Introduction

The Sapir-Whorf hypothesis (a.k.a. the Whorfian hypothesis) concerns the relationship between language and thought. Neither the anthropological linguist Edward Sapir (b. 1884–d. 1939) nor his student Benjamin Whorf (b. 1897–d. 1941) ever formally stated any single hypothesis about the influence of language on nonlinguistic cognition and perception. On the basis of their writings, however, two proposals emerged, generating decades of controversy among anthropologists, linguists, philosophers, and psychologists. According to the more radical proposal, linguistic determinism, the languages that people speak rigidly determine the way they perceive and understand the world. On the more moderate proposal, linguistic relativity, habits of using language influence habits of thinking. As a result, people who speak different languages think differently in predictable ways. During the latter half of the 20th century, the Sapir-Whorf hypothesis was widely regarded as false. Around the turn of the 21st century, however, experimental evidence reopened debate about the extent to which language shapes nonlinguistic cognition and perception. Scientific tests of linguistic determinism and linguistic relativity help to clarify what is universal in the human mind and what depends on the particulars of people’s physical and social experience.

General Overviews and Foundational Texts

Writing on the relationship between language and thought predates Sapir and Whorf, and extends beyond the academy. The 19th-century German philosopher Wilhelm von Humboldt argued that language constrains people's worldview, foreshadowing the idea of linguistic determinism later articulated in Sapir 1929 and Whorf 1956 (Humboldt 1988). The intuition that language radically determines thought has been explored in works of fiction such as Orwell's dystopian fantasy 1984 (Orwell 1949). Although there is little empirical support for radical linguistic determinism, more moderate forms of linguistic relativity continue to generate influential research, reviewed from an anthropologist’s perspective in Lucy 1997, from a psychologist’s perspective in Hunt and Agnoli 1991, and discussed from multidisciplinary perspectives in Gumperz and Levinson 1996 and Gentner and Goldin-Meadow 2003.


Edited volume containing position papers for and against linguistic relativity. Includes reviews of some of the experimental studies that revived widespread interest in the Sapir-Whorf hypothesis at the beginning of the 21st century.


Edited volume containing position papers for and against linguistic relativity. A cross-section of Whorfian research in anthropology, psychology, and linguistics at the end of the 20th century.


Humboldt argues that language determines one’s world view.


A critical review of 20th-century Whorfian research, in which the authors sketch proposals for several studies that were brought to fruition.
by other researchers over the ensuing two decades.

A review focusing on the various ways in which the Whorfian question was approached empirically during the 20th century.

Fictitious account of a totalitarian state in which language is used to control thought.

Sapir states the view that language shapes one’s worldview, subsequently called linguistic determinism.

The definitive collection of Whorf's writings, some posthumously published.

**Journals**

For further reading, the following journals have a record of publishing important papers on the relationship between language and thought, reporting research in anthropology (e.g., *American Anthropologist*), psychology (e.g., *Cognitive Psychology*; *Journal of Experimental Psychology: General*; *Psychological Science*), and interdisciplinary cognitive science (e.g., *Cognition*; *Cognitive Science*; *Language and Cognition*; *Trends in Cognitive Sciences*).

**American Anthropologist.**

**Cognition.**
An international journal of cognitive science. Peer reviewed. Published by Elsevier. Publishes theoretical and experimental papers on all aspects of cognition. Several influential debates relevant to the Sapir-Whorf hypothesis have appeared here, in the form of a series of independently submitted papers.

**Cognitive Psychology.**
Published by Elsevier. Peer reviewed. Publishes empirical, theoretical, and tutorial papers; methodological articles; and critical reviews. Focuses on empirical articles that provide major theoretical advances in the study of cognition.

**Cognitive Science**
Journal of the Cognitive Science Society. Peer reviewed. Published by Wiley-Blackwell. Promotes scientific interchange among researchers in disciplines comprising the field of cognitive science, including artificial intelligence, linguistics, anthropology, psychology, neuroscience, philosophy, and education.

**Journal of Experimental Psychology: General.**
Anti-Whorfian Literature

During the latter half of the 20th century, some of the most memorable writing about the Whorfian hypothesis was by its opponents. Leading figures in linguistics (Chomsky 1973), philosophy (Fodor 1985), and psychology (Pinker 1994) appear to have been vying to see who could denounce the notion of linguistic relativity the most emphatically, or the most humorously (Pullum 1991). Even as studies accumulated that caused some scholars to reexamine the Whorfian question in the 21st century, others remained convinced that the Sapir-Whorf hypothesis was flawed in principle (Bloom and Keil 2001), or that the empirical support was weak (Gleitman and Papafragou 2005, Munnich and Landau 2003) or uninteresting (Pinker 2007).

Critique of the Sapir-Whorf hypothesis and the data supporting it.

An “extremely skeptical” (p. x) assessment of Adam Schaff’s pro-Whorfian stance by an eminent linguist.

Statement about the relations among language, cognition, and perception by an eminent philosopher of mind, who states that he “hate[s]” the notion of relativity (linguistic, cultural, etc. [p. 5]).

Review and critique of Whorfian research by one of its most outspoken opponents (Gleitman).

A review and skeptical evaluation of Whorfian research, including some research from the early 21st century.

An eminent psychologist argues that the Whorfian hypothesis is “wrong, all wrong” (p. 57).

A broad and engaging book on relationships between language and thought, which includes a skeptical appraisal of Whorfian research in the early 21st century.

A humorous and insightful critique of linguistic data and argumentation for the Whorfian hypothesis.

**Pro-Whorfian Literature**

Nearly all pro-Whorfian writing begins with an acknowledgment of the controversy surrounding the Sapir-Whorf hypothesis, and of the limitations of previous Whorfian research. In the following books and articles, the authors provide reasons why decades of controversy should not cause readers to reject the notion of linguistic relativity, marshaling theoretical arguments, empirical data, and experimental methods that provide new answers to long-debated questions. Carroll’s introduction to *Language, Thought, and Reality* (Carroll 1956) provides an overview of Whorf’s life and work. Lee 1996 provides a thorough exegesis and reanalysis of writing by Whorf and his critics. Levinson 2003 seeks to sort out some of the sources of real or perceived disagreement among pro- and anti-Whorfian researchers. Schaff 1973 applies Whorfian arguments to philosophy and political theory. Lenneberg 1953, Slobin 1996, and Casasanto 2008 each introduce new theoretical perspectives and innovative methods for testing the Whorfian hypothesis. Boroditsky 2003 briefly reviews some of the experimental studies conducted by psychologists and linguistic anthropologists around the turn of the 21st century.

A brief review of Whorfian research by a leading scholar, focusing on the first wave of experimental studies that marked a resurgence of interest in the Sapir-Whorf hypothesis at the beginning of the 21st century.

An overview of Whorf’s life and work.

A refutation of anti-Whorfian arguments in Pinker 1994 (cited under Anti-Whorfian Literature) and other works, and a demonstration of methods that make it possible to conduct fully nonlinguistic tests of the Sapir-Whorf hypothesis, thus escaping the circularity of earlier Whorfian research (see Pullum 1991, cited under Anti-Whorfian Literature).

A historical overview and reanalysis of Whorf and his critics.

A critique of ethnolinguistic research in the first half of the 20th century, followed by an attempt to propose a more fruitful methodology for conducting Whorfian research.
Empirical Tests of the Whorfian Hypothesis

For years, the Whorfian controversy was fueled by a dearth of relevant empirical evidence. Although a large literature documented differences among the grammars and lexicons of the world’s languages, these data are not sufficient to support Whorfian claims: In order to establish whether people who talk differently also think differently, it is necessary to show that linguistic differences correspond to different behavior on some measure of nonlinguistic cognition or perception. The circularity of Whorfian claims that were based on linguistic data alone drove many scholars to dismiss the Whorfian hypothesis as logically flawed and empirically unsupported. More than half a century after the deaths of Sapir and Whorf, however, methodological advances have given rise to a body of evidence supporting some version of linguistic relativity: Numerous aspects of cognition and perception appear to depend, in part, on aspects of people’s linguistic experience. Cross-linguistic differences in grammatical or lexical patterns have been reported to influence mental representations in a variety of conceptual domains (e.g., Time, Space, Motion, Color). Debate continues about how to interpret these empirical data with respect to theories of language and mind.

TIME

Time is one of the conceptual domains Whorf analyzed in a controversial essay, in which he noted differences between the Hopi language and so-called Standard Average European languages and suggested that there must be corresponding differences in their speakers’ concepts (Whorf 1956). Since then, time has become a frequent test bed for the Whorfian hypothesis. In particular, researchers have focused on how different languages use different spatial metaphors for time. Several studies suggest that people also think about time spatially, conceptualizing temporal order (Boroditsky 2001, Núñez and Sweetser 2006) or duration (Casasanto, et al. 2004) as predicted by the particular space-time metaphors found in their languages. Beyond influences of spoken language, other studies show that the orthography used by literate members of a culture can influence the way they conceptualize time (Casasanto and Bottini 2010; Fuhrman and Boroditsky 2010; Ouellet, et al. 2010; Tversky, et al. 1991).

The first experimental study showing that speakers of English and Mandarin Chinese interpret temporal expressions in language differently, suggesting that English and Mandarin speakers tend to use different spatial schemas to conceptualize events in time.

By using psychophysical tasks with nonlinguistic stimuli and responses, this study provides the first evidence that speakers of different languages conceptualize time differently, even when they are not using language.

Whereas cross-linguistic studies show correlations between writing habits and mental representations of time, this training study establishes a causal relationship between orthography and temporal concepts and shows that exposure to a new orthography can reverse the left-to-right spatial mapping of time in people’s minds.


Evidence that the left-right mapping of time reverses between English and Hebrew speakers, consistent with the orthographies of these languages.


Evidence that speakers of Aymara, a South American language, gesture forward for the past and backward for the future, consistent with spoken metaphors in their language.


Evidence that the left-right mapping of time reverses between Spanish and Hebrew speakers, consistent with the orthographies of these languages.


The first study to show that people implicitly associate earlier times with the left and later times with the right of body-centered space, and that the direction in which time flows in their minds covaries with the direction of reading and writing in their language.


Whorf argues, on the basis of the grammar and lexicon of Hopi, that Hopi speakers must have markedly different conceptions of time from speakers of “Standard Average European” languages.

**SPACE**

Spatial cognition is a frequent test bed for the Sapir-Whorf hypothesis. Several studies examine cross-linguistic differences in the use of spatial frames of reference and their effects on spatial reasoning (see Levinson and Brown 1994; Majid, et al. 2004 for Whorfian claims; see Li and Gleitman 2002 for counterargument). Related studies investigate how spatial words, such as “right” and “left,” enable speakers of languages that have these terms available to navigate their spatial environment, using strategies that are unavailable to people and animals who lack these words (Hermer and Spelke 1994; Hermer-Vazquez, et al. 1999). Other studies test whether cross-linguistic differences in lexicalized spatial categories (e.g., “in,” “on”) correspond to differences in the way speakers of different languages conceptualize relationships between objects (Bowerman and Choi 2001; Hespos and Spelke 2004; McDonough, et al. 2003).


Evidence that children learn to use different spatial concepts as they acquire languages that lexicalize different spatial relationships.
Evidence that, like animals, young children who lack use of the terms “left” and “right” are limited in their strategies for spatial navigation.

Evidence that when the use of spatial terms like “left” and “right” is blocked, adults are limited in their strategies for spatial navigation, like children and animals. This study establishes a causal role for language in spatial reasoning.

Evidence for spatial conceptual categories that precede language development.

An argument against the universality of body-based frames of reference in language and thought.

An experimental study on the basis of which the authors question earlier conclusions about the influence of spatial language on spatial reasoning.

A concise review of experiments testing the universality of spatial frames of reference in cognition.

Evidence that, although a range of spatial relational categories are available to infants, some of these categories become harder to use in adults whose language does not encode them.

**MOTION**
Across languages, syntax requires the same information to be packaged differently. Due to this syntactic packaging, information that is obligatory in one language may be optional in another. Slobin 1996 proposed that this should lead to differences in the way speakers of different languages experience the world and mentally represent their experiences, at least while they are using language (a proposal known as the “thinking for speaking” hypothesis, elaborated in Slobin 2003). The initial test bed for Slobin’s proposal was the syntactic encoding of motion information. Whereas some languages focus speakers’ attention on the path that people or objects travel (e.g., entering vs. exiting), others focus attention on the manner of motion (e.g., running vs. rolling; see Talmy 1991 for a full explanation of this linguistic distinction). Numerous studies have tested whether this difference has consequences for speakers’ nonlinguistic mental representations of motion events. Some studies offer no support for the proposal that syntax influences “thinking for speaking” about motion events (Papafragou, et al. 2002). Other studies support the hypothesis (Choi and Bowerman 1991, Kita and Özyürek 2003), and investigate the circumstances under which this Whorfian effect is found (Gennari, et al. 2002; Papafragou, et al. 2008).

Evidence that the languages children are exposed to can influence their mental representations of motion events.
Evidence that grammar affects motion-event representations after people verbally describe the event, supporting a weak version of linguistic relativity.

Reports cross-linguistic differences in the encoding of path and manner information in spontaneous co-speech gestures.

An eyetracking study that provides insight into the time course over which event processing is affected by language.

Experiments that find no influence of grammar on event representation.

A statement of Slobin’s original hypothesis about effects of motion language on event cognition.

A refinement of Slobin’s hypothesis (see Slobin 1996, cited under Pro-Whorfian Literature).

Explains the linguistic basis for the proposed cross-linguistic difference in event representation.

**NUMBER**

The capacity to mentally represent exact numbers of objects larger than three or four appears to be uniquely human, and appears to depend in part on exposure to a counting system in language. Carey 2004 sketches a mechanism by which learning to count enables children to construct the concept of number, augmenting their innate cognitive capacities (see Rips, et al. 2006 for a counterargument).
The role of language in the acquisition of number concepts has been tested in members of Amazonian tribes that lack extensive counting systems in their languages (Frank, et al. 2008; Gordon 2004; Pica, et al. 2004; but see Casasanto 2005). Especially compelling are tests of numerical cognition in deaf people who are members of numerate culture but do not know how to count (Spaepen, et al. 2011). Number is one cognitive domain in which claims for strong linguistic determinism are still viable: It appears that if people are not exposed to counting numbers in language, they do not develop cognitive capacities that most Western adults take for granted (e.g., the ability to distinguish six objects from seven). Beyond its role in the acquisition of number, it appears that language is used online as people perform mental arithmetic (Spelke and Tsivkin 2001).
A proposal that learning number words, initially as a meaningless word game, enables children to acquire number concepts.

A critique of the methods and strong Whorfian conclusions of Gordon 2004.

A study of the Pirahãs’ numerical abilities that addresses some of the concerns raised about Gordon 2004 and clarifies the facts of the Pirahã number lexicon.

A study claiming evidence for strong linguistic determinism, on the basis of experiments on the numerical abilities of the Pirahã, an Amazonian group with no words for exact numbers.

Evidence for a dramatic difference in numerical abilities between speakers of a language with a full counting list and speakers of a language with few number words.

An argument against the proposal in Carey 2004 that children use language to learn number concepts.

Evidence that deaf people who are not exposed to number words as children do not acquire number concepts, despite living in a numerate society. This study strengthens the claim that language (as opposed to other aspects of culture) is crucial for the development of numerical cognition.

Evidence of interactions between language and mental arithmetic.

COLOR

Some of the earliest attempts to test the Sapir-Whorf hypothesis used the domain of color as a test bed. Experiments have led to numerous conflicting pro-Whorfian and anti-Whorfian claims, sometimes by the same researcher analyzing different data sets (see Brown and Lenneberg 1954 and Brown 1976) or by different researchers analyzing the same data set (see Heider 1972 and Davidoff, et al. 1999). Debates continue about the extent to which color categories in language influence color categories in cognition and perception. A focus of debate has been the role of language in producing categorical perception of colors. Although the wavelength spectrum is continuous, mental color space appears to be divided into discrete categories, such that people are generally more efficient at discriminating colors across category boundaries than within a category (e.g., it is easier to distinguish a green dot from a blue dot than to discriminate two blue dots, even when the pairs of dots differ by the same number of wavelengths). Some scholars have argued that these category boundaries are universal (Berlin and Kay 1969). Yet experiments suggest that mental color boundaries are at least partly determined by language (see Kay and Kempton 1984; Özgen and Davies 2002; Regier and Kay 2009; Thierry, et al. 2009).
conditions under which language influences color discrimination and the cognitive mechanisms by which language has its effects remain areas of active research.

Berlin, Brent, and Paul Kay. 1969. *Basic color terms: Their universality and evolution*. Berkeley: Univ. of California Press.  An ethnographic study that reduces all of the variability in the color vocabularies of the world's languages to eleven basic terms.


Özgen, Emre and Ian R. L. Davies. 2002. Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General* 131:477–493.  Evidence that categorical perception effects can be induced through training. The authors propose and validate an account of how lexical color terms shape nonlinguistic color categories in the mind, via perceptual learning.

Regier, Terry, and Paul Kay. 2009. Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences* 13.10: 439–446.  A concise review of studies showing effects of language on color judgments primarily for stimuli presented in the right visual hemifield, and therefore processed by the left hemisphere of the brain.

Thierry, Guillaume, Panos Athanasopoulos, Alison Wiggett, Benjamin Dering, and Jan-Rouke Kuipers. 2009. Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences* 106.11: 4567–4570.  Using electrophysiological measures, the authors demonstrate language-specific sensitivity to color categories, providing the first clear evidence of influences of linguistic color categories on color perception; because the study used an implicit measure of unconscious
It may seem surprising that mental representations of concrete objects should be susceptible to influences of language, but this appears to be the case. According to many people’s intuitions, our conceptualizations of things like puddles, pots, and ponies should depend solely on our perceptual experiences with these things, but several lines of research challenge this intuition. One set of studies investigates how grammatically marking a distinction between objects and substances influences people’s construal of ambiguous stimuli (Lucy and Gaskins 2001; Imai and Gentner 1997; but see Barner, et al. 2010). Related studies investigate effects of nominal classifiers (i.e., particles that classify objects according to their shape) on object perception and categorization (Srinivasan 2010, Gao and Malt 2009). Still other studies examine effects of grammatical gender on people’s concepts of inanimate objects. These studies suggest that using a masculine gender marker, such as the French “le” or a feminine marker like “la,” can cause speakers to attribute more masculine or feminine qualities to inherently sexless objects like toasters and teapots (Boroditsky, et al. 2003; Sera, et al. 2002; Vigliocco, et al. 2005). The mechanisms by which these aspects of grammar affect object concepts, and the circumstances under which Whorfian effects can be observed, are still under investigation.

A brief review and skeptical evaluation of studies suggesting that count/mass syntax influences people’s object representations.

Evidence that grammatical gender affects people’s conceptual representations of objects.

Extensive exploration of Chinese classifiers and of their observed and potential cognitive consequences.

Evidence that cross-linguistic differences in the count/mass distinction in grammar can influence children’s categorization of objects and substances.

Evidence that cross-linguistic count/mass syntax influences adults’ construal of objects and substances.

Evidence of effects of grammatical gender on categorization in Italian but not in German, and an exploration of these positive and negative findings using connectionist modeling.

Evidence that nominal classifiers influence object processing in a visual search task.


**COUNTERFACTUAL REASONING**

Some languages, like Chinese, lack the subjunctive mood, which is used in other languages like English to express counterfactual scenarios (e.g., If it were to rain, then I would take an umbrella). Because counterfactuals are harder to express in Chinese than in English, Bloom 1981 posited that Chinese speakers would have more difficulty understanding counterfactual aspects of narratives than speakers of languages like English. Although this proposal appeared to be supported initially, Au 1983 challenged Bloom’s claim, suggesting that Bloom’s results were due to awkward translations, and providing experimental evidence that when tested in Chinese, Chinese bilinguals performed better on tests of counterfactual understanding than American subjects did when tested in English.


**THEORY OF MIND**

Learning to use certain syntactic structures may help children develop theory of mind (i.e., the capacity to think about thinking and to understand what other people know, want, or believe). On one proposal, using complement clauses (e.g., he said that the apple was an orange) plays a special role in developing the capacity to reason about false beliefs (De Villiers and Pyers 1997, Lohmann and Tomasello 2003; but see Slade and Ruffman 2005). Another proposal suggests a similar effect of using evidential morphology. Evidential markers on verbs indicate the source of the speaker’s knowledge (e.g., whether you saw something yourself or only heard about it from another person), and are obligatory in some languages, such as Turkish and Korean. The empirical record shows some results that suggest using evidential morphology may hasten the development of theory of mind (Papafragou 2002) and other results that challenge this proposal (Papafragou, et al. 2007).


than age-matched speakers of English (a language with no evidential markers), supporting a proposed role for evidential morphology in the acquisition of theory of mind.


CAUSATION

Across languages, there are different ways of describing causal relationships between actors and actions. In English, it is possible to say that the key opened the lock. In Korean and other languages, however, such a statement would sound strange because keys are inanimate and cannot cause locks to open without an animate agent (Wolff, et al. 2009; Wolff, et al. 2010). In Spanish, a common way to express the idea that she broke the vase would be se rompió el florero, literally “the vase broke itself” (Fausey and Boroditsky 2011). Such differences in the linguistic coding of causation appear to influence how speakers of different languages perceive and remember events and how people assign blame for the effects of people’s actions (Fausey and Boroditsky 2010).


