Meaning and Motor Action

Daniel Casasanto  
(djc@psych.stanford.edu)  
Sandra Lozano  
(scl@psych.stanford.edu)

Department of Psychology, Stanford University  
Jordan Hall, Bldg. 420, Stanford, CA 94043

Abstract

Two experiments investigated relations between manual motor actions and the meanings of abstract words. In Experiment 1, participants told stories with literal or metaphorical spatial content while moving marbles from one box to another continually, in a specified direction (i.e., up, down, right, or left). Although marble movements were irrelevant to storytelling they influenced verbal fluency dramatically, increasing fluency relative to baseline when the direction of movement was congruent with the spatial schema implied by the accompanying speech (e.g., upward movement during ‘the temperature went up’ or ‘my grades got better’), and decreasing verbal fluency when movement was incongruent with speech (e.g., upward movement during ‘my teams rank fell’ or ‘the price got cheaper’).

Experiment 2 showed complementary ‘motor-meaning congruity’ effects in a non-communicative setting. Participants performed a marble moving task while incidentally reading irrelevant abstract words. Marble movements were fastest when the direction of movement was congruent with the spatial schema implied by the word (e.g., upward for genius; downward for gloomy). Furthermore, motor-meaning congruity during word presentation strongly predicted subsequent memory for incidentally encoded words. Together, these motor-meaning congruity effects suggest that motor representations play a role in instantiating abstract concepts that transcend any role they may play in communicating these concepts.

Keywords: Metaphor; Motor-Meaning Congruity Effect; Gesture; Space; Embodied Cognition.

Introduction

Metaphor and Embodiment

Abstract concepts like time, truth, or value present a challenge for any ‘embodied’ theory according to which thoughts are perceptual simulations: how can we perceptually simulate things we can never perceive? Cognitive Linguists offer a potential answer. According to Conceptual Metaphor Theory (Lakoff & Johnson, 1999), linguistic metaphors (e.g., a long time, a high grade, a deep thought) show that many of our abstract ideas are structured in terms of a few simpler concepts grounded directly in perceptuo-motor experience (e.g., experience with physical motion, force, and space). This idea is supported by an impressive body of linguistic theory and data (e.g., Clark, 1973; Jackendoff, 1983; Sweetser, 1991; Talmy, 1988), and more recently, by a growing body of experimental psychological data (e.g., Boroditsky, 2000; Casasanto & Boroditsky, In Press; Casasanto & Lozano, 2006; Meier & Robinson, 2004; Torralbo, Santiago, & Lupiáñez, 2006).

Within Cognitive Linguistics, Embodiment Theory and Conceptual Metaphor Theory are assumed to be mutually inextricable (Lakoff & Johnson, 1999). Yet, ‘embodied cognition’ researchers in other subfields of Cognitive Science have marginalized the role of metaphor in the mental representation of abstract concepts (e.g., Barsalou, 1999; Barsalou & Wiemer-Hastings, 2005; Prinz, 2002). Barsalou (1999), for example, argues that “a direct, non-metaphorical representation of an abstract domain is essential,” and proposes that “perceptual symbol systems can represent all abstract concepts directly” (pg. 600). This is somewhat surprising, as Conceptual Metaphor seems a likely component of the solution to the problem of mentally representing the imperceptible via perceptual simulations. To the extent that perceptuo-motor schemas constitute the content of abstract concepts, these concepts can be instantiated by the same neural and mental structures that simulate perception and action in the physical world.

Some skepticism about connections between Conceptual Metaphor and Embodiment may be warranted, however, because the first behavioral studies that explored links between metaphor and motor action were difficult to interpret. Glenberg & Kaschak (2002) tested Glenberg’s Indexical Hypothesis, according to which language understanding entails three processes: first, words are mapped to perceptual symbols (Barsalou, 1999). Next, physical affordances are derived from these symbols. Following Gibson (1979), Glenberg & Kaschak define affordances as “potential interactions between bodies and objects” (pg. 558). Finally, physical affordances are “meshed under the guidance of syntactic constructions” (pg. 559). The result is a cognitive simulation of the actions described in language which, by hypothesis, requires activation of the same neural structures that guide real action. Glenberg & Kaschak found that participants were fastest to decide that a sentence was sensible when their response required a hand movement in the direction that was implied by the sentence. This was true not only for sentences describing “concrete transfer” (e.g., movement away from the body following You give Liz the toy) but also for sentences describing “abstract transfer” (e.g., movement away from the body following You told Liz the story).

The latter result is frequently cited as evidence that abstract concepts rely on cognitive simulations of physical action (e.g., Barsalou & Wiemer-Hastings, 2005), yet this conclusion should be treated as tentative, for at least two
In the present experiments, we sought to address some of the limitations of these previous experiments, and also to expand the scope of inquiry in several ways. Whereas previous studies have focused on activation of perceptuo-motor schemas during language comprehension, Experiment 1 explored connections between action, concepts, and language during speech production. Schema-congruity effects of action on language production and of language production on action were measured independently. Experiment 2 tested whether spatio-motor schemas are activated automatically, even when single words referring to abstract concepts are read incidentally, in a non-communicative setting.

**Experiment 1: Moving Stories**

Experiment 1 builds upon previous experiments examining spontaneous gestures during storytelling (Casasanto & Lozano, 2006). In one experiment, pairs of participants took turns reading and retelling brief stories to one another. Each participant retold three types of stories:

1. **Literal Spatial Language stories (LSL)**
   - (e.g., my rocket went higher…)
2. **Metaphorical Spatial Language stories (MSL)**
   - (e.g., my grades went higher…)
3. **Non-Spatial Language stories (NSL)**
   - (e.g., my grades got better…)

Literal Spatial Language (LSL) stories described physical objects and events oriented along either a horizontal or vertical axis, and directed either upward, downward, right, or left. Metaphorical Spatial Language (MSL) stories described non-spatial phenomena that are nevertheless commonly expressed using spatial metaphors implying both orientation and direction (e.g., looking forward to the future; the price went lower). Non-Spatial Language (NSL) stories were identical to the MSL stories, except that all metaphorical spatial language was replaced with non-spatial language conveying nearly the same meaning (e.g., anticipating the future; the price got cheaper).

Although participants were not instructed to gesture, their spontaneous gestures were overwhelmingly congruent with the spatial schemas implied by the target phrases in the stories they told. Overall, 87% of gestures matched our *a priori* predictions about their orientations and directions based on the spatial trajectories implied by the stories. This was true not only in the Literal Spatial Language condition (e.g., upward gesture accompanying “the rocket went higher”), but also in the Metaphorical Spatial Language condition (e.g., upward gesture accompanying “my grades went higher”; rightward gesture accompanying “looking forward to the future”), and critically in the Non-Spatial Language condition, as well (e.g., upward gesture accompanying “my grades got better”; rightward gesture accompanying “anticipating the future”). Schema congruity did not differ across conditions: even when participants used no spatial language, their gestures revealed that they formed spatio-motor representations corresponding to the abstract ideas they expressed.
A second study began investigating the functional significance of these gestures. A new group of participants told the same stories while holding down buttons on a keyboard with both hands (supposedly to activate the left and right microphone channels that recorded their voices). Comparison of verbal disfluency between the first and second groups showed that participants who were prevented from gesturing produced disfluencies at a much higher rate than participants who were allowed to gesture. This increase was found during target clauses with literal and metaphorical spatial content, but not during non-target clauses with no spatial content, which served as a control. Importantly, this effect did not differ across LSL, MSL, and NSL conditions: gesture prevention led to similar increases in disfluency for targets with literal and metaphorical spatial content -- even when no spatial language was used.

The present experiment was designed to investigate why speech with literal and metaphorical spatial content is more fluent when speakers are allowed to gesture than when they’re not, and to determine how effects of gesture prevention on speech fluency inform theories of concepts and word meaning. Broadly speaking, it is possible that gestures could benefit speech production either because of their *semiotic function* or because of their *physical form*. On this first possibility, in Peirce’s (1960) terms both iconic and metaphorical gestures can be considered to be *representamen* that denote *objects*, in much the same way that words do. An upward gesture can denote the concept of *rising* in the sender’s mind (whether literal or metaphorical rising) similarly to the way the words ‘rising’ or ‘going up’ can. It is possible that speakers describe spatial concepts more fluently when gesturing because producing one representamen (i.e., the gesture) facilitates production of another representamen (i.e., the word or phrase) that denotes the same object.

Alternatively, iconic and metaphorical gestures may facilitate speech with spatial content because the motor programs that generate directed hand and arm movements also activate or reinforce the spatio-motoric schemas that partly constitute concrete and abstract word meanings. If so, then the ‘active ingredient’ in these gestures is motor action, per se; speech is facilitated to the extent that the motor programs underlying gestures reinforce the spatio-motor schemas underlying word meanings.

It is not possible to distinguish these alternatives based on existing data: during free gesturing the *semiotic function* and the *physical form* of gestures are conflated, and during traditional gesture prevention experiments they are both eliminated. Thus, we created a dual-task experiment to distinguish effects of the semiotic and the purely motoric components of gesture on speech production. We administered a version of the storytelling task described above which prevented participants from making meaningful gestures, but also forced them to execute directed motor actions. While telling stories, participants continually transferred marbles between pairs of boxes that were stacked on top of one another and positioned on the right and left of the computer screen in front of them. Marble movements were timed by a metronome, and were either congruent or incongruent with the spatial schemas implied by the target ideas in the stories (e.g., ‘the price went lower’). We compared the rate of verbal disfluencies produced during target clauses accompanied by schema congruent vs. schema incongruent marble movements. Additionally, we planned comparisons between data from the current experiment and data from previous experiments in which the same stories were told under free gesturing and traditional gesture prevention conditions.

**Methods**

**Materials and procedure** 16 Stanford University undergraduates were recruited in pairs, and participated in exchange for course credit. Participants took turns telling brief stories (50-100 words). Each story contained one of 3 types of language (Literal Spatial Language, Metaphorical Spatial Language, or Non-Spatial Language, see examples above), and the target items in each story implied motion or
extension in one of 4 directions (upward, downward, right, or left). After telling a warm-up story, each participant told 6 target stories: 2 LSL, 2 MSL, and 2 NSL. Each story contained between 3 and 7 targets. The order of stories was randomized, and content of stories was counterbalanced so that each pair of participants received only one version of each metaphorical story: either MSL or NSL (i.e., one pair of Ss would receive the MSL story about ‘grades going higher’, another pair of Ss would receive the minimally paired NSL story about ‘grades getting better’). (For further details regarding stories, see Casasanto & Lozano, 2006.)

The marble apparatus was constructed out of Xerox paper boxes and packing tape. Participants told the stories while transferring marbles with both hands between boxes that were stacked on top of each other, and positioned on the right and left of the computer screen in front of them. They were instructed to deposit one marble with each hand into the appropriate box at the instant that they heard a metronome tick, once every 2000 ms (this rate was selected after pilot Ss found they were unable to reliably complete faster marble movements). Each pair of participants was assigned to make either upward, downward, leftward, or rightward movements, so their marble movements were always in the same direction. Since each participant told stories with target items implying motion or extension in all four directions, marble movements were congruent with some of their target clauses and incongruent with others. Stories and marble directions were counterbalanced across participants.

**Data coding and analysis** Marble movements were classified as either congruent with the spatial schema implied by a target clause (correct axis and correct direction: e.g., upward movement during a target clause implying upward motion), incongruent (correct axis but incorrect direction: e.g., upward movement during a target clause implying downward motion), or orthogonal (incorrect axis: e.g., upward movement during target clauses implying leftward motion). Due to space limitations, the analyses reported here will focus on congruent and incongruent trials.

Participants’ video recordings were surveyed to make sure they followed instructions. Their audio recordings were transcribed and parsed into clauses that contained target ideas (Target Clauses) and those that did not (Non-Target Clauses). Two independent coders determined whether each clause contained a verbal disfluency, and recorded the types of any disfluencies found (i.e., repeats, repairs, fillers, and insertions). We computed the number of Target Clauses and Non-Target Clauses that contained disfluencies, and disfluency rates were compared for target phrases accompanied by schema congruent and schema incongruent marble movements. Inter-coder agreement exceeded 99%. Finally, we computed the rate of success with which participants executed a schema congruent and incongruent marble movement (i.e., delivered a marble to the appropriate box, or dropped it while trying).

**Results and Discussion**

Overall, on average speakers produced disfluencies during 62% of target clauses accompanied by schema incongruent marble movements, but only produced disfluencies during 1% of target clauses accompanied by schema congruent movements (difference of means = 61%, t(15)=33.50, p<.000001; Fig. 1 columns a and d). Disfluency rates were nearly identical across all language conditions, both for target clauses (M\textsubscript{MSL}=62%, M\textsubscript{MSL}=63%, M\textsubscript{NSL}=62%) and for non-target clauses (M\textsubscript{MSL}=6%, M\textsubscript{MSL}=6%, M\textsubscript{NSL}= 6%). Planned comparisons of these results with results from previous experiments revealed that disfluency rates for target clauses varied significantly as a function of the presence or absence of gestures, and of the congruity of the motor actions that accompanied storytelling (F(3,4)=1277.64, p<.000001). Storytellers produced disfluencies at a higher rate during schema incongruent marble movements than during gesturing prevention via button holding (difference of means =.17, t(42)=36.76, p<.000001; Fig. 1 columns a and b), and produced disfluencies at a significantly lower rate during schema congruent marble movements than during free gesturing (difference of means = .08, t(42)=5.25, p<.000001; Fig. 1 columns c and d).

**Figure 1.** Proportion of target clauses containing verbal disfluencies under different storytelling conditions. Bars indicate s.e.m.

Not only did marble movements affect participants’ ability to tell stories, but storytelling also affected their ability to complete marble movements. Overall, participants dropped 4% of the marbles they attempted to move. The rate of marble dropping did not differ across language conditions (M\textsubscript{MSL}=4%, M\textsubscript{MSL}=4%, M\textsubscript{NSL}= 4%). Strikingly, all marble dropping occurred during target clauses accompanied by schema incongruent marble movements. A sign test showed that the total number of marbles dropped during schema incongruent movements was significantly greater than the number dropped during schema congruent movements (45 vs. 0; p=5.86x10\textsuperscript{-14}). When their hand movements were incongruent with the literal or metaphorical spatial content of their speech, our participants were quite literally losing their marbles.

These results strongly support the proposal that motor actions, per se, are the active ingredient in the cognitive function of literal and metaphorical gestures, and suggest that their physical form can influence speech fluency,
independent of any semiotic function. Additionally, these data show that conguity effects between motor action and word meaning extend beyond language comprehension to language production. Furthermore, influences of language and action are bidirectional: using language with literal or metaphorical spatial content influences the execution of simple, irrelevant motor actions – and vice versa.

**Experiment 2: Marbles, Metaphor, & Meaning**

Experiment 1 showed that metaphorical gestures – and even directed motor actions with no semiotic value – facilitated speech about abstract concepts. This was true even when the target words had no literal spatial meaning (e.g., *better, hotter*), supporting the claim that the abstract concepts these words refer to are partly constituted by spatio-motor image schemas. Yet, on a skeptical interpretation, these results only show that motor actions influence the kinds of mental representations people form for the purpose of communicating (i.e., they influence the process Slobin (1987) called *thinking for speaking*). Experiment 2 tested whether motor actions play a role in accessing or constituting abstract word meanings more generally, even in non-communicative contexts.

In another ‘motor-meaning conguity’ task, participants incidentally read abstract words referring to metaphorically spatialized concepts while moving marbles into boxes positioned so as to require schema congruent or inconguent motor actions. We reasoned that if spatio-motor schemas are automatically activated when people see abstract words – even if they’re not planning to speak the words or communicate their content, and even if their meaning is irrelevant to the task – then schema conguity should facilitate the marble movements themselves (an on-line effect), and should also facilitate subsequent memory for the stimulus words (an off-line effect).

**Methods**

**Materials and Procedure** 18 Stanford University students participated in exchange for course credit. Individual participants were seated in front of a computer screen. Stacked next to the screen on both the right and left there were three large boxes. The top box was red, and the bottom box was blue (or vice versa, counterbalanced across subjects). The middle box was white, and was filled with hundreds of marbles. 48 words appeared briefly (2000 ms) in the center of the computer screen, one at a time, in either blue or red font. The words were nouns and adjectives referring to abstract entities that have no spatial instantiation in the world, but which are often associated with vertical spatial metaphors (e.g., *wealthy, poor, virtuous, evil, joy, disgust*, etc.), and which subjects in a previous norming study spatialized accordingly (e.g., placing *wealthy* above *poor, virtuous* above *evil*, etc.) Participants were instructed that as soon as each word appeared, they should move one marble with each hand into the box corresponding to the color of the word’s font, as quickly as possible. Unknown to the participants, half of their marble movements were schema-conguent (e.g., if the word *wealthy* appeared in blue and the blue box was on top) while the other half were schema-inconguent (e.g., if the word *wealthy* appeared in red and the red box was on bottom). Word order was randomized, and the assignment of font colors to words was counterbalanced across subjects. A video camera placed behind the participant recorded their marble movements. The color of the words was their only attribute that was relevant for performing the marble task. Participants were not instructed to read or remember the words. After completing the marble moving task, they performed a surprise old/new recognition memory task during which all words presented previously as cues for marble movements were presented again in black font, randomly intermixed with an equal number of distractor words matched with targets for length, frequency, valence, and concreteness.

**Results and Discussion**

Participants were dramatically faster to make schema congruent marble movements than schema inconguent marble movements, both when analyzed by subjects (difference of means = 308 ms, $t(17)=13.50, p<.00000001$; Fig. 2a) and by items (difference of means = 304 ms, $t(94)=14.00, p<.0000001$). We found an equally dramatic effect of schema conguity on recognition memory. On average, participants correctly recognized 94% of words incidentally encoded during schema congruent marble movements, but were at chance recognizing words encoded during schema inconguent movements (54% correct recognition; difference of means = 0.40, $t(17)=10.99, p<.00000001$; Fig. 2b). Finally, even though participants were only required to process stimulus words superficially according to their font color, words that required schema inconguent movements caused participants to lose their marbles significantly more often than words that required schema congruent movements (12 vs. 1, $p<.003$).

![Figure 2. Experiment 2 results. 2a (left) Time to make schema congruent and schema inconguent marble movements. 2b. (right) Correct recognition for words seen during schema congruent and schema inconguent marble movements (chance performance = 50%). Bars indicate s.e.m.](image-url)
action and language that we observed in Experiment 1 are not restricted to communicative contexts; the relationship between meaning and motor action extends beyond thinking for speaking.

**General Discussion and Conclusions**

Despite the close association between Conceptual Metaphor and Embodiment Theory in the Cognitive Linguistics literature, embodied cognition researchers in other areas of Cognitive Science have pursued alternative accounts of how abstract concepts are structured, and of how they might be implemented by Perceptual Symbol Systems (Barsalou, 1999; Prinz, 2002). Barsalou and colleagues, for example, have proposed that the embodied meanings of abstract concepts may emerge from the “situations” in which they are used. While this hypothesis may eventually be validated by behavioral tests, so far it is supported only by descriptive analyses and in-principle arguments. In outlining this proposal, Barsalou and Wiemer-Hastings (2005) wrote that they would “not attempt to provide evidence that modality-specific systems represent abstract concepts” (pg. 129). Yet, this is exactly what embodiment theorists *must do* if Perceptual Symbols theory is to remain viable. Abstract concepts are not esoteric oddities that can be bracketed while we investigate ‘ordinary’ concepts (e.g., the 11th O.E.D. reports that the most frequent noun in the English language is *time*). Arguably, the majority of our words and concepts are abstract, so it’s quite problematic for Perceptual Symbols theory if its proponents cannot provide evidence that modality-specific systems can represent them.

The primary objection to Metaphor Theory that Barsalou and colleagues raise is that abstract concepts must have some non-metaphorical structure of their own, or else there would be nothing for metaphorical structures to map onto (Barsalou & Wiemer-Hastings, 2005, pg. 133-134), therefore metaphor theory offers an incomplete explanation. But far from leveling a devastating critique, they are simply (and accurately) describing the source domain - target domain relation that is the essence of metaphor theory. Interenceptive representations in abstract domains like *time*, *love*, or *anger* are further structured by motoric and perceptual representations in more concrete domains like *space*, *motion*, and *force*. Importantly, both the interoceptive representations of target domains and the perceptuo-motor representations of source domains are of the sort that could plausibly be ‘embodied’ in Perceptual Symbols. As such, Conceptual Metaphor may be Embodiment Theory’s best hope for accommodating abstract concepts, thereby avoiding non-starter status.

In the present study, we demonstrate automatic activation of spatio-motor schemas that appear integral to the meanings of dozens of abstract concepts. Much remains unknown about how perceptuo-motor representations structure target concepts, and Metaphor Theory may never provide a complete answer. As Barsalou and colleagues propose, the contexts in which we instantiate abstract concepts may contribute to their form and content. But Metaphor Theory appears to provide *part* of the answer, and no theory of abstract concepts -- embodied or otherwise -- will be complete unless it predicts the kind of relations between meaning and motor action that we demonstrate here.

**Acknowledgements**

This research was supported in part by NRSA post-doctoral fellowship #F32MH072502 to D.C.

**References**


