8 Embodiment, simulation and meaning

Benjamin Bergen

1 Introduction

Approaches to meaning differ in ways as fundamental as the questions they aim to answer. The theoretical outlook described in this chapter, the *embodied simulation* approach, belongs to the class of perspectives that ask how meaning operates in real time in the brain, mind, and body of language users. Clearly, some approaches to meaning are better suited to this question than others. Mechanistic models that bridge levels of analysis—from the brain and its computations to the functions and behaviours they support—and that recruit the convergent tools of empirical cognitive science are particularly well equipped. The embodied simulation approach is an example of this type of approach.

The fundamental idea underlying the embodied simulation hypothesis is a remarkably old one. It's the notion that language users construct mental experience of what it would be like to perceive or interact with objects and events that are described in language. Carl Wernicke described the basic premise as well as anyone has since (and with remarkably little modern credit, as Gage and Hickok (2005) point out). Wernicke wrote, in 1874:

> The concept of the word "bell," for example, is formed by the associated memory images of visual, tactual and auditory perceptions. These memory images represent the essential characteristic features of the object, bell.
>
> (Wernicke 1977 [1874]: 117)

This is the essence of simulationism. Mental access to concepts involves the activation of internal encodings of perceptual, motor, and affective—that is, *modality-specific*—experiences. This proposal entails that understanding the meaning of words involves activating modality-specific representations or processes. Wernicke came to this notion through his work on localization of cognitive functions in the brain, and as a result, it should be no surprise that he had a very clear view of what the neural substrate of these "memory images" would be and where it would be housed:

> the memory images of a bell [. . .] are deposited in the cortex and located according to the sensory organs. These would then include the acoustic imagery aroused by the sound of the bell, visual imagery established by means of form and color, tactile imagery

> acquired by cutaneous sensation, and finally, motor imagery gained by exploratory movements of the fingers and eyes.
>
> (Wernicke 1977 [1885–1886]: 179)

In other words, the same neural tissue that people use to perceive in a particular modality or to move particular effectors would also be used in moments not of perception or action but of conception, including language use. This, in words now 130 years old, is the embodied simulation hypothesis.

Naturally, this idea has subsequently been developed in various ways. Part of its history involves some marginalization in cognitive science, especially starting in the 1950s with the advent of symbolic approaches to cognition and language. If the mind is a computer, and a computer is seen as a serial, deterministic, modular symbol system, then there is no place for analog systems for perception and action to be reused for higher cognitive functions like language and conceptualization.

But more recent history has seen substantial refinement of the embodied simulation hypothesis, on three fronts. First, cognitive psychologists came to the idea because of the so-called "symbol grounding" problem (Harnad 1990). In brief, the problem is this: if concepts are represented through symbols in the mind, these symbols must somehow be grounded in the real world, or else they don't actually mean anything. For instance, if mental symbols are only defined in terms of other mental symbols, then either there must be core mental symbols that are innate and serve as the basis for grounding meaning (see e.g. Fodor (1975)), or symbols must relate to the world in some meaningful way. Otherwise, they are ungrounded and meaningless. This is a hard problem, and as a result, some cognitive psychologists began to suggest that perhaps what people are doing during conceptualization doesn't involve abstract symbol manipulation, but rather manipulation of representations that are like action and perception in kind (Barsalou 1999). In essence, perhaps the way out of the symbol grounding problem is to get rid of the distance (or "transduction" as Barsalou et al. 2003 argue) between perception and action on the one hand and the format of conceptual representation on the other. (For a more complete account of transduction, see Chapter 2 on internalist semantics.)

A second branch of work that pointed towards embodied simulation came from *cognitive semantics*. This is an approach to analytical linguistics that aims to describe and explain linguistic patterning on the basis of conceptual and especially embodied individual knowledge, experience, and construal (Croft and Cruse 2004). Cognitive semanticists argue that meaning is tantamount to conceptualization—that is, it is a mental phenomenon in which an individual brings their encyclopedic experience to bear on a piece of language. Making meaning for a word like *antelope* involves activating conceptual knowledge about what antelopes are like based on one's own experience, which may vary across individuals as a function of their cultural and idiosyncratic backgrounds. The idea of embodied simulation dovetails neatly with this encyclopedic, individual, experiential view of meaning, and cognitive semanticists (see Chapter 5) were among the early proponents of a reinvigorated embodied simulation hypothesis.

And finally, action-oriented approaches to robotics and artificial intelligence pointed to a role for embodied simulation in language. Suppose your goal is to build a system that is able to execute actions based on natural language commands. You have to build dynamic motor control structures that are able to control actions, and these need to be selected and parameterized through language. In such a system, there may be little need for abstract symbols to represent linguistic meaning, except at the service of driving the motor actions. But the very same architecture required to enact actions can also be used to allow the system to also understand language even when not actually performing actions. The theory of meaning that grew from this work, and its name, *simulation semantics* (Feldman and Narayanan 2004), is one implementation of the embodied simulation hypothesis.

In the past decade, embodied simulation has become a bona fide organized, self-conscious enterprise with the founding of a regular conference, the Embodied and Situated Language Processing workshop, as well as publication of several edited volumes (Pecher and Zwaan 2005) and books (Pulvermüller 2003; Bergen 2012). It's important to note that none of these approaches view simulation as necessary or sufficient for all meaning construction—indeed, one of the dominant ongoing research questions is precisely what functional role it performs, if any. The varied simulationist approaches merely propose simulation as part of the cognitive toolkit that language users bring to bear on dealing with meaning in language.

2 Current research on simulation

While most current work on simulation in linguistic meaning-making is empirical, as the review in this section will make clear, this empirical work is motivated by introspective and logical arguments that something like simulation might be part of how meaning works in the first place.

One such argument derives from the symbol grounding problem, mentioned in the previous section. Free-floating mental symbols have to be grounded in terms of something to mean anything. One thing to tether symbols to is the real world—symbol-world correspondences allow for truth-conditional semantics (Fodor, 1998; see Chapter 1). Another thing to ground symbols in is other symbols—inspired, perhaps, by Wittgenstein's proposal that meaning is use (Wittgenstein 1953). On this account, exemplified by distributional semantic approaches like HAL (Lund and Burgess 1996) and LSA (Landauer et al. 1998), to know the meaning of a symbol, you need only know what company it keeps. However, as Glenberg and Robertson (2000) demonstrate, these word- or world-based approaches to grounding both fail to make correct predictions about actual human processing of language.

Another argument is based on parsimony of learning, storage, and evolution. Suppose you're a language learner. You have perceptual and motor experiences in the world, which are processed using specific brain and body resources that are well tuned and appropriately connected for these purposes. To reuse these same systems in a slightly different mode seems more parsimonious than would be transducing the patterns of activation in these systems into some other representational format (abstract symbols, for instance) that would need to recapitulate a good deal of the same information in a different form. The same argument goes for subsequent storage—storing two distinct versions of the same information in different formats could potentially increase robustness but would decrease parsimony. And similarly, over the course of evolution, if you already have systems for perceiving and acting, using those same systems in a slightly different way would be more parsimonious than introducing a new system that represents transduced versions of the same in a different format.

Finally, from introspection, many people are convinced that something like simulation is happening because they notice that they have experiences of imagery (the conscious and intentional counterpart of simulation) while processing language. Processing the words *pink elephant* leads many people to have conscious visual-like experiences in which they can inspect a non-present visual form with a color that looks qualitatively like it's pink and has a shape that looks qualitatively like that of an elephant, from some particular perspective (usually from the right side of the elephant).

But each of these arguments has its weaknesses, not least of which is that they can't inform the pervasiveness of simulation, the mechanisms behind it, or the functions it serves. To address these issues, a variety of appropriate empirical tools have been brought to bear on the question, ranging from behavioural reaction time experiments to functional brain imaging.

2.1 Behavioural evidence

The largest body of empirical work focusing on simulation comes from behavioural experimentation. For the most part, these are reaction time studies, but there are also eye-tracking and mouse-tracking studies that measure other aspects of body movement in real time as people are using language. Generally, these behavioural studies aim to infer whether people are constructing simulations during language use, and if so what properties these simulations might have, what factors affect them, and at what point during processing they're activated.

Reaction time studies of simulation generally exhibit some version of the same basic logic. If some language behaviour, say understanding a sentence, involves activating a simulation that includes certain perceptual or motor content, then language on the one hand and perception or action on the other should interact. For instance, when people first process language and then have to subsequently perceive a percept or perform an action that's compatible with the implied or mentioned perceptual or motor content, they should be faster to do so than when the percept or action is incompatible. For example, processing a sentence about moving one's hand toward one's body (like *Scratch your nose!*) leads to faster reactions to press a button close to the body. Conversely, sentences about action away from the body (like *Ring the doorbell!*) lead to faster responses away from the body (Glenberg and Kaschak 2002). Similarly, a sentence that describes an object in a vertical orientation (like *The toothbrush is in the glass*) leads to faster responses to an image of that vertical object, while sentences about objects in a horizontal orientation (like *The toothbrush is in the sink*) lead to faster processing of horizontal images of the same object (Stanfield and Zwaan 2001).

Compatibility effects like these demonstrate that language processing primes perceptual and motor tasks, in ways that are specifically sensitive to the actions or percepts that language implies. Similar designs have demonstrated that comprehension primes not only the direction of action and orientation of objects, but also the effector used (hand, foot, or mouth), hand shape, direction of hand rotation, object shape, direction of object motion, visibility, and others (see Bergen (2012) for a review).

One of the most interesting features of this literature is that there are various experiments in which the priming effect appears—superficially—to reverse itself. For example, Richardson et al. (2003) found that language about vertical actions (like *The plane bombs the city*) lead to slower reactions to circles or squares when they appear along the vertical axis of a computer monitor (that is directly above or below the center of the screen) while language about horizontal actions (like *The miner pushes the cart*) lead to slower reactions along the horizontal axis. Other experiments have reported findings of this same type (Kaschak et al. 2005; Bergen et al. 2007).

At the surface, this might seem problematic, but a leading view at present is that these two superficially contradictory sets of findings are in fact consistent when the experimental designs are considered closely. In fact, they may reveal something important about the neural mechanisms underlying the different effects. Richardson et al.'s (2003) work is a good case study. In their experiment, a circle or square was presented on the screen with only a slight delay after the end of the preceding sentence (50–200msec). With the time it takes to process a sentence, this meant that the participant was still processing the linguistic stimulus when the visual stimulus was presented. So the two operations—comprehending the sentence and perceiving the shape—overlapped. That's design feature number one. Second, the objects described in the sentences in this study (such as bombs or carts) are visually distinct from the circles and squares subsequently presented. That is, independent of where on the screen they appeared, the mentioned objects did not look visually like the mentioned objects. Other studies that find interference effects (Kaschak et al. 2005; Bergen et al. 2007; Yee et al. 2013) have the same design features—the language and the visual stimulus or motor task have to be dealt with simultaneously, and in addition, the two tasks are non-integrable—they involve the same body-part performing distinct tasks (Yee et al. 2013) or distinct visual forms (Kaschak et al. 2005).

Interference findings like these are often interpreted as suggesting that the two tasks (language use on the one hand and perception or motor control on the other) use shared neural resources, which cannot perform either task as efficiently when called upon to do two distinct things at the same time. By contrast, compatibility effect studies, like those that present language followed at some delay by an action or image that matches the implied linguistic content, do not call on the same resources to do different things at the same time, and as a result, do not induce interference but rather facilitation of a matching response.

One major weakness of reaction time studies like these is that they present a perceptual stimulus or require a physical action that matches the linguistic content or not. This raises the concern that it might only be this feature of the experimental apparatus that induces simulation effects. That is, perhaps people only think about the orientation of toothbrushes in the context of an experiment that systematically presents visual depictions of objects in different orientations. Perhaps the experiment induces the effects.

One way to methodologically circumvent this concern is with the use of eye-tracking. Several groups have used eye-tracking during passive listening as a way to make inferences about perceptual processes during language processing. For instance, Spivey and Geng (2001) had participants listen to narratives that described motion in one direction or another while looking at a blank screen, and while the participants believed the eye-tracker was not recording data. The researchers found that the participants' eyes were most likely to move in the direction of the described motion, even though they had been told that this was a rest period between the blocks of the real experiment. Another study (Johansson at al. 2006) first presented people with visual scenes and then had them listen to descriptions of those scenes while looking at the same scene, looking at nothing, or looking at nothing in the dark. They found that people's eye movements tracked with the locations of the mentioned parts of the scene. Both studies suggest that even in the absence of experimental demands to attend to specific aspects of described objects, actions, and scenes, people engage perceptual processes. This is consistent with the idea that they perform simulations of described linguistic content, even when unprompted by task demands.

2.2 Imaging

Behavioural evidence provides clues that people may be activating perceptual and motor knowledge during language use. Brain imaging research complements these findings, by allowing researchers to ask where in the brain there is differential activity when people are using language of one type or another. Modern models of functional brain organization all include some degree of localization of function—that is, to some extent there is neural tissue in certain locations that performs certain computations that contribute differently to cognition than other neural tissue does. For example, there are parts of the occipital lobe, such as primary visual cortex, that are involved in the calculation of properties from visual stimuli, and parts of the frontal lobe, like primary motor cortex and premotor cortex, that are involved in controlling motor action. Brain scanning, such as functional magnetic resonance imaging (fMRI), permits a measure to be taken of where activity is taking place in the brain—in the case of fMRI, this is blood flow, which increases as a consequence of neuronal firing. By comparing the fMRI signals obtained while people are performing different tasks, it's possible to localize differences in how the brain responds across those tasks.

Dozens of brain imaging studies have looked at what happens in people's brains when they're presented with language that has different semantic content, and the findings are relatively clear. When people are processing language about motor actions, there's an increased signal in motor areas, as compared with language not about motor actions. This signal in the motor system observed during motor language processing is weaker than when people are actually moving their bodies, and overlaps but may not be fully co-extensive with the area in which a signal is observed while people are performing intentional imagery of motor actions (Willems et al. 2010). But the signal is present even when people are not asked to think deeply about the meanings of the sentences they're presented with. Similarly, language that describes visual scenes leads to an increased signal coming from the brain's vision system. For instance, language about motion leads to increased activity in the medial temporal lobe, which houses a region implicated in the visual processing of motion (Saygin et al. 2010).

Brain imaging techniques are often criticized for their limitations—for instance, the subtractive approach they typically adopt doesn't afford insight into the actual function of the implicated brain areas, their temporal resolution is often poor (with fMRI, it's on the order of seconds) and they can only be used to compare grossly contrasting stimuli, like language about motion versus static language. But they are a critical component of a methodologically triangulating approach to meaning. Behavioural methods can reveal details of timing and functional interaction between different cognitive mechanisms, but they can only indirectly reveal anything about location, which is imaging's strength. To complete the story, we need to know not only where and when, but also how. And that's what we'll turn to next.

2.3 Neuropsychology

Traditionally, the best type of evidence on what functions certain brain and body systems are used for—that is, what precisely they do for a particular behaviour—comes from localized brain damage. When damage to a particular part of the brain is accompanied by a cognitive impairment, but damage to some other area does not lead to the same cognitive impairment, that suggests that the first brain region but not the second is mechanistically involved in that particular cognitive behaviour. This logic—known as a *dissociation*—is the form of evidence that first led pathologists like Paul Broca and Carl Wernicke to be able to pair brain regions involved in language use with hypothesized functions in the nineteenth century. And it has been used consistently since then as the gold standard for localization of function in the human brain.

Despite the appeal of neuropsychological studies, they have clear limitations. Because it's not possibly to ethically induce brain lesions in humans, researchers are restricted to those brain insults that occur naturally due to stroke, traumatic brain injury, etc., which often leave damage that is distributed across the brain. As a result, neuropsychological studies often include participants who have lesions to similar or overlapping brain regions, but no two lesions will be the same. In addition, most patients are studied some time after the injury, which means that there will have been brain changes over the course of recovery that obscure the organization at the time of damage.

These limitations notwithstanding, there have been a few dissociation studies focusing on language use that differentiated across language with different content. For instance, Shapiro et al. (2005) showed that damage to the left temporal cortex, a region implicated in visual recognition of objects, often leads patients to lose the ability to access nouns describing physical objects. But damage to the left frontal cortex, an area dedicated to motor control, tends to lead to difficulties with verbs describing actions. This evidence is suggestive that parts of the brain used for perceiving objects and performing actions also underlie meanings of words.

2.4 Transcranial magnetic stimulation

Although it's not possible to lesion living human brains, techniques are available that temporarily disrupt activity in a particular region. Transcranial Magnetic Stimulation (TMS) is one of these—it involves the application of a strong electromagnet to the scalp, inducing a field that enters the cerebral cortex and modifies neuron behaviour locally. TMS is often used as a transient proxy for lesion studies—it can be applied to a specific brain area, and if doing so impairs some behaviour (and if applying TMS to other brain regions does not impair this same behaviour), then this is effectively a dissociation.

Several studies have reported on the result of applying TMS to specific brain regions during language processing, in the hope of determining whether access to motor or perceptual processes plays a mechanistic role in using language about action or percepts. For instance, Shapiro et al. (2001) found that applying TMS to motor regions interferes with producing verbs but not nouns. Verbs often describe actions, while nouns are less likely to do so. So it could be that the reason people have more trouble producing verbs when motor cortex is interfered with is that accessing knowledge about how to move the body is part of the machinery people use to produce verbs, more so than for nouns.

2.5 Adaptation

The inferential logic of dissociations is so useful that there's even some behavioural work that tries to harness it, but without relying on naturally incurred brain damage or artificially induced brain changes from applied magnetic fields. Instead, it's based on *adaptation*. When people perform certain behaviours continuously for a long enough duration, the neural structures that they rely on come to be less active over time. A classic example of this is motion adaptation effects—when you look at something moving in one direction for long enough, and then look away, the world appears to be moving in the opposite direction (the *waterfall illusion*). Recent studies have attempted to use adaptation paradigms to knock out certain brain structures and potentially also language processing capacities, to determine what functional role these structures play in language use.

For instance, Glenberg et al. (2008) had people move 600 beans by hand in a single direction—either towards or away from their body. They then had to make judgements about sentences that described motion towards or away from their body. And surprisingly, people were slower to make judgements about sentences that described motion in the *same* direction in which they had just moved the 600 beans. In other words, first adapting to a particular action makes people take longer to process language about a similar action.

2.6 Computational modeling

So there's a good deal of experimental evidence that simulation—or something like it—occurs during language use and even some indication that it might play a functional role. An experimental approach is one way to address the viability of the simulation hypothesis. In cognitive science, similar claims are often assessed in a second way as well, through computational modeling. By implementing a proposed model of how the proposed mechanism would work—in this case, simulation—and then observing how a system that incorporates that mechanism behaves, it's possible to ask a slightly different question: would this mechanism do what it's supposed to do? That is, a model can provide a practical proof-of-concept. Another side benefit of computational implementation is learning in detail about other mechanisms that would need to be in place for a system to actually do whatever it is supposed to do—for instance, understand language.

There has been some modeling work using simulation as part of a language comprehension system. The most extensively developed uses a particular model of grammar (Embodied Construction Grammar—Bergen and Chang 2005; 2013) as an interface between language processing and simulation. There's been a good deal of work in this paradigm, implementing simulation using dynamic computational control schemas and deploying those schemas to compute and propagate inferences for both literal and metaphorical language (Narayanan 1997).

Among the things that have been hammered home from modeling efforts like this one is that it's critical for the language processing system to implement an organizational interface between linguistic form and simulation. The fundamental issue is that to simulate usefully, a system has to know what to simulate. And to generate instructions to simulation, that system will need to take cues not only from the linguistic input but also context and world knowledge. An example might help clarify this point. Suppose you're dealing with a verb like *jump*. There are different simulations appropriate to different types of jumping. If *jump* appears in a sentence like *The high-jumper jumped*, that jumping will be different from *The triple-jumper jumped*. And both will be very different from *The train jumped* or *The cat jumped*. For a simulation to appropriately reflect the inferred intent of the speaker/writer, it has to be sensitive to contextual factors, and this implicates a process of assembling the cues from language and extralinguistic context. Different theorists have labeled this assembly process as "meshing" (Kaschak and Glenberg 2000) or "analysis" (Bergen and Chang 2005).

3 Toward a simulation semantics

Empirical evidence suggests that processing words and sentences leads to perceptual and motor simulation of explicitly and implicitly mentioned aspects of linguistic content. This could have consequences for theories of semantics, but not without raising a series of fundamental questions. What and how might words contribute to simulation? Does this differ as a function of properties of the word? (Perhaps words with more concrete referents like *dog* involve simulation differently from ones with more abstract referents, like *God.*) How do linguistic elements other than words contribute to and constrain meaning, including linguistic elements like idioms or grammatical constructions? And for that matter, what are the effects of context, including the linguistic and physical environment as well as the current belief and knowledge state of the language user? Current work attempting to develop a simulation-based theory of meaning has made some limited progress on these questions.

3.1 Limits of simulation

Based on the evidence presented in the previous section, it might be tempting to equate simulation with meaning—perhaps words are associated with specific but varied sensorimotor experiences, and the internal re-enactment of these experiences constitutes the meaning of those words (see e.g. Pulvermüller (2003)). But while there is indeed a great deal of evidence for such associations, this simplest account would fail to capture fundamental aspects of meaning.

To begin with, many words have syncategorematic meaning properties—their meaning depends on the words they co-occur with. The sensorimotor features that *big* contributes to a described scene are non-overlapping in *big dog* versus *big run*, for instance, not to mention more abstract uses like *big headache* or *big problem.* An even bigger problem is presented by well-known examples like *safe*, which evokes a relatively abstract frame and doesn't specify roles within it. A *safe beach* can be one on which one is unlikely to be in danger (for instance, from a dangerous shore break), but it can also be a beach that itself is not in danger (perhaps from erosion). Not only is it hard to pin down shared sensorimotor properties of these different uses of *safe*; more troublingly, the comprehender has to decide whether the noun it modifies is the thing that threatens safety of something else (and if so, what) or is itself under threat (and again, if so, of what).

These complications aren't limited to parade examples like these. As discussed in the previous section, even words with clear, low-level sensorimotor associations like *jump* display rampant polysemy and vagueness. And this can in principle only be navigated using contextual information, from low-level grammatical details to high-level expectations generated from situations of use. As a consequence, it appears to be that in general, meaning construction is an active process that involves the interplay of knowledge of context, encyclopedic knowledge, and prior expectations. All of this conspires to constrain what goes into the content of simulation. None of this is compatible with an account in which word-associated simulation *is* the meaning of a word.

Instead, there appear to be a variety of processes at play as people attempt to construct meaning from perceived language. (The emphasis in the literature on simulation has been placed on comprehension, since this is far easier to study than language production, and I continue with that bias here, though there's certainly much to say about simulation in language production; see, for instance, Hostetter and Alibali (2008), Sato (2010), and Parrill et al. (2013).) Simulation may play a role in meaning. But what role is still to be ascertained. And its role may vary as a function of factors like the comprehender's relation to the language. (Is it highly conventional? Is the comprehender paying close attention? Is she listening or reading to deeply understand, to act, or to get through the page as quickly as possible?)

Processes other than simulation are clearly at play, including superficial statistical ones—some words are more likely to invite a particular interpretation in some contexts than others. Knowledge about how grammar constrains combinatorial semantics is also in play. And beyond these are still higher-level pragmatic issues, like what a language user can infer about why a speaker would use a particular word or produce a particular utterance, what's presupposed, and implied, and so on.

3.2 Linguistic and extralinguistic sources of simulation

Simulation can be systematically affected by words, especially open-class words. There's now ample evidence that a noun or verb contributes content to be simulated. For instance, one study (Bergen et al. 2007) manipulated just whether a noun or a verb was associated with up or down, as in the examples below (the up- or down-associated words are italicized). The presence of just one such word in a sentence affected participants' ability to subsequently perceive a circle or square in the upper or lower part of the screen, which is interpreted as a consequence of location-specific simulation.

(1) a The *sky* darkened.
 b The *grass* glistened.
 c The chair *toppled*.
 d The mule *climbed*.

Much more generally, the presence of an open-class word can often be solely responsible for the appearance of a category of thing or event in simulation—a sentence that includes *toothbrush* will result in simulation of a toothbrush and possibly brushing of teeth.

But context—both linguistic and physical context—can affect the content of simulation. We've already seen that the words surrounding a given content word can affect the details of the simulation that corresponds to what it denotes—the orientation of a simulated toothbrush, for instance, depends on whether it's described in language as being found in the sink or a cup. There's also clear evidence that physical context affects simulation as well. It's been shown, for example, that when a person has seen an object in a particular configuration (for example, a vertical versus a horizontal toothbrush), their subsequent language-driven simulation processes are affected—it takes longer for someone who has previously seen a horizontal toothbrush in an unconnected task to read a sentence that implies a vertical toothbrush (Wassenburg and Zwaan 2010).

So at the minimum, simulation is sensitive to physical and lexical context. But it also appears that grammatical context plays a role in configuring simulation. Consider the difference between an action described using progressive versus perfect aspect, or active versus passive voice. Grammatical choices like these needn't change the details of the action itself, but they might change the simulation language users produce in response to them. Recent work shows that progressive aspect increases simulation of the nucleus of a described event (Bergen and Wheeler 2010), while perfect aspect augments endstate simulation (Madden and Zwaan 2003). Similarly, active voice might increase and passive voice might suppress simulation from the perspective of the agent (Bergen 2012). These effects of grammar constitute second-order contributions to simulation—they dictate not what to simulate but how to simulate it—what perspective to adopt or what part of the simulated event to focus on.

And it also appears that the goals of the language user or task-specific demands imposed on them affect simulation. Simply stated, people can approach language with different degrees of attention to semantic details. In experimental setups where they know that after a sentence, they'll be asked to decide whether an image depicts something mentioned in the sentence, people appear to activate more perceptual detail in simulation, as compared with situations where they are not asked to compare images with sentence content (Rommers et al. 2013). So the language user's orientation towards language—what they expect to need to do with it—can affect the richness of the perceptual simulations they construct (see Chapter 29 on semantic processing).

3.3 Simulation and abstract concepts

There are a number of obvious limits to what simulation might do for language use. Perhaps chief among these, simulation seems best suited to representing concrete perceptual or motor concepts, and less well suited for abstract concepts. Even if we're willing to stipulate that there's a case to be made that something in the semantics of *cat* or *jump* involves perceptual or motor simulation, the case seems much harder to make for *justice* or *truth* or *value*. Acknowledging this concern, some theorists have argued that concepts live on a cline from concrete to abstract, where the most abstract will have no use for simulation (Dove 2009; Chatterjee 2010).

But there are other perspectives worth discussing. The first, articulated by Barsalou and Wiemer-Hastings (2005), among others, argues that abstract concepts may be fleshed out through simulations of the situations in which those concepts are grounded. For instance, perhaps understanding the concept denoted by *truth* involves simulating the common experience of holding in mind a belief about the world, which is confirmed by observation of the world. For example, my experience of it being true that I'm writing this chapter on a laptop involves holding both the belief in mind that I'm writing on a laptop and the perceptual observation of the laptop in front of me. That match between belief and observation might ground my experience of what *truth* denotes. The experience is both situated and perceptually grounded in that it involves perceptual predictions and perceptual observations, and perhaps understanding the word *truth* involves running a simulation of a relevant prediction-perception match. Other senses of *truth* might involve not belief but assertions about the world, but the general form of the story would be similar. In sum, abstract concepts might be grounded through perceptual and motor simulations of their situations of use.

A second, more deeply explored view sees abstract concepts as grounded metaphorically in concrete ones (Lakoff and Johnson 1980). There's a long and rich tradition of linguistic research on metaphorical language—it's well documented that we frequently and conventionally speak about abstract concepts in concrete terms—in terms of physical objects (***doling out*** *justice*; *a* ***morsel*** *of truth*) or physical space (***high*** *value*). The embodied metaphor hypothesis proposes that we similarly conceive of abstract concepts in terms of concrete concepts (Gibbs 2006). That is, perhaps accessing the meanings of *justice* or *truth* involves brain systems dedicated to representing physical objects, and perhaps understanding the meaning of *value* involves brain systems dedicated to perceiving changes in vertical position.

Psychologists of language have been pursuing this hypothesis since the 1990s, and neuroscientists began joining the action about ten years later, along with social psychologists. To summarize these distinct literatures briefly, at present we know certain things. First, processing metaphorical language primes the actual source domain (for instance, reading about someone *blowing his stack* primes heat; Gibbs et al. 1997) and the reverse is also true (performing a grasping action speeds processing of *grasp an idea*; Wilson and Gibbs 2007). Second, brain systems specialized for perception and action sometimes do and sometimes do not become measurably active when people passively read metaphorical sentences like *He kicked the habit* (compare Aziz-Zadeh and Damasio 2008 with Tettamanti et al. 2005). One of the factors at play appears to be familiarity—less familiar metaphorical expressions induce greater activation in perception and action systems than more familiar expressions do (Desai et al. 2012). And finally, even in the absence of metaphorical language, there appear to be deep-seated connections that show up in real-world judgements and decision making. For instance, just as people who are kind and generous are often described metaphorically as *warm*, holding a warm coffee cup leads people to judge others as more kind and generous (Williams and Bargh 2008). We metaphorically describe morality as cleanliness and people not only show an increased preference for cleansing products after describing an immoral act they've performed, but also act as though the moral transgression has been washed away after physically cleaning their hands (Zhong and Liljenquist 2006). In sum, if not incontrovertible, there are several strands of research that converge on the conclusion that there are conceptual connections across domains that are connected through linguistic metaphor.

Clearly, there's much more to know about meaning and metaphorical and abstract language. And there are hints that a full account of how it works might involve simulation. Some might take this as a case of taking one of the most glaring weaknesses of a simulation-based account and turning it into a strength.

4 Current and future directions

Many questions remain about how simulation relates to meaning, and in this section I highlight several areas of continuing research.

4.1 Functions of simulation

While there is substantial evidence of simulation during language use, less is known about exactly what function it serves, if any. What role, mechanistically, does simulation serve in meaning-making?

A variety of possibilities have been entertained in the literature. Perhaps simulation serves as a representational format for semantics—that is, perhaps Mentalese or the Language of Thought is in fact articulated in terms of perceptual and motor simulations. Another possible use for simulation is to generate inferences. For instance, connecting the dots between *A colony of nudists moved in next door to Tristan* and the subsequent sentence *He decided to build a wall* involves a logical connecting step. Understanding the causal connection between the two propositions might involve simulating the visual details. Another possibility is that people might simulate to prepare to act. Upon hearing the word *scalpel*, a comprehender's belief about what's intended by that utterance might modulate the simulation they perform. If they believe they're being asked to hand someone the object, then they might prepare to do so by simulating manual interactions with it. But if they're merely being asked to point to the image of the object in a book, they might simulate that pointing action instead.

It's also possible that, while pervasive, simulation plays no functional role in meaning, as argued, for instance, by Pavan and Baggio (2013). On this account (Mahon and Caramazza 2008), simulation effects are epiphenomenal: meaning is constructed using abstract symbols that have no perceptual or motor substance, but which are connected to motor and perceptual systems such that activation spreads to those systems during language use. So what appear to be simulation effects are merely the consequence of non-functional downstream activation that results from and doesn't affect or play a role in meaning-making.

This deflationary account provides a useful counterweight to the enthusiasm generated by a stream of apparently confirmatory results indicating that systems for perception, motor control, and affect are also engaged during language use. It highlights the sorts of inference that can be reasonably drawn from studies using different methodologies. Facilitatory priming and brain imaging results cannot adjudicate between a causal and epiphenomenal role for simulation in meaning. But other methods discussed above, especially neuropsychological dissociation, magnetically induced impairments, and behavioural adaptation, do have the potential to address the functional role of simulation.

At present, however, those types of study are only starting to be applied systematically to different language functions to ask precisely what simulation does for meaning and under what circumstances. One tantalizing question remains—if simulation isn't necessary or sufficient for meaning, then what other processes are involved? How do people make meaning without simulation, and how does simulation coordinate with other meaning-related processes?

4.2 Simulation and gesture

If simulation involves generating an internal facsimile of actual perception and action, then how does it relate to the actual perception and action that people engage during communication through gesture? A number of papers have recently proposed that the production and perception of paralinguistic gesture might relate to simulation in a variety of ways. For one, gesture of certain types might be an external manifestation of simulation, enacted either when simulation surpasses some threshold for actual action (Hostetter and Alibali 2008) or for targeted communicative purposes—to affect the comprehension, including the simulated content, of the understander (Marghetis and Bergen 2014). On either of these accounts, gesture might be a tool that can be used to indirectly measure the presence or nature of simulation (Parrill et al. 2013). On the language perceiver's side, perceived gestures might indeed affect ongoing simulation, but the reverse is also possible—ongoing or prior simulation might affect how particular gestures are interpreted. In short, there may be ample room for productive work studying interactions between simulation and gesture.

4.3 Simulation and individual differences

One key component of the simulation account is that the capacity to simulate specific perceptual, motor, or affective experiences in response to language is the product of the individual's experiences. It follows that different experiential histories will produce different capacities and tendencies to simulate. This variation could in principle arise within a culture as the result of idiosyncratic or systematic intra-cultural variation or interculturally as a function of those precise dimensions along which culturally contingent experiences differ.

A second way that people might differ with respect to simulation involves individual differences, driven not by culturally contingent experience but differential cognitive styles and cognitive capacities. People with greater capacity for verbal working memory or a preference for verbal-auditory processing might simulate differently from people whose preferences and capacities lie more in the visual domain (Bergen 2012).

Promising early findings suggest that both experience and cognitive style predict differences in how language users simulate. For instance, cultural practice (for example, how many hands to use when transferring small objects to social superiors) produces motor simulation differences in Americans versus Koreans (Dennison and Bergen 2010). Experience playing hockey affects the use of the motor system while people are listening to descriptions of hockey actions (Beilock et al. 2008). And visualizers show greater visual simulation effects during language processing than verbalizers (Dils and Boroditsky 2010).

4.4 Limits of simulation

It's uncontroversial that there is a lot about meaning that simulation cannot account for. Metalinguistic intuitions—for instance, the feeling that the word *one* means precisely and not approximately one—are not easily dealt with through simulation alone. Simulation is clearly not sufficient for meaning, any more than visual perception is sufficient for what we know about objects. Work thus far has focused on expanding what we know about when simulation happens and what it does, while relatively less scrutiny has attended to what has to surround it—what other mechanisms must be in place alongside and integrated with simulation to account for all that humans do with respect to meaning. These appear to be productive directions in which the field is headed.

Further reading

Barsalou, L. W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660. This seminal article defines simulation and its potential role in language and other cognitive functions.

Bergen, B. 2012. *Louder than words: the new science of how the mind makes meaning*. New York: Basic Books. This book summarizes work on simulation and language, with sections on grammar, timing, abstract concepts, and individual and cross-linguistic differences.

Pecher, D. and Zwaan, R. A. (eds) 2005. *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge: Cambridge University Press. This edited volume presents a collection of chapters from leaders in the field on core theoretical topics in embodied cognition and language.

Pulvermueller, F. 2002. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press. This book proposes a unified theory of embodied grounding of language, based on well-understood principles of neurocomputation.

References

Aziz-Zadeh, L. and Damasio, A. 2008. Embodied semantics for actions: findings from functional brain imaging. *Journal of Physiology (Paris)* 102: 35–39.

Barsalou, L. W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660.

Barsalou, L. W., W. Kyle Simmons, A. K. Barbey and C. D. Wilson 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences* 72: 84–91.

Barsalou, L.W. and K. Wiemer-Hastings 2005. Situating abstract concepts. In D. Pecher and R. Zwaan (eds), *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thought*. New York: Cambridge University Press, 129–163.

Beilock, S. L., I. M. Lyons, A. Mattarella-Micke, H. C. Nusbaum and S. L. Small 2008. Sports experience changes the neural processing of action language. In *Proceedings of the National Academy of Sciences USA*, 105: 13269–13273.

Bergen, B. K. 2012. *Louder than Words: The New Science of how the Mind Makes Meaning*. New York: Basic Books.

Bergen, B. 2007. Experimental methods for simulation semantics. *Methods in Cognitive Linguistics* 18: 277.

Bergen, B. and N. Chang 2005. Embodied construction grammar in simulation-based language understanding. In Jan-Ola Östman and Miriam Fried (eds), *Construction Grammars: Cognitive Grounding and Theoretical Extensions*. Amsterdam: John Benjamins, 147–190.

Bergen, B. and N. Chang 2013. Embodied construction grammar. In Thomas Hoffmann and Graeme Trousdale (eds), *Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, 168–190.

Bergen, B. K., S. Lindsay, T. Matlock and S. Narayanan 2007. Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science* 315: 733–764.

Bergen, B. and K. Wheeler 2010. Grammatical aspect and mental simulation. *Brain and Language* 112: 150–158.

Chatterjee, A. 2010. Disembodying cognition. *Language and Cognition* 2: 79–116.

Croft, W. and D. A. Cruse 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Dennison, Heeyeon and Benjamin Bergen 2010. Language-driven motor simulation is sensitive to social context. *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, 901–906.

Desai, R. H., J. R. Binder, L. L. Conant, Q. R. Mano and M. S. Seidenberg 2012. The neural career of sensorimotor metaphors. *Journal of Cognitive Neuroscience* 239: 2376–2386.

Dils, A. T. and L. Boroditsky 2010. A visual motion aftereffect from understanding motion language. In *Proceedings of the National Academy of Sciences USA*, 107: 16396–16400.

Dove, G. 2009. Beyond perceptual symbols: a call for representational pluralism. *Cognition* 1103: 412–31.

Feldman, J. and S. Narayanan 2004. Embodied meaning in a neural theory of language. *Brain and Language* 892: 385–392.

Fodor, J. 1975. *The Language of Thought*. New York: Crowell.

Fodor, J. A. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.

Gage, Nicole and Gregory Hickok 2005. Multiregional cell assemblies, temporal binding and the representation of conceptual knowledge in cortex: a modern theory by a classical neurologist, Carl Wernicke. *Cortex* 41: 823–832.

Gibbs, R. W. 2006. Metaphor interpretation as embodied simulation. *Mind and Language* 213: 434–458.

Gibbs, R. W., J. M. Bogdanovich, J. R. Sykes and D. J. Barr 1997. Metaphor in idiom comprehension. *Journal of Memory and Language* 37: 141–154.

Glenberg, A. M. and M. P. Kaschak 2002. Grounding language in action. *Psychonomic Bulletin and Review* 93: 558–565.

Glenberg, A. M. and D. A. Robertson 2000. Symbol grounding and meaning: a comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language* 43: 379–401.

Glenberg, A. M., Sato, M. and L. Cattaneo 2008. Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology* 18: R290–R291.

Harnad, S. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42: 335–346.

Hostetter, A. B. and M. W. Alibali 2008. Visible embodiment: gestures as simulated action. *Psychonomic Bulletin and Review* 153: 495–514.

Johansson, R., J. Holsanova and K. Holmqvist 2006. Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science* 306: 1053–1079.

Kaschak, M. P. and A. M. Glenberg 2000. Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language* 43: 508–529.

Kaschak, M. P., C. J. Madden, D. J. Therriault, R. H. Yaxley, M. Aveyard, A. A. Blanchard and R. A. Zwaan 2005. Perception of motion affects language processing. *Cognition* 94: B79–B89.

Lakoff, G. and M. Johnson 1980. *Metaphors We Live By*. Chicago, IL: University of Chicago Press.

Landauer, T. K., P. W. Foltz and D. Laham 1998. An introduction to latent semantic analysis. *Discourse Processes* 25: 259–284.

Lund, K. and C. Burgess 1996. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments and Computers* 28: 203–208.

Madden, C. J. and R. A. Zwaan 2003. How does verb aspect constrain event representations? *Memory and Cognition* 31: 663–672.

Mahon, B. Z. and A. Caramazza 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology (Paris)* 1021: 59–70.

Marghetis, T. and B. Bergen 2014. Embodied meaning, inside and out: the coupling of gesture and mental simulation. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva H. Ladewig, David McNeill and Sedinha Tessendorf (eds.), *Body-Language-Communication*. New York: Mouton de Gruyter.

Narayanan, S. 1997. *KARMA: Knowledge-based active representations for metaphor and aspect*. Unpublished doctoral dissertation. University of California, Berkeley.

Parrill, F., B. Bergen and P. Lichtenstein 2013. Grammatical aspect, gesture and conceptualization: using co-speech gesture to reveal event representations. *Cognitive Linguistics* 241: 135–158.

Pavan, A. and G. Baggio 2013. Linguistic representations of motion do not depend on the visual motion system. *Psychological Science* 242: 181–188.

Pecher, D. and R. A. Zwaan 2005. *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking*. Cambridge: Cambridge University Press.

Pulvermüller, F. 2003. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press.

Richardson, D. C., M. J. Spivey, L. W. Barsalou and K. McRae 2003. Spatial representations activated during real-time comprehension of verbs. *Cognitive Science* 27: 767–780.

Rommers, J., A. S. Meyer, F. Huettig and J. Rommers 2013. Object shape and orientation do not routinely influence performance during language processing. *Psychological Science* 24: 2218–2225.

Sato, M. 2010. Message in the "body": effects of simulation in sentence production (Doctoral dissertation, University of Hawaii).

Saygin, A. P., S. McCullough, M. Alac and K. Emmorey 2010. Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *Journal of Cognitive Neuroscience* 2211: 2480–2490.

Shapiro, K., F. M. Mottaghy, N. O. Schiller, T. D. Poeppel, M. O. Fluss, H. W. Muller, A. Caramazza and B. J. Krause 2005. Dissociating neural correlate for verbs and nouns. *Neuroimage* 24: 1058–1067.

Shapiro, K. A., A. Pascual-Leone, F. M. Mottaghy, M. Gangitano and A. Caramazza 2001. Grammatical distinctions in the left frontal cortex. *Journal of Cognitive Neuroscience* 136: 713–720.

Spivey, M. and J. Geng 2001. Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research* 65: 235–241.

Stanfield, R. A. and R. A. Zwaan 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12: 153–156.

Tettamanti, M., G. Buccino, M. C. Saccuman, V. Gallese, M. Danna, P. Scifo, F. Fazio, G. Rizzolatti, S. F. Cappa and D. Perani 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 17: 273–281.

Wassenburg, S. I. and R. A. Zwaan 2010. Readers routinely represent implied object rotation: the role of visual experience. *The Quarterly Journal of Experimental Psychology* 639: 1665–1670.

Wernicke, C. 1977 [1874]. Der aphasische Symptomenkomplex: eine psychologische Studie auf anatomischer Basis. In G. H. Eggert (ed.), *Wernicke's Works on Aphasia: A Sourcebook and Review*. The Hague: Mouton, 91–145.

Wernicke, C. 1977 [1885–1886]. Einige neuere Arbeiten ueber Aphasie. In G. H. Eggert (ed.), *Wernicke's Works on Aphasia: A Sourcebook and Review*. The Hague: Mouton, 31–40.

Willems, R. M., I. Toni, P. Hagoort and D. Casasanto, 2010. Neural dissociations between action verb understanding and motor imagery. *Journal of Cognitive Neuroscience* 22: 2387–2400.

Williams, L. E. and J. A. Bargh 2008. Experiencing physical warmth influences interpersonal warmth. *Science* 322: 606–607.

Wilson, N. L. and R. W. Gibbs, Jr. 2007. Real and imagined body movement primes metaphor comprehension. *Cognitive Science* 31: 721–731.

Wittgenstein, L. 1953. *Philosophical Investigations*. G. E. M. Anscombe and R. Rhees (eds.), G. E. M. Anscombe (trans.). Oxford: Blackwell.

Yee, E., E. Chrysikou, E. Hoffman and S. L. Thompson-Schill 2013. Manual experience shapes object representation. *Psychological Science* 246: 909–919.

Zhong, C. B. and K. Liljenquist 2006. Washing away your sins: threatened morality and physical cleansing. *Science* 313: 1451–1452.

Related topics

Chapter 2, Internalist semantics; Chapter 5, Cognitive semantics; Chapter 7, Categories, prototypes and exemplars; Chapter 26, Acquisition of meaning; Chapter 29, Semantic processing.