Comprehension – Behavioral Studies: The Sapir-Whorf Hypothesis

HISTORY AND CONTROVERSY

The Sapir-Whorf hypothesis refers to the claim that features of one's language affect one's thoughts (linguistic determinism), so that different languages foster different conceptual systems (linguistic relativity). This lemma reviews experimental tests of this hypothesis in Chinese.

The hypothesis is so named because of notions expressed in Sapir (1929) and Whorf (1939) (cf. Hill and Mannheim 1992). In its first modern application to Chinese (cf. Kwan 2001), Hockett (1954) cites the greater emphasis on manner in Chinese than in English verbs of holding and breaking, as well as the lack of number marking in Chinese, but he provides no evidence that these differences affect thought.

The Sapir-Whorf hypothesis remains controversial (Gumperz and Levinson 1996), in part because researchers differ in their views on its content and value. Pinker (2007) distinguishes ten distinct interpretations of linguistic determinism, among them "thinking for speaking" (Slobin 1996), in which speakers habitually encode concepts in ways that make them easier to translate into speech. However, it has proven difficult to demonstrate empirically that such conceptual habits are active in the absence of linguistic processing. Putatively nonlinguistic tasks may still trigger the use of language as a silent mental tool (e.g. activating plausible category names). Nevertheless, if language structure does affect behavior, this fact has scientific and practical importance regardless of the mechanism.

Linguistic relativity poses the additional challenge of inferring causation from correlation. Speakers of different languages also differ in non-linguistic experience, which may influence both language and concepts (Ji et al. 2004). For example, Ross et al. (2002)

found that speaking Chinese led Chinese-English bilinguals to emphasize group memberships in self-descriptions more than when speaking English: each language primed the social assumptions of the associated culture, despite the lack of linguistic features marking them.

EXPERIMENTAL TESTS OF THE SAPIR-WHORF HYPOTHESIS IN CHINESE

Experiments have tested potential linguistic influences on cognition along a continuum from lexical to grammatical.

In a study testing for lexical effects, Brown (1986) found that English speakers draw causal inferences from verbs matching their associated adjectives (e.g. subject-oriented *attract*/*attractive* vs. object-oriented *abhor*/*abhorrent*). However, when the materials were translated into Chinese, which does not have such adjective associates, speakers in Hong Kong had the same causal interpretations. These results support universal constraints on verbal semantics, not linguistic relativity. Frank et al. (2000) were led to similar conclusions regarding types of shame distinguished lexically in Chinese but not English (e.g. 羞耻 *xiūchǐ* vs. 慚愧 *cánkuì*). English-speaker descriptions of various shame-related scenarios fit categories corresponding to those lexicalized in Chinese, again supporting universal semantics.

In another lexical study, Hoffman et al. (1986) gave English monolinguals and Chinese-English bilinguals descriptions fitting an English or Chinese personality schema (e.g. the same person may be described as 世故 *shìgù* in Chinese or less efficiently in English as worldly and somewhat reserved). Novel descriptions derived from schemas in the testing language tended to be misidentified as having already appeared, showing that the schemas were mentally active. However, the authors admit that these results are consistent with language influencing behavior in the task rather than thought per se.

Moving into semi-closed class vocabulary, Huang (1999) examined two time systems in Chinese. Chinese months are named sequentially after numbers (e.g. 十二月 *shí'èr yuè* 'December', literally 'twelve month'), but the 節氣 *jiéqì* system uses twenty-four lexically distinct units (e.g. 大雪 *dàxuě* 'heavy snow'). Speakers fluent in both systems judged if a given month or *jiéqì* unit occurs three or five units before or after another. While *jiéqì* judgments were slower when units were further apart in time, month judgments were equally fast, as if participants used arithmetic. These results again seem to show language being used as a tool in the task.

Chinese commonly uses a vertical metaphor for time (e.g. 上個月 *shàng ge yuè* 'previous month', literally 'up a month'), whereas English favors a horizontal metaphor (e.g. 'the month before'). To see if the vertical time metaphor has truly restructured Chinese cognition, Boroditsky (2001) tested speakers exclusively in English. Bilingual Chinese-English speakers (but not English monolinguals) made faster 'earlier'/'later' judgments about the relative order of months after making overtly spatial judgments on vertical primes.

However, these effects could not be replicated, in either Chinese (Chen 2007) or English (January and Kako 2007); Boroditsky et al. (2011) concede their own replication difficulty. Moreover, Chen (2007) notes that even in Chinese, horizontal time metaphors far outnumber vertical ones, and January and Kako (2007) ask why brief training caused even English speakers to show vertical priming, when years of English experience barely influenced the Chinese-English bilinguals.

Boroditsky et al. (2011) (see also Miles et al. 2011) tested the same claim with a new task. Participants saw a sequence of two pictures (e.g. photographs of the same person at different ages) and judged whether the second was 'earlier' or 'later'. Chinese but not English speakers were faster if the 'earlier' response key was above the 'later' key. Similarly, Furhman

et al. (2011) found that Chinese speakers were more likely than English speakers to point above or below a reference point to indicate relative time.

Yet ascribing cognitive effects solely to the vertical metaphor is difficult given that cognition is also affected by the vertical orientation of Chinese text (Chan and Bergen 2005). The effect of text orientation on temporal judgments was confirmed by Chen and O'Seaghdha (forthcoming), who found that only participants from Taiwan (where vertical text is common) behaved as described in Boroditsky et al. (2011); participants from China (where vertical text is rarer) behaved like the English-speaking participants.

Chinese classifiers represent another semi-closed lexical class. Zhang and Schmitt (1998) found that Chinese speakers, but not English speakers, judged noun pairs as more similar if they shared the same classifier. In a more ambitious study, Saalbach and Imai (2007) presented Chinese and German speakers with pairs of nouns related taxonomically (categorically), thematically (relationally), or via shared classifier. Although two out of four tasks showed stronger classifier effects for Chinese speakers, the Germans also grouped classifier-related nouns together, suggesting that classifiers encode universal semantic features.

Although shape is a universal semantic feature, experience with shape classifiers may make Chinese speakers more sensitive to it. Kuo and Sera (2009) found that in classifying objects, Chinese speakers were influenced by shape more than English speakers, though both groups primarily classified taxonomically and functionally. Imai et al. (2010) found similar results for Chinese and German children in a classification task, but in other tasks, both groups favored either shape or taxonomic categories.

Such classifier effects seem to require linguistic processing. Gao and Malt (2009) had Chinese speakers memorize nouns in sentence contexts. Noun recall fell more into classifier-defined clusters (compared with English speakers), but only when nouns had

appeared in sentences with overt classifiers. Huang and Chen (2011) replicated this pattern in a different task. In an eye-tracking study, Huettig et al. (2010) found that Mandarin speakers looked at pictures of objects sharing the same classifier as an object named in an auditorily presented sentence, but only when the classifier was also presented. Tsang and Chambers (2011) report related results in Cantonese.

One function of Chinese classifiers is to individuate nouns, leading some to propose that all Chinese nouns are inherently mass. Consistent with this, Li et al. (2009a) found that Chinese speakers tended to group novel entities with same-material substances more often than English speakers, who tended to classify them as objects. However, this effect appeared only in linguistic tasks (interpreting names or deixis), not in cognitive tasks (rating objecthood/substancehood). Similarly, Barner et al. (2009) found that Chinese-English bilinguals made material-based generalizations only when tested in Chinese. Further undermining the notion that all Chinese nouns are mass, Li et al. (2009b) found that Mandarin-learning infants had the cognitive capacity to recognize plurality even before acquiring the plurality marker 們 *men*.

Unlike English, Chinese numbers encode a decimal system (e.g. 十二 *shí'èr* 'ten-two' for 'twelve', 二十 *èrshí* 'two-ten' for 'twenty'). Miura et al. (1988) found that this decimal system affected how Chinese children counted with blocks representing tens and ones, as compared with English-speaking children. Geary et al. (1996) replicated these results. However, not only are they consistent with language being used as a tool (including overt counting), but they have minimal practical implications: Wang and Lin (2009) found no difference in computation ability between Chinese and American students, while in other areas of mathematics the Chinese advantage increases over years of schooling, suggesting roles for teaching or learning strategies.

Moving to grammatical morphemes, Chen and Su (2011) studied the effect of gender,

marked in English but not Chinese (他 *tā* traditionally represents both 'he' and 'she', with the character 她 *tā* 'she' a recent invention). English participants identified the biological gender of people in stories, whether by listening or reading, more quickly and accurately than Chinese participants. While this discovery is intriguing (especially given the sharply distinguished gender roles in traditional Chinese culture), note that all pronouns were presented overtly, making it unnecessary to posit Chinese-specific gender-encoding habits independent of Chinese language processing.

Another grammatical difference between English and Chinese is that only the former marks tense. Chen et al. (2012) asked Chinese speakers to judge whether pictures showing one stage of an action matched Chinese sentences marked for prospective (準備要 *zhǔnbèi yào*), progressive (正在 *zhèngzài*), or completive aspect (剛剛 *gānggāng*). Accuracy was higher for participants with greater English proficiency; monolingual English speakers, given the task in English with tense markings, were most accurate. The authors conclude that experience with English tense makes it easier to decompose events into stages. However, the authors acknowledge that participants may have described the pictures to themselves (language as a tool), and that broader cross-cultural differences in encoding event structure (Ji et al. 2004) may have played a role.

Finally, in one of the most notorious applications of the Sapir-Whorf hypothesis in any language, Bloom (1981) claimed that the lack of overt counterfactual marking in Chinese (cf. the English subjunctive) made it difficult for Chinese speakers to interpret if-then structures with false premises. However, Chinese speakers actually process counterfactuals quite reliably (Feng and Yi 2006) and Bloom's effects have repeatedly failed to replicate using more natural materials (Au 1983, Liu 1985, Wu 1994). Culture again seems to play a role: Lardiere (1992) found that speakers of Arabic, which explicitly marks counterfactuals, also balked at Bloom's counterfactuals, perhaps, Lardiere speculates, because of the prominence

of rote learning in both Chinese and Arabic educational traditions.

References

Au, Terry Kit-Fong, "Chinese and English counterfactuals: The Sapir-Whorf hypothesis revisited", *Cognition* 15/1-3, 1983, 155-187.

Barner, David, Shunji Inagaki and Peggy Li, "Language, thought, and real nouns", *Cognition* 111, 2009, 329-344.

Bloom, Alfred H., *The Linguistic Shaping of Thought: A Study in the Impact of Language on Thinking in China and the West*. Lawrence Erlbaum Associate, 1981.

Boroditsky, Lera, "Does language shape thought? Mandarin and English speakers' conceptions of time", *Cognitive Psychology* 43, 2001, 1-22.

Boroditsky, Lera, Orly Fuhrman and Kelly McCormick, "Do English and Mandarin speakers think about time differently?", *Cognition* 118/1, 2011, 123-129.

Brown, Roger, "Linguistic relativity", in: Stewart H. Hulse, ed., *One Hundred Years of Psychological Research in America: G. Stanely Hall and the Johns Hopkins Tradition*, The Johns Hopkins University Press, 1986, 241-276.

Chan, Ting-Ting and Benjamin Bergen, "Writing direction influences spatial cognition", in: Bruno G. Bara, Lawrence W. Barsalou and Monica Bucciarelli, eds., *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, Mahwah, NJ: Lawrence Erlbaum Associates, 2005, 412-417.

Chen, Jenn-Yeu, "Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001) ", *Cognition 104*/2, 2007, 427-436.

Chen, Jenn-Yeu and Padraig G. O'Seaghdha, "Do Mandarin and English speakers think about

time differently? Review of existing evidence and some new data", *Journal of Chinese Linguistics*, Forthcoming.

Chen, Jenn-Yeu and Jui-Ju Su, "Differential sensitivity to the gender of a person by English and Chinese speakers", *Journal of Psycholinguistic Research* 40/3, 2011, 195-203.

Chen, Jenn-Yeu, Jui-Ju Su, Chao-Yang Lee and Padraig G. O'seaghdha, "Linguistically directed attention to the temporal aspect of action events in monolingual English speakers and Chinese-English bilingual speakers with varying English proficiency", *Bilingualism: Language and Cognition*, 15/2, 2012, 413-421.

Feng, Gary, and Li Yi, "What if Chinese had linguistic markers for counterfactual conditionals? Language and thought revisited", *CogSci 2006: Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum Publishers, 2006.

Frank, Harry, O. J. Harvey and Karen Verdun, "American responses to five categories of shame in Chinese culture: A preliminary cross-cultural construct validation", *Personality and Individual Differences* 28/5, 2000, 887-896.

Fuhrman, Orly, Kelly McCormick, Eva Chen, Heidi Jiang, Dingfang Shu, Shuaimei Mao and Lera Boroditsky, "How linguistic and cultural forces shape conceptions of time: English and Mandarin time in 3D", *Cognitive Science*, 35, 2011, 1305-1328.

Gao, Ming Y. and Barbara C. Malt, "Mental representation and cognitive consequences of Chinese individual classifiers", *Language and Cognitive Processes* 24/7, 2009, 1124-1179.

Geary, David C., C. Christine Bow-Thomas, Fan Liu and Robert S. Siegler, "Development of arithmetical competencies in Chinese and American children: Influence of age, language and schooling", *Child Development* 67/5, 1996, 2022-2044.

Gumperz, John J. and Stephen C. Levinson, "Introduction to Part I", in: John J. Gumperz and

Stephen C. Levinson, eds., *Rethinking Linguistic Relativity*, Cambridge, UK: Cambridge University Press, 1996, 21-36.

Hill, Jane H. and Bruce Mannheim, "Language and world view", *Annual Review of Anthropology* 21, 1992, 381-406.

Hockett, Charles F., "Chinese versus English: an exploration of the Whorfian theses", in: Harry Hoijer, ed., *Language in Culture: Conference on the Interrelations of Language and Other Aspects of Culture*, University of Chicago Press, 1954, 106-123.

Hoffman, Curt, Ivy Lau and David R. Johnson, "The linguistic relativity of person cognition: An English-Chinese comparison", *Journal of Personality and Social Psychology* 51/6, 1986, 1097-1105.

Huang, Shuping and Jenn-Yeu Chen, "The effects of numeral classifiers and taxonomic categories in Chinese speakers' recall of nouns", in: Laura Carlson, Christoph Hoelscher and Thomas F. Shipley, eds., *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 2011, 3199-3204.

Huang, Wei, "Reasoning about conventional time as a function of conventional time systems", *Memory and Cognition* 27/6, 1999, 1080-1086.

Huettig, Falk, Jidong Chen, Melissa Bowerman and Asifa Majid, "Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing", *Journal of Cognition and Culture* 10, 2010, 39-58.

Imai, Mutsumi, Henrik Saalbach, and Elsbeth Stern, "Are Chinese and German children taxonomic, thematic, or shape biased? Influence of classifiers and cultural contexts", *Frontiers in Psychology*, 1, 2010, Article 194, 1-10.

January, David and Edward Kako, "Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001) ", *Cognition* 104/2, 2007, 417-426.

Ji, Li-Jun, Zhiyong Zhang, and Richard E. Nisbett, "Is it culture or is it language? Examination of language effects in cross-cultural research on categorization", *Journal of Personality and Social Psychology* 87/1, 2004, 57-65.

Kuo, Jenny Yi-chun and Maria D. Sera, "Classifier effects on human categorization: The role of shape classifiers in Mandarin Chinese", *Journal of East Asian Linguistics* 18/1, 2009, 1-19.

Kwan, Tze-Wan, "Wilhelm Von Humboldt on the Chinese language: Interpretation and reconstruction", *Journal of Chinese Linguistics* 29/2, 2001, 169-242.

Lardiere, Donna, "On the linguistic shaping of thought: Another response to Alfred Bloom", *Language in Society* 21/2 1992, 231-251.

Li, Peggy, Yarrow Dunham and Susan Carey, "Of substance: The nature of language effects on entity construal", *Cognitive Psychology* 58/4, 2009a, 487-524.

Li, Peggy, Tamiko Ogura, David Barner, Shu-Ju Yang and Susan Carey, "Does the conceptual distinction between singular and plural sets depend on language?", *Developmental Psychology* 45/6, 2009b, 1644 -1653.

Liu, Lisa Garbern, "Reasoning counterfactually in Chinese: Are there any obstacles?", *Cognition* 21/3, 1985, 239-270.

Miles, Lynden K., Lucy Tan, Grant D. Noble, Joanne Lumsden and C. Neil Macrae, "Can a mind have two time lines? Exploring space-time mapping in Mandarin and English speakers", *Psychonomic Bulletin and Review*, 18, 2011, 598-604.

Miura, Irene T., Chungsoon C. Kim, Chih-Mei Chang and Yukari Okamoto, "Effects of language characteristics on children's cognitive representation of number: Cross-national comparisons", *Child Development* 59/6, 1988, 1445-1450.

Pinker, Seven, *The Stuff of Thought: Language as a Window into Human Nature*, New York: Penguin Books, 2007.

Ross, Michael, W. Q. Elaine Xun and Anne E. Wilson, "Language and the bicultural self", *Personality and Social Psychology Bulletin* 28, 2002, 1040-1050.

Saalbach, Henrik and Mutsumi Imai, "Scope of linguistic influence: Does a classifier system alter object concepts?" *Journal of Experimental Psychology: General* 136/3, 2007, 485-501.

Sapir, Edward, "The status of linguistics as a science", *Language* 5/4, 1929, 207-214.

Slobin, Dan I., "From 'thought and language' to 'thinking for speaking'", in: John J. Gumperz and Stephen C. Levinson, eds., *Rethinking Linguistic Relativity*, Cambridge: Cambridge University Press, 1996, 70-96.

Tsang, Cara and Craig G. Chambers, "Appearances aren't everything: Shape classifiers and referential processing in Cantonese", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37/5, 2011, 1065-1080.

Wang, Jian and Emily Lin, "A meta-analysis of comparative studies on Chinese and US students' mathematics performance: Implications for mathematics education reform and research", *Educational Research Review* 4/3, 2009, 177-195.

Whorf, Benjamin Lee, "The relation of habitual thought and behavior to language", in: John B. Carroll, ed., *Language, thought and reality,* MA: MIT Press, 1956.

Wu, C. H.-F., "*If Triangles Were Circles,...": A Study of Counterfactuals in Chinese and in English*, Taipei: Crane Publishing Company, 1994.

Zhang, Shi and Bernd Schmitt, "Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations", *Journal of Experimental Psychology: Applied* 4/4, 1998, 375-385.

James Myers

Summary

The Sapir-Whorf hypothesis refers to the claim that features of one's language affects one's thoughts (linguistic determinism), so that different languages foster different conceptual systems (linguistic relativity). This lemma reviews experimental tests of this hypothesis in Chinese.

Index terms

Cantonese

classifiers

count/mass

counterfactuals

gender

linguistic determinism

linguistic relativity

numbers

plurality

time words