Casasanto, D. (2016). Linguistic relativity. In N. Riemer (Ed.), Routledge handbook of semantics (pp. 158-174). New York: Routledge.

9

Linguistic relativity

Daniel Casasanto

1 Introduction

According to the theory of linguistic relativity, language shapes the way people think; as a result, speakers of different languages may think differently, in predictable ways. This proposal, often associated with the writings of Benjamin Whorf (1956; see also von Humboldt (1988); Sapir (1929)), has generated decades of controversy among linguists, psychologists, philosophers, and anthropologists. Many scholars believe the theory of linguistic relativity to be "wrong, all wrong" (Pinker (1994: 57); see also Bloom and Keil 2001), and some have sought to convince readers that research on the "Whorfian" question is pseudoscience (e.g. Pullum 1991). But, for a moment, imagine you are unaware of the history of this embattled idea. Imagine you are coming to the question, "Does language influence how we think?" for the first time, and reasoning from first principles. You might posit that: (1) Language is a nearly omnipresent part of the context in which we use our minds. (2) Thinking depends on context. (3) Therefore, language could have pervasive influences on how and what we think.

Does it? If so, *how* does language shape thought? Which parts of language influence which aspects of cognition or perception, and by what mechanisms? For much of the twentieth century, scholars despaired of answering these questions, pointing to two stumbling blocks for linguistic relativity. Some suggested that the "Whorfian" question was inherently circular, because tests of the influence of language on thought were contaminated by participants' use of language during the tests (Gleitman and Papafragou 2013; Pullum 1991). Others argued that, after decades of trying, proponents of linguistic relativity had only managed to produce evidence of "weak" and "banal" effects of language on thought (Pinker 1994: 65; see also McWhorter 2014). Yet, in the early twenty-first century, theoretical and experimental advances have reinvigorated efforts to understand how language influences our cognition and perception, and how people who speak different languages think differently as a consequence.

This chapter will focus on studies that have begun to address the two perennial concerns about research on linguistic relativity mentioned above. One set of studies overcomes the problem of circularity by showing patterns of behavior that differ as a function of linguistic experience, but which cannot be explained by the use of language during the test. The second set of studies overcomes concerns about the magnitude of the impact language can have on cognition, suggesting that words found in some languages but not others can radically transform people's minds, and have reshaped the world we live in.

To focus on studies that address these persistent causes of skepticism, I will omit discussion of a great deal of influential research on linguistic relativity that preceded them, including the many studies on color categories by Paul Kay and colleagues (e.g. Kay and Regier 2006), studies on ontological categories by John Lucy and colleagues (e.g. Lucy 1996), and studies on spatial frames of reference by Stephen Levinson and colleagues (e.g. Levinson 2003). Also, space precludes discussion of all of the variants of linguistic relativity that have been articulated; Whorf, himself, never stated any unitary "Whorfian hypothesis" (for discussion see Kay and Kempton (1984); Lucy (1992)).

2 Are Whorfian effects limited to effects of "thinking for speaking"?

Dan Slobin (1996: 75–76) proposed an influential version of linguistic relativity, arguing that language should affect cognition during the process of encoding our thoughts into words, or "thinking for speaking":

> There is a special kind of thinking that is intimately tied to language – namely, the thinking that is carried out, on-line, in the process of speaking. [. . .] "Thinking for speaking" involves picking those features of objects or events that . . . are readily encodable in the [speaker's] language.

Different languages may guide speakers to specify different kinds of information in their speech. Therefore, different languages may cause their speakers to activate different information in memory when recounting the same episode, or to highlight different information about the perceptible world when inspecting or describing the same scene.

Slobin (1996) gives the example that English and Spanish bias their speakers to specify different kinds of information in the verb when describing motion events. English motion verbs tend to specify a *manner* of motion (e.g. *running*, *flying*, *rolling*), whereas Spanish verbs more often specify a *path* of motion (e.g. *entering*, *exiting*). Although English typically includes path information in other parts of speech (e.g. in prepositions, as in "the bird flew *down* from the nest"), Spanish and other "path languages" like Korean often omit manner information altogether.

Slobin's PhD student, Kyung-ju Oh, tested whether differences in the linguistic encoding of path and manner can influence Koreans' and US English speakers' memories for motion events (Oh 2003). Monolingual participants, tested in their home countries, watched videos of people performing activities like strolling out of a building or trudging along a path, and described what they saw. Later, they received a surprise memory test probing details of these events. English and Korean speakers did not differ in their memory for path-relevant details, consistent with the salient encoding of path information across both languages, nor did they differ in their memory for control aspects of the events (e.g. the color of an actor's shirt). Yet the English speakers remembered more manner-relevant details than the Korean speakers, consistent with manner being more frequently encoded in the English speakers' event descriptions than in the Koreans'.

Oh's experiment showed a correlation between the way people talked about events and the way they remembered them later, suggesting that the way people code their experiences into language can influence performance on subsequent nonlinguistic tasks. Yet, on a skeptical interpretation of these results, it is possible that speakers used language to reconstruct the events from memory when asked to recall them, reactivating verbal descriptions consciously or unconsciously. If so, although Oh's memory task was "nonlinguistic," the results could be considered to be a type of thinking-for-speaking effect, provided that "speaking" includes covertly encoding experiences into words (i.e. speaking to one's self). Consistent with this possibility, subsequent studies suggest that language-specific differences in event representation disappear when participants perform a concurrent verbal task that interferes with using language online to encode the events (but not when they perform a concurrent non-verbal task; Trueswell and Papafragou 2010).

This pattern of results is not limited to tests of how language influences event representation. Across a variety of experimental paradigms, effects of language on cognition or perception are readily explained as effects of overt or covert thinking for speaking. Some of the most thoroughly studied effects of language on thought are extinguished under verbal interference conditions (e.g. effects of color vocabulary on color judgements; Winawer et al. (2007)). Such results have led some researchers to conclude that linguistic relativity effects are limited to thinking-for-speaking effects (cf. Slobin 2003), and even to argue that effects of language on thought may be strictly limited to circumstances that encourage participants to use language strategically to perform a task (Papafragou et al. 2008).

Some scholars have speculated that when people are *not* thinking for speaking, linguistic relativity effects should disappear (Clark 2003; Papafragou et al. 2007; Landau et al. 2010). For example, Eve Clark predicted that if truly nonlinguistic tests of linguistic relativity could be devised, their results should differ dramatically from the results of thinking-for-speaking-driven experiments:

> [W]e should find that in tasks that require reference to representations in memory that don't make use of any linguistic expression, people who speak different languages will respond in similar, or even identical, ways. That is, representations for nonlinguistic purposes may differ very little across cultures or languages.
>
> (2003: 22)

Clark added:

> Of course, finding the appropriate tasks to check on this without any appeal to language may prove difficult.
>
> (2003: 22)

3 Beyond thinking for speaking: Whorfian psychophysics

Casasanto and colleagues developed a strategy for meeting Clark's challenge: the "Whorfian psychophysics" paradigm. Psychophysics, one of the oldest branches of experimental psychology, is the study of how precisely organisms can encode, discriminate, or reproduce simple physical stimuli (e.g. the length of a line or the brightness of a flash of light). In our studies, differences between languages predicted differences in their speakers' mental representations, but these predictions were tested using psychophysical stimulus-reproduction tasks with nonlinguistic stimuli and responses.

The first set of experiments built upon a task that Casasanto and Boroditsky (2008) developed to investigate relationships between mental representations of time and space. In the original experiments, English speakers saw objects that varied in their spatial or temporal extents. In the canonical version of the task, lines of different spatial lengths and durations "grew" gradually across a computer screen, and disappeared when they had reached their maximum extent in both space and time. Participants were then asked to reproduce either the spatial or temporal extent of the stimulus by clicking the mouse to indicate the beginning and ending points of either the spatial or temporal interval. Participants were unable to ignore the spatial dimension of the stimuli when reproducing the temporal dimension: for lines of the same average duration, those that traveled a longer distance were judged to take a longer time, and those that traveled a shorter distance to take a shorter time. Numerous versions of this task showed the effect to persist over variations in perceptual, attentional, and mnemonic factors, in children and adults (see Bottini and Casasanto (2013) for a review). It appears that the tendency to conceptualize duration in terms of spatial distance is a robust aspect of temporal cognition – at least for English speakers.

3.1 Time in one or three dimensions

Like many other languages, English tends to describe duration in terms of one-dimensional spatial length (e.g. a *long* time, like a *long* rope; Alverson (1994); Evans (2004)). This uni-dimensional mapping has been assumed to be universal: a consequence of the unidirectional flight of time's arrow, and of universal aspects of our bodily interactions with the environment (Clark 1973). It is hard to avoid using uni-dimensional spatial metaphors when talking about the durations of events in English. Try replacing the word "long" in the phrase "a *long* meeting" with a synonym. Words like *lengthy*, *extended*, *protracted*, or *drawn out* would suffice – all of which express time in terms of linear extent.

In contrast with English speakers, however, Greek speakers tend to express duration in terms of volume or amount (e.g. a *lot of* time (*tr. poli* ora), like a *lot of* water (*tr. poli* nero)). Rather than "a *long* night," Greek speakers would say "a *big* night" (*tr. megali* nychta) to indicate that the night seemed to last a long time. Greek speakers *can* express duration in terms of linear extent, just as English speakers can make use of volume or amount expressions, but volume metaphors are more frequent and productive in Greek, whereas linear extent metaphors are more frequent and productive in English (Casasanto et al. 2004; Casasanto 2008; 2010).

Does the tendency to talk about duration in terms of one-dimensional or three-dimensional space influence the way people tend to think about it? To find out, in one set of experiments we gave English and Greek speakers a pair of nonlinguistic psychophysical tests of their ability to estimate duration in the presence of irrelevant length or volume information (Casasanto et al. 2004; Casasanto 2008; 2010). In the length interference task, participants were asked to reproduce the durations of lines that gradually extended across the screen while trying to ignore the lines' spatial length, as described above. In the volume interference task, participants reproduced the durations for which they saw a container gradually filling up, while trying to ignore the container's fullness.

As before, English speakers had difficulty screening out interference from spatial distance when estimating duration: lines that traveled a longer distance were mistakenly judged to take a longer time than lines of the same duration that traveled a shorter distance. But their time estimates were relatively unaffected by irrelevant volume information. Greek speakers showed the opposite pattern. They had more difficulty screening out interference from volume, so fuller containers were judged to remain on the screen for more time than emptier containers, but their judgements were relatively unaffected by the spatial extent of lines. The pattern of distance and volume interference in these nonlinguistic psychophysical tasks reflected the relative prevalence of distance and volume metaphors for duration in English and Greek. Similar patterns were found in speakers of Indonesian (a "distance language") vs. speakers of Spanish (a "volume language"; Casasanto et al. 2004).

In these experiments, participants were informed before each trial whether they would need to reproduce the spatial or temporal dimension of the stimulus. Is it possible that these results were due to participants covertly labeling the relevant dimension of the stimuli as they perceived or reproduced it? If so, these experiments would be subject to the same inferential limitations of previous studies whose results can be explained by the online use of language: an effect of thinking for covert speaking. Fortunately, this skeptical possibility is ruled out by the design of the experiments. We cannot know whether participants *tried* to label the stimuli (e.g. using words like "long" and "short" covertly), but we can definitively rule out the possibility that the observed effects were due to such verbal labeling. In each experiment there were nine different levels of duration (the durations ranged from 1 to 5 seconds, increasing in 500-millisecond increments), which were fully crossed with nine different levels of length or volume (e.g. the lengths ranged from 100 to 500 pixels, increasing in 50-pixel increments). Due to the crossing of these levels (i.e., pairing each level of space with each level of time), space and time were orthogonal: there was no correlation between the spatial and temporal magnitudes of the stimuli.

Because of this feature of the experimental design, it is impossible that labeling the relevant dimension could produce the predicted effect of interference from the irrelevant dimension – which varied orthogonally. Consider, for example, a participant who labeled all of the long-duration lines "long" and the short duration lines "short" during the time estimation trials. This labeling strategy, if it affected time estimates at all, would only work *against* the effect of spatial length on time estimation, given that spatial length was orthogonal to time. Even if participants attempted to label durations of the stimuli as "long" or "short" (etc.), the experimental design ensured that the predicted effects of space on time estimation occurred in spite of this labeling strategy, not because of it.

3.2 Beyond a language-thought correlation

The cross-linguistic comparison between Greek and English speakers shows a correlation between temporal language and temporal thinking. Can language play a causal role in shaping nonlinguistic time representations? To test whether using volume metaphors in language can change the way people think about duration, the experimenters trained English speakers to use Greek-like metaphors for time (Casasanto 2008; 2010). After about 20 minutes of exposure to these new metaphors, the effect of irrelevant volume information on English speakers' nonlinguistic duration estimates was statistically indistinguishable from the effect found in native Greek speakers. Together, these data show that people who use different temporal metaphors in their native languages conceptualize time the way they talk about it, even when they are not using language. Furthermore, linguistic experience can play a causal role in shaping mental representations of time. Producing or understanding spatio-temporal language like a Greek speaker, even for a few minutes, can cause English speakers to think about time differently, using a different kind of spatial scaffolding.

3.3 Alternative spatial metaphors for musical pitch

The Whorfian psychophysics paradigm used to establish cross-linguistic differences in temporal thinking has been extended to probe language-based differences in people's mental representations of musical pitch. Like English, Dutch describes pitches as "high" (*hoog*) or "low" (*laag*), but this is not the only possible spatial metaphor for pitch. In Farsi, high pitches are "thin" (*nāzok*) and low pitches are "thick" (*koloft*). Dutch and Farsi speakers' performance on nonlinguistic pitch reproduction tasks reflects these linguistic differences (Dolscheid et al. 2013). Participants were asked to reproduce the pitch of tones that they heard in the presence of irrelevant spatial information: lines that varied in their height (height interference task) or their thickness (thickness interference task). Dutch speakers' pitch estimates showed stronger cross-dimensional interference from spatial height, and Farsi speakers' from the thickness of visually presented stimuli. This effect was not explained by differences in accuracy or in musical training between groups. When Dutch speakers were trained to talk about pitches using Farsi-like metaphors (e.g. a tuba sounds *thicker* than a flute) for 20–30 minutes, their performance on the nonlinguistic thickness interference task became indistinguishable from native Farsi speakers'. Experience using one kind of spatial metaphor or another in language can have a causal influence on nonlinguistic pitch representations.

These space-pitch interference studies used a similar design to the space-time interference studies described above: nine levels of space were crossed with nine levels of pitch. Therefore, the experimental design rules out the possibility that labeling pitches with spatial words (e.g. "high" or "low") could produce the observed effects; on the contrary, using such verbal labels for pitch during the task could only work against the predicted effects of height or thickness on pitch reproduction. To underscore the point that differences between Dutch and Farsi speakers' pitch reproduction were not caused by using language online (i.e., by thinking for speaking to one's self), Dolscheid et al. (2013) asked Dutch speakers to do the height interference task while performing a secondary verbal suppression task that prevented them from encoding the stimuli verbally. If the effect of height on pitch were driven by covertly labeling the stimuli using spatial words, then it should disappear under verbal interference. However, we hypothesized that this effect was not due to online use of spatial metaphors for pitch in language, but rather to the activation of an implicit association between nonlinguistic representations of space and pitch in memory: a *mental metaphor* (Casasanto 2010; Lakoff and Johnson 1999). If so, the effect of height on pitch should persist under verbal interference. Consistent with this prediction, the effect of height on pitch reproduction in Dutch speakers was equally strong with and without concurrent verbal suppression.

3.4 The role of language in shaping mental metaphors

What role does spatial language play in shaping nonlinguistic representations of time and pitch? Is language creating cross-domain associations, or is linguistic experience modifying pre-linguistic mental metaphors? Pre-linguistic infants appear to intuit a link between more duration and more spatial extent (Srinivasan and Carey 2010), and also between more duration and more size (Lourenco and Longo 2011). Thus, both the distance-duration mapping that is most prevalent in English and the volume-duration mapping that is most prevalent in Greek may be present pre-linguistically. Likewise, infants as young as four months old are sensitive to the height-pitch mapping found in Dutch-speaking adults (but not in Farsi-speaking adults), and also to the thickness-pitch mapping found in Farsi-speaking adults (but not in Dutch-speaking adults; Dolscheid et al. (2014)). There is no need, therefore, to posit that using linguistic metaphors causes people to construct these mappings *de novo*.

Together, these infant and adult data support a developmental story with two chapters. First, children represent duration via a family of spatial mappings, which includes mappings from both spatial length and volume. Likewise, they represent pitch via mappings from both height and thickness. These initial mappings may be universal, based either on innate cross-domain correspondences (Walker et al. 2010) or on early-learned correlations between source and target domains in children's experience with the physical world (Lakoff and Johnson 1999). The distance-duration and volume-duration mappings could be learned by observing that more time passes as objects travel farther distances and as quantities accumulate in three-dimensional space. Height-pitch mappings could be learned from seeing (or feeling) the larynx rise and fall as people produce higher and lower pitches with their voices. Thickness-pitch mappings could be learned from observing the natural correlation between the size of an object or animal and the sound that it makes (imagine the sound made by banging on a soda can vs. an oil drum).

Later, linguistic experience modifies these pre-linguistic source-target mappings. Suppose each time speakers use a linguistic metaphor like "a *long* meeting" or "a *high* soprano" they activate the corresponding mental metaphor. Repeatedly activating one source-target mapping instead of another (e.g. height-pitch instead of thickness-pitch) should strengthen the activated mapping and, as a consequence, weaken the competing mapping via competitive learning (Casasanto 2008; Dolscheid et al. 2013). This process of strengthening one spatial mapping during language use, at the expense of the alternative spatial mapping, may explain how universal space-time and space-pitch mappings in infants become language-specific mappings in adults.

On this account, our mental metaphors are structured hierarchically (Casasanto and Bottini 2014). Specific mappings, conditioned by linguistic experience, are selected from families of mappings conditioned by relationships between source and target domains in the natural world. This hierarchical structure may help to explain how source-target mappings can be important for our representations of target domains but also surprisingly flexible. For example, perhaps Dutch speakers could be trained to think like Farsi speakers so quickly because they did not have to *learn* the thickness-pitch mapping during their 20–30 minutes of using Farsi-like linguistic metaphors. Rather, this linguistic training strengthened the association between thickness and pitch that was present in participants' minds from infancy (as indicated by data from Dutch four-month-olds), but which had been weakened as a consequence of their frequent use of height-pitch metaphors in language.

One prediction of this account (called Hierarchical Mental Metaphors Theory; Casasanto and Bottini (2014)) is that specific source-target mappings should be easy to activate via linguistic training so long as they are members of a family of nonlinguistic source-target mappings encoded in our minds (over either phylogenetic or ontogenetic time) on the basis of observable source-target correspondences in the world. Mappings that are not members of a pre-linguistically established family – and that do not reflect correlations between source and target domains in the natural world – should be relatively hard to activate via training, because these mappings would need to be created, not just strengthened.

In a test of this prediction, Dutch speakers were trained to use a thickness-pitch mapping that is the reverse of the mapping found in Farsi, and in the natural world: thin=low and thick=high. These "reverse-Farsi"-trained participants received the same amount of training as the participants trained to use the Farsi-like mapping. Whereas Farsi-like training had a significant effect on participants' nonlinguistic pitch representations, reverse-Farsi training had no effect (Dolscheid et al. 2013). Thus, brief linguistic experience caused Dutch participants to use the thickness-pitch mapping that reflects correlations between thickness and pitch in the world (and is evident in pre-linguistic infants). Yet the same amount of linguistic experience was not effective at instilling the opposite thickness-pitch mapping, which has no obvious experiential correlates, and is therefore not predicted to be among the pre-linguistically established space-pitch mappings.

3.5 Summary of Whorfian psychophysics findings

It is possible to test Whorfian questions without using words, thereby escaping the circularity that can result from using language as a source of hypotheses about the mind and also as a means to test these hypotheses. Furthermore, it is possible to construct experiments that show influences of language on nonlinguistic mental representations: representations that are "nonlinguistic" insomuch as (a) they are found, in some form, in pre-linguistic infants, (b) they are continuous and difficult to describe adequately using the lexical categories available in ordinary non-technical speech, and (c) they can be activated without using language overtly or covertly, and persist in the presence of a verbal suppression task.

In the cases of space-duration and space-pitch mappings, it appears that the role of linguistic experience is to modify the strength of pre-linguistically available cross-domain associations. As a result of using one kind of verbal metaphor or another repeatedly, speakers who rely on different metaphors in language subsequently activate different mental representations of time and pitch, scaffolded by different kinds of spatial representations. These results provide the first evidence that using language can have *offline effects* on speakers' mental representations – that linguistic relativity effects are not limited to thinking-for-speaking effects.

4 When are the effects of language on thought "important"?

The psychophysical studies reviewed above suggest that experience using language can cause people to form systematically different mental representations, even if they are not using language online, at the moment they form them. These studies address one longstanding concern about linguistic relativity effects: circularity. Yet they do not fully address a second longstanding concern: are these language-induced cognitive differences *important*? Some researchers have concluded that they are not: either because the effects are context-dependent (Gleitman and Papafragou 2013), or because cross-linguistic differences do not appear to radically change the way speakers of different languages perceive or understand their world (McWhorter 2014). This conclusion appears to emerge from three common beliefs about linguistic relativity research or about how our minds work, more broadly – beliefs that bear reexamination.

4.1 Belief no. 1: if mental representations are flexible, they must not be very important

Some researchers suggest that because effects of language on thought "are malleable and flexible," they "do not appear to shape core biases in . . . perception and memory" (Trueswell and Papafragou (2010), writing about effects of language on motion event representation). Indeed, all of the Whorfian effects mentioned so far are flexible: influences of motion descriptions on event representations and influences of color words on color judgements are modulated by verbal interference. Influences of spatial metaphors on nonlinguistic representations of duration and pitch can be rapidly changed by new patterns of experience using linguistic metaphors. However, the fact that these representations are flexible and context-dependent does not make them unimportant: if this were the case, then *all mental representations* would be "unimportant."

In the twentieth century, cognitive scientists took seriously the notions that our minds contain a mental dictionary of word meanings (e.g. Johnson-Laird 1987) and a mental encyclopedia of concepts (e.g. Pinker 1999). These technical metaphors have a misleading entailment. Entries in a dictionary or an encyclopedia are essentially unchanging: once written, the entries are simply looked up subsequently. Any aspect of a concept or word meaning that can change, then, is often deemed peripheral: not part of the true, "core" entry in our dictionary or encyclopedia. Yet the idea that thoughts are fixed entities in a mental repository, which are simply accessed when needed, is incompatible with what we know about brains. Thoughts (i.e. concepts, percepts, feelings, word meanings) are instantiated in brains, and brains are always changing; therefore, thoughts are always changing. Although space prohibits elaboration of this argument (see Casasanto and Lupyan (2015); Spivey (2007)), it should be unsurprising that the mental representations people form during linguistic relativity experiments are context-dependent; our mental representations are flexible and context-dependent – arguably, without exception (see Besner et al. (1997); James (1890); and see Chapter 2).

4.2 Belief no. 2: the goal of relativity research is to demonstrate differences between minds, therefore relativity effects are only important if they show radical differences

If the primary goal of relativity research were to show that people with different experiences think differently, then the more dramatic the between-group differences were, the more important the results would be. For many researchers, however, demonstrating cross-linguistic differences in thinking is a means to an end – not an end in itself.

The most fundamental goal of linguistic relativity research is to determine whether and how language shapes thought: that is, whether language merely reflects our conceptualizations of the world, or whether it also contributes to those conceptualizations. If the latter, then language can provide part of the answer to myriad questions about how cognition and perception develop, and how they change throughout the lifetime on various timescales. Comparing ways of thinking across language groups, to determine whether people who talk differently also think differently, is one powerful way to work toward the more basic goal of determining how language shapes thought. Tests of linguistic relativity can advance our scientific understanding of the origins and structure of our knowledge *whether or not* they show that members of different groups form radically different mental representations when presented with the same stimuli.

4.3 Belief no. 3: there's no evidence that cross-linguistic differences can produce radical differences between minds

Before the twenty-first century, there may have been no evidence for any *single* cross-linguistic difference that produces radical differences in speakers' thoughts. Yet dramatic differences in thinking could arise from the combination of many subtle differences between languages. So far, cross-linguistic differences in thinking have been reported across many fundamental domains of human experience. Language not only influences representations of *motion events* (e.g. Oh 2003), *duration* (e.g. Casasanto et al. 2004), *color* (e.g. Thierry et al. 2009; Winawer et al. 2007), and *pitch* (Dolschied et al. 2013), but also representations of *spatial relationships* (e.g. Levinson and Brown 1994; McDonough et al. 2003), *concrete objects* (e.g. Boroditsky et al. 2003; Lucy and Gaskins 2001; Srinivasan 2010), *theory of mind* (e.g. Lohmann and Tomasello 2003; Papafragou 2002, cf., Papafragou et al. 2008), *causation* (e.g. Fausey and Boroditsky 2011; Wolff et al. 2009), and *number* (e.g. Frank et al. 2008; Spelke and Tsivkin 2001).

Some aspects of language may have pervasive effects on thought. For example, applying grammatical gender to nouns can influence mental representations of their referents (e.g. Boroditsky et al. 2003). Naming an object like a table with a masculine noun (*il tavolo* in Italian) or a feminine noun (*la table* in French) can influence how people conceptualize these objects, causing speakers of different languages to endow the objects with stereotypically masculine or feminine qualities. The influence of gender on each object may be subtle, but since gender is applied to every noun (and often reiterated on verbs and adjectives), the collective effect of arbitrarily sexing every nameable object in one's lexicon could be substantial – whether or not such an effect is available to introspection. Moreover, when the pervasive effects of gender on object representations are combined with the effects of many other aspects of language on many other aspects of thought (such as those listed above), the aggregate of their individual effects could be that speakers of different languages tend to activate manifestly different conceptions of the same objects and events. It may be shortsighted, therefore, to dismiss apparently non-radical effects of language on thought as unimportant, theoretically or practically.

Furthermore, twenty-first-century research on language and thought provides evidence of at least one example of a radically mind-altering effect of language on thought. Language appears to transform our minds by playing a crucial role in creating the domain of *large exact number*, causing the thoughts entertained in some language communities to be incommensurable with the thoughts entertained in others. This cognitive transformation, in turn, has wrought immeasurable changes on the world as we know it.

4.4 Effects of language that transform mind and world

Imagine showing an adult a tray with four apples on it, taking it away, then showing them another tray with five apples on it and asking which tray held more apples: could they tell you? The answer appears to be: only if their language provides words for "four" and "five."

Children are not born with the capacity to represent "large" exact quantities, meaning quantities greater than three, nor do they develop this capacity through universal aspects of physical and social experience. They develop the capacity to mentally represent "exactly four," or "exactly seventeen," studies suggest, only if they are exposed to a list of counting numbers in their language. It is easy to take the existence of a verbal count list for granted, since they are found in the languages used by all modern, industrialized cultures. Yet counting systems like ours are recent and rare in human history, and are still unknown to people in many cultures (Dehaene 1999).

Human infants start out with two systems for representing quantity that they share with non-human animals: a system for individuating small collections of objects (up to three), and a system for representing and comparing large approximate quantities (Carey 2004; 2009; Dehaene 1999; Feigenson et al. 2004). The first *parallel individuation* system allows children to determine whether a box contains two toys or three, but does not enable them to reliably distinguish three toys from four. The second *approximate number* system allows them to distinguish larger collections of objects from one another so long as their ratio is sufficiently large. Infants under six months old can only discriminate quantities if their ratio is 1:2 (Feigenson et al. 2004). Eventually, the approximate number system becomes attuned to closer ratios, but discrimination performance remains probabilistic even in adults who rely on this system, and never attains the precision needed to reliably discriminate nine objects from ten, or to entertain an idea like "exactly 78." Separately and together, the parallel individuation and approximate number systems lack the power to represent number as we know it, or even to represent positive integers – something that seems to come naturally to Western children from their earliest years of school.

Carey (2004; 2009) posited that learning the list of counting words in their native language is what allows children to exceed the representational capacities of their primitive number systems, over a lengthy developmental process. Children first learn to say the numbers in order as a word game (typically the numbers one through ten in Western cultures), much like they learn nonsensical nursery rhymes: they can recite the numbers in order, but they do not understand them. Through having the spoken numbers matched with fingers or collections of objects, children start to understand that each number word refers to a precise quantity. First they learn that "one" refers to one object. Weeks or months later they learn that "two" refers to two objects. At this stage of being a "two-knower," children asked for "two marbles" can deliver the correct number, but when asked for "three marbles" or "four marbles" they will respond with some small collection of marbles greater than two. Eventually, after a period as "three-knowers," children induce how counting works; number becomes a productive system, where each number word in the count list refers to a unique numerosity, and successive number words refer to numerosities that differ by exactly one. (Acquiring the semantics of natural language quantifiers may contribute to the acquisition of small exact number concepts (Barner et al. 2007). This claim is distinct from the claim discussed here, that a verbal count list is essential for the acquisition of large exact number concepts.)

Some details of the process by which children learn to map number words to numerosities remain unclear, but there is now compelling evidence that language is essential for the acquisition and use of large exact number concepts. It appears that in the absence of a count list in language, people do not develop the capacity to enumerate exact quantities greater than three. Initial evidence for this radical claim came from a study by Peter Gordon (2004), who tested the numerical abilities of an Amazonian people known as the Pirahã. The Pirahã have no words for exact numbers. They quantify collections of things using the terms *hói* (about one), *hoí* (about two), and *baágiso* (many; see Frank et al. 2008). When Gordon (2004) asked members of the Pirahã tribe to perform a set of simple counting tasks their responses suggested a surprising lack of numerical competence. In one task Gordon asked each of his Pirahã participants to watch him drop up to nine nuts into a can. He then withdrew the nuts one by one and asked the participant to indicate when the can was empty. If the participants were counting nuts, this task would be easy. Yet the participants were unable to evaluate the number of nuts correctly. Some participants responded incorrectly even when there were only two or three nuts in the can. About half of the participants responded incorrectly when there were four nuts in the can, and the majority responded incorrectly when there were more than four nuts. On the basis of results like these, Gordon (2004: 498) concluded that "the Pirahã's impoverished counting system truly limits their ability to enumerate exact quantities when set sizes exceed two or three items," thus advancing two radical claims: (1) The Pirahã are unable to mentally represent exact numbers greater than three and (2) They lack this representational capacity because they lack number words. The first claim challenges the universality of humans' basic numerical competence, and the second supports a radical version of linguistic relativity, often called linguistic determinism: the limits of one's vocabulary can determine the limits of one's conceptual repertoire.

Although both of Gordon's (2004) claims have been supported by subsequent studies, initially they were strongly criticized, not only by opponents of linguistic relativity (Gelman and Gallistel 2004) but also by its proponents (Casasanto 2005; Frank et al. 2008). Limitations of the data from this pioneering study preclude drawing any conclusions about the Pirahã's numerical competence, and limitations of the experimental design preclude drawing conclusions about a causal role for language in shaping numerical abilities.

Concerning the data, the Pirahã not only failed tasks that rely on exact enumeration, like the nuts-in-a-can task, they also failed simpler tasks that can be completed without drawing on exact number representations greater than one. In one task, Gordon (2004) placed a stick on a flat surface to delineate the experimenter's side from the participant's side. He then arranged batteries in a row on his side of the partition and asked the participants to arrange batteries on their side so as to "make it the same" (p. 497). In another version of this task, the experimenter drew lines on a piece of paper, and asked participants to copy them. Crucially, these and other similar tasks do not require exact enumeration. It would be possible to create a matching array of batteries, for example, by aligning them spatially, placing batteries on the surface in one-to-one correspondence with the example array. The fact that the Pirahã did not do this suggests that they did not understand the goal of the task. It is informative that the Pirahã did not respond randomly; the number of batteries (or lines or nuts) in their responses were correlated with the number of items in the stimuli, suggesting that they were producing an *approximate match*, using their approximate number system. They appear to have been performing a different task from the one the experimenter intended.

The fact that the Pirahã failed to create matching arrays even on tasks that could be completed successfully without exact enumeration renders their failures on the other exact-number-requiring tasks completely uninterpretable. The one-to-one matching tasks served as a *manipulation check*: an experimental condition that allows the experimenter to verify that the tasks are working as planned (e.g. that the participants understand the instructions and the goal of the task). If participants fail the manipulation check, then their failures on the conditions of interest cannot be interpreted with respect to the experimental hypothesis. Put simply, if participants fail the manipulation check, the experiment is a failure, and it teaches us nothing. In this case, the Pirahã rarely gave the correct answer on tasks that could be solved by one-to-one matching or spatial alignment. Therefore, we cannot interpret their failures on tasks that require exact enumeration as evidence that they lack large exact number concepts.

Fortunately, this weakness of the data was addressed by a subsequent study. Michael Frank and colleagues (Frank et al. 2008) returned to the Pirahã and performed five tasks that were similar to Gordon's: two tasks could be solved via one-to-one matching, and therefore served as manipulation checks. The remaining three tasks could only be performed correctly by creating a representation of an exact number. This time, the data provided clear evidence that the Pirahã understood the tasks; they passed the manipulation checks, matching the number of objects that the experimenter presented with almost perfect accuracy for arrays of up to ten objects. Yet they failed to reproduce the correct number of objects when the tasks required exact enumeration (e.g. nuts in a can). As in Gordon's (2004) data, the Pirahã's responses approximated the target number, but participants rarely gave the correct answer for quantities greater than four. Since they passed the manipulation checks, these failures cannot be explained away as failures to understand the task. Rather, these data show that the capacity to represent "exactly four" as distinct from "exactly five" is not a human universal, validating Gordon's (2004) first claim and supporting a foundational assumption of Carey's (2004; 2009) model of how large exact number competence is acquired through language.

Yet for both Gordon's (2004) and Frank and colleagues' (2008) study, limitations of the experimental design preclude any claims about a causal role for language in shaping number concepts. These studies each reported data from only one group of participants, but implicitly the studies were cross-cultural comparisons between the Pirahã and Western adults (who were shown subsequently to perform almost perfectly on these simple enumeration tasks (Frank et al. 2012)). Therefore, the studies used a *quasi-experimental* design: participants were not randomly assigned to treatment and control groups, as they would be in a study using a *randomized-controlled* experimental design (i.e. participants were not assigned to be "treated" with number words or not). Quasi-experiments, no matter how well executed, cannot support causal inferences; they can only show correlations. In these studies, the experimental design was capable of showing a correlation between people's number vocabularies and their capacity for exact enumeration. The studies were not capable of showing a causal role for language in shaping number representations; thus, they were not capable, even in principle, of supporting a claim for linguistic determinism.

It is often challenging to implement randomized-controlled experiments in the field, and researchers must settle for constructing correlational studies so as to rule out alternative explanations for the correlations they observe. In the case of the studies by Gordon (2004) and by Frank et al. (2008), there was a looming alternative to the suggestion that language shapes numerical abilities; perhaps *culture* shapes numerical abilities. Cultures that have a counting system in language differ from cultures that lack such a system in many ways, making it difficult to isolate the role of language. The Pirahã results suggest that keeping track of large exact quantities is not critical for getting along in Pirahã society. In the absence of any environmental or cultural demand for exact enumeration, perhaps the Pirahã never developed this representational capacity – and consequently, they never developed the words (Casasanto 2005). On this view, perhaps being part of a numerate society is what drives the development of exact number concepts in individuals' minds: not language.

This skeptical possibility was addressed in a study by Elizabet Spaepen and colleagues (Spaepen et al. 2011) who tested numerical competence in Nicaraguan homesigners: deaf individuals who do not have access to any working model of language, oral or manual, and who have developed an idiosyncratic set of gestures to communicate, called *homesigns*. Although deprived of language, these homesigning adults are nevertheless functional members of a numerate society, and are exposed to opportunities and motivations to enumerate things exactly. Nevertheless, the homesigners could not consistently match target sets greater than three. Like the Pirahã, the homesigners were capable of approximating the target numbers (e.g. extending approximately the same number of fingers as the experimenter had shown them), but not matching them exactly. Thus, even when integrated into a numerate society, individuals do not spontaneously develop representations of large exact numerosities without input from a conventional language with a counting system.

Together, these findings provide compelling evidence that the capacity to represent exact numerosities greater than three is not a human universal. They also strongly suggest that a counting system in language may be necessary for the development of large exact number concepts. The data, to date, are correlational and cannot demonstrate a causal role for language in the development of number concepts, but no credible alternative has been advanced in light of the evidence that enculturation in a numerate society is not sufficient to drive the development of numeracy in individuals (Spaepen et al. 2011).

People exposed to number words develop a new representational capacity, and can entertain thoughts that are unthinkable by people who lack this linguistic experience: not only thoughts like, "What's the 12th digit of Pi?" but also thoughts like, "I'll have a dozen eggs" or "Take four steps forward." The capacity to represent and manipulate exact numbers is fundamental to the world as we know it, underlying the science and engineering that produced the buildings we live in, the medicines we take, the cars we drive, and the computers at our fingertips. Without exposure to a count list in language, it appears that the large exact number representations in our minds would not exist; without large exact number, our modern, technological world would not exist. As such, it seems reasonable to suggest that large exact number provides an example of one conceptual domain in which language has a dramatic and transformative effect on thought (cf. Pinker 1994; Bloom and Keil 2001; McWhorter 2014).

5 Varieties of linguistic relativity effects

How does language shape thought, over what timescale, by what mechanism, and how dramatic are the effects? There are no single answers to these questions. There are many parts of language, many aspects of cognition and perception, and many possible ways in which they can interact. This chapter illustrated how the grammatical packaging of information about motion events can *direct attention* to different aspects of the perceptible world, influencing what people remember about their experiences, at least so long as they can encode these experiences in words. Using different spoken metaphors can *strengthen some implicit associations in memory while weakening others*, leading to differences in the mental representation of time and musical pitch that can be found even when people are prevented from using language. Linking nouns with gendered determiners can *highlight certain features of objects or ideas*, making these features more salient for speakers of one language than for speakers of another, thus changing mental representations of nouns' referents in ways that may be subtle but pervasive. Finally, by serving as placeholders in an ordered sequence, number words can help to *create a new representational capacity*: one that radically changes our mind and world. When researchers arrive at different answers to questions about linguistic relativity this may be due, in part, to their examining different kinds of language-thought interactions, which operate over different timecourses, by different mechanisms.

Acknowledgements

This research was supported in part by grants from Junta de Andalucía, the European Regional Development Fund (P09-SEJ-4772), and the Spanish Ministry of Economy and Competitiveness (PSI2012-32464), and by a James S. McDonnell Foundation Scholar Award (220020236).

Further reading

Carey, S. 2004. Bootstrapping and the origin of concepts. *Daedalus* 133, 59–68. A brief overview of Susan Carey's proposal that number words play a critical role in the development of number concepts.

Casasanto, D. 2008. Who's afraid of the Big Bad Whorf? Cross-linguistic differences in temporal language and thought. *Language Learning* 58: 177, 63–79. A rebuttal to an influential critique of linguistic relativity.

Dolscheid, S., S. Shayan, A. Majid, and D. Casasanto 2013. The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychological Science* 24: 613–621. An illustration of "offline" effects of language on thought, showing that previous linguistic experience can influence subsequent thinking.

Frank, M. C., D. L., Everett, E. Fedorenko, and E. Gibson 2008. Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition* 108: 819–824. Evidence that an indigenous Amazonian group has neither exact number words nor exact number concepts.

Whorf, B. L. 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf.* Edited by John B. Carroll. Cambridge, MA: MIT Press. A collection of Benjamin Lee Whorf's writings on relationships between language and thought.

References

Alverson, H. 1994. *Semantics and Experience: Universal Metaphors of Time in English, Mandarin, Hindi, and Sesotho*. Baltimore: Johns Hopkins University Press.

Barner, D., D. Thalwitz, J. Wood, and S. Carey 2007. On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science* 10: 365–373.

Besner, D., J. A. Stolz and C. Boutilier 1997. The Stroop effect and the myth of automaticity. *Psychonomic Bulletin and Review* 42: 221–225.

Bloom, P. and F. C. Keil 2001. Thinking through language. *Mind and Language* 164: 351–367.

Boroditsky, L., L. A. Schmidt and W. Phillips 2003. Sex, syntax, and semantics. In D. Gentner and S. Goldin-Meadow (eds) *Language in Mind: Advances in the Study of Language and Thought*. Cambridge, MA: MIT Press, 61–79.

Bottini, R. and D. Casasanto, 2013. Space and time in the child's mind: metaphoric or ATOMic? *Frontiers in Cognition.* doi: 10.3389/fpsyg.2013.00803

Carey, S. 2004. Bootstrapping and the origin of concepts. *Daedalus* 133: 59–68.

Carey, S. 2009. *The Origin of Concepts.* New York: Oxford University Press.

Casasanto, D. 2005. Crying "Whorf!" *Science* 307: 1721–1722.

Casasanto, D. 2008. Who's afraid of the Big Bad Whorf? Cross-linguistic differences in temporal language and thought. *Language Learning* 58 Suppl 1: 63–79.

Casasanto, D. 2010. Space for thinking. In V. Evans and P. Chilton (eds) *Language, Cognition and Space: State of the Art and New Directions.* London: Equinox Publishing, 453–478.

Casasanto, D. and L. Boroditsky 2008. Time in the mind: using space to think about time. *Cognition* 1062: 579–593.

Casasanto, D., L. Boroditsky, W. Phillips, J. Greene, S. Goswami, S. Bocanegra-Thiel et al. 2004. How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. In K. Forbus, D. Gentner, and T. Regier (eds) *Proceedings of the 26th Annual Conference Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates, 575–580.

Casasanto, D. and R. Bottini 2014. Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General* 1432: 473–479.

Casasanto, D. and G. Lupyan 2015. All concepts are ad hoc concepts. In E. Margolis and S. Laurence (eds) *The Conceptual Mind: New Directions in the Study of Concepts.* Cambridge: MIT Press, 543–566.

Clark, E. 2003. Languages and representations. In D. Gentner and S. Goldin-Meadow (eds) *Language in Mind: Advances in the Study of Language and Thought*. Cambridge: MIT Press, 17–23.

Clark, H. H. 1973. Space, time, semantics and the child. In T. E. Moore (ed.) *Cognitive Development and the Acquisition of Language*. New York: Academic Press, 27–63.

Dehaene, S. 1999. *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford University Press.

Dolscheid, S., S. Hunnius, D. Casasanto, and A. Majid 2014. Prelinguistic infants are sensitive to space-pitch associations found across cultures. *Psychological Science.* doi: 10.1177/0956797614528521.

Dolscheid, S., S. Shayan, A. Majid, and D. Casasanto 2013. The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychological Science* 245: 613–621.

Evans, V. 2004. *The Structure of Time: Language, Meaning and Temporal Cognition*. Amsterdam: John Benjamins.

Fausey, C. M. and L. Boroditsky 2011. Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin and Review* 18: 150–157.

Feigenson, L., S. Dehaene, and E. Spelke 2004. Core systems of number. *Trends in Cognitive Sciences* 8: 307–314.

Frank, M. C., D. L. Everett, E. Fedorenko, and E. Gibson 2008. Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition* 108: 819–824.

Frank, M. C., E. Fedorenko, P. Lai, R. Saxe, and E. Gibson 2012. Verbal interference suppresses exact numerical representation. *Cognitive Psychology* 64: 74–92.

Gelman, R. and C. R. Gallistel 2004. Language and the origin of numerical concepts. *Science* 306: 441–443.
Gennari, S. P., S. A. Sloman, B. C. Malt, and W. Fitch 2002. Motion events in language and cognition. *Cognition* 83: 49–79.
Gleitman, L. and A. Papafragou 2013. Relations between language and thought. In D. Reisberg (ed.) *Handbook of Cognitive Psychology*. New York: Oxford University Press, 504–523.
Gordon, P. 2004. Numerical cognition without words: evidence from Amazonia. *Science* 306: 496–499.
James, W. 1890. *Principles of Psychology*. Vol. 1. New York: Holt.
Johnson-Laird, P. N. 1987. The mental representation of the meaning of words. *Cognition* 25: 189–211.
Kay, P. and W. Kempton 1984. What is the Sapir-Whorf hypothesis? *American Anthropologist* 86: 65–79.
Kay, P. and T. Regier 2006. Language, thought and color: recent developments. *Trends in Cognitive Sciences* 10: 51–54.
Lakoff, G. and M. Johnson 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Chicago: University of Chicago Press.
Landau, B., B. Dessalegn, and A. Goldberg 2010. Language and space: momentary interactions. In V. Evans and P. Chilton (eds) *Language, Cognition and Space: The State of the Art and New Directions. Advances in Cognitive Linguistics Series.* London: Equinox, 51–78.
Levinson, S. C. 2003. *Space in Language and Cognition: Explorations in Cognitive Diversity.* Vol. 5. Cambridge: Cambridge University Press.
Levinson, S. C. and P. Brown 1994. Immanuel Kant among the Tenejapans: anthropology as empirical philosophy. *Ethos* 22: 3–41.
Lohmann, H. and M. Tomasello 2003. The role of language in the development of false belief understanding: a training study. *Child Development* 74: 1130–1144.
Lourenco, S. F. and M. R. Longo 2011. Origins and development of generalized magnitude representation. In S. Dehaene and E. M. Brannon (eds.) *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought*. London: Academic, 225–244.
Lucy, J. A. 1992. *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
Lucy, J. A. (ed.) 1996. *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
Lucy, J. A. and S. Gaskins 2001. Grammatical categories and the development of classification preferences: a comparative approach. In M. Bowerman and S. C. Levinson (eds) *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 257–283.
McDonough, L., S. Choi, and J. M. Mandler 2003. Understanding spatial relations: flexible infants, lexical adults. *Cognitive Psychology* 46: 229–259.
McWhorter, J. H. 2014. *The Language Hoax: Why the World Looks the Same in Any Language*. Oxford: Oxford University Press.
Oh, K. J. 2003. *Language, Cognition, and Development: Motion Events in English and Korean*. Doctoral dissertation, University of California, Berkeley.
Papafragou, A., C. Massey, and L. Gleitman 2002. Shake, rattle, 'n' roll: the representation of motion in language and cognition. *Cognition* 84: 189–219.
Papafragou, A., J. Hulbert and J. Trueswell 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 10: 155–184.
Papafragou, A., P. Li, Y. Choi, and C. H. Han 2007. Evidentiality in language and cognition. *Cognition* 103: 253–299.
Pinker, S. 1994. *The Language Instinct: How the Mind Creates Language*. New York: Harper.
Pinker, S. 1999. *Words and Rules: The Ingredients of Language*. New York: Basic Books.
Pullum, G. K. 1991. *The Great Eskimo Vocabulary Hoax and Other Irreverent Essays on the Study of Language*. Chicago: University of Chicago Press.
Sapir, E. 1929. The status of linguistics as a science. *Language* 5: 207–214.
Slobin, D. I. 1996. From "thought and language" to "thinking for speaking." In J. J. Gumperz and S. C. Levinson (eds) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press, 70–96.

Slobin, D. I. 2003. Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner and S. Goldin-Meadow (eds) *Language in Mind: Advances in the Study of Language and Thought*. Cambridge: MIT Press, 157–192.
Spaepen, E., M. Coppola, E. S. Spelke, S. E. Carey, and S. Goldin-Meadow 2011. Number without a language model. *Proceedings of the National Academy of Sciences* 10: 3163–3168.
Spelke, E. S. and S. Tsivkin 2001. Language and number: a bilingual training study. *Cognition* 78: 45–88.
Spivey, M. 2007. *The Continuity of Mind*. Oxford: Oxford University Press.
Srinivasan, M. and S. Carey 2010. The long and the short of it: on the nature and origin of functional overlap between representations of space and time. *Cognition* 11: 217–241.
Srinivasan, M. 2010. Do classifiers predict differences in cognitive processing? A study of nominal classification in Mandarin Chinese. *Language and Cognition* 2: 177–190.
Thierry, G., P. Athanasopoulos, A. Wiggett, B. Dering, and J. R. Kuipers 2009. Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences* 106: 4567–4570.
Trueswell, J. C. and A. Papafragou 2010. Perceiving and remembering events cross-linguistically: evidence from dual-task paradigms. *Journal of Memory and Language* 63: 64–82.
von Humboldt, W. 1988. *On Language: The Diversity of Human Language-Structure and its Influence on the Mental Development of Mankind*. Cambridge: Cambridge University Press.
Walker, P., J. G. Bremner, U. Mason, J. Spring, K. Mattock, A. Slater, and S. P. Johnson 2010. Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science* 21: 21–25.
Whorf, B. L. 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf.* Edited by John B. Carroll. Cambridge: MIT Press.
Winawer, J., N. Witthoft, M. C. Frank, L. Wu, and L. Boroditsky 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Science* 104: 7780–7785.
Wolff, P., G. H. Jeon, and Y. Li 2009. Causers in English, Korean, and Chinese and the individuation of events. *Language and Cognition* 1: 167–196.

Related topics

Chapter 2, Internalist semantics; Chapter 5, Cognitive semantics; Chapter 7, Categories, prototypes and exemplars; Chapter 8, Embodiment, simulation and meaning; Chapter 25, The semantics of lexical typology; Chapter 26, Acquisition of meaning.