of matrices, Mantel, 1967). When the matrices were pooled by summing distances across all four groups, the estimated reliability of the combined sorting data increased to 0.98, demonstrating that the four groups behave as if from a single population (consistent with sorting in other domains; Ameel, Storms, Malt, & Sloman, 2005; Kronenfeld et al., 1985; Malt et al., 1999).

Video clip names

To determine lexicalization patterns of the languages, we tabulated the naming responses for each language group. We regularized variations in the surface form of similar responses. For instance, for a given clip, in English, many participants might have said that the woman is *walking* but a few might have said *doing a walk* and some even just *walks* or *a walk*. We counted as the same name all surface forms containing the same root word(s) labeling a manner of movement. Details of the counting guidelines are given in Appendix B.

Table 1
Between-group correlations of physical similarity sorting matrices.

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

Note: Pearson correlation values were calculated using the Mantel test for correlation of matrices, based on 1000 permutations. All correlations are significant at $p < .0009$.

Table 2
Names dominant for at least one clip for a language.

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				

Besides regularizing surface forms, we had to consider what portion of the complete responses to look at. For instance, an English response might be *walking*, *fast walking*, or *walking lightly with her arms swinging*. Because the goal is to understand patterns of lexicalization across languages, our interest was in conventional basic-level names. The modifiers in the second two cases create ad hoc descriptions that are combinatorily infinite, so phrases of this type were counted as instances of *walk*, not as separate names. We did, however, include two-word responses that constituted fixed, conventional, and common names. For Japanese, in particular, limiting consideration to single words risks excluding certain common action labels. Japanese makes use of a construction consisting of an action noun plus the light verb *suru* (‘do’) for some actions. Such phrases are similar to English *doing a jump* or *doing a march step*, except that some are fixed, conventional, and common (e.g., *kenken suru* for English *hop*). We considered each clip’s “dominant” name to be whatever name was produced most often across the participants in the language group. Clips for which fewer than 30% of respondents agreed on a name were considered “mixed”. Table 2 provides the names dominant for at least one clip for each language.² The number of such names ranges from 5 for Spanish to 14 for English. Appendix C provides the dominant name for each clip for each language along with the proportion of speakers who produced that name.

Similarity matrices from naming

Pairwise name similarity matrices for each language group were created to use as input to name-based multidimensional scaling analyses. Many of the clips did not

² Differences in name inventory should not be taken to imply that speakers of the different languages necessarily differ in the degree to which they notice differences among the actions presented. The full responses to stimuli show that speakers of all the languages are, in various ways, making more distinctions among the actions than those reflected in the basic level names. For instance, Spanish speakers often used modifiers to discriminate among cases of *saltar* done on one foot versus two (distinguished with *hop* vs. *jump* for English speakers), and speakers of all the languages used location modifiers to point out the actions done in place. The high level of agreement in the physical sorting also indicates that speakers of all languages noticed similar distinctions among the actions.

Table 3
Between-group correlations of name similarity matrices.

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

Note: Pearson correlation values were calculated using the Mantel test for correlation of matrices, based on 1000 permutations. All correlations are significant at $p < .0009$.

produce 100% name consensus within a language group, so even if one language has different dominant names for two clips (e.g., *walk* and *stroll*), and another gives them the same name (e.g., *caminar*), some speakers of the first language may have produced the same name for both (e.g., *walk*), pointing to a perceived similarity between them. This name variation was used to create name similarity matrices by first assigning, 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 (following the name coding criteria described in Appendix B). As for the sorting data, this procedure produces 630 pairs. The name similarity matrix for each language group was constructed by summing the entries for each of the pairs across the participants in that language group. This method is similar to using confusion matrices as similarity data (e.g., Rothkopf, 1957; Shepard & Chang, 1963).

The within-group reliabilities of the name similarity matrices were .94 for Americans, .86 for Dutch, .96 for Spanish, and .94 for Japanese participants (estimated using the split-half technique followed by the Spearman–Brown formula). These values show that groups were internally consistent. When the matrices were pooled by summing distances across all four groups, the estimated reliability of the combined name similarity data increased to 0.98. Parallel to the sorting data, we 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 3 shows that these correlations are all significant, indicating that the patterns of name use share substantial commonalities.

Constraints

In this section, we first examine the perceived physical similarity structure of the domain and then, more broadly, what properties or dimension values constrain naming patterns.

Structure in perceived physical similarity

We used the physical similarity sorting data to evaluate to what extent there is perceived physical structure in this domain beyond the walk-run discontinuity observed by Malt et al. (2008), and what kinds of physical features create perceived structure. To do this, we created a visual representation of the data using multidimensional scaling (MDS). Given the high estimated reliability of the combined data, we combined the data of all four language groups (procedure MDS, SAS Institute, 1999) in a stacked

similarity matrix consisting of 36 columns (the clips) and 4×36 (language \times clips) rows. The group matrices were stacked (entered one after another) rather than summed to allow the program to compute different weights for each language group for the dimensions extracted. Allowing this possibility is more important for the naming solution presented below (where the matrices differ more), but for comparison it was preferable to use the same algorithm. (Weights obtained for dimension 1, dimension 2: English 1.12, 0.86; Dutch 1.09, 0.90; Spanish, 1.05, 0.94; Japanese 1.14, 0.84). The two-dimensional solution shown in Fig. 2 (stress value = 0.22) is presented for ease of viewing and because the greatest decline in stress values was obtained moving from one to two dimensions. (For 36 objects $p < .01$ that this stress value results from random data; Sturrock & Rocha, 2000.) Labels refer to clip names bestowed by the experimenters, not responses by participants, and are provided to help the reader interpret the solution. To help identify what physical elements of the actions were salient to observers, Additive Tree clusters (Corter, 1982; Sattath & Tversky, 1977) were drawn on the solution (stress value, formula 1 = 0.07). Addtree imposes fewer constraints on the data than many other clustering algorithms³ and permits a more faithful representation of the data (Sattath & Tversky, 1977). We provide only the top three levels of clusters for ease of viewing.

The horizontal dimension of Fig. 2 corresponds largely to the basic biomechanical distinction of pendulum motion with one foot on the ground at all times vs. elastic impact-and-recoil motion. The two top-level clusters separate the clips almost entirely on this basis. The only exceptions are the walking-in-place and marching-in-place actions that fall into the cluster containing mainly impact-and-recoil actions, apparently by sharing the in-place feature with running-, hopping-, and jumping-in-place. The overall salience here of the basic biomechanical distinction of pendulum motion vs. impact-and-recoil, reproduced without any contextual emphasis a treadmill might place on biomechanics, confirms the importance of this distinction in perception of human locomotion (Malt et al., 2008; see also Slobin, 2010) and documents structure across the breadth of this domain.

The second- and third-level clusters, along with the vertical dimension, indicate sensitivity to isolated physical dimensions as well as further reflecting the basic biomechanical distinction. The second-level clusters separate in-place from forward actions on the impact-and-recoil side, and higher-knees from the straighter-leg ones on the pendulum side. The third-level clusters reinforce the basic biomechanical distinction by separating non-bouncy in-place actions (walking-in-place and marching-in-place) from the true impact-and-recoil in-place ones (running-, hopping-, and jumping-in-place). They also separate forward running actions (running, jogging, trotting) from forward impact-and-recoil actions with more exaggerated bouncing (hopping, jumping, leaping, etc.). On the left, the clusters also separate the timid, slow pendulum

³ Addtree allows intra-cluster distances to exceed inter-cluster ones, and objects outside a cluster do not have to be equidistant from all objects inside it.

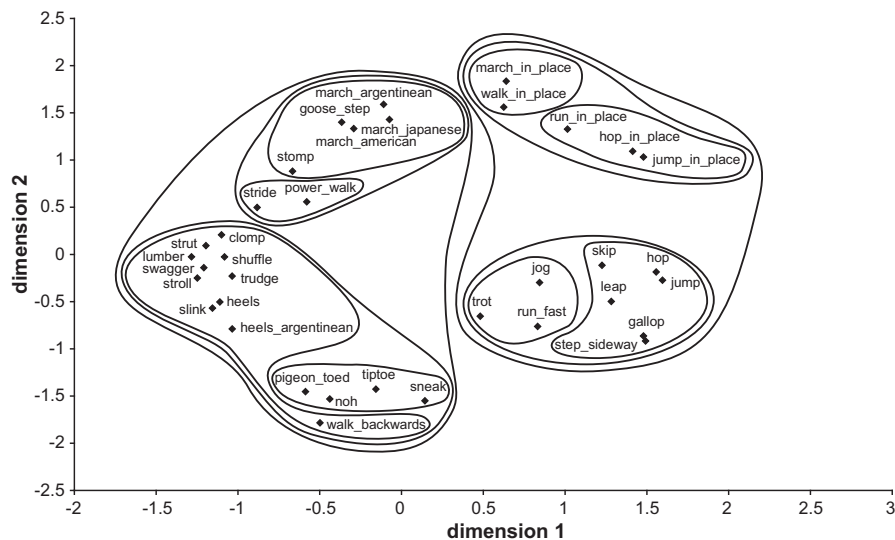


Fig. 2. MDS of aggregate physical similarity matrix.

motions from more typical ones, and those from faster, more aggressive ones at the top. (Values on an overall dimension of speed and aggression seem to reverse between left and right, with slower/less aggressive ones appearing on the bottom left and top right.) Thus perceivers are sensitive to properties of the physical movements beyond the basic biomechanical distinction, including forward vs. in-place motion, degree of bounciness, leg/knee position, and speed/aggressiveness.

Although there is a substantial gap between clusters within the right-hand top-level cluster (separating the in-place motions from forward-moving ones), other sub-clusters picked out by Addtree are less spatially distinct. This domain may constitute a case besides artifacts where partial structure exists: Once we look below the top-level biomechanical distinction, discontinuities in similarity space are less evident. (See Vulchanova et al., 2012, for a similar suggestion for action clips with less representation of human upright locomotion but encompassing motions such as caterpillars crawling and koalas climbing.) The limited structure for household containers (Malt et al., 1999) apparently does not come about only because artifact properties can be combined in unlimited ways. Even for locomotion, constrained by the ways bodies can move across surfaces, only partial structuring exists.

Commonalities in naming patterns and their bases

We now ask whether there are commonalities in the naming patterns, and, if so, to what dimensions or properties they are linked. To do this, we created an MDS solution based on the four languages' name similarity matrices. These matrices reflect to what extent each pair of actions received the same name across participants of a language group, as described earlier. Because MDS can only discover, not create, coherence in data, a solution that fits the combined data well will confirm commonalities in the naming patterns. Because this solution is derived from observed naming patterns, it will reflect all contributions to their commonalities, not only physical similarities. It will allow us to see to what extent dimensions apparent in the

physical sorting (isolated physical properties such as speed and knee height, and the biomechanical property clusters) are important to the naming patterns. Critically, it also lets us assess whether other dimensions such as effort or goals are important in structuring the naming space. Further, we can establish whether naming patterns disregard perceived physical structure, and whether the properties involved in named groupings shift between the more central, commonly experienced parts of the domain and those that are less important.

We carried out the scaling as before using a stacked name similarity matrix consisting of 36 columns (the clips) and 4×36 (language \times clips) rows. (Weights for dimension 1, dimension 2: English 1.08, 0.92; Dutch 1.04, 0.95; Spanish, 1.00, 1.00; Japanese 1.13, 0.85). Fig. 3 shows the two-dimensional solution (stress value of 0.25; for 36 objects, $p < .01$ that this stress value results from random data; Sturrock & Rocha, 2000). Three levels of Addtree clusters (stress value, formula 1 = 0.04) are imposed, as for the physical similarity solution. Labels again refer to clip names given by the experimenters, not participants.

This solution has clear similarities to the physical similarity sorting solution. The first-level clusters observe the basic biomechanical distinction by neatly separating the impact-and-recoil motions from the pendulum-based ones, with the exception of the trot clip falling just into the pendulum cluster. In fact, this solution produces a neater separation of actions along the biomechanical dimension than the physical similarity sort did because here, the walking-in-place and marching-in-place actions appear with other pendulum actions instead of affiliating with the impact-and-recoil in-place actions. The placement of the trot clip at the edge of the pendulum cluster (and separated out by a second-level cluster) may reflect the pull of the two biomechanical alternatives: The clip provides little evidence of both feet being off the ground at once (causing some respondents to use a name for pendulum motions), but it shows a distinctly bouncy motion (causing other respondents to use a name associated with impact-and-recoil motions).

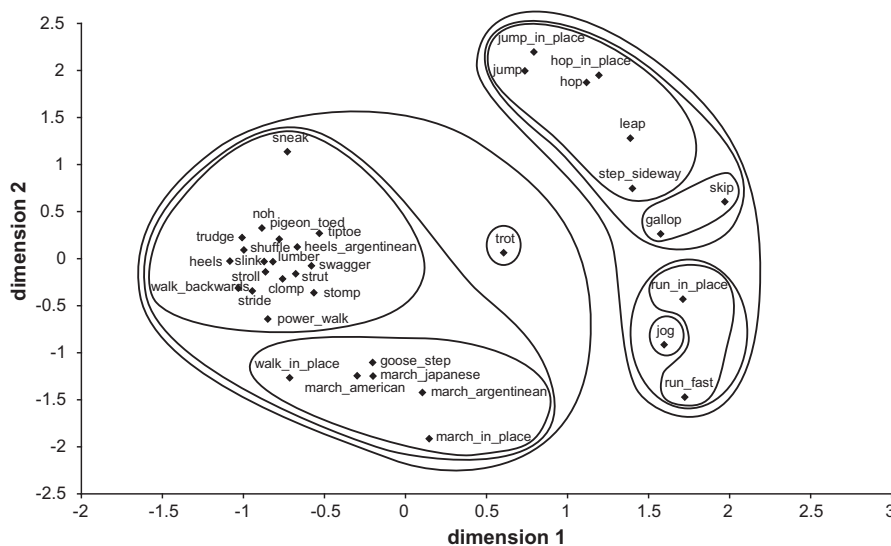


Fig. 3. MDS of aggregate name similarity matrix.

Additional elements of the solution also parallel the physical naming solution. The y-axis seems to reflect speed and aggressiveness, with slower/less aggressive actions toward the top and faster/more aggressive ones towards the bottom. On the right, the second-level clusters separate the three actions with a modest impact-and-recoil from those with more pronounced impact-and-recoil. The third-level clusters on the left largely separate pendulum actions done with higher knees from those with straighter legs. (The goose-step clip, in which knees are high but the legs are straight, clusters with the other high-knee actions). On the right, the third-level clusters separate a leisurely run (the clip with the experimenter-given label of *jog*) from two faster running-type actions and those from the more exaggerated bouncing motions of leaping, hopping, and jumping. Thus, lexicalized distinctions reflect some properties of the physical movements beyond the basic biomechanical distinction, just as judgments of physical similarity did. As in the physical sorting solution, these include forward vs. in-place motion, degree of bounciness, leg/knee position, and speed/aggressiveness.

Do effort and functions or goals contribute to how actions cluster by name here? Effort may be correlated to some extent with speed and aggressiveness (with greater speed/aggression requiring greater effort), and also with the basic biomechanical distinction (with impact-and-recoil actions being more effortful than pendulum ones). However, it is not clear that effort contributes independently to explaining how the actions are grouped. For instance, the actions called *walk* in English vary from less to more energy-intensive (and probably overlap in effort with actions called *march*), and the more vigorous marches probably overlap in effort with the less energetic impact-and-recoil actions. As for goals – such as relaxing, exercising, hurrying, acting purposefully, playing, displaying attitudes – none clearly maps onto either x- or y-axes of the solution. There is also no obvious dichotomy of goals that can be mapped onto the first-level clusters we have attributed to the biomechanical distinction. If function or goals cluster actions by name, then perhaps this is to be found

in finer-grained levels of clustering. Inspection of the clusters, however, suggests that they generally cross-cut goals by mixing actions that are in place and moving forward (or backwards), are both vigorous and less so; are for exercise or for getting somewhere; are associated with positive attitudes or with negative, and so on. The only cluster not more readily accounted for in terms of physical properties is the separation of skipping and galloping motions from other impact-and-recoil actions. This separation may be isolating the more child-like or playful actions from the others. The clustering of marches could also be construed as drawing on their military context, since it includes both straight-legged and bent-leg forms. However, this same cluster existed in the physical similarity solution, which suggests that it is based more on physically-grounded speed/aggressiveness along with knee position. Thus, only the clustering of skipping and galloping unambiguously suggests an influence of goals.

In sum, the coherence of the name similarity solution confirms shared elements in the naming patterns. The interpretability of the solution largely in terms of the same dimensions apparent in the physical similarity solution, and absence of interpretability in terms of effort or goals, argues that physical similarities are the primary driver of the named distinctions of each language. The basic biomechanical distinction reflecting correlated properties abstracted over time and space seems to be most prominent in guiding naming distinctions. Other isolated physical dimensions serve as secondary features for creating named distinctions. Some of these features, such as speed, endure across the different phases of a stride; others, such as angle of the knee, may change across the stride but entail only one region of the body. They are relatively easily observed, but, perhaps counter-intuitively, are less important. The only apparent contribution of non-physical properties is the element of child-likeness or playfulness separating skipping and galloping from other impact-and-recoil actions.

Given that the naming solution reflects the basic biomechanical distinction at the top level, there is no evidence

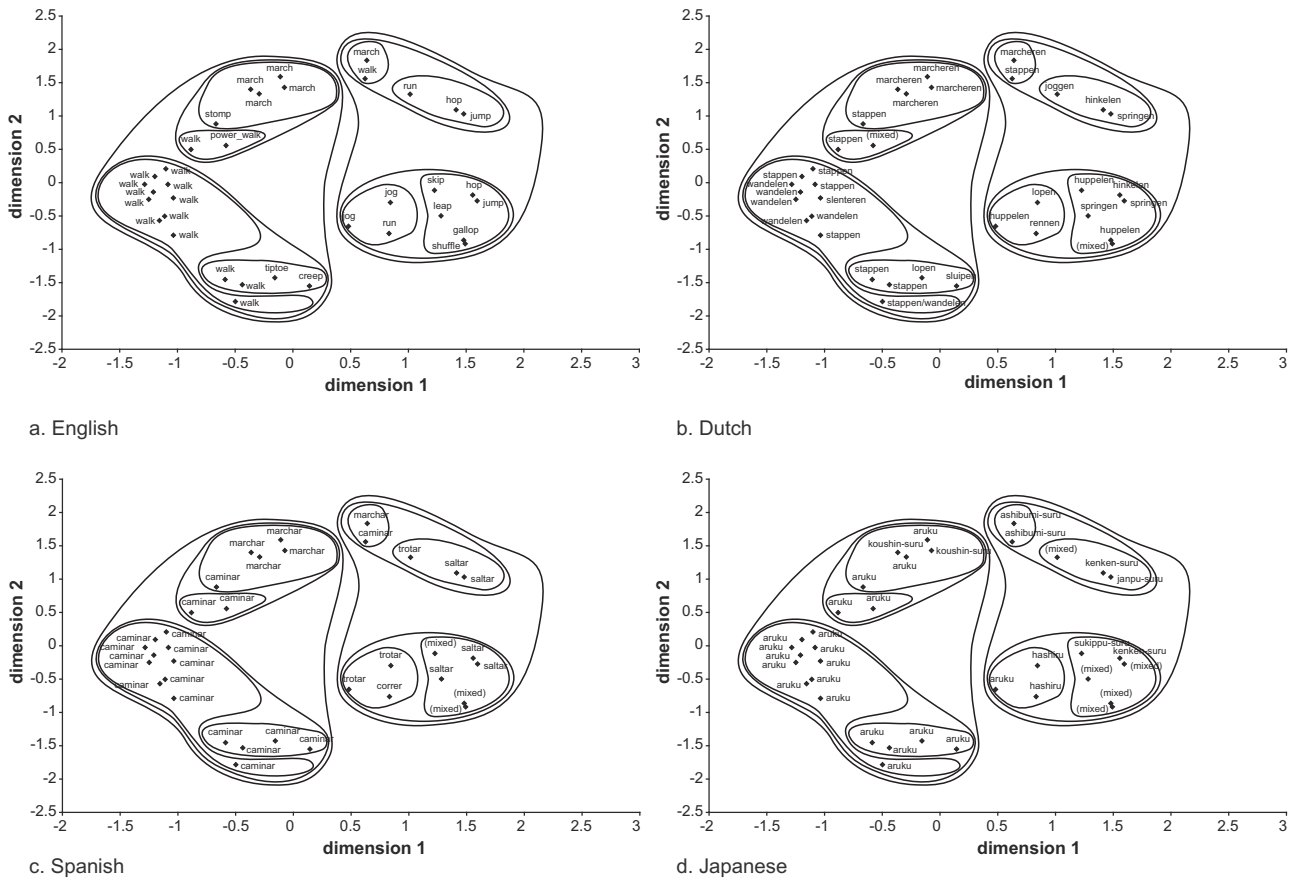


Fig. 4. Mapping of names for each language to the aggregate physical similarity solution.

that languages create named groupings of actions on the basis of features that cross-cut the physical similarity structure (such as effort or goals). That is, where languages *do* create named groupings, they respect salient physical similarity. However, there is some evidence for the languages ignoring an element of physical structure *without* creating groupings on alternative bases. The group naming data do not show lexical differentiation of actions in place vs. moving forward, although this difference was salient in physical similarity sorting. The distinction cross-cuts the biomechanical one, and in naming, the biomechanical distinction dominates. The relatively low frequency of the actions in question may also contribute to the lack of lexical differentiation. However, frequency cannot be the sole determinant of what is lexicalized, since some named distinctions are made for rare gaits (e.g., *hop*, *skip*, *tiptoe*), and the same is true for rare actions of all sorts (e.g., *bellyflop*; *levitate*; *self-immolate*; *impeach*; Dutch *klunen* meaning to walk on land with ice skates on). Furthermore, walking- and running-in place have become fairly common in exercise contexts (on treadmills and in home exercise programs on DVD) for at least some of our language groups. The differing numbers of dominant verbs across the four languages also argues against frequency of the actions providing a full explanation of what is lexicalized and what is not.

We had raised the question of whether the property types that matter to named distinctions shift between the parts of the domain of most central and lesser impor-

tance to everyday life. As noted, walking and running can be considered the central parts, given their high familiarity and frequency of use. The preceding discussion suggests no major shift from physical to non-physical features across any parts of the domain. We further examined the clips according to the clusters in the name similarity MDS solution. The clips in the top left-hand cluster are mostly called *walk*, *wandelen*, *caminar*, and *aruku* (by English, Dutch, Spanish, and Japanese, respectively) and those in the lower right-hand cluster are mostly called *run*, *rennen*, *correr*, and *hashiru* (see Fig. 4 below; also Appendix C for names given to clips by individual languages). We took these two clusters to constitute the more central part (21 clips), and all other clips (15) the less central part. Based on this division, the clustering of skipping and galloping, and hence the possible contribution of a playfulness element to naming, occurs in the peripheral part. This outcome is consistent with the idea that special-purpose considerations are more likely in the peripheral part. However, any such contribution appears to be small here.

Variation

We now consider where variations across the languages lie and what their source may be.

As a broad first measure to see how well each language reflects the aggregated physical similarity and name similarity, we correlated the name similarity matrix for each language with the two group matrices. Correlations of

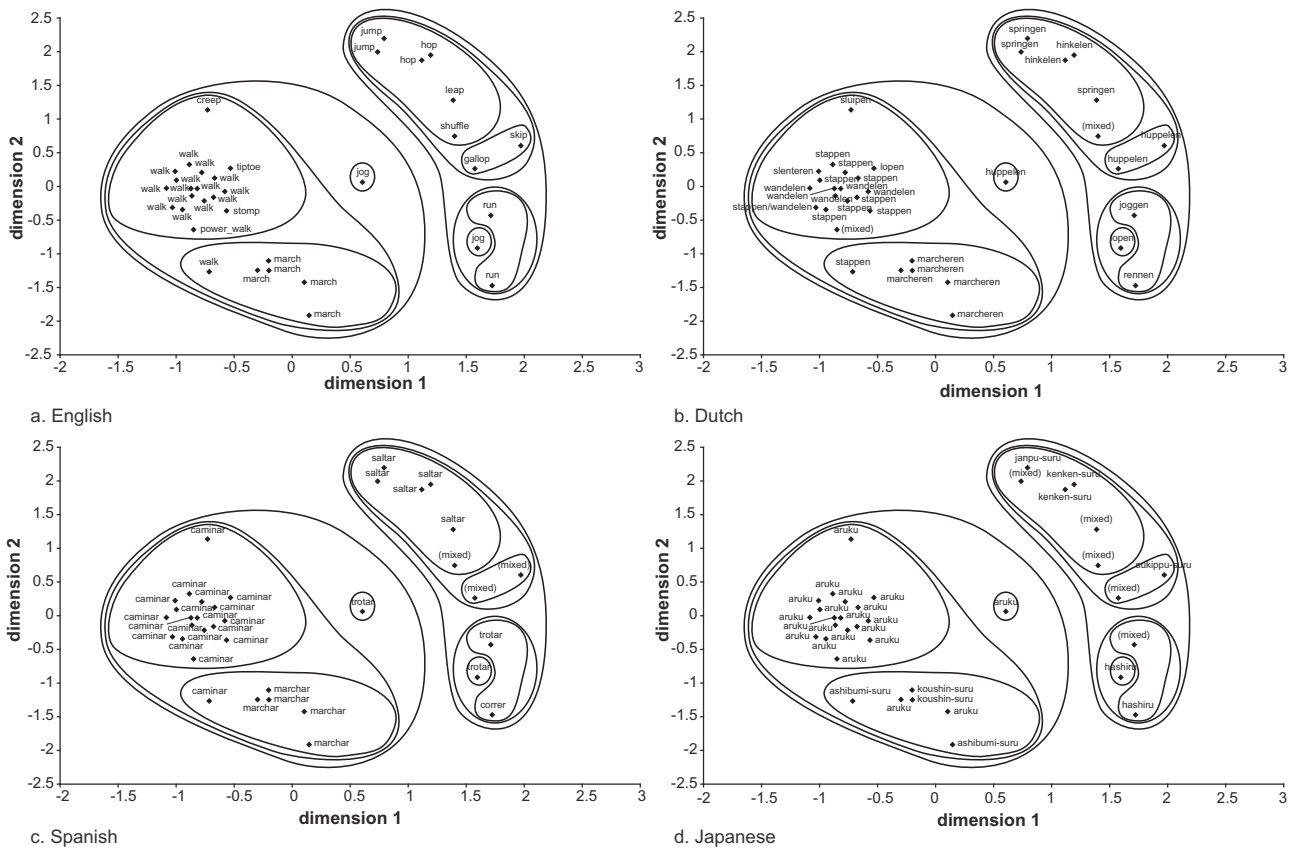


Fig. 5. Mapping of names for each language to the aggregate name similarity solution.

the name similarity matrices with the group physical similarity matrix were 0.57, 0.61, 0.64, and 0.55 for English, Dutch, Spanish, and Japanese respectively. English, Dutch, and Japanese are lower than Spanish ($p < .03, .055,$ and $.005,$ respectively), most likely reflecting the fact that they make more named distinctions among clips that are close in physical similarity space. Although the differences are significant given the large N (630 cells of the matrices), the values are all in the same ballpark, with 31–41% of the variance accounted for. For name similarity, because the aggregated matrix is, itself, based on the naming patterns of these four languages, it is logically impossible to find that all of the individual languages deviate substantially from it. However, one could deviate from the rest. Correlation values were 0.92, 0.85, 0.93, and 0.90 for English, Dutch, Spanish, and Japanese respectively. All values are significantly larger than for physical similarity ($p < .001$), with 72–87% of variance accounted for, demonstrating that aggregated naming captures variance in the individual naming patterns not captured by physical sorting. Dutch is significantly lower than the other three ($ps < .02$ or greater), probably reflecting the fact that it makes a distinction (*stappen* vs. *wandelen*) among actions close in name similarity space in the other languages.

To look at how individual languages' named distinctions relate to dimensions of the group physical similarity and name similarity, we then mapped, for each language, the dominant name for each clip (as given in Appendix C) onto the solutions in Figs. 2 and 3. These mappings are shown in Figs. 4 and 5, respectively. Names on stimulus

points here are the dominant names of each language group (not the experimenter-given names). These mappings allow us to assess on what basis the individual languages make named distinctions, including whether non-physical properties are used by some languages and, if so, in what parts of the domain.

Do individual languages use the same bases for named distinctions?

All four languages broadly respect the major perceived distinctions of the physical similarity solution, as shown by the tendency of each name of a language to be confined to a single cluster of this solution. However, each language does have some violations of these groupings, shown in instances of a name appearing outside of the cluster it mainly occurs in. To quantify the number of violations for each language, we first counted the number at the top level of clustering, then setting those instances aside, counted any additional violations at the second level, and then setting those aside, counted any additional ones at the third level. English had 8 and Dutch 9; Spanish had 12 and Japanese, 11. The slightly higher numbers for Spanish and Japanese were due to each having two third-level clusters that were not differentiated from each other by name. Thus, all the languages use physical similarity to the same extent where they *do* make distinctions, although some make finer distinctions. (We will return later to differences in how many distinctions are made.)

Consistent with broad agreement in the extent of use of physical distinctions, the specific violations are to some

extent shared. There are, however, a few cases of differential use of physical properties. English, Dutch, and Spanish each label the walking-in-place and marching-in-place actions by names otherwise restricted to the left-hand cluster, whereas Japanese does not, showing Japanese sensitivity to the in-place physical feature. (Japanese does not, however, have a single name separating all in-place actions from forward ones; it treats this distinction as secondary to the basic biomechanical one.) Japanese also has one instance of *aruku* appearing on the right although the rest appear on the left, and Dutch has one instance of *lopen* within each first-level cluster. We have noted (unpublished data) that Belgian Dutch uses *lopen* mainly for impact-and-recoil actions, while Netherlands Dutch uses it mainly for pendulum-type actions. The word form seems to have drifted across meanings in the dialects and is not as fully linked to the biomechanical distinction as other verbs are (consistent also with the use of German *laufen*; Jessen, in press). The Japanese *aruku* violation occurs for a clip that may be physically ambiguous (the trot clip discussed earlier), suggesting greater flexibility of *aruku* to encompass bounciness. And, as noted above, English and Dutch partially differentiate the lower left-hand third level cluster from the upper one whereas Spanish and Japanese do not. In short, some modest differences exist in use of physical features secondary to the basic biomechanical distinction.

Turning to the mapping of names to aggregated name similarity, this solution does a better job of capturing the naming patterns of individual languages. This outcome is expected given the correlations above and the fact that the solutions are derived directly from their combined naming data. There are, however, some violations of the shared naming solution. Counting as before, English had two violations, Dutch and Spanish three each, and Japanese, four, indicating that dimensions revealed by combined name similarity work about equally well for each of them. In each case, the number is less than for physical similarity, reinforcing the notion that naming patterns have some shared structure not isomorphic to the physical similarity sorting solution.

As with the mapping onto the physical similarity solution, some violations are shared across languages. The only language-idiosyncratic cases are consistent with those identified in the physical similarity solution concerning Dutch *lopen* and Japanese *aruku*. Again, *lopen* may be less strictly tied to the biomechanical distinction than other walking- and running-type verbs, and *aruku* may have a range of application broader than other walking-type verbs.

Taking both mappings together, all languages seem to use physical similarity to a similar extent where they do draw named distinctions, although some lexicalize distinctions that others do not. No language is an obvious outlier to the shared name structure, indicating no major differences in use of properties including non-physical ones.

Differences in numbers of named distinctions: Added features

The most notable difference across the languages is in the number of named groupings, ranging from 5 for

Spanish to 14 for English. We can look within these differences to see whether some types of dimensions are subject to more usage variation than others. This discussion will be based on mappings to the name similarity solution, which best shows the actions sharing each name in each language.

Spanish provides the fewest named contrasts. On the impact-and-recoil side, it contrasts faster and slower running by name, and separates those from actions with more pronounced bouncing. To this set, English, Dutch, and Japanese all add a distinction between one-footed and two-footed actions with pronounced bouncing. Only English gives four further unique names to the impact-and-recoil actions of *leap*, *shuffle*, *gallop*, and *skip*, but the other languages each distinguish some by name and have no agreed-upon name for the others, indicating that their names for other impact-and-recoil actions do not extend over these. These distinctions imply attention to details of the bouncing actions: alternating feet, sideways movement, the little inter-stride bounce in skipping, and absence of the feet passing each other in galloping. As noted earlier, given that skipping and galloping cluster in the name similarity solution despite being fairly different in other ways, the child-like or playful association of these actions may be involved in separating them by name from others with more pronounced bouncing. To the extent that this factor matters, all the languages seem to observe it by not extending their names for other pronounced bouncing actions to skipping and galloping. On this side, then, languages largely differ in how many secondary physical features they draw on, with minor but shared attention to the child-like or playful qualities of skipping and galloping.

On the pendulum side, Spanish distinguishes only walking-type from marching-type actions by name (which differ on speed/aggressiveness and knee angle), while the other languages have additional contrasts. The most notable added contrast is that the actions called *walk*, *caminar*, and *aruku* by English, Spanish, and Japanese, respectively, are divided between *wandelen* and *stappen* in Dutch. *Stappen* appears to reflect a more “steppy” movement (higher lifting) of the foot. English also distinguishes *creep*, *tiptoe*, *stomp*, and *power-walk* from *walk*, and Dutch distinguishes some of these as well *slenteren* (labeled *trudge* by the experimenters). For the most part the distinctions seem to be based on physical properties such as speed, forcefulness, and part of the foot contacting the ground, though some element of attitude or goal may be included in distinguishing *creep* and *stomp*. The other main difference is that within the marching actions (called *march*, *marcheren*, and *marchar* by English, Dutch, and Spanish, respectively), Japanese also distinguishes two in-place actions that differ in having a straighter vs. more bent leg and two emphatic military actions also contrasting on straighter vs. more bent leg. On the pendulum side, then, some of the languages again use physical details that other languages ignore. English and Dutch may also observe here some features linked to attitude or goals that the other two do not.

In sum, the conclusions about types of features used derived from the aggregated naming solution hold true

for individual languages. Physical features including the basic biomechanical distinction and additional isolated physical features are used by all four languages in creating named distinctions, with minor use of other types of features also shared. Languages virtually always individually respect salient physical structure where they do create names, although they may not all lexicalize the same secondary distinctions. However, minor refinement is needed. Dutch has a single word (*lopen*) that may to some extent cross-cut the basic biomechanical distinction. Thus, ignoring strong structure is rare but not impossible.

Typological influences on number of named distinctions

Our previous data (Malt et al., 2008) showed that the biomechanical discontinuity between walking and running was encoded in languages that predominantly express path in the main verb (Spanish, Japanese) as well as those that predominantly express manner in the main verb (English, Dutch). However, English and Dutch made finer distinctions within the basic gaits than Spanish and Japanese. The current data extend both observations across a larger range of actions. All four languages observed the basic biomechanical distinction, but English and Dutch had more dominant names than Spanish and Japanese. The similarity of English and Dutch may derive in part from their shared history, although they have diverged enough within this domain to differ in the number of terms and some of their meanings. However, Spanish and Japanese are from different families, and both show fewer names. This fact supports the idea that verb typology impacts lexicalization of shared perceptual distinctions.

Despite similar numbers of dominant names for Spanish and Japanese, there is one major difference in the nature of their responses. The mean level of participant agreement on the dominant name for a given clip (see Appendix C) is 74% for Spanish, whereas it is 60% for Japanese. The Japanese figure is similar to those for English and Dutch speakers (60% and 52%, respectively), who had many more manner verbs to choose among. The differences are significant, $F(3,105) = 7.75$, $p < .001$, with pairwise contrasts showing that only Spanish differs significantly from the others, $ps < .01$. Simpson's D , a measure of diversity that takes into account both the number of different verbs produced for each clip and the consistency with which they were produced,⁴ shows the same result. Averaging the D value across the 36 clips of each language yielded values of .44, .36, .63, and .46 for English, Dutch, Spanish, and Japanese, respectively, with a significant difference among them, $F(3,105) = 9.11$, $p < .001$. Pairwise contrasts show that only Spanish differs significantly from the other languages, $ps < .01$.

The discrepancy between Spanish and Japanese arises because Japanese has many different ways of describing action in multi-word phrases. Besides monolexemic manner verbs, there are conventional verbs composed of a

noun followed by *suru* ["do"], as described earlier. In addition, this noun + *suru* construction is used flexibly to create novel verb phrases as needed. Japanese speakers can also create novel manner verb phrases in other ways, such as using mimetics (sound symbolic words) in combination with *suru* or a conventional verb, and combining two verb phrases with a particle. Where Japanese speakers had low agreement in naming a clip, it was typically not because they showed diversity in the conventional manner verb chosen (as was often the case for English and Dutch) but because manner was being expressed through varied phrases of these different sorts. Spanish does not have such a range of grammatical devices for expressing motion, and participants relied on their conventional verbs for all the locomotion stimuli, thereby producing high agreement on most clips.

Language differences in the central vs. peripheral parts of the domain

Last, we can ask if there is more variation across the individual languages in naming patterns for parts of the domain of lesser importance to everyday life. If so, then greater variation could result from either lesser perceived structure or more variation in the types of properties used. However, previous analyses found no major property shifts across parts of the domain for any of the languages, eliminating the second possibility. We therefore focus on the structure question here.

As before, we considered clips falling into the top left-hand cluster and lower right-hand cluster of the group name similarity MDS solution to be the more central part, and all other clips the less central part. To determine if cross-language agreement differs between the two parts, we correspondingly divided the individual name similarity matrices and examined the correlations between each language pair for the more and less central parts using Mantel tests for the correlation of matrices (Mantel, 1967). Tables 4a and 4b report the results. The average

Table 4a

Correlations of name similarity matrices for 21 actions in the central portion.

	English	Dutch	Spanish
Dutch	0.83		
Spanish	0.68	0.67	
Japanese	0.67	0.61	0.91

Note: Pearson correlation values were calculated using the Mantel test for correlation of matrices, based on 1000 permutations. All correlations are significant at $p < .001$.

Table 4b

Correlations of name similarity matrices for 15 actions in the peripheral portion.

	English	Dutch	Spanish
Dutch	0.80		
Spanish	0.68	0.49	
Japanese	0.82	0.91	0.58

Note: Pearson correlation values were calculated using the Mantel test for correlation of matrices, based on 1000 permutations. All correlations are significant at $p < .001$ except Dutch–Spanish, significant at $p < .005$.

⁴ Simpson's D is calculated as $D = (n_i - 1) / (N(N - 1))$, where n_i is the total number occurrences of a particular verb for a clip and N is the total number of all verbs for a clip. D varies between 0 and 1.

pairwise correlations between languages for the central and peripheral parts are almost identical (72.8 for central; 71.3 for peripheral), with three language pairs higher for the central part, two higher for the peripheral part, and one pair tied. Thus, there is not more variation across languages in the peripheral part of the domain.

The lack of a consistent trend indicates that languages observe perceived structure in naming regardless of whether it occurs in more frequently experienced and talked about parts of the domain or not. To the extent that part of the domain matters, its influence may interact with typological differences. Across languages, there were ten clips that elicited less than 30% name agreement (those marked as “mixed” in the figures): 2 for Dutch, 3 for Spanish, and 5 for Japanese. Nine of these fall into the peripheral portion of the domain, including all of the Spanish and Japanese ones, suggesting that languages with fewer words to cover a domain may skimp on names for things less likely to be talked about.

General discussion

Summary of results

We examined how 36 diverse instances of upright human locomotion are named in English, Dutch, Spanish, and Japanese in order to understand constraints and variation in naming patterns. We first used multidimensional scaling to determine the structure of perceived physical similarities among the instances. The physical MDS solution showed a top-level biomechanical distinction between pendulum-like and impact-and-recoil motions, reflecting constellations of properties abstracted across a stride. Variation on more isolated aspects of the physical movements, such as speed/aggressiveness and angle of the knee, created a distribution of instances within the two major groupings in which some form clusters but others were more scattered across the space. The latter outcome demonstrates that even for locomotion, limited by the ways human bodies can move across surfaces, only partial domain structuring can exist. However, the partial structuring is embedded within a strong higher-level division of the domain.

An MDS solution on naming responses to the same stimuli found shared elements in naming patterns across the four languages, interpretable largely in terms of the same basic biomechanical distinction apparent in the physical similarity solution, along with some shared use of the more isolated physical properties reflected in it. Goal-related influences on naming were small, and there was no evidence for use of other inferred properties such as effort. There was also no evidence of languages cross-cutting strong perceived structure with names, and no evidence for differential treatment of parts of the domain of less importance to everyday life. However, all languages ignored one element of structure in the physical similarity solution – the separation of in-place from forward motions. These actions formed a cluster in the sorting data but were not distinctly lexicalized in any of the languages in our sample.

As implied by the preceding point, there was consistency across the languages in that physical features accounted for most named distinctions, although there were differences in some of the specific features used. The most notable differences, though, were in how many different verbs the languages utilized, especially in the peripheral part of the domain, and in the level of agreement speakers showed in naming each clip. These differences are consistent with an influence of language typology on naming patterns.

Implications for understanding constraints on word meaning

The current results help identify whether commonalities in naming patterns reflect, in part, a world that presents itself in “a series of discontinuities whose structure and content are seen by all human beings in essentially the same ways. . .” (Berlin, 1992, p. 9). The perceived structure in the physical sorting similarity solution adds evidence for strong domain-wide structure outside of plants and animals (e.g., Berlin, 1992; Hunn, 1977) and body parts (Majid et al., 2006). Several unrelated domains have now been shown to contain structure that can feed into naming patterns. The present finding suggests that more may be identified.

Despite the salient biomechanical distinction, some of the instances within each top-level biomechanically-based cluster were spread relatively continuously through the physical similarity space of each cluster. This situation of one major perceptual distinction with lesser ones embedded within it (see also Vulchanova et al., 2012) may differ from the case of household containers (Malt et al., 1999), where it is not clear that the physical similarity space bifurcates so neatly. Thus partial structuring may play out in different ways in different domains, and these patterns of structuring may foster different patterns of similarity and differences across languages.

Our data add to observations that constellations of properties abstracted over time and space are salient in perception of domains and critical in determining naming patterns. Co-occurrence or correlation information is automatically extracted from the environment under many circumstances (e.g., McNorgan, Kotack, Meehan, & McRae, 2007; Peterson & Beach, 1967; Wattenmaker, 1993), even by infants (Aslin, Saffran, & Newport, 1999). Correlated properties may create perceptually distinct clusters of entities that serve as an important constraint in naming patterns. Recognizing such correlations in naming may create lexical categories that are maximally useful across circumstances because they convey information about many attributes (Berlin, 1992; Hunn, 1977).

The primacy of correlated properties in naming helps address whether perceived structure inevitably serves to ground meaning, and whether named groupings pick out the same set of properties as those salient non-linguistically. The named groupings do closely reflect distinctions present in the physical similarity sorting. However, there is little impact of the perceptually salient distinction of forward vs. in-place motion on the named distinctions. Isolated properties may be more likely to be overlooked in naming if they are orthogonal to salient groupings

capturing larger constellations of properties. Further evidence for this argument lies in the fact that the languages do not lexicalize distinctions based on speed alone, which is also an isolated physical property reflected in the scaling solutions and observable in all instances of forward locomotion.

This point may also explain the minimal impact of other isolated properties such as effort or goals on the naming patterns for locomotion. Likewise, it can explain the consistency we found in the use of physical features, especially correlated property clusters, across all parts of the domain and across the four languages. Besides the terms elicited by our stimuli, languages may have additional terms that do capture properties like speed or aggressiveness, as in English *barge* and *bolt*. Nevertheless, the current data show that languages have vocabulary linked to the physical properties of gaits, regardless of whether the gaits are frequent or rare. This finding parallels the predominance of taxonomic over utilitarian categories in folkbiology (e.g., Berlin, 1992; Hunn, 1977).

Despite the importance of correlated property clusters, a strong tendency is not the same as an inviolable constraint. The data showed that Belgian Dutch *lopen* crosses the biomechanical boundary (see also Jessen, *in press*, on German *laufe*). Japanese *aruku*, used mostly for pendulum-type actions, also showed some flexibility in its application across the boundary. As with most aspects of language, more relevant than declaring a constraint absolute is determining its place among possible constraints. Informal queries of speakers of diverse other languages suggest that a labeled distinction between walking and running is common. Systematic studies of additional languages are needed to look beyond this contrast and to probe the boundary conditions for applying the terms. Such data will provide a broader perspective on the hierarchy of constraints suggested by the current data.

Implications for understanding variation in naming patterns

A major question about diversity is whether languages vary among themselves in the nature of properties attended to within a domain, and if so, where does this variation occur. Consistent with the preceding points, the four languages share a reliance on physical features. Variation occurred mainly on which and how many verbs were used (see Vulchanova et al., 2012, for a related conclusion for other sorts of motion events). This finding argues against the possibility that languages routinely embody radically different construals of the world, in the sense of encoding different features of a domain in words.

This conclusion has implications for the debate about the extent to which children's word learning can build upon pre-linguistic perceptual and conceptual knowledge (see, e.g., discussions in Bowerman & Levinson, 2001). Variations in naming patterns across languages dictate that any pre-linguistic perceptual or conceptual segmentation of the world must be only a stepping stone to the language being learned. However, the current data argue that observed structure in the world can provide a building block setting parameters within which variation can occur. In

some, and possibly many, domains, the specifics of word meaning in a language will rely on dimensions or constellations of properties that are salient even for young infants (see also Parish-Morris, Pruden, Ma, Hirsh-Pasek, & Golinkoff, 2010).

One source of variation may come from other characteristics of the languages. The number of named distinctions that emerged in our four languages aligned with whether the language tends to express path or manner of movement in the main verb. More broadly, the types of devices a language uses for one purpose can impact other features of a language. For instance, Wolff, Jeon, Klettke, and Li (2010) observe that some languages prohibit inanimate nouns from serving as subjects of causal verbs. In particular, languages marking grammatical role by word order are more flexible in their sentence subjects than ones marking it by case (see also Hawkins, 1985). In a similar vein, Beavers, Levin, and Tham (2009) argue that whether manner is expressed in the main verb or satellite is influenced by what other morpho-syntactic devices the languages have for encoding aspects of motion events. Malt, Sloman, Gennari, Shi, and Wang (1999) and Malt, Gennari, and Imai (2010) suggested that the large Spanish vocabulary for household containers may reflect, in part, the availability of productive suffixes that create nouns from other nouns. The impact of typological differences among languages on lexical patterns warrants more investigation (see also Slobin, 2004).

We have also argued previously that cultural and linguistic historical processes may contribute to diversity in naming patterns for artifacts (Malt, 2010; Malt, Gennari, & Imai, 2010; Malt et al., 1999). For instance, the order in which objects enter a culture can vary, as well as the names available for extension. As each generation builds on the previous one's naming patterns, different language groups can spawn substantially different naming patterns. Such processes are less likely to matter in domains where exemplars are relatively uniform across cultures and time. Still, some historical processes may affect naming patterns for domains like locomotion. For instance, since the end of World War II, thousands of loan words from English have entered Japanese (e.g., Kay, 1995). Our clips called *skip* and *jump* in English were *sukippu-suru* and *janpu-suru* in Japanese, both adoptions from English. Spanish speakers, not using loan words, called both actions *saltar*. Language contact, and the factors influencing whether a community accepts words from a contact language (e.g., Trudgill, 2002), may contribute to variation in virtually any domain.

Finally, some domains may not contain anything like sets of correlated properties. For instance, colors vary in hue, intensity, and saturation, but the dimensions vary independently. Where there are no perceived discontinuities in stimulus space, one might expect greater variation across languages despite using a shared feature set. However, it is difficult to compare the degree of naming diversity across domains based on studies using different languages and numbers of languages, and different sized stimulus sets that sample the domain differently. Domain comparisons within a common framework are needed.

Conclusion

One theoretical possibility is that word meanings and patterns of naming vary across languages without constraint. Evidence for shared tendencies within domains argues against this possibility. However, it still leaves open the possibility that each domain plays out in its own way without yielding generalizations about the nature of constraints on development of naming patterns and the forces that create variation. The current data support the alternative that there are general principles across domains explaining the emergence of commonalities and differences. They move forward an understanding of what those principles are.

Acknowledgments

This work was supported by funds from the National Science Foundation (#BCS-1057855) to Barbara Malt, the Netherlands Organization for Scientific Research (NWO Vici grant) and the Max Planck Gesellschaft to Asifa Majid, and MEXT KAKENHI (#15300088, #22243043, #23120003) and Keio University (Keio Gijuku Academic Development Funds and Keio Gijuku Mori Memorial Research Fund) to Mutsumi Imai. Eef Ameel was supported on a postdoctoral fellowship by the Fund for Scientific Research-Flanders (FWO-Vlaanderen). We thank Dan Slobin for helpful comments on an earlier draft and Adam Darlow, Kiri Lee, Yo Matsumoto, Kyoko Ohara, Steven Sloman, and Phillip Wolff for useful discussion. Kristine Schuster and Hiroyuki Okada developed the web-based naming experiment, Naoaki Tsuda created the sorting program, and Stephanie Sterrett served as walkway actor. Ludy Cilissen, Lisa Guest, Celina Hayes, and Erin Howard assisted with data collection and analysis.

An earlier analysis of some of the data was reported in Malt, B. C., Gennari, S., & Imai, M. (2010). Lexicalization patterns and the world-to-words mapping. In B. C. Malt and P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 29–57). New York: Oxford University Press. Some of the current results vary slightly from those reported there due to implementation of more formal name coding criteria (see Appendix B).

A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jml.2013.08.003>.

References

- Alexander, R. M. (2002). *Principles of animal locomotion*. Princeton: Princeton University Press.
- Ameel, E., Malt, B. C., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language*, 58, 262–285.
- Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53, 60–80.
- Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98, 409–429.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1999). Statistical learning in linguistic and nonlinguistic domains. In B. MacWhinney (Ed.), *The emergence of language* (pp. 359–380). Mahwah, NJ: Lawrence Erlbaum.
- Beavers, J., Levin, B., & Tham, S. W. (2009). The typology of motion expressions revisited. *Journal of Linguistics*, 46, 331–377.
- Bennett, M. B. (1992). Empirical studies of walking and running. In R. McN. Alexander (Ed.), *Comparative and environmental physiology: Mechanics of animal locomotion* (Vol. 11, pp. 141–165). Berlin: Springer-Verlag.
- Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.
- Boster, J. S. (1987). Agreement between biological classification systems is not dependent on cultural transmission. *American Anthropologist*, 8, 914–919.
- Boster, J. S. (1994). The successive pile sort. *Cultural Anthropology Methods*, 6, 7–8.
- Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.
- Bowerman, M. (2005). Why can't you "open" a nut or "break" a cooked noodle? Learning covert object categories in action word meanings. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 209–243). Mahwah, NJ: Erlbaum.
- Bowerman, M., & Levinson, S. C. (Eds.). (2001). *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.
- Bulmer, R. (1967). Why is the cassowary not a bird? A problem of zoological taxonomy among the Karam of the New Guinea Highlands. *Man*, 2, 5–25.
- Carter, J. E. (1982). ADDTREE/P: A PASCAL program for fitting additive trees based on Sattath & Tversky's ADDTREE program. *Behavior Research Methods and Instrumentation*, 14, 353–354.
- Crystal, D. (1987). *The Cambridge encyclopedia of language*. Cambridge, England: Cambridge University Press.
- Diedrich, F. J., & Warren, W. H. (1995). Why change gaits? Dynamics of the walk-run transition. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 183–202.
- Garner, W. R. (1974). *The processing of information and structure*. Hillsdale, NJ: Lawrence Erlbaum.
- Gennari, S., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj, II (Ed.), *Language development: Language, thought and culture* (Vol. 2, pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.
- Hawkins, J. A. (1985). *A comparative typology of English and German: Unifying the contrast*. Austin: University of Texas Press.
- Hunn, E. (1977). *Tzeltal folk zoology: The classification of discontinuities in nature*. New York: Academic Press.
- Jameson, K. A. (2005). Culture and cognition: What is universal about the representation of color experience? *Journal of Cognition and Culture*, 5, 293–347.
- Jessen, M. (in press). Semantic categories in the domain of motion verbs by adult speakers of Danish, German, and Turkish. *Linguistik online*.
- Kay, G. (1995). English loanwords in Japanese. *World Englishes*, 14, 67–76.
- Kay, P., Berlin, B., Maffi, L., & Merrifield, W. (1997). Color naming across languages. In C. L. Hardin & L. Maffi (Eds.), *Color categories in language and thought*. Cambridge: Cambridge University Press.
- Kemler Nelson, D. G., Egan, L. C., & Holt, M. (2004). When children ask *What is it?* what do they want to know about artifacts? *Psychological Science*, 15, 384–389.
- Khetarpal, N., Majid, A., & Regier, T. (2009). Spatial terms reflect near-optimal spatial categories. In N. Taatgen et al. (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (pp. 2396–2402).
- Kronenfeld, D. B., Armstrong, J. D., & Wilmoth, S. (1985). Exploring the internal structure of linguistic categories: An extensionist semantic view. In J. W. D. Dougherty (Ed.), *Directions in cognitive anthropology* (pp. 91–113). Urbana: University of Illinois Press.
- Lynch, E. B., Coley, J. D., & Medin, D. L. (2000). Tall is typical: Central tendency, ideal dimensions and graded category structure among tree experts and novices. *Memory & Cognition*, 28, 41–50.
- Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, 109, 235–250.
- Majid, A., Enfield, N. J., & van Staden, M. (2006). Parts of the body: Cross-linguistic categorisation. *Special Issue of Language Sciences*, 28.

- Majid, A., Gullberg, M., Van Staden, M., & Bowerman, M. (2007). How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. *Cognitive Linguistics*, 18, 179–194.
- Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology*, 29, 85–148.
- Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *WIREs Cognitive Science*, 4, 583–597.
- 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. <http://dx.doi.org/10.1111/j.1467-9280.2008.02074.x>.
- Malt, B. C., Gennari, S., & Imai, M. (2010). Lexicalization patterns and the world-to-words mapping. In B. C. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 29–57). Oxford: New York.
- Malt, B. C. (2010). Naming artifacts: Patterns and processes. In B. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 1–38). Amsterdam: Elsevier.
- 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., & Wolff, P. (2010). *Words and the mind: How words capture human experience*. Oxford: New York.
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27, 209–220.
- McNorgan, C., Kotack, R. A., Meehan, D. C., & McRae, K. (2007). Feature–feature causal relations and statistical co-occurrences in object concepts. *Memory & Cognition*, 35, 418–431.
- Medin, D. L., & Atran, S. (1999). *Folkbiology*. Cambridge, MA: MIT Press.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108, 155–184.
- Parish-Morris, J., Pruden, S., Ma, W., Hirsh-Pasek, K., & Golinkoff, R. (2010). A world of relations: Relational words. In B. C. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 219–242). New York: Oxford University Press.
- Pavlenko, A., & Malt, B. C. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian-English bilinguals. *Bilingualism: Language and Cognition*, 14, 19–45.
- Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, 68, 29–46.
- Ratneshwar, S., Barsalou, L. W., Pechmann, C., & Moore, M. (2001). Goal-derived categories: The role of personal and situational goals in category representations. *Journal of Consumer Psychology*, 10, 147–157.
- Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, 104, 1436–1441.
- Rogers, T. T., & McClelland, J. L. (2004). Semantic cognition: A parallel distributed processing approach.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Rosenberg, S., & Kim, M. P. (1975). The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489–502.
- Rothkopf, E. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, 53, 94–101.
- Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition*, 118, 45–61.
- SAS Institute Inc (1999). *SAS STAT® users's guide 8*. Cary, NC: SAS Institute Inc.
- 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.
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (Vol. 2, pp. 219–257). Hillsdale, NJ: Erlbaum.
- Slobin, D. (2010). Linguistic relativity and determinism: Whence and whither? Plenary talk presented at the German Linguistics Society, Berlin, Germany.
- Slobin, D. (1996). Two ways of travel: Verbs of motion in English and Spanish. In M. Shibatani & S. Thompson (Eds.), *Grammatical constructions: Their form and meaning*. Oxford: Clarendon Press.
- Sturrock, K., & Rocha, J. (2000). A multidimensional scaling stress evaluation table. *Field Methods*, 12, 49–60.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description: Grammatical categories and lexicon* (Vol. 3, pp. 56–149). Cambridge: Cambridge University Press.
- Trudgill, P. (2002). *Sociolinguistic variation and change*. Washington, DC: Georgetown University Press.
- Vulchanova, M., Martinez, L., & Vulchanov, V. (2012). Distinctions in the linguistic encoding of motion: Evidence from a free naming task. In M. Vulchanova & E. Van der Zee (Eds.), *Motion encoding in space and language*. Oxford University Press.
- Wattenmaker, W. D. (1993). Incidental concept learning, feature frequency, and correlated properties. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 203–222.
- Wolff, P., Jeon, G.-H., Klettke, B., & Li, Y. (2010). Force creation and possible causers across languages. In B. C. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 93–110). New York: Oxford University Press.