

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Do Classifier Categories Structure our Concepts?

Permalink

<https://escholarship.org/uc/item/8r45826p>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN

1069-7977

Authors

Imai, Mutsumi
Saalbach, Henrik

Publication Date

2005

Peer reviewed

Do Classifier Categories Structure our Concepts?

Henrik Saalbach (saalbach@mpib-berlin.mpg.de)

Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

Mutsumi Imai (imai@sfc.keio.ac.jp)

Keio University at Shonan-Fujisawa, 5322 Endo, Kanagawa, Japan

Abstract

Whether and to what extent our conceptual structure is universal is of great importance for our understanding of the nature of human concepts. Two major factors that might affect our concepts are language and culture. In this research, we tested whether these two factors affect our concepts of everyday objects in any significant ways. For this purpose we compare adults of three cultural/language groups—Chinese, Japanese, and German—on similarity judgment and property induction. In particular, we tested whether classifier categories influence the conceptual structure of speakers of classifier languages. Some classifier effect was found, but only for Chinese speakers in similarity judgment. Our overall results indicate that the global structure of our concepts is similar across different culture/language groups.

Keywords: Concepts, Cross-cultural differences, Linguistic relativity, Classifiers, Thematic relations

Introduction

One of the key questions in the literature of human concepts is to what extent they are universal across different cultures and language groups (e.g., Bailenson, Shum, Atran, Medin & Coley, 2002; Berlin, 1992; Imai & Gentner, 1997; Imai & Mazuka, 2003; Rosch, 1978; Malt, 1995). In this research, we tested whether language and culture affect our concepts of everyday objects in any significant ways. For this purpose, we compared adults of three cultural/language groups - Chinese, Japanese, and Germans - on similarity judgment and property induction.

Comparison of Chinese, Japanese and German speakers is particularly interesting for examining the influence of language on our concepts because Chinese and Japanese are classifier languages. Numeral classifiers are somewhat similar to English quantifiers such as *a piece of*, *a portion of*. The important difference between English quantifiers and numeral classifiers is that while the former are only used for quantifying mass nouns (with numerals being used directly with count nouns, e.g. *two cars*), in Chinese and Japanese, numeral classifiers must be applied to all nouns when quantifying them, including clearly individuated objects such as cars, computers, and even humans.

Like nouns, classifiers linguistically categorize entities in the world. However, the lexical organization of classifiers is very different from that of nouns. While the noun lexicon is organized by taxonomic relations, the classifier lexicon is organized around semantic features such as animacy, shape,

dimensionality, size, functionality, and flexibility. Categories made by classifiers often crosscut taxonomic categories, although functional classifiers in part overlap with them. For example, nouns classified with *hon*, a Japanese classifier for long, thin things (and things that are metonymically or metaphorically related to long, thin things), include pens, baseball bats, home-runs, bananas, carrots, ropes, necklaces, wires, and telephone calls (Lakoff, 1987). *Tiao*, a Chinese classifier for long and flexible things, even crosses the animal and non-animal ontological boundary, including fish, dogs, rivers, roads, pants, and more, in the set of the things it classifies. An extremely interesting question is whether classifier categories are an integral part of conceptual structures in speakers of a classifier language.

Zhang and Schmitt (1998) addressed this issue. They tested English speakers and Mandarin-Chinese speakers on a similarity judgment task and found that Chinese speakers in fact rated pairs of objects as more similar than English speakers did when the objects were drawn from the same classifier class. Although Zhang and Schmitt's results may be interpreted as evidence for *a version of* the Sapir-Whorf hypothesis, it is important to note that these results do not tell us whether Chinese speakers have significantly different conceptual structures than English speakers do, as we could not judge whether Chinese speakers in fact organize their concepts around classifier categories. Do Chinese speakers rely on classifier membership more heavily than taxonomic or thematic relations in grouping objects, judging similarity, or making inductive inference of novel properties? If this is indeed the case, we can comfortably conclude that the conceptual structure of Chinese speakers is qualitatively different from that of English speakers, and an endorsement for the strong version of the Sapir-Whorf hypothesis. However, if we find that, rather than giving precedence to classifier class membership, Chinese speakers organize their concepts around taxonomic or thematic relations in the way that native speakers of European languages do, we must qualify the impact of the classifier system on the speakers' conceptual structure.

To explore this issue, in Experiment 1, we tested speakers of Mandarin-Chinese and German on similarity judgment as well as on property inference. The participants were presented with pairs of everyday objects bearing different kinds of relations. The first type of pairs were related taxonomically, and the second type were related thematically. The pairs of the third type were drawn from the same classifier class in Chinese, but were not related taxonomically or thematically (e.g., *fish* and *ropes*) and the fourth type had no relation and served as a control (see table 1). In this

design, not only were we able to test whether the classifier system affects Chinese speakers' judgment of similarity and/or inductive inference of properties, but also whether Chinese speakers rely on classifier membership more strongly than two of the other major relations underlying our concepts.

How Chinese and Germans utilize thematic relations in similarity judgment and property induction is also of great interest in evaluating a proposal that has attracted much attention in the recent literature of cross-cultural cognition. Nisbett and his colleagues have proposed that philosophy, values and customs that have been nursed in a culture throughout its history leads to a "culturally specific" style of cognition (Nisbett, 2003). In his empirical work, Nisbett focused on the comparison between East Asians and Westerners. Characterizing the former as "holistic", and the latter as "analytic," Nisbett argued that while East Asians tend to view the environment as a unified whole and pay much attention to relations that tie elements in the environment, Westerners tend to focus individual elements of the environment separately. Based on this scheme, Nisbett and colleagues have made a specific prediction regarding the conceptual structure of East Asians and Westerners: East Asians, with their predisposition to see a scene or event as a whole, are expected to categorize the world around thematic relations; Westerners, with their focus on properties of individual objects, are expected to categorize the world by taxonomic relations. Ji, Zhang, and Nisbett (2004) in fact reported that monolingual Chinese people showed a preference for "relational" groupings while European Americans tended to group things "categorically". However, Lin and Murphy (1999) demonstrated that even educated European American young adults sometimes show a preference for categorizing objects based on thematic relations over taxonomic relations. It is thus extremely interesting to see how Chinese and German participants in our study utilize taxonomic and thematic relations in similarity judgment and property induction tasks.

Experiment 1

In Experiment 1, we compared native Mandarin-Chinese speakers and native German speakers on similarity judgment and inductive inference of a novel property.

Method

Participants. Thirty-seven Chinese undergraduates from Beijing and 38 German undergraduates from Berlin participated in this study. The Chinese students were all native speakers of Mandarin Chinese, and the German students were all native speakers of German.

Materials. The structures of the stimuli used in this study and in Experiment 2 are shown in Table 1. The stimuli for this experiment were drawn from Set Type 1. In Set Type 1, the stimuli were 12 quintuplets of objects, each consisting of one standard and four targets. The first target type was from the same classifier class as the standard item, but was not related to it either taxonomically or thematically (e.g., flower-cloud: CH_CLS). The second type was taxonomically related to the standard (e.g., flower-tree: TAX), and the third type was thematically related to the standard (e.g., flower-vase: THEME). Neither the TAX nor THEME items belonged to the same classifier class as the standard item. The fourth type served as a control condition, unrelated taxonomically or thematically, and with objects from different classifier classes (CO). Twelve sets of quintuplets were constructed. The stimulus sets were prepared on the basis of the results of a pretest with native Chinese speakers.

The item pairs were presented in a questionnaire. The twelve sets were split into two groups (Item Group A and B). The questionnaire booklet included both a similarity judgment task and a property induction task and was prepared in such a way that Group A sets were used for the similarity judgment and Group B sets were used for property induction in one version of the booklet, and vice versa in the other version. In both versions, however, similarity judgment was presented before property inference because similarity judgement is more susceptible to influence from a previous task.

Participants were randomly assigned to either version of the booklet. Item pairs within each version of the booklet were arranged in random order for both similarity judgment and property inference. The participants judged the similarity of each object pair on a rating scale of 1 (not similar at all) to 7 (very similar). In the property inference task, following Lin and Murphy (1999), we used the property "carry the same bacteria," as most of our items were artifacts. The participants were asked to judge the likelihood that the two objects would carry the same bacteria on a rating scale of 1 (not likely at all) to 7 (very likely).

Table 1. Structure of the stimuli used for Experiments 1 & 2 with a sample set for each set type.

	Standard	CH/JP CLS+Tax	TAX	THEME	CH/JP CLS	CH CLS	JP CLS	CO
SetType1	flower	-	(tree)	(vase)	-	cloud	-	cup
SetType2	bed	table	chair		-	-	-	-
SetType3	bone	-	-	-	tube	-	-	platter
SetType4	bus	-	-	-	-	-	TV	hat

Experiment 1: Set Type1 only (including the items in parenthesis); Experiment 2: Set Type 1-4. However, the items in parenthesis were not used

Procedure. The participants received the booklet in groups. They were instructed to go through the questionnaire carefully at their own pace, and to rely on their intuition.

Results

Results are reported separately for similarity judgement and property induction.

Similarity Judgment. Figure 1 shows the mean similarity ratings for the four target types. The effect of the target type was tested on a repeated measure ANOVA for each culture separately. For both cultures, the effect for Target Type was highly significant, $F(3,108) = 126.70$ for Chinese, and $F(3,111) = 113.17$ for Germans, both $ps < .01$. In both cultures, similarity of the four targets was rated in the order of the TAX, THEME, CH_CLS, and Control pairs (CO). Post-hoc pairwise comparisons were carried out with Bonferroni corrections. For both cultures, the similarity ratings for the four target types were all significantly different from one another, all $ps < .01$.

We then tested whether the magnitude of the effect for each target relation differed across the two cultures. For this test, we obtained difference scores (TAX_diff, THEME_diff, CH-CLS_diff) by subtracting the Control scores (CO) from each of the TAX, THEME, and CH-CLS scores in order to adjust the difference in the baseline across the two cultures, and we used these difference scores as the dependent variables for the analysis (see Table 2).

A 3 (target type) X 2 (culture) repeated measure

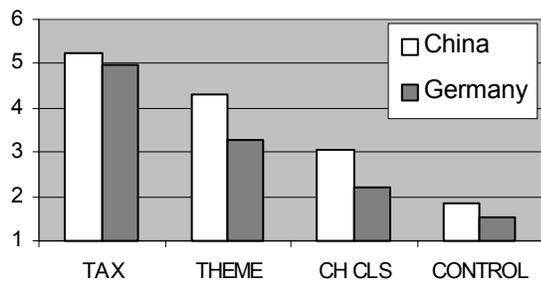


Figure 1: Mean similarity ratings for each target type in each culture in Experiment 1

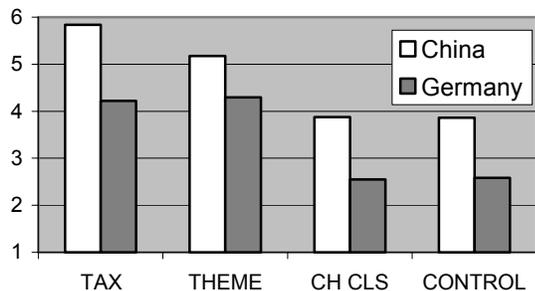


Figure 2: Mean likelihood ratings for property induction for each target type in each culture in Experiment 1

Table 2: Difference scores for each target type in each culture in Experiment 1

Culture	N	TAX_diff		THEME_diff		CH CLS_diff	
		Sim	Prop	Sim	Prop	Sim	Prop
China	37	3.40	1.98	2.45	1.31	1.22	0.01
Germany	38	3.42	1.69	1.72	1.70	0.63	-0.12

Sim: similarity judgment; Prop: property induction

analysis of variance revealed main effects for target type, $F(2,146) = 154.94, p < .01$, and culture, $F(1,73) = 4.49, p < .05$, as well as a significant interaction effect, $F(2,146) = 3.99, p < .05$. Separate ANOVAs for each of the three rating scores revealed a significant cross-cultural difference on CH-CLS_diff, $F(1,73) = 9.59, p < .01$, as well as on THEME_diff, $F(1,73) = 5.17, p < .05$.

Property Induction. Figure 2 shows the mean rating scores for the likelihood of the pairs sharing the same property for each target type. As for similarity judgment, the effect of the target type was tested within each culture. For both cultures, the effect for Target Type was highly significant, $F(3,102) = 58.71$ for Chinese, and $F(3,102) = 51.47$ for Germans, both $ps < .01$. A post hoc test revealed that for the Chinese group, all the pairwise comparisons (Bonferroni corrected) were significant (all $ps < .01$) except for the difference between CH CLS and Control. For the German group, there was no significant difference between Taxonomic and Thematic or between CH CLS and Control. Other comparisons turned to be significant, $ps < .01$.

As we did for similarity judgment, we conducted a 3 (target type) X 2 (culture) on the difference scores (see Table 2). No effect of culture or the interaction involving this factor was found.

Discussion

Several important findings should be noted. First, we found that both taxonomic and thematic relations are important organizers of adults' conceptual structure in two very different language/culture groups. This finding provides strong support for the view that thematic relations, along with taxonomic relations, are an important part of our conceptual structures, even for educated young adults (e.g., Lin & Murphy, 1999).

At the same time, some support for linguistic relativity and for Nisbett's cultural theory was found, as (1) Chinese people's similarity judgment for pairs drawn from the same classifier classes was higher than those of the Germans, and (2) the Chinese gave higher similarity ratings for thematically related object pairs than the Germans did. However, these effects need to be qualified both in terms of the range and magnitude: For both cultures, neither the cultural nor the linguistic effect was large enough to change the relative order of the preference among the four target types. Also, the thematic effect and classifier effect in Chinese people were not found in property inference. In fact, in both cultures, the

rated likelihood of two objects from the same classifier class sharing the same property was no different from that for two unrelated objects. Interestingly, in similarity judgment, not only the Chinese but also the Germans rated the same-classifier pairs higher than the control pairs. This suggests that even speakers of a non-classifier language can detect an inherent similarity between objects belonging to the same classifier class. However, this inherent similarity may be magnified for speakers of the classifier language. Taken together, we can say that whether or not their language has the classifier system, people can detect semantic features underlying the classifier system, but they know that these features are not very useful for inductive inference of a novel property.

In summary, we did find “some” evidence for linguistic relativity and for the cultural difference view advanced by Nisbett (2003). However, these effects are rather local and subtle, as, at a global level, people from two very different linguistic and cultural backgrounds showed strikingly similar performance both in similarity judgment and property induction.

Given that there was some classifier effect for Chinese speakers on similarity judgment, we wished to examine whether the classifier effect would be replicated for a larger set of stimuli and if it would also be found for speakers of another classifier language, Japanese. As in the case of Chinese classifiers, Japanese classifiers are structured around semantic features such as animacy, shape, and functionality, although, unlike Chinese, Japanese classifiers are not sensitive to flexibility. However, there are also clear differences in the grammatical function of classifiers between the two languages. Chinese classifiers must be used not only in numeral phrases (e.g. one [classifier] table) but also in phrases with demonstratives (e.g. this [classifier] table). By contrast, Japanese classifiers are only used with numerals, and are not used as determiners or demonstratives. Given this structural difference between Chinese and Japanese, it is interesting to see whether the classifier effect on similarity judgment is found not only for Chinese speakers but also for Japanese speakers.

Experiment 2

In Experiment 2, speakers of two classifier languages, Chinese and Japanese, and speakers of a non-classifier language, German, were tested on similarity judgment and property inference. In this study, we focused on the influence of classifiers. Thus, we designed the stimuli in such a way that the classifier effect could be examined more finely than in Experiment 1. Specifically, we tested the classifier effect in three situations: (1) object pairs belonging to the same classifier class both in Chinese and Japanese; (2) object pairs belonging to the same classifier class in Chinese but not in Japanese; (3) objects pairs belonging to the same classifier class in Japanese but not in Chinese. We also tested whether we could see the classifier effect even when two objects are drawn from the same taxonomic category. For this purpose, we contrasted pairs in which two objects share both

taxonomic category membership and classifier membership in both Chinese and Japanese (e.g., bed and table) to pairs in which the two objects only shared taxonomic category membership (e.g., bed and chair).

Method

Participants. Thirty-nine Chinese undergraduates from Beijing, 35 German undergraduates from Berlin, and 40 Japanese undergraduates from the Tokyo area participated in this study.

Material and Procedure. Stimulus set types 1- 4 in Table 1 were used but the items in parenthesis were not included for this experiment. Type 2 sets contrasted pairs in which the two objects shared both taxonomic category membership and classifier class membership (e.g., bed and table: CH/JP_CLS+TAX) with pairs in which the two objects belonged to the same taxonomic category but to different classifier classes. (e.g., bed and chair: TAX). Type 3 sets were used to contrast the object pairs from the same classifier class in both Chinese and Japanese (e.g., bone and tube: CH/JP_CLS) with the pairs sharing no relation (e.g., bone and platter: Control). Type 1 sets, which were also used for Experiment 1, contrasted object pairs sharing classifier membership in Chinese only (CH_CLS) with pairs sharing no relation. Type 4 sets were used to contrast object pairs from the same classifier class in Japanese but not in Chinese (JP_CL) with pairs sharing no relation. As in Experiment 1, the standards and the same classifier targets were verified by a pretest.

As in Experiment 1, each participant received a questionnaire containing both similarity judgment and property inference. As in Experiment 1, “carrying the same bacteria” was used for the property inference. The structure as well as the format of the questionnaire was the same as that used in Experiment 1: Item pairs of all set types were divided into two groups (A and B) within each set type and two versions of the questionnaire were prepared, one using Group A pairs for similarity judgment and Group B pairs for property inference, the other using Group B pairs for similarity judgment and Group A pairs for property inference. The procedure of Experiment 2 was identical to that of Experiment 1.

Results

Similarity Judgment. Figure 3 shows the average rating score for each of the six target types in each culture. We first examined the classifier effects within each culture. The contrast between the pairs sharing both taxonomic category membership and classifier membership (in both Chinese and Japanese) and the pairs sharing the taxonomic category membership only revealed an advantage of sharing classifier membership for all three cultures, all $ps < .01$. The pairs sharing classifier membership in Chinese and Japanese were rated higher than the controls in all cultures as well, all $ps < .01$. The pairs sharing classifier membership only in Chinese were significantly different from the corresponding control pairs in Chinese, $p < .01$, and German, $p < .05$, but not

in Japanese, $p > .1$. The pairs sharing classifier membership only in Japanese were rated more similar than the corresponding control pairs in all three cultures, all $ps < .01$.

We then tested whether the magnitude of the classifier effect for each type of contrast differed across the three cultures. For this test, as in Experiment 1, we used the difference scores to adjust the difference in the scores for the control pairs across three cultures. Again, we obtained the difference scores by subtracting the rating scores for the corresponding control pairs from each target. For example, CH/JP_CLS+TAX difference score was obtained by calculating CH/JP_CLS+TAX – TAX in Item Set Type 2, and JP_CLS difference score was obtained by subtracting the mean score of the control pairs (in Item Set Type 4) from the mean score of JP_CLS pairs. Difference scores across each type of contrast in each culture are shown in Table 3.

A cross-cultural difference was found for the pairs sharing classifier membership in both Chinese and Japanese (CH/JP CLS difference score) and for the pairs sharing classifier membership only in Chinese (CH-CLS difference score), $F(2,111)=6.86$, and $F(2,111)=5.09$, respectively, both $ps < .01$. For the former, post-hoc pairwise comparisons (Tukey) revealed a significant difference between Chinese and German, and Chinese and Japanese, both $p < .01$. For the CH_CLS difference score, the difference between Chinese and German was marginally significant $p < .08$, and the difference between Chinese and Japanese was significant, $p < .01$. No cross-cultural difference was found on the

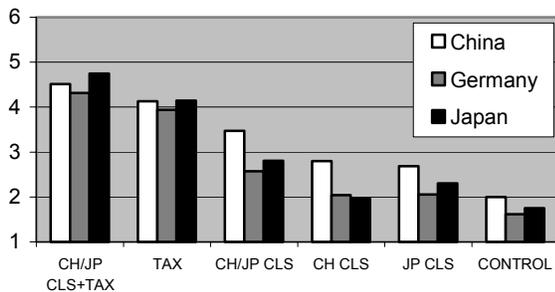


Figure 3: Mean similarity ratings for each target type in each culture in Experiment 2

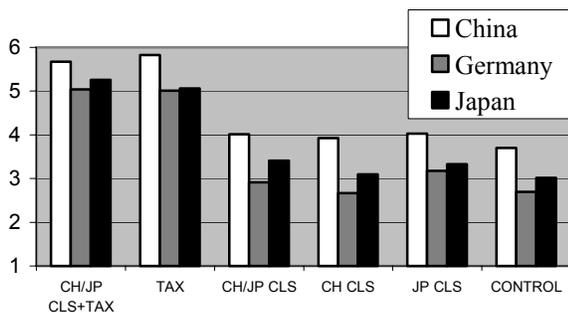


Figure 4: Mean likelihood ratings for each target type in each culture in Experiment 2

Table 3: Difference scores for each classifier contrast in Experiment 2

	CH/JP CLS+		CH/JP		CH		JP	
	TAX_diff		CLS_diff		CLS_diff		CLS_diff	
	Sim	Prop	Sim	Prop	Sim	Prop	Sim	Prop
CH	0.38	0.16	1.57	0.34	0.92	0.27	0.45	0.26
GR	0.38	0.03	0.83	0.18	0.59	0.18	0.37	0.03
JP	0.60	0.19	0.91	0.36	0.48	0.21	0.40	0.19

Sim: similarity judgment; Prop: property induction

CH/JP_CLS+TAX difference score or on the JP_CLS difference score.

Property Induction. Figure 4 shows the average rating score for the property induction task for each target type in each culture. Again, the classifier effects were tested within each culture by paired t-tests testing the difference between each of the classifier targets against the corresponding controls. Unlike the case with similarity judgment, the classifier effect was only found for the pairs in which the target object was from the same classifier class as the standard in both Chinese and Japanese (CH/JP CLS). In all three cultures, the likelihood score for this target type was higher than the control, all $ps < .01$. In the other three types of contrasts (i.e., CH/JP_CLS+TAX-TAX, CH_CLS – Control, and JP_CLS – Control), there was no advantage for shared classifier class membership in any of the three cultures.

We again tested the effect of culture for each of the four difference scores given in Table 3. Unlike the case with similarity judgement, in none of the difference scores was there an effect for culture, all $ps > .05$.

Discussion

The classifier effect found in the similarity judgment task in Experiment 1 for Chinese speakers was replicated with a larger stimuli set. In this study, the Chinese participants gave higher similarity rating scores for the objects from the same classifier class in Chinese than German and Japanese speakers did, although this classifier effect was not found when the two objects were taxonomically related to start with. These results again suggest that taxonomic relations are a primary factor organizing our concepts, and within this constraint, the classifier system influences our construal of similarity among objects. We also replicated the finding from Experiment 1 that classifier categories are not utilized in inference of novel properties.

Another important finding from Experiment 2 is the lack of the classifier effect in Japanese speakers. As stated earlier, the link between objects and corresponding classifiers may be weaker for Japanese than for Chinese speakers because classifiers are associated with objects only when objects are enumerated in Japanese, while in Chinese, classifiers also have the function analogous to English determiners. This suggests that, in considering the influence of language, how often (or how habitually) a given linguistic category is used in

the language is important in addition to whether the linguistic categories are present in a given language.

General discussion

This research examined whether classifier categories affect our conceptual structures in significant ways. In a way, we tested linguistic relativity in two forms, to see (1) whether classifier categories function as organizers of our concepts and categories (cf. Lakoff, 1987); (2) whether speakers of classifier languages attend to features underlying classifier categorization more strongly than speakers of non-classifier languages in certain cognitive contexts (cf. Hunt & Agnoli, 1991). We found some support for linguistic relativity in (2) but not in (1). In both Experiments 1 and 2, Chinese people showed higher similarity ratings for objects from the same classifier class in their language than German people did, which supports linguistic relativity. At the same time, this effect must be qualified because the magnitude of the effect as well as the cognitive domain it applies to was limited. First, the classifier effect was found in similarity judgment but not in property inference. Second, more importantly, speakers of classifier and non-classifier languages showed strikingly similar performance in both similarity judgment and property induction at a global level: In Experiment 1, both Chinese and Germans gave the highest ratings for objects holding the taxonomic relation, then to objects that were thematically related, then to objects from the same classifier class in Chinese. It is also important to note that even Germans could see some similarity in object pairs from the same classifier class, presumably because these objects shared a semantic feature such as shape, size, or function. In Experiment 2, the classifier effect on similarity judgment was replicated for Chinese people, but it was not found in Japanese speakers. This suggests that, in addition to the presence of certain linguistic categories, how frequently the linguistic categories are used in the language is important when we explore issues in linguistic relativity. It is possible, however, if we employ a paradigm that assesses more automatic, lower levels of cognition (e.g., a priming task), we may find the classifier effect even for Japanese speakers. Further examination is also required whether the classifier effect is found in property induction if we employed a different task, e.g., asking the likelihood of sharing a blank property.

The results of Experiment 1 in this research also speak to two important proposals in the literature of concepts and cross-cultural cognition. The fact that Chinese people gave higher ratings for thematically related objects than Germans is consistent with Nisbett's proposal (2003; Ji et al., 2004) that East Asians weigh thematic relations more heavily than Westerners. At the same time, however, this finding also needs to be qualified in that Chinese people relied on taxonomic relations more heavily than thematic relations in both similarity judgment and property induction just like Germans. Thus, we found little evidence that Chinese people's concepts and categories were organized differently from Germans. Our results provide support for recent views that thematic relations are an important and integral part of concepts not only in children but also in adults (Lin & Murphy, 1999).

To conclude, the results of the two experiments suggest that the overall structure of our concepts is similar across different cultures and language groups, because the structure of the world and objects place strong constraints on how we perceive them and relations among them (Malt, 1995). At the same time, language and/or cultural traditions and customs affect our concepts and cognition, but the role of these factors seems to be secondary rather than primary in our conceptual structure (see also Imai & Mazuka, 2003).

Acknowledgement

This research was supported by Japan Society of the Promotion of Science (JSPS). We are deeply indebted to Zhou Xiaolin, Elsbeth Stern, Sachiko Saitoh, Zhu Liqi, and Huangfu Gang for help for data collection and discussion.

References

- Bailenson, J., Shum, M., Atran, C., Medin, D. & Coley, J. (2002). A bird's eye view: biological categorization and reasoning within and across cultures. *Cognition*, 84, 1-53.
- Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.
- Hunt, E & Agnoli, F. (1991). The Whorfian hypothesis: a cognitive psychology perspective. *Psychological Review* 98(3): 377-389
- Imai, M. & Gentner, D. (1997). A crosslinguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, 62, 169-200.
- Imai, M. & Mazuka, R. (2003). Re-evaluation of linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind: Advances in the issues of language and thought*. Cambridge, MA: MIT Press.
- Ji, L.J., Zhang, Z., & Nisbett, R.E. (2004) 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), 57-65.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lin, E. L. & Murphy, G. L. (1999). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, 130, 3-28.
- Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology*, 29, 85-148.
- Nisbett, R. E. (2003). *The geography of thought: how Asians and Westerners think differently and why*. New York: Free Press.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B.B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum.
- Zhang, S. and Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, 4, 375-385.