This article was downloaded by: [Keio University]
On: 26 July 2011, At: 05:51
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/plcp20

The relation between linguistic categories and cognition: The case of numeral classifiers

Henrik Saalbach [a] & Mutsumi Imai [b]
[a] Institute for Behavioral Sciences, ETH Zurich, Zurich, Switzerland
[b] Department of Environmental Information, Keio University at Shonan-Fujisawa, Endo, Kanagawa, Japan

Available online: 26 Jul 2011

To cite this article: Henrik Saalbach & Mutsumi Imai (2011): The relation between linguistic categories and cognition: The case of numeral classifiers, Language and Cognitive Processes, DOI:10.1080/01690965.2010.546585

To link to this article: http://dx.doi.org/10.1080/01690965.2010.546585

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

LANGUAGE AND COGNITIVE PROCESSES
0000, 00 (00), 1–48

Psychology Press
Taylor & Francis Group

The relation between linguistic categories and cognition: The case of numeral classifiers

Henrik Saalbach[1] and Mutsumi Imai[2]

[1]Institute for Behavioral Sciences, ETH Zurich, Zurich, Switzerland
[2]Department of Environmental Information, Keio University at Shonan-Fujisawa, Endo, Kanagawa, Japan

Downloaded by [Keio University] at 05:51 26 July 2011

The classifier grammar system categorises things in the world in a way that is drastically different from the way nouns do. Previous research revealed amplified similarity among objects belonging to the same classifier category in Chinese speakers, but how this amplified classifier similarity effect arises was still an open question. The present research was conducted to address this question. For this purpose, we compared speakers of Chinese, Japanese (classifier languages), and German (nonclassifier language) on a range of cognitive tasks including similarity judgements, property induction, and fast-speed word-picture matching. Although Chinese and Japanese classifier systems are similar in their semantic structures, classifier classes for nouns are marked more systematically in Chinese than in Japanese. The amplified classifier similarity effect was found in Chinese but not in Japanese speakers. We explore the nature of the amplified classifier similarity effect and propose an explanation for how it may arise.

Correspondence should be addressed to Henrik Saalbach, Institute for Behavioral Sciences, ETH Zurich, UNO C14 8092 Zurich, Switzerland. E-mail: henrik.saalbach@ifv.gess.ethz.ch

This research was supported by fellowships from the Japan Society of the Promotion of Science (JSPS) and the Friedrich-Ebert-Stiftung awarded to Saalbach and a Ministry of Education grant-in-aid for Scientific Research awarded to Imai (#15300088 and # 18300089), as well as Keio Gijuku Academic Development Funds and Keio Gijuku Mori Memorial Research Fund awarded to Imai. We are deeply indebted to Zhou Xiaolin, Elsbeth Stern, Sachiko Saitoh, Lennart Schalk, Shu Hua, Zhu Liqi, and Huangfu Gang for their help with data collection and discussion. We also thank Hiroki Okada for writing the program to conduct Experiments 4 and 5, Miriam Leuchter for her advices and help concerning data analysis, and Hella Beister for their thoughtful comments on earlier versions of this manuscript.

The two authors equally contributed to the paper and the order of the authorship is arbitrary.

© 2011 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
http://www.psypress.com/lcp DOI: 10.1080/01690965.2010.546585

Keywords: Concepts; Linguistic relativity; Classifiers; Similarity; Inductive reasoning.

The relation between language and thought has been one of the central issues in the cognitive sciences. Many researchers have asked whether there is a relation between linguistic categories and people's concepts and cognitive processes (Boroditsky, Fuhrman, & McCormick, 2011; Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996; Hunt & Agnoli, 1991; Li & Gleitman, 2002; Lucy, 1992; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Malt & Wolff, 2010; Regier & Kay, 2009; Roberson, Davidoff, Davies, & Shapiro, 2005). In this paper, we explore whether categorisation systems of numeral classifiers affect the structure of object concepts in speakers of classifier languages.

The grammar system of classifiers[1] differs from the count-mass grammar system or the gender grammar system in that they usually classify nouns into over 100 classes according to their inherent semantic features. Numeral classifiers serve to provide the noun with a counting unit just as quantifiers in languages with a count-mass grammar do. However, unlike in English, where only mass nouns require a unit of quantification, classifier languages require numeral classifiers for all nouns when quantifying them, including what seem to be clearly individuated objects (such as "chair").

Importantly, categories created by classifiers largely crosscut categories created by nouns. In particular, while the noun lexicon is structured hierarchically around taxonomic relations, the system of classification by classifiers is usually organised around semantic features such as animacy, shape, function, size, rigidity, and social importance (e.g., Adams & Conklin, 1973; Allan, 1977; Craig, 1986; Croft, 1994; Denny, 1979; Downing, 1996; Gomez-Imbert, 1996; Senft, 1996) and does not have hierarchical structures as the noun lexicon does. Take the Chinese system of numeral classifiers, for example: Chinese classifiers are subdivided by a single semantic feature or a combination of a few semantic features, including animacy, shape, dimensionality, size, and rigidity. For example, *tou* is a classifier typically used for big animals like cows, elephants, rhinos etc., and *zhi* is typically used for small animals like cats and rabbits. *Zhang*, a classifier for flat things, includes tables, beds, blankets, paper, faces, and tickets.

Given that a classifier system divides the world in ways very different from those of taxonomic categories, the issue of whether or not there is a relation between a classifier system and speakers' conceptual representation and

[1] Although there are different types of classifier languages (e.g., Aikhenvald, 2000; Allan, 1977; Grinevald, 2000; Senft, 2000), our focus in this paper is on numeral classifier systems only. Thus, the terms "classifiers" and "numeral classifiers" will be used synonymously in the present article.

Downloaded by [Keio University] at 05:51 26 July 2011

Downloaded by [Keio University] at 05:51 26 July 2011

cognitive processes is indeed worth clarifying. Some researchers assumed that classifier categories reflect speakers' conceptual representation (Lakoff, 1987). However, recent empirical research has shown that linguistic categories do not necessarily reflect nonlinguistic conceptual categories (e.g., Kay & Regier, 2006; Malt, Sloman, Gennari, Shi, & Wang, 1999; Malt et al., 2008). Thus, it is worthwhile to empirically investigate whether there is any systematic relation between classifier categories and speakers' conceptual categories.

Provided that there is a relation between the classifier system and speakers' representation of objects, there are at least two different possibilities as to how this is achieved. One possibility—a strong version—is that this relation is strong enough to provide a way of organising object concepts in addition to taxonomic or thematic relations, resulting in categories that are nonexistent in speakers of nonclassifier languages. If this is the case, we may be able to conclude that classifier categories and speakers' conceptual representations are strongly related and that in the presence of objects from the same classifier category, speakers of a classifier language and speakers of a nonclassifier language can be expected to behave very differently in almost all cognitive activities, including category formation, memory, similarity judgement, and, most importantly, inductive reasoning. Furthermore, if the classifier system provides a way of organising object concepts, we may expect classifier relations to be automatically accessed in online processing just like taxonomic relations.

A second possibility—a weak version—is that speakers of a classifier language do not actually "organise" their concepts by classifier category membership. In other words, speakers do not consider that classifier category membership implies deep conceptual similarity that can be useful for inductive inference, as taxonomic category membership does. Rather, the experience of using classifiers to linguistically categorise objects may heighten their awareness of the similarity underlying classifier categories and, as a consequence, their sensitivity to this similarity without necessarily inducing them to infer a deeper similarity, as is the case with taxonomic relations.

There is only a handful of studies that have examined the relation between a classifier system and speakers' object concepts (e.g., Gao & Malt, 2009; Huettig, Chen, Bowerman, & Majid, 2010; Zhang & Schmitt, 1998). Zhang and Schmitt (1998), for example, found that Chinese speakers rated pairs of every-day objects belonging to the same classifier category as more similar than native English speakers did, whereas ratings of the different-classifier pairs did not differ between the two groups. On the basis of these results Zhang and Schmitt concluded that classifier categories strongly affect conceptual categories and, thus, argued for a strong version of the classifier effect.

Downloaded by [Keio University] at 05:51 26 July 2011

The results of other studies on classifiers tend to be more consistent with the weak version. For example, Gao and Malt (2009) found effects of classifier categories on the storage and retrieval of information from memory. Chinese speakers, after memorising nouns presented (either with or without a classifier) in random order, reproduced these items in a clustered structure which resembled classifier categories. This effect, however, was limited to so-called *well-defined* classifier categories—classifier categories where one or more features fully define the scope of the category—and to those cases in which the item was presented with the corresponding classifier within a sentence. Similarly, in a very recent eye-movement study, Huettig and colleagues (Huettig et al., 2010) tested whether classifier relations are reflected by patterns of eye movements. They tracked Mandarin speakers' eye-gaze when they were hearing a sentence containing a classifier-noun phrase or a noun phrase without a classifier. Participants were presented with four different objects, including one object from the same classifier category as the noun in the sentence, each object being accompanied by an auditory stimulus. Huettig and colleagues found that participants' eye-gaze indeed shifted to the same-classifier object when the noun was presented with the classifier, but failed to do so when the target noun was presented without a classifier.

Saalbach and Imai (2007) provided further evidence for a weak version of the classifier effect in Chinese speakers. Importantly, they also specified the magnitude of the classifier relations and the scope of the classifier effect. They specifically examined the relation between the classifier system and cognition by conducting a range of cognitive tasks including categorisation, similarity judgements, inductive inference of novel properties as well as fast-speed word-picture matching with Chinese and German speakers. Stimuli were designed to allow for an assessment of the relative importance of the classifier relation (membership of the same classifier category, e.g., flower-cloud) as compared with two other major conceptual relations, i.e., taxonomic relations (e.g., flower-tree) and thematic relations (e.g., flower-vase).

Chinese and German speakers both rated two objects that belonged to the same classifier category in Chinese (but were not taxonomically or thematically related) as more similar than two objects that did not share the classifier relation, but in Chinese speakers, the magnitude of this effect was larger. Thus, a language-specific classifier effect was indeed found. However, the effects of classifier relations were much less prevalent than those of taxonomic and thematic relations in these Chinese speakers. Also, the scope of this language-specific classifier effect was limited to the similarity judgement and inductive reasoning tasks where little prior knowledge except for the conceptual relation between the target and the test objects could be recruited for making the inference; when asked to draw an inductive inference about carriers of bacteria, i.e., a case that allowed for causal reasoning based on background knowledge, Chinese speakers did not rely on the classifier relation any more strongly than German speakers did. More importantly, no classifier effect was found in speeded word-picture matching in Chinese speakers, which suggests that classifier relations are not automatically activated when a word is accessed, in sharp contrast to taxonomic and thematic relations. On the basis of these results, Saalbach and Imai (2007) concluded that the classifier system in Chinese amplifies the inherent similarity underlying the semantics of classifier categories.

Previous research (Gao & Malt, 2009; Huettig et al., 2010; Saalbach & Imai, 2007; Zhang & Schmitt, 1998) has consistently demonstrated that Chinese speakers' object concepts were indeed to some degree affected by the classifier system of their language. However, several issues remained unsolved. First of all, can the classifier effect be generalised to other classifier language? To date, all the studies showing the classifier effect tested speakers of Mandarin Chinese. We thus do not know whether the same classifier effect is observed in other languages with a classifier system. This question leads to another important question, that is, the conditions under which the magnified classifier similarity effect arises. Is the mere presence of a classifier system in a language sufficient to produce such an effect? If a notable cognitive effect due to a given grammatical classification system is identified in one language, it is tempting to generalise the effect to other languages with a similar grammatical system (e.g., Imai & Gentner, 1997; Lucy, 1992; Mazuka & Friedman, 2000). However, this assumption may itself be worth testing.

In fact, previous research suggests that the influence of gender classification systems on speakers' object concepts varies across different languages with grammatical gender systems. In particular, Vigliocco and colleagues (Vigliocco, Vinson, Paganelli, & Dworzynski, 2005) tested whether randomly chosen pairs of object nouns belonging to the same gender class were rated as more similar than pairs of object nouns from different grammatical gender classes. They found the same-gender similarity effect for Italian speakers but not for German speakers. They attributed this to the fact that German has three gender classes rather than two, reasoning that the semantic mapping between sex and grammatical gender is less transparent in German than in Italian.

In the present research, we thus examined whether the magnitude and the scope of the classifier effect differs between Chinese speakers and Japanese speakers. Japanese is also a classifier language, but there are substantial differences between the classifier systems in Chinese and Japanese, as will be described below. Clarifying whether the same classifier similarity effect is

Downloaded by [Keio University] at 05:51 26 July 2011

observed in Japanese speakers will provide an insight into the condition under which the classifier similarity effect arises.

Differences between the Chinese and Japanese classifier systems

In which respect does the Japanese classifier system differ from the Chinese system? Here, both semantic and syntactic aspects should be considered. It is extremely difficult to determine how similar (or how different) the *overall semantic structure* of the Chinese and the Japanese classifier systems is. In many ways, Japanese and Chinese classifier systems are similar. As stated earlier, classifier classes have conceptual meanings in that they are defined by concrete semantic features of the object referent, such as animacy, shape, function, etc. (cf. Croft, 1990). Thus, objects belonging to the same classifier category share one (or sometimes more than one) of these features. For example, members of the *hon* category in Japanese and the *tiao* category in Chinese share the semantic feature of "having a one-dimensionally extended shape (i.e., long-thin shape)", and similarity among members of these classifier categories was detected even by speakers of a nonclassifier language (Saalbach & Imai, 2007). In both languages, there is a large within-language variability with respect to category size and coherence (e.g., Denny, 1979; Downing, 1996; Erbaugh, 1986; Matsumoto, 1985). Some categories, shape classifiers in particular, cover a very broad range of things and show low category coherence since they include members from many different taxonomic categories and many members whose semantic basis for membership is not very transparent. For example, the range of objects covered by *tiao* is very broad, including both animals and inanimate objects (e.g., road, jumping rope, river, snake, or fish) as well as abstract nouns without any visible physical features (e.g., life). *Hon*, is also very broad and includes a rather wide range of concrete things from many different taxonomic categories (e.g., road, jumping rope, river, pencils, bats, and carrot) as well as abstract nouns and metonymically extended concepts (a number of matches in sports, messages, and phone calls). Both languages, on the other hand, have classifiers that are only used for very specific kinds. For example, the Chinese classifier *shou* is used to count poems, songs, and music. Japanese has a similar counter-part, *shu*, which is used to count poems and songs, but not music (because there is a specific classifier for counting pieces of music in Japanese). The category sizes of these classifiers are small, and category coherence is high.

One marked semantic difference between Chinese and Japanese concerns the treatment of animals. Japanese classifiers never cross the boundary between animals and nonanimals. Even the generic classifier *ko* cannot be used for counting live animals. In Chinese, in contrast, even though it has

Downloaded by [Keio University] at 05:51 26 July 2011

Downloaded by [Keio University] at 05:51 26 July 2011

animal classifiers, the boundary between animals and nonanimals is not as rigid as in Japanese, and several frequently used classifiers (e.g., *tiao*, *ge*) include both animal and nonanimal members. Concerning the conceptual distinction between animals and nonanimals, we might thus say that the semantic structure of the Japanese classifier system is clearer than its Chinese counterpart. However, it is not possible to determine whether classifier categories in one language are *in general* more compact and coherent than classifier categories in another language, nor do different approaches to the animal vs. nonanimal distinction alone warrant the conclusion that the Japanese classifier system is semantically more clearly structured than the Chinese system.

It is beyond the scope of the present work to pin down exactly how the diversity of classifiers with respect to category size and coherence interacts with the classifier effect. This will have to be explored by future research. For the purpose of the present research, it seems legitimate to say that the semantic structures of the Chinese and the Japanese classifier systems are *comparable*, with many semantic properties in common at a global level.

In contrast to the similarity between the Chinese and Japanese classifier systems with respect to various semantic properties, the two languages do differ with respect to the grammatical role of classifiers and, consequently, their function in discourse. For example, Chinese classifiers must be used not only in numeral phrases (e.g., [numeral + classifier] table) but also in phrases with demonstratives (e.g., this [numeral + classifier] table) (the numeral after the demonstrative is often dropped, however). In contrast, Japanese classifiers are only used with numerals and are not used in constructions with demonstratives. Furthermore, in Chinese, a classifier functions as a rough equivalent to an indefinite article, while in Japanese, classifiers are only used when it is pragmatically important to specify the number of things in discourse. For example, as an equivalent to the English phrase: "I have a cat", Chinese speakers are most likely to say: "wo (I) yang (raise) yi (one) zhi (classifier for small animals) mao (cat)". In contrast, Japanese speakers are most likely to say: "watashi (I) wa (topic-marking particle) neko (cat) wo (Accusative) katte (raise)-iru (state)". Here, the information "one" is not verbalised unless this information is pragmatically important, e.g., when saying: "I have only one cat, but not two", in response to the question: "Do you have two cats?"

These inter-language differences should result in a much higher frequency of classifier use in Chinese than in Japanese. In other words, the amount of exposure to the association between a noun and its classifier category may be significantly larger in Chinese speakers than in Japanese speakers.

To confirm this intuition, we compared texts that have been translated into Chinese and/or Japanese, using two sources. In the first source, Lamarre

(2008) compared the frequency of the classifier construction in Chinese and Japanese, respectively, using the translation of the same original text (Chapter 4 of *Harry Potter and the Chamber of Secrets,* Rowling, 1999). She reported that the frequency of the classifier construction in the Chinese translation was four times as high as in the Japanese translation (82 vs. 19 tokens).

In the second source, we compared the frequency of classifier use in a Japanese novel and in its Chinese translation, using the Chinese-Japanese parallel corpus (Beijing Center for Japanese Studies, 2003). The material we chose was the novel *Bocchan* (Master Daring), by Soseki Natsume (1964). In the original Japanese text, there were 111 classifier counts.[2] In the Chinese translation, there were 405 counts. Thus, 294 classifier tokens were added in the course of translating the original Japanese text to Chinese. On closer examination, there were 58 cases in which a classifier was used with "one" ("ichi") in the Japanese original. In the Chinese translation, there were 156 cases of "one" ("yi") with a classifier construction. When the number was "two" or "three", there were 21 classifier counts in Japanese and 53 in Chinese. In the Chinese translation, classifier counts were 175 in the "demonstrative + classifier + noun" construction (e.g., "Zhe (this) zhang (classifier) weirenzhuang (document)"). However, in the original Japanese text, all these cases were simple "demonstrative + noun" constructions without a classifier. This survey thus revealed that classifiers are used roughly four times as often in Chinese as in Japanese, which is consistent with our structural analysis of the Chinese and Japanese classifier systems.

Given this difference between Chinese and Japanese classifier systems, it is important to see whether the amplified classifier similarity effect found in Chinese speakers in our previous research (Saalbach & Imai, 2007) is also found in Japanese speakers. This in turn will provide important insights into the situations under which the (language-specific) amplified classifier effect may arise. If the effect was found to be weaker in Japanese than in Chinese speakers, this would suggest that the amplified classifier similarity effect arises as a function of the strength of the association, i.e., of how frequently

Downloaded by [Keio University] at 05:51 26 July 2011

[2]The number of classifier counts does not reflect an exhaustive count of all classifier uses. Because the Japanese-Chinese parallel corpus was not tagged, we had to conduct a manual search for cases of classifier uses. We thus started out from the Chinese translation, as we expected to find a much larger number of classifier counts in Chinese than in Japanese. It was not possible to search for all possible classifier types, as there are over 100 classifiers. We thus searched for the classifiers used in this research, as they are frequent and representative classifiers used with the numerals "one" "two" "three" and the demonstratives *zhe* (this) and *na* (that). We then examined the corresponding part of the original Japanese text to see if, in the original, a classifier was used in the same cases. Thus, classifier uses with numerals other than "one", "two", and "three" as well as uses with classifiers that were not included among the stimuli in this research were not included in the count.

nouns occur with classifiers. On the other hand, semantics of classifier categories may reflect Japanese (as well as Chinese) speakers' conceptual representation, as Lakoff (1987) argued. If this is the case, the amplified classifier effect should be observed not only in Chinese but also in Japanese speakers.

OVERVIEW OF THE STUDIES

We conducted five experiments using similarity judgements, two versions of an inductive reasoning task (induction of a blank property and induction of a known property), and two versions of fast-speed word-picture matching tasks with Chinese, Japanese, and German speakers (see the section *Tasks and predictions* below). As in Saalbach and Imai (2007), speakers of German, a nonclassifier language, served as a control group.

As noted above, although the Japanese and the Chinese classifier systems both rely on largely overlapping semantic features, the category members included in the corresponding categories in each of these languages only partially overlap. For example, *mai* in Japanese and *zhang* in Chinese are both associated with flat things. However, while "table" is included in the *zhang* category in Chinese, presumably because of the flatness of the surface, *mai* is not used for "table" in Japanese; instead, "table" is usually counted with *dai,* the classifier for machines and other large functional objects. As a result, it is possible to select pairs of objects that belong to the same classifier category only in Japanese or only in Chinese, and pairs of objects that belong to the same classifier category in both Japanese and Chinese. We thus designed the stimulus material so as to have object pairs represent one of the four types of classifier relations: (1) nouns belonging to the same classifier category only in Chinese; (2) nouns belonging to the same classifier category only in Japanese; (3) nouns belonging to the same classifier in both Chinese and Japanese; (4) nouns belonging to the same classifier category in both Chinese and Japanese and also belonging to the same taxonomic category. The third type of pairs allowed us to test the classifier effect in Japanese and Chinese speakers using the very same objects. The object pairs of this type were particularly important because they provided an opportunity to evaluate the potential amplified similarity effect in Chinese speakers: Even if awareness of similarity was found to be stronger in Chinese speakers than in Japanese speakers with respect to object pairs of the first type (belonging to the same classifier category in Chinese but not in Japanese), it would be difficult to rule out the possibility that this difference was due to some idiosyncratic properties of the stimuli. However, if with respect to the same object pairs the amplified similarity effect due to classifier category membership was weaker in Japanese speakers than in Chinese speakers,

Downloaded by [Keio University] at 05:51 26 July 2011

the difference in the classifier effect could be attributed to the structural difference between Chinese and Japanese. The fourth type of pairs was included to see whether sharing the same classifier category provided an additional similarity over and above taxonomic relations.

As shown in Table 1, one object served as the target in each type of classifier relations. For each target item of Type 1–3 pairs (same classifier in Chinese only, same classifier in Japanese only, and same classifier in both languages), a control item was selected to serve as a baseline for the potential classifier effect. The control item had no semantic relation to the target. In other words, it was neither taxonomically nor thematically related to the target, nor did it belong to the same classifier category. For each target item of Type 4 pairs (same classifier in both languages and taxonomic relation), a taxonomic item was selected to serve as a baseline for the classifier effect within this type. The taxonomic item was an object that was taxonomically related to the target but did not share the same classifier.

Tasks and predictions

In the similarity judgement task, Chinese, Japanese, and German speakers were presented with two objects which shared one of the four types of classifier relations, or two objects which were members of different classifier categories. Participants were asked to judge the similarity between the two objects on a rating scale of 1 (very dissimilar) to 7 (very similar). In the inductive reasoning task with an unknown property, speakers from the three language groups were presented with the same stimuli as in the similarity judgement task, and asked to rate the likelihood for the two objects in each

TABLE 1
Structure of the stimuli used for Experiments 1–3 with a sample set for each stimulus type

Type of same_classifier pair	*Target item*	*Same-classifier item*	*Control item*	*Taxonomic item*
Type 1 (C-cls)	Flower	Cloud	Cup	–
Type 2 (J-cls)	Bus	TV	Hat	–
Type 3 (C/J–cls)	Bone	Tube	Platter	–
Type 4 (C/J-cls-tax)	Bed	Table	–	Chair

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

Downloaded by [Keio University] at 05:51 26 July 2011

pair to share an unknown but important property, again using a rating scale of 1 (not likely at all) to 7 (very likely).

First, we expected to replicate our previous findings (Saalbach & Imai, 2007) for Chinese speakers: they should rate the two objects sharing the classifier in Chinese (same_classifier Pair Type 1 and 3) as more similar and as more likely to share an unknown property than German speakers do, although it is unclear whether the classifier relation further amplifies similarity when the two objects are already taxonomically related. Second, and most important for the purpose of the present research, we wanted to verify whether the language-specific classifier effect was also found in Japanese speakers and if so, whether the effect would be as strong as in Chinese speakers. As discussed above, if the effect varied as a function of how frequently classifiers appear with nouns, we might find a weaker effect or possibly no effect at all in Japanese speakers.

Based on Saalbach and Imai's (2007) results, we did not expect the amplified classifier effect to occur in this inductive reasoning task in which prior causal knowledge was available. Nonetheless, it is worth replicating the previous results with more extensive stimuli and also with Japanese participants, especially since the conclusion of Saalbach and Imai (2007) was based on a null result.

To further explore the conditions under which classifier similarity effects arise, we conducted two versions of a fast-speed word-picture matching. In both versions, participants of the three language groups saw a cue consisting of one of the objects from a pair, which was orthographically presented on a computer screen. They then saw the other object of the pair (target), which was pictorially presented. Participants were asked to judge as quickly as possible whether the target matched the cue. The previous study by Saalbach and Imai (2007) had not found any priming effect due to the classifier relation when bare nouns were used as cues. However, it is possible that when the noun is accompanied by the classifier, some online activation of classifier relations occurs. As mentioned above, Huettig and colleagues (Huettig et al., 2010) indeed found classifier effects on eye-movement only when classifier categories were explicitly invoked. When hearing a noun with its classifier, Chinese speakers' eye gaze shifted to the object which belonged to the classifier category they had just heard. However, when the noun was presented without a classifier, speakers looked equally to the same-classifier object and to the different-classifier distracters (for similar results for grammatical gender, see Vigliocco, Vinson, Indefrey, Levelt, & Hellwig, 2004). In the present research, we thus conducted the word-picture matching task in two versions: a version in which the cue word was presented alone, without a classifier, and a version in which the cue word was presented in the classifier phrase (e.g., "yi (one) ge (classifier) pinguo (apple)"). Given the

Downloaded by [Keio University] at 05:51 26 July 2011

results of previous research, a classifier priming effect could be expected to occur in the phrase condition but not in the bare noun condition.

EXPERIMENT 1: SIMILARITY JUDGEMENTS

Using more extended stimuli than Saalbach and Imai (2007), Experiment 1 examined whether the magnitude of the effect of classifier categories on similarity judgements would be the same in Chinese and in Japanese speakers.

Method

Participants

Thirty-nine Chinese undergraduates from Beijing, 40 Japanese undergraduates from the greater Tokyo Metropolitan area, and 35 German undergraduates from Berlin participated in this study. The participants in this and the following experiments were native speakers of Mandarin Chinese, Japanese, and German, respectively, with no reported history of speech or reading difficulties. In all the three language groups, most of the participants were undergraduates majoring in psychology, but there were also students from other majors. Participants of all groups had learned English in school for at least six years, but reported that they did not routinely use English outside the classroom, and their English proficiency was not high enough for them to be characterised as bilinguals. Participants were paid for participation. Participants' demographic backgrounds were the same for all of the studies reported in this paper within each language group, and were comparable across the three language groups. None of the participants in this and the following experiments participated in more than one experiment.

Downloaded by [Keio University] at 05:51 26 July 2011

Materials (common to studies 1–3)

We constructed 44 item triplets, each consisting of a target item, an item from the same classifier category, and a control item or a taxonomic item (see below). From each triplet, we constructed a same_classifier pair and either a control pair (Type 1–Type 3) or a same_taxonomic pair (Type 4). As described in the overview, there were four types of *same_classifier pairs* (see Table 1 for the structure of the stimuli). The first type included a target and an object which belonged to the same classifier category as the target object only in Chinese (same_classifier pair Type 1). The second type consisted of a target object and an object belonging to the same classifier category as the target only in Japanese (same_classifier pair Type 2). The third type consisted of a target and an object belonging to the same classifier category as the target in both Chinese and Japanese (same_classifier pair Type 3). In

Type 1–3 sets, the two objects (i.e., the target and the same-classifier item) did not belong to the same taxonomic category. For Type 1–3 sets, each triplet included a control item—an item which had no conceptual relation to the target. Thus, a *control pair* consisted of the target and the control item.

In the fourth type of the same_classifier pairs, the target and the same_classifier cue were also taxonomically related. In other words, the same_classifier item belonged to the same classifier category as the target in both Chinese and Japanese, and at the same time, the target and this same-classifier item were taxonomically related. To assess the classifier effect over and above the similarity due to a taxonomic relation, we included an additional taxonomic item for this type—an item which was taxonomically related to the target but had no classifier relation with it—and constructed a *taxonomic pair* to serve as a baseline to the same_classifier pair.

For selecting the items, we conducted a pre-study with native speakers of Mandarin Chinese and native speakers of Japanese who did not participate in any of the subsequent studies. First, a group of graduate students at Peking University and a group of graduate students at Keio University selected a set of prominent classifiers from a Chinese classifier dictionary (Guo, 2002) and a Japanese classifier dictionary (Iida, 2004), and a total of one hundred unambiguous objects associated with the selected classifiers. Object names were then randomly arranged in a questionnaire booklet and distributed to 10 native Mandarin-Chinese speakers from the Beijing region and 10 native Japanese speakers from the Tokyo region. They were instructed to write down the most appropriate classifier for each noun. The nouns for which at least 8 out of 10 participants agreed on the most appropriate classifier were retained. Finally, the remaining nouns from the first round as well as some newly chosen ones were again arranged in a questionnaire booklet and distributed to four graduate students at the Chinese Academy of Sciences (China) and at Keio University (Japan) for a final check. For all nouns in the booklet, at least three of the four judges agreed on what was the most appropriate classifier.

From the final Chinese and Japanese lists of object names, we selected those which fitted into our design. We selected 10 pairs of nouns from the same classifier category in Japanese (Type 2 same_classifier pairs), 10 pairs of nouns from the same classifier category in both Chinese and Japanese (Type 3 same_classifier pairs), and 10 pairs of objects from both the same Chinese and Japanese classifier category and taxonomic category (Type 4 same_classifier pairs). For the Type 1 set (same classifier only in Chinese), we chose the stimuli set consisting of 14 pairs of objects used in the Saalbach and Imai (2007) study, the selection of which had been made in the same way as the selection of the Type 2–4 pairs. One item of each pair was assigned to serve as the target, the other one as the same_classifier item. Assisted by

Downloaded by [Keio University] at 05:51 26 July 2011

native speakers, we selected a corresponding control and a taxonomic item for each target on common-sense grounds (see Appendix 1 for the complete list of stimuli).

It should be noted that there are over 100 classifier categories in both Chinese and Japanese. Among them, there are several subtypes with respect to semantic structures. According to Gao and Malt (2009), these are: "Well-defined" classifier categories, "Prototype" classifier categories, and "Arbitrary" classifier categories.[3] Among these subtypes, the classifiers most frequently used by far are those with a gradient prototype structure and a fairly broad category boundary, with members including many different taxonomic categories in both Japanese and Chinese. As we wished to draw same_classifier object pairs only from frequently used common classifier categories with no large overlap with taxonomic categories, the classifier categories for object pairs for the current research were mostly of this type, i.e., those with a gradient structure.

Materials and procedure of Experiment 1

Questionnaire booklets were prepared using 44 of the selected same-classifier pairs from Type 1 through Type 4 sets as well as the corresponding control or taxonomic pairs. Each item pair was presented with a scale of 1 (very dissimilar) to 7 (very similar). The two objects were shown orthographically (i.e., as a word) in the bare noun form (i.e., without a classifier in Japanese and Chinese, and without a gender article in German). Participants were asked to judge the similarity between the two objects in each pair. They were instructed to go through the questionnaire carefully and at their own pace, and to rely on their intuition. We prepared three versions of the booklet, each with a different randomised order. Participants in all three language groups randomly received one of these versions of the booklet.

Results

We will first briefly report the pattern of results within each language group in order to see whether and how the different classifier relations were related to the similarity ratings of the three language groups. We will then report the comparison of the effect of classifier relations in the three language groups in order to see whether speakers of classifier languages (Chinese and Japanese) gave higher similarity ratings to objects sharing the same classifier in their

[3] However, we do not always agree with Gao and Malt's classifications of the classifiers. What appears "arbitrary," especially for non-native speakers of the language, could have covert semantic meanings, which native speakers could unconsciously detect. Thus, in our view, it is difficult to clearly distinguish their "prototype" and "arbitrary" classifiers. The three types are better thought of as being on the same continuum.

Downloaded by [Keio University] at 05:51 26 July 2011

Downloaded by [Keio University] at 05:51 26 July 2011

own language than speakers of a nonclassifier language (German) did. Specifically, by comparing the differences between the ratings for objects from the same Chinese classifier category (Type 1 and Type 3 pairs) and the controls, we tested whether Chinese speakers were more sensitive than German speakers to the similarity underlying classifier categories in Chinese. Likewise, to examine whether Japanese speakers were more sensitive than German speakers to the similarity underlying classifier categories in Japanese, we compared Japanese and German speakers with regard to the difference between ratings for objects from the same classifier category in Japanese (Type 2 and Type 3 pairs) and their corresponding controls.

Finally, we tested whether sensitivity to the similarity underlying classifier categories differed between Chinese and Japanese speakers. For this purpose, we compared the difference between the ratings for objects sharing the same classifiers in both Chinese and Japanese (Type 3 pairs) and the corresponding controls across the two language groups.

To test the hypothesised cross-linguistic interaction effects, we used Hierarchical Linear Modelling (HLM), Raudenbush & Bryk, 2002), as this allowed us to simultaneously assess whether the effects hold true on the item level (same vs. different classifier) and on the subject level (Chinese vs. German, Japanese vs. German, and Chinese vs. Japanese). To test the effects that involved comparisons across different item sets (e.g., when testing the overall differences between same-classifier and different classifier item pairs across Types 1–4), Multi-Level Modelling was not suitable. For such cases, we conducted repeated-measure ANOVAs using participants as the unit of analyses.

As in Saalbach and Imai (2007), we found that Chinese speakers tended to give higher similarity ratings than speakers of Japanese and German in all conditions including the control, $F(2, 110) = 6.99$, $p < .01$, $\eta_p^2 = .113$. In order to adjust this baseline differences between subjects, we standardised similarity ratings and used the z-scores for the subsequent analysis, as in Saalbach and Imai (2007). However, for easier readability, we will report the means of the raw scores in the text. Table 2 shows the average standardised similarity ratings (z-scores) for each of the four types of same_classifier pairs as well as for the corresponding control or taxonomic pairs in each language.

The overall response pattern was very similar across the three language groups. Participants in all three language groups rated two objects from the same classifier category (across all types) as more similar than the corresponding control or taxonomic pairs [Chinese: 3.31 vs. 2.52, $F(7, 266) = 94.65$, $\eta_p^2 = .731$; Japanese: 2.83 vs. 2.22, $F(7, 273) = 194.73$, $\eta_p^2 = .833$, and German: 2.64 vs. 2.13, $F(7, 238) = 142.45$, $\eta_p^2 = .807$, all ps $< .01$.] Furthermore, speakers of all three languages showed higher ratings for taxonomically related pairs (average of same_classifier pairs and taxonomic pairs in Type 4 sets) than for taxonomically unrelated same_classifier pairs (average of

TABLE 2
Mean standardised-scores for each item pair type in each language in Experiment 1 (Similarity judgements)

Language	*N*	*Item pair type*	*Type 1 C-cls*	*Type 2 J-cls*	*Type 3 C/J-cls*	*Type 4 C/J-cls-tax*
Chinese	39	Same_classifier	−0.03	−0.24	0.25	1.02
		Control	−0.56	−0.43	−0.64	–
		Taxonomic	–	–	–	0.87
Japanese	40	Same_classifier	−0.28	−0.31	0.07	1.42
		Control	−0.57	−0.47	−0.54	–
		Taxonomic	–	–	–	1.02
German	35	Same_classifier	−0.19	−0.33	0.08	1.31
		Control	−0.58	−0.55	−0.55	–
		Taxonomic	–	–	–	1.13

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

Downloaded by [Keio University] at 05:51 26 July 2011

same_classifier pairs of Types 1–3) [Chinese: 4.49 vs. 2.87, $F(1, 38) = 89.55$, $\eta_p^2 = .702$, Japanese: 4.53 vs. 2.16, $F(1, 39) = 284.99$, $\eta_p^2 = .880$; German: 4.19 vs. 2.08, $F(1, 34) = 165$, $\eta_p^2 = .830$, all ps $< .01$] and control pairs [Chinese: 4.49 vs. 1.90, $F(1, 38) = 358, 29$, $\eta_p^2 = .904$, Japanese: 4.53 vs. 1.56, $F(1, 39) = 595.65$, $\eta_p^2 = .939$, German: 4.19 vs. 1.49, $F(1, 34) = 618.53$, $\eta_p^2 = .948$, all ps $< .01$].

Thus, speakers of Chinese and Japanese as well as speakers of a nonclassifier language (German) showed sensitivity to the similarity underlying classifier categories. Throughout the paper, we will refer to this sensitivity by the term of *classifier similarity effect.* To test whether the magnitude of the classifier similarity effect is larger in speakers of a classifier language than in speakers of a nonclassifier language, we then conducted a set of Hierarchical Linear Models. Specifically, we tested whether Chinese or Japanese speakers, as compared to German speakers, showed an *amplified classifier similarity effect* for the object pairs taken from the same classifier category in their own language. Table 3 summarises the parameter estimates and statistics for the Hierarchical Linear Models for Experiment 1.

Was there an amplified classifier similarity effect in Chinese speakers? We first tested whether Chinese speakers and German speakers differed in their ratings for Chinese same-classifier pairs by testing the effect of Language (Chinese vs. German) on the contrast relevant to the same_classifier pairs in

TABLE 3
Estimated for fixed effects and variance components in the Hierarchical Linear Models for Experiment 1 (Similarity judgement)

		Fixed effects		*Variance components*	
Contrasts		*Estimate (SE)*	T	*Estimate*	χ^2
Type 1 & $3_{(C\ vs.\ G)}$	Intercept	−0.230 (0.046)	−4.93**	0.139	501.98**
	Classifier × language	−0.289 (0.091)	−3.168**	0.117	489.89**
Type 1 & $2_{(J\ vs.\ G)}$	Intercept	−0.384 (0.040)	−9.544**	0.105	504.47**
	Classifier × language	0.021 (0.074)	0.285	0.108	550.22**
Type $1_{(C\ vs.\ J)}$	Intercept	−0.148 (0.047)	−3.104**	0.113	204.99**
	Classifier × language	−0.336 (0.138)	−2.445*	0.114	241.33**
Type $4_{(C\ vs.\ G)}$	Intercept	1.037 (0.064)	16.069**	0.422	948.70**
	Classifier × language	0.009 (0.091)	0.107	0.462	759.76**
Type $4_{(J\ vs.\ G)}$	Intercept	1.015 (0.077)	13.020**	0.408	623.58**
	Classifier × language	−0.213 (0.084)	−2.529*	0.407	635.99**

Note: Type 1 = same classifier in Chinese and Japanese; Type 2 = same classifier only in Japanese; Type 3 = same classifier only in Chinese; Type 4 = same classifier in Chinese and Japanese and taxonomically related; C = Chinese; J = Japanese; G = German.
*p <.05; **p <.01.

Downloaded by [Keio University] at 05:51 26 July 2011

Chinese (the average of Type 1 and Type 3 same_classifier pairs vs. the average of the corresponding controls [Type 1 control pairs and Type 3 control pairs]). As predicted, we found a significant Contrast × Language interaction effect: the magnitude of the classifier similarity effect was larger for Chinese speakers (3.08 vs. 1.79) than for German speakers (2.20 vs. 1.49).

Was there an amplified classifier similarity effect in Japanese speakers? To test whether Japanese speakers' construal of similarity was also amplified for object pairs sharing the same classifier in Japanese, we tested whether the contrast between Japanese same_classifier pairs (the average of Type 2 and Type 3 same_classifier pairs) and the corresponding control pairs (the average of Type 2 control pairs and Type 3 control pairs) interacted with Language (Japanese vs. German). Here, no Contrast × Language interaction was revealed (Japanese speakers: 2.24 vs. 1.59; German speakers: 2.11 vs. 1.50).

Comparison of the classifier similarity effect in Japanese and Chinese speakers. We then tested whether the structural differences between the Chinese and the Japanese classifier systems resulted in a difference in the size of the classifier similarity effect. For this purpose, we examined whether the contrast between Type 3 same_classifier pairs and control pairs differed between Chinese and Japanese speakers (objects were drawn from the same classifier categories in both Chinese and Japanese). The significant Contrast × Language interaction was found, indicating that the classifier similarity effect was indeed larger in Chinese (3.35 vs. 1.71) than in Japanese speakers (2.57 vs. 1.53).

Was there a language-specific amplified classifier effect in Chinese and Japanese speakers when object pairs were taxonomically related? Finally, we tested whether classifiers amplified similarity over and above the similarity due to taxonomic relations and whether this effect differed across languages. To test this, we contrasted ratings for the same_classifier pairs against the taxonomic pairs in Type 4 sets and tested whether the effect of Contrast interacted with Language (Chinese vs. German and Japanese vs. German). In the Type 4 set, the two objects in the same_classifier pairs were taxonomically related and belonged to the same classifier category in both Japanese and Chinese, whereas the two objects in the taxonomic pairs were taxonomically related but did not share the same classifier category membership. Unexpectedly, for Type 4 sets, the Contrast × Language interaction did not reach the level of significance when Chinese was compared to German. However, when Japanese was compared to German, the interaction reached significance (Chinese: 4.61 vs. 4.36; Japanese: 4.85 vs. 4.20; German: 4.33 vs.4.06).

In summary, speakers of all three languages rated two objects sharing the same Chinese or Japanese classifiers as more similar than two objects from different classifier categories. A comparison of this effect between classifier and nonclassifier languages indicates that the classifier similarity effect is magnified in Chinese speakers, but not in Japanese speakers.

Discussion

The results showed that speakers of Chinese, Japanese, and German gave highest similarity ratings to the taxonomically related item pairs, followed by the same_classifier pairs, and the unrelated control pairs. Consistent with previous research (Saalbach and Imai, 2007), we found that speakers of the two classifier languages as well as speakers of the nonclassifier language rated the same_classifier object pairs as more similar than the control pairs. These findings further support the thesis that there is an inherent similarity underlying classifier categories, and that it can be detected even by speakers of a nonclassifier language.

Also consistent with previous research by Saalbach and Imai (2007), we found that the classifier similarity effect was amplified in Chinese speakers, as compared to German speakers, when two objects shared the same classifier in Chinese. In contrast, no such amplified classifier similarity effect was found in Japanese speakers' responses to pairs of objects which shared the same classifier in Japanese except for pairs which shared taxonomic relations in addition to same-classifier relations. This finding

Downloaded by [Keio University] at 05:51 26 July 2011

was unexpected, and should be further explored. One way to do this could be to1 see if this effect would consistently arise in the following experiments, as well, especially in the inductive inference task with a blank property (Experiment 2), where the same amplified similarity effect was identified by Saalbach and Imai (2007). We will further discuss this effect in the General Discussion.

In any case, when considering the object pairs sharing the same classifier in both Chinese and Japanese, we found that the magnitude of the classifier similarity effect was significantly larger in Chinese speakers than in Japanese speakers. This pattern of findings suggests that the relation between classifier categories and cognition revealed in previous studies conducted with Chinese speakers (Saalbach & Imai, 2007; Imai, Saalbach, & Stern, 2010; Zhang & Schmitt, 1998) cannot be simply generalised to speakers of other classifier languages. We next examined whether classifier categories affect inductive reasoning about object properties and, if so, whether this effect is, again, larger in Chinese than in Japanese speakers.

Downloaded by [Keio University] at 05:51 26 July 2011

EXPERIMENT 2: INDUCTIVE REASONING WITH A BLANK PROPERTY

Supporting the inductive generalisation of properties is seen as the basic function of categories (e.g., Murphy, 2002). If classifier categories are important categories within the mental apparatus of speakers of a classifier language, there should be properties that can be projected across objects that share the same classifier. We thus tested Japanese and Chinese speakers on an inductive reasoning task in which participants were asked to judge how likely it was for a property of the target object to be generalised to each of the test objects. In previous research, Saalbach and Imai (2007) found different results depending on the nature of the property. Inference of a blank (unknown) property may involve higher sensitivity to classifier relations than inference of a known property because speakers have no other bases for induction than the similarity underlying classifier relations. In contrast, inference of a known property is heavily constrained by particular background knowledge about the property, and in this case, the language-specific amplified similarity due to classifier category membership may be superseded by knowledge. We thus conducted two kinds of the property induction task, one with a blank property (the present experiment) and one with a known property (Experiment 3).

Method

Participants

Twenty-two Chinese undergraduates, 25 Japanese undergraduates, and 25 German undergraduates participated in this study.

Materials and procedure

The stimulus set and its arrangement in the questionnaire, as well as the item order randomisation procedure, were the same as in Experiment 1. For each item pair, participants saw the following question: "Property X is an important property for [object 1]. If [Object 1] has property X, how likely is it that [Object 2] also has property X?" Participants were asked to judge the likelihood on a rating scale of 1 (not likely at all) to 7 (very likely). [Object 1] was the target object, and [Object 2] was either an object from the same classifier category or an unrelated object. Participants were instructed to go through the questionnaire carefully and at their own pace, and to rely on their intuition.

Results

As in Experiment 1, we adjusted the baseline differences between subjects by standardising likelihood ratings within subjects, and used the *z*-scores for the subsequent analysis, but will report the means of the raw scores in the text. Table 4 shows the mean standardised likelihood ratings for each of the four types of the same_classifier pairs as well as for the corresponding control and taxonomic pairs in each language group.

As in Experiment 1, we first examined the pattern of results within each language and then compared the patterns across the three languages. Again, the overall response pattern was strikingly similar across the three language groups: Speakers of the three language groups all judged that two objects from the same classifier category (Types 1–4) were more likely to share a common property than two objects from different classifier categories [Chinese: 4.15 vs. 3.55, $F(7, 147) = 127.57$, $\eta_p^2 = .895$; Japanese: 3.61 vs. 3.22, $F(7, 168) = 51.13$, $\eta_p^2 = .680$, and German: 3.48 vs. 3.05, $F(7, 168) = 267.94$, $\eta_p^2 = .918$, all ps $< .01$]. Furthermore, all three groups gave higher likelihood ratings for taxonomically related pairs (the average of the same_classifier pairs and the taxonomic pairs in Type 4 sets) than for taxonomically unrelated same_classifier pairs (the average of the same_classifier pairs of Types 1–3) [Chinese: 5.90 vs. 3.61, $F(1, 21) = 178.32$, $\eta_p^2 = .895$, Japanese: 5.36 vs. 3.03, $F(1, 24) = 90.43$, $\eta_p^2 = .790$; German: 5.47 vs. 2.84, $F(1, 24) = 329.18$, $\eta_p^2 = .932$; all ps $< .01$], and they showed higher ratings for taxonomically related pairs than for control pairs (the average of the control

Downloaded by [Keio University] at 05:51 26 July 2011

Downloaded by [Keio University] at 05:51 26 July 2011

TABLE 4
Mean standardised-scores for each item pair type in each language in Experiment 2 (Induction of a blank property)

Language	*N*	*Item pair type*	*Type 1* *C-cls*	*Type 2* *J-cls*	*Type 3* *C/J-cls*	*Type 4* *C/J-cls-tax*
Chinese	22	Same_classifier	−0.10	−0.23	0.04	1.08
		Control	−0.62	−0.53	−0.61	–
		Taxonomic	–	–	–	1.25
Japanese	25	Same_classifier	−0.24	−0.19	−0.05	1.02
		Control	−0.43	−0.47	−0.37	–
		Taxonomic	–	–	–	0.99
German	25	Same_classifier	−0.17	−0.32	−0.13	1.25
		Control	−0.57	−0.53	−0.57	–
		Taxonomic	–	–	–	1.33

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

pairs of Types 1–3) [Chinese: 5.90 vs. 2.72, $F(1, 21) = 372.43$, $\eta_p^2 = .947$; Japanese: 5.36 vs. 2.50, $F(1, 24) = 122.55$, $\eta_p^2 = .837$; and German: 5.47 vs. 2.21, $F(1, 24) = 1039.13$, $\eta_p^2 = .978$; all ps $< .01$].

Thus, the pattern within each language group suggests that (1) speakers tend to think that objects belonging to the same classifier category (across all types of same_classifier pairs) are more likely to share an unspecified but important property than objects belonging to different classifier categories, regardless of whether or not the speakers' language has a classifier system; and (2) speakers of all three language groups think that taxonomically related objects are more likely to share a blank property than taxonomically unrelated objects are, regardless of whether or not the object belongs to the same classifier category.

To test whether the magnitude of the classifier similarity effect was larger in speakers of a classifier language than in speakers of a nonclassifier language, a set of Hierarchical Linear Models was carried out, as in Experiment 1, for the effects we were interested in. The parameter estimates and statistics for all models of Experiment 2 are shown in Table 5.

Was there an amplified classifier similarity effect in Chinese speakers' inductive inference of a blank property? As in Experiment 1, we first tested whether Chinese speakers' inductive inference of a blank property was affected more strongly by classifier relations than that of speakers of a nonclassifier language (i.e., German). For this purpose, we set up a contrast

TABLE 5
Estimates for fixed effects and variance components in the Hierarchical Linear Models for Experiment 2 (Induction of a blank property)

		Fixed effects		*Variance components*	
Contrasts		*Estimate (SE)*	*T*	*Estimate*	χ^2
Type 1 & $3_{(C\ vs.\ G)}$	Intercept	–0.299 (0.061)	–4.857**	0.168	720.81**
	Classifier × language	–0.145 (0.077)	-1.878^{+}	0.146	680.80
Type 1 & $2_{(J\ vs.\ G)}$	Intercept	–0.401 (0.056)	–7.068**	0.145	441.79**
	Classifier × language	0.020 (0.089)	0.228	0.143	445.37**
Type $1_{(C\ vs.\ J)}$	Intercept	–0.182 (0.068)	–2.674*	0.174	224.99**
	Classifier × language	–0.236 (0.137)	-1.726^{+}	0.175	242.14**
Type $4_{(C\ vs\ G)}$	Intercept	1.128 (0.055)	20.216**	0.136	507.11**
	Classifier × language	0.075 (0.065)	1.152	0.127	471.63**
Type $4_{(J\ vs\ G)}$	Intercept	1.001 (0.053)	18.607**	0.121	274.26**
	Classifier × language	–0.047 (0.063)	–0.736	0.123	272.76**

Note: Type 1 = same classifier in Chinese and Japanese; Type 2 = same classifier only in Japanese; Type 3 = same classifier only in Chinese; Type 4 = same classifier in Chinese and Japanese and taxonomically related; C = Chinese; J = Japanese; G = German.
$^{+}p < .10$; $^{*}p < .05$; $^{**}p < .01$.

for testing the magnitude of the difference between Chinese same_classifier pairs (the average of Type 1 & Type 3 same_classifier pairs) and their corresponding control pairs (the average of Type 1 control pairs and Type 3 control pairs) to see whether Language interacted with this contrast. The Contrast × Language interaction was marginally significant ($p = .06$), suggesting that the magnitude of the classifier similarity effect tended to be larger in Chinese (3.73 vs. 2.68) than in German speakers (2.94 vs. 2.19).

Was there an amplified classifier effect in Japanese speakers' inductive inference of a blank property? To examine whether an analogous amplified classifier effect was found in Japanese speakers' inductive inference of an unknown property, we compared Japanese and German speakers' likelihood ratings for the same_classifier pairs (a contrast between the average of Type 2 & Type 3 same_classifier pairs and the average of Type 2 & Type 3 control pairs). In contrast to Chinese speakers, no Contrast × Language interaction was revealed in this task. Thus, the classifier similarity effect for objects belonging to the same Japanese classifiers was not amplified in Japanese speakers (Japanese: 3.12 vs. 2.50; German: 2.81 vs. 2.22).

Was there a difference between the magnitudes of the classifier similarity effect in Chinese as compared to Japanese speakers? We next examined whether the magnitude of the classifier similarity effect in Type 3 sets, in which the same_classifier pairs belonged to the same classifier in both

Downloaded by [Keio University] at 05:51 26 July 2011

Downloaded by [Keio University] at 05:51 26 July 2011

Chinese and Japanese, differed between speakers of Chinese and speakers of Japanese. This was again done by testing the Contrast (same_classifiers pairs and control pairs in Type 3 sets) × Language (Chinese vs. Japanese speakers) interaction. The interaction effect was marginally significant ($p = .08$), suggesting that the relation between classifier categories and inductive reasoning about a blank property tends to be stronger in Chinese speakers (3.87 vs. 2.68) than in Japanese speakers (3.26 vs. 2.54).

Was there a language-specific amplified classifier effect in Chinese and Japanese speakers when object pairs were taxonomically related? Finally, we tested whether classifier relations increased the effect of taxonomic relations on (blank) property inductions in a way that was language specific. To do this, we contrasted the same_classifier pairs (i.e., two objects that were taxonomically related and also belonged to the same classifier category) and the corresponding taxonomic pairs (i.e., two objects that were taxonomically related but belonged to different classifier categories) in Type 4 sets, and tested whether there were significant Contrast × Language (Chinese vs. German and Japanese vs. German) interactions. This was not the case (Chinese: 5.76 vs. 6.05, Japanese: 5.34 vs. 5.39, and German: 5.40 vs. 5.55). Importantly, unlike in Experiment 1, no amplified classifier similarity effect was found.

In summary, the pattern of results in the blank property induction task resembled the pattern found in the similarity judgement task (Experiment 1). First, participants from all three language groups gave higher ratings for (Chinese or Japanese) same_classifier pairs than for pairs consisting of objects from different classifier categories. Second, the classifier effect tended to be amplified in Chinese but not in Japanese speakers. We will discuss these findings together with the results from Experiment 3.

EXPERIMENT 3: INDUCTIVE REASONING WITH A KNOWN PROPERTY

Experiment 3 examined whether there was any language-specific amplified classifier similarity effect on inductive reasoning with regard to a property for which some background knowledge could be accessed. Saalbach and Imai (2007) had found no classifier effect in Chinese speakers in this case. If this result was generalisable to different stimuli, the amplified classifier effect should not be observed in either Chinese or Japanese speakers.

Method

Participants

Thirty-four Chinese undergraduates, 39 Japanese undergraduates, and 34 German undergraduates participated in this study.

Materials and procedure

Stimulus set and procedure were the same as in Experiment 2, except that participants were asked to make likelihood judgements about two objects "carrying the same bacteria" instead of sharing a blank property. Specifically, participants saw the question: "How likely is it that [object 1] and [object 2] carry the same bacteria?" for each object pair, together with a rating scale of 1 (not likely at all) to 7 (very likely).

Results

As in the previous experiments, we adjusted the baseline differences between subjects by standardising likelihood ratings within subjects, and used the *z*-scores for the subsequent analysis, but will report the means of the raw scores in the text.

As shown in Table 6, the pattern of likelihood ratings within each language group paralleled the one found in the induction task using a blank property: In all three language groups, same_classifier pairs (Types 1–4) were

TABLE 6
Mean standardised-scores for each item pair type in each language in Experiment 3 (Induction of a known property)

Language	*N*	*Item pair type*	*Type 1* *C-cls*	*Type 2* *J-cls*	*Type 3* *C/J-cls*	*Type 4* *C/J-cls-tax*
Chinese	34	Same_classifier	−0.23	−0.16	−0.20	0.87
		Control	−0.34	−0.30	−0.41	–
		Taxonomic	–	–	–	1.04
Japanese	39	Same_classifier	−0.23	−0.26	−0.19	1.08
		Control	−0.38	−0.35	−0.40	–
		Taxonomic	–	–	–	0.97
German	34	Same_classifier	−0.36	−0.12	−0.31	1.11
		Control	−0.47	−0.27	−0.38	–
		Taxonomic	–	–	–	1.12

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

Downloaded by [Keio University] at 05:51 26 July 2011

rated as more likely to "carry the same bacteria" than pairs of objects from different classifier categories [Chinese: 4.39 vs. 4.28, $F(1, 33) = 3.20$, $p < .1$, $\eta_p^2 = .088$; Japanese: 3.66 vs. 3.43, $F(1, 38) = 15.95$, $\eta_p^2 = .296$; German: 3.44 vs. 3.31, $F(1, 33) = 7.46$, $\eta_p^2 = .185$; all ps $< .01$], which again suggests that the similarity underlying classifier categories can be detected, and is used for drawing an inductive inference, not only by speakers of a classifier language but also by speakers of a none-classifier language. As expected, taxonomically related pairs (the average of same_classifier pairs and taxonomic pairs in Type 4 sets) received significantly higher ratings than taxonomically unrelated same_classifier pairs (the average of the same_classifier pairs of Types 1–3) [Chinese: 5.96 vs. 3.93, $F(1, 33) = 204.83$, $\eta_p^2 = .862$; Japanese: 5.20 vs. 3.11, $F(1, 38) = 324.56$, $\eta_p^2 = .896$; German: 5.23 vs. 2.85, $F(1, 33) = 628.10$, $\eta_p^2 = .950$, and control pairs (the average of the control pairs in the Types 1–3) Chinese: 5.96 vs. 3.66, $F(1, 33) = 286.38$, $\eta_p^2 = .897$; Japanese: 5.20 vs. 2.88, $F(1, 38) = 500.67$; German: 5.23 vs. 2.66, $F(1, 33) = 679.04$, $\eta_p^2 = .954$; all ps $< .01$].

As in the analyses for the previous experiments, we tested whether the magnitude of the classifier effect was larger in speakers of a classifier language than in speakers of a nonclassifier language by carrying out a set of Hierarchical Linear Models for the contrasts we were interested in. Table 7 shows the models' parameter estimates and statistics.

Was there a magnified classifier similarity effect in Chinese speakers' inductive reasoning about bacteria? In contrast to the blank property induction task (Experiment 2), no trend towards an amplified effect of Chinese classifiers (i.e., difference between Chinese same_classifier pairs and their corresponding control pairs; Type 1 vs. control and Type 3 vs. control) was found in Chinese speakers as compared to German speakers for the "bacteria" property (Chinese: 3.91 vs. 3.61, German: 2.72 vs. 2.57).

Was there a magnified classifier effect in Japanese speakers' inductive reasoning about bacteria? As in Chinese speakers, no effect due to Japanese classifier category membership was found in Japanese speakers (Type 2 & Type 3 same_classifier pairs) as compared to German speakers (Japanese: 3.12 vs. 2.87, German: 2.94 vs. 2.75).

Was there a difference between the magnitudes of the classifier similarity effect in Chinese as compared to Japanese speakers? As in Experiments 1 and 2, we then compared the magnitude of the classifier effect in Chinese speakers and in Japanese speakers with respect to the pairs where the two objects belonged to the same classifier category in both languages by testing the Contrast (Type 3 same_classifier pairs vs. Type 3 control pairs) and

Downloaded by [Keio University] at 05:51 26 July 2011

TABLE 7
Estimates for fixed effects and variance components in the Hierarchical Linear Models for Experiment 3 (Induction of a known property)

		Fixed effects		*Variance components*	
Contrasts		*Estimate* (SE)	T	*Estimate*	χ^2
Type 1 & $3_{(C\ vs.\ G)}$	Intercept	−0.239 (0.075)	−3.181**	0.369	1253.58**
	Classifier × language	0.070 (0.062)	1.126	0.295	1006.97**
Type 1 & $2_{(J\ vs.\ G)}$	Intercept	−0.372 (0.065)	−5.701**	0.295	1042.78**
	Classifier × language	−0.027 (0.063)	−0.432	0.299	1044.49**
Type $1_{(C\ vs.\ J)}$	Intercept	−0.169 (0.074)	−2.265*	0.364	585.75**
	Classifier × language	−0.007 (0.043)	−0.165	0.338	547.75**
Type $4_{(C\ vs.\ G)}$	Intercept	0.940 (0.063)	14.904**	0.237	502.85**
	Classifier × language	0.148 (0.078)	1.880^{+}	0.209	446.18**
Type $4_{(J\ vs\ G)}$	Intercept	0.752 (0.064)	11.732**	0.261	506.91**
	lassifier × language	−0.120 (0.077)	−1.551	0.265	507.34**

Note: Type 1 = same classifier in Chinese and Japanese; Type 2 = same classifier only in Japanese; Type 3 = same classifier only in Chinese; Type 4 = same classifier in Chinese and Japanese and taxonomically related; C = Chinese; J = Japanese; G = German.
$^{+}p<.10$; $^{*}p<.05$; $^{**}p<.01$.

Downloaded by [Keio University] at 05:51 26 July 2011

Language (Chinese vs. Japanese) interaction. As expected, unlike Experiments 1 and 2, Chinese speakers did not show stronger sensitivity to classifier relations than Japanese speakers in this task (Chinese: 3.89 vs. 3.56; Japanese: 3.18 vs. 2.82).

Was there a magnified classifier similarity effect in Chinese and Japanese speakers' inductive reasoning about bacteria when two objects were taxonomically related? The difference between likelihood ratings for taxonomically related pairs from the same classifier category (Type 4) and taxonomically related pairs from different classifier categories did not differ between the Japanese and the German groups (Japanese: 5.31 vs. 5.09, German: 5.22 vs. 5.24). There was a marginally significant effect in the likelihood ratings of the Chinese as compared to those of the German speakers ($p=.06$). However, this effect is due to a reverse pattern in the Chinese speakers. They gave higher likelihood ratings for the taxonomically related pairs from different classifier categories than for pairs that had both taxonomic and classifier relations (5.79 vs. 6.12, respectively).

In summary, although speakers from all three language groups rated pairs of objects from the same classifier category as more likely to "carry the same bacteria" than pairs of objects from different classifier categories, there was no language-specific amplified classifier effect in this experiment, in contrast to Experiment 2.

Downloaded by [Keio University] at 05:51 26 July 2011

Discussion

Experiments 2 and 3 revealed that speakers of all three languages gave higher likelihood ratings for same-classifier pairs than for control pairs, whether or not they spoke a classifier language, and whether or not a given classifier relation was present in their own language. This is consistent with the results of previous research showing that the semantics underlying classifier categories is detectable even by speakers of a nonclassifier language and increases the likelihood of property generalisation (Saalbach & Imai, 2007). However, inter-language differences in the magnitude of the classifier similarity effect were only found for inductive inference about an unknown (blank) property. As for the similarity judgement task, we found an amplified classifier similarity effect in Chinese but not in Japanese speakers; furthermore, concerning the item pairs belonging to the same classifier category in both languages, the magnitude of the classifier similarity effect was significantly larger in Chinese than in Japanese speakers.

Why was there a language-specific classifier effect (in Chinese speakers) for inference about a blank property but not for inference about the carrier of the same bacteria? In contrast to the bacteria case, inference about an unknown property does not allow subjects to access specific background knowledge; hence, subjects resorted to overall similarity as a basis for inference. Experiment 1 had revealed that Chinese speakers showed an amplified similarity effect based on the classifier relation in Chinese, which was directly reflected in their inductive inference about a blank property. In contrast, when participants were able to use background knowledge in terms of factors emphasising the likelihood for objects to carry the same bacteria, they no longer needed to resort to similarity as indicated by classifier relations; hence the amplified classifier effect diminished.

The absence of an amplified classifier effect in inductive inferences about the bacteria property does not mean that this finding holds true also for inductive inferences about any known property. It has been suggested that the patterns of inductive inferences depend on the kind of the properties to be projected (e.g., Nisbett, Krantz, Jepson, & Kunda, 1983). Thus, it may be possible that an amplified classifier effect occurs in Chinese speakers if the property is in some way semantically related to the classifier categories. Most of the classifiers we used in the present research are defined by shape or functional properties, which are not likely to promote inductive inferences about the carrying of bacteria. The results suggest that speakers of Chinese and Japanese both knew this, and did not rely on the same_classifier relations in this particular case.

Downloaded by [Keio University] at 05:51 26 July 2011

EXPERIMENT 4: SPEEDED WORD-PICTURE MATCHING TASK

Experiments 4 and 5 examined whether the relation between classifier categories and cognition extended to cognitive processes that are unconsciously and automatically evoked. It is widely known that recognising an object involves the activation of other objects which are part of the same semantic network. A great many studies using the so-called semantic priming paradigm have shown inhibition or facilitation effects when two conceptually related objects were presented as a sequence. Yokosawa and Imai (1997), for example, found strong priming effects for conceptual relations in speeded word-picture and picture-picture matching tasks. In their experiments, Japanese participants saw a cue that was presented either orthographically or pictorially, followed by a briefly presented target picture. The participants were asked to judge, as fast as possible, whether the word or picture cue represented the same object as the target picture. Yokosawa and Imai found that, regardless of whether the cue was a word or a picture, participants were more delayed in detecting a mismatch between the cue and the target when the cue and the target objects were taxonomically related (e.g., rabbit, dog), or thematically related (e.g., rabbit, carrot), than when they were unrelated (e.g., rabbit, hammer).

Saalbach and Imai (2007) used this (negative) semantic priming method to test whether objects of the same classifier category would be activated in the online cognitive process in speakers of Chinese. Replicating Yokosawa and Imai's (1997) results, they found priming effects for both taxonomically and thematically related cue-target pairs. However, in this task, no priming effect due to the same_classifier relation was found: Chinese speakers' latencies were not longer when the cue (word) and the target (picture) belonged to the same classifier category than when they did not share the classifier category membership; nor were Chinese speakers' response latencies for the same_classifier pairs longer than those of German speakers.

The absence of a priming effect due to same_classifier relations in the previous research may suggest that the presentation of an isolated noun does not automatically invoke other members of its classifier category. It is possible that without explicit classifier invocation, the link among objects in terms of classifier category membership is not strong enough to spread activation online (cf. Huettig et al., 2010).

Viewed in this light, a semantic priming effect due to classifier category membership may be only obtained when the noun is presented in a classifier phrase. On the other hand, we cannot rule out the possibility that the absence of the classifier priming effect in Saalbach and Imai (2007) was due to idiosyncratic properties of the stimuli. In the present research, we thus conducted two versions of the fast-speed word-picture matching task: a task in which the cue noun was presented in isolation (this experiment) and a task in which the cue noun was embedded in a classifier-noun phrase (Experiment 5).

Method

Participants

Twenty-five Chinese, 23 Japanese, and 24 German speakers, all undergraduate students, participated in this study.

Materials and procedure

The same stimuli sets as in Experiments 1–3 were used for this study. Furthermore, additional taxonomic items and control items were included for Experiments 4 and 5: In Type 1–3 sets, an object that was taxonomically related but belonged to a different classifier category was added for each target object; in a Type 4 set, an object without either a taxonomic or a classifier relation to the target was added (see Table 8 for the structure of the stimuli). These manipulations were done to ascertain whether the semantic priming procedure itself would work throughout Type 1–4 sets. If this was the case, delayed rejection of taxonomically related pairs should occur as compared to unrelated control pairs. Thus, in case no delay due to the same classifier relation was found, we would be able to determine whether the null result was due to the absence of the classifier effect or to methodological problems.

There were three items representing three types of conceptual relations for each target object: same-classifier relation, taxonomic relation, and unrelated. The cue was presented orthographically (i.e., a word), and the target object was presented pictorially. The pictures (simple black and white line

Downloaded by [Keio University] at 05:51 26 July 2011

TABLE 8
Structure of the stimuli used for Experiments 4–5 with a sample set for each stimulus type

Type of Same_classifier pair	*Target item*	*Same-classifier item*	*Control item*	*Taxonomic item*
Type 1 (C-cls)	Flower	Cloud	Cup	Tree
Type 2 (J-cls)	Bus	TV	Hat	Boat
Type 3 (C/J–cls)	Bone	Tube	Platter	Muscle
Type 4 (C/J-cls-tax)	Bed	Table	Wire	Chair

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

drawings) were derived from the Snodgrass and Vanderwart picture database (1980), when available. For items not available from the Snodgrass and Vanderwart database, we had an artist drawing pictures in the same style. Each word (in the respective language) and each picture was presented in black on a white background in the centre of the screen. Each participant went through a total of 284 test trials, half of which required a positive response (cue and target matched), while the other half required a negative response (cue and target did not match). Each target picture appeared only once in the positive trials. To avoid presenting the target pictures too many times, the rest of the positive trials consisted of filler items. In order to prevent participants from repeatedly resorting to the same type of response in the case of filler trials, we also included a set of negative filler trials using the same target pictures as in the positive filler trials. The cue-target pairs were presented in a random order, with the constraint that the same target could not appear more than once within a three-trial window.

Participants were instructed that they would see a word followed by a short presentation of a picture and were asked to judge, as fast as possible, whether or not the word and the picture matched. Participants first saw a fixation cross for 1,500 ms, followed by a 1,000 ms presentation of the cue word. Another fixation cross appeared and stayed for 500 ms. The target picture was then presented for 200 ms, followed by a pattern mask that remained until the subjects gave their response. After the response, a new trial started with the 1,500 ms presentation of the fixation cross. Before starting the test trials, participants received 15 practice trials to guarantee their correct understanding of the task and of the keys assigned to "Yes" and "No". During the practice sessions, error responses were indicated by beeps. During the test session, however, no feedback was provided.

Downloaded by [Keio University] at 05:51 26 July 2011

Results and discussion

Error responses were low (less than 10% in all three language groups), and the distributions were extremely skewed, with most of the data points concentrating near 0. We thus focused on the analysis of response latencies. Filler trials, positive trials (in which the word cue and the target picture matched), and trials including error responses were excluded from the analysis of the response latencies.

As shown in Table 9, Japanese participants were faster to respond than Chinese and German speakers across all stimulus types [Chinese: 377 ms; German: 338 ms; Japanese: 281 ms), $F(2, 69) = 11.07$, $p < .01$, $\eta_p^2 = .241$]. As in Saalbach and Imai (2007), we adjusted the baseline differences across the three language groups by standardising response latencies within each participant, and used the standardised scores for the subsequent analysis, but will report the means for the raw scores in the text.

TABLE 9
Mean standardised-scores for each item pair type in each language in Experiment 4 (Word-picture priming)

Language	*N*	*Item pair type*	*Type 1* *C-cls*	*Type 2* *J-cls*	*Type 3* *C/J-cls*	*Type 4* *C/J-cls-tax*
Chinese	25	Same_classifier	−0.17	−0.06	0.21	0.31
		Control	−0.16	−0.24	−0.12	−0.25
		Taxonomic	0.24	0.23	0.17	0.04
Japanese	23	Same_classifier	−0.10	−0.13	0.10	0.35
		Control	−0.10	−0.30	−0.12	−0.12
		Taxonomic	0.16	0.18	0.08	0.12
German	24	Same_classifier	−0.05	−0.13	0.19	0.13
		Control	−0.26	−0.30	−0.10	−0.21
		Taxonomic	0.32	−0.02	0.32	0.24

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

Downloaded by [Keio University] at 05:51 26 July 2011

Within-language analyses revealed that, as expected, all three groups were slower to respond to the taxonomically related cue/target pairs than to the taxonomically unrelated same_classifier pairs [Chinese: 397 ms vs. 370 ms, $F(1, 24) = 20.09$, $\eta_p^2 = .455$; Japanese: 293 ms vs. 277 ms, $F(1, 22) = 6.79$, $\eta_p^2 = .235$; German: 367 ms vs. 341 ms, $F(1, 23) = 16.37$, $\eta_p^2 = .414$, all ps $<.05$] and the taxonomically unrelated control pairs, [Chinese: 397 ms vs. 354 ms, $F(1, 24) = 68.23$, $\eta_p^2 = .740$; Japanese: 293 ms vs. 268 ms, $F(1, 22) = 46.57$, $\eta_p^2 = .680$; German: 367 ms vs. 316 ms, $F(1, 23) = 61.74$, $\eta_p^2 = .729$, all ps $<.05$]. Furthermore, speakers of all three languages showed longer response latencies for cue/target pairs from the same classifier category (averaged across Type 1–4) than for pairs from different classifier categories, [Chinese: 381 ms vs. 360 ms, $F(1, 24) = 16.78$, $\eta_p^2 = .411$; Japanese: 285 ms vs. 276 ms, $F(1, 22) = 10.80$, $\eta_p^2 = .330$; German: 347 ms vs. 328 ms, $F(1, 23) = 14.97$, $\eta_p^2 = .395$, all ps $<.01$]. These results indicate that speakers of Chinese, Japanese, and German were all sensitive to the similarity underlying classifier categories in the task accessing online cognitive processing.

To test the language-specific amplified classifier effects, we again conducted a set of Hierarchical Linear Models. Table 10 shows the parameter estimates and statistics of the models.

Was there an amplified classifier similarity effect in Chinese speakers? To examine whether classifier relations in Chinese yielded a larger Priming effect

TABLE 10
Estimates for fixed effects and variance components in the Hierarchical Linear Models for Experiment 4 (Word-picture priming)

Contrasts		*Fixed effects*		*Variance components*	
		*Estimate (*SE*)*	T	*Estimate*	χ^2
Type 1 & $3_{(C\ vs.\ G)}$	Intercept	0.168 (0.075)	2.229*	0.335	1250.94**
	Classifier × language	−0.072 (0.064)	−1.120	0.332	1228.83**
Type 1 & $2_{(J\ vs.\ G)}$	Intercept	−0.168 (0.073)	−2.306	0.294	1049.29**
	Classifier × language	0.023 (0.076)	0.307	0.246	890.32**
Type $1_{(C\ vs.\ J)}$	Intercept	−0.182 (0.068)	−2.674	0.174	224.99**
	Classifier × language	0.084 (0.061)	1.373	0.303	519.16**
Type $4_{(C\ vs.\ G)}$	Intercept	0.351 (0.079)	4.400**	0.336	422.41**
	Classifier × language	−0.332 (0.116)	−2.860**	0.330	413.16**
Type $4_{(J\ vs.\ G)}$	Intercept	0.032 (0.075)	0.428	0.286	386.63**
	Classifier × language	−0.129 (0.124)	−1.041	0.241	334.62**

Note: Type 1 = same classifier in Chinese and Japanese; Type 2 = same classifier only in Japanese; Type 3 = same classifier only in Chinese; Type 4 = same classifier in Chinese and Japanese and taxonomically related; C = Chinese; J = Japanese; G = German.
$^{+}p < .10$; $^{*}p < .05$; $^{**}p < .01$.

in Chinese speakers than in German speakers, we contrasted the same_ classifier pairs in Chinese (average of Type 1 and Type 3 same_classifier pairs) with response latencies for the corresponding control pairs (average of Type 1 and Type 3 control pair) and tested the effect of the Contrast × Language (Chinese vs. German) interaction. Unlike Experiments 1 and 2, no language-specific effect of classifier relations on response latencies was found. The magnitude of the classifier effect in Chinese speakers (i.e., the contrast between object pairs belonging to the same classifier category in Chinese and the corresponding control pairs, 373 ms vs. 364 ms) was not larger than in German speakers (347 ms vs. 318 ms, respectively).

Was there an amplified classifier similarity effect in Japanese speakers? We then contrasted latencies for Japanese same_classifier pairs (average of Type 2 and Type 3 same_classifier pairs) with latencies for corresponding control pairs (average of Type 2 and 3 control pairs) and tested whether there was a significant Contrast × Language (Japanese vs. German) interaction. This effect was not significant. Thus, again, there was no amplified classifier effect in Japanese speakers (Same classifier in Japanese: 280 ms vs. control: 266 ms) as compared to German speakers (345 ms vs. 322 ms).

Was there a difference between the magnitudes of the classifier similarity effect in Chinese and Japanese speakers? We next tested whether the

Downloaded by [Keio University] at 05:51 26 July 2011

classifier similarity effect differed between Chinese and Japanese speakers with regard to the same object pairs. As in the earlier experiments, we set up a contrast pitting the same_classifier pairs in Type 3 sets against the Type 3 control pairs and tested whether there was a significant Contrast × Language (Chinese vs. Japanese) interaction effect. Again, this interaction was not significant, suggesting that the magnitude of the classifier similarity effect did not differ between Japanese (Same classifier: 287 ms vs. control: 274 ms) and Chinese speakers (Same classifier: 387 ms vs. control: 370 ms).

Was there a language-specific amplified classifier similarity effect in Chinese and Japanese speakers when pairs also had taxonomic relations? Finally, we tested whether the magnitude of the classifier effect differed between the three language groups when object pairs were taxonomically related. For this purpose, we contrasted the Type 4 same_classifier pairs against the Type 4 control pairs and tested the Contrast × Language interaction. This interaction turned out to be significant when contrasting Chinese and German speakers (Chinese: 415 ms vs. 375 ms; German: 359 ms vs. 359 ms). However, no amplified classifier effect was found in Japanese speakers (307 ms vs. 301 ms).

In summary, except for the same_classifier pairs in Type 4 sets, no language-specific classifier effect was identified either in Chinese or Japanese speakers, replicating the results by Saalbach and Imai (2007). We then tested whether a language-specific classifier priming effect would be obtained when the noun was presented in a classifier phrase. As there are no classifiers in German, we conducted Experiment 5 only with Chinese and Japanese speakers.

Downloaded by [Keio University] at 05:51 26 July 2011

EXPERIMENT 5: SPEEDED PHRASE-PICTURE MATCHING TASK

Method

Participants

Twenty-four Chinese undergraduates and 25 Japanese undergraduates participated in this study.

Materials and procedure

The material and procedure of this experiment were the same as in Experiment 4, with one exception: In this case, the cue to be presented was a classifier phrase (instead of a bare noun). Each noun phrase consisted of the numeral "one", a classifier, and a noun (e.g., one [CLASSIFIER] table). As in Experiment 4, participants were instructed to respond as fast and as accurately as possible, by pressing one of the keys assigned to "Yes" or "No", respectively, to indicate whether the word (presented in the classifier phrase) and the target picture matched.

Results

As in Experiment 4, the proportion of error responses was low in both language groups (less than 10% in both language groups), with an extremely skewed distribution. We therefore concentrated on analyses of response latencies. As in Experiment 4, filler trials, positive trials (in which the word cue and the target picture matched) and error trials were excluded from the analysis of response latencies.

Table 11 shows the mean standardised response latencies for each of the three language groups for all stimulus types. As in Experiment 4, we used the standardised scores for the analyses to adjust differences between the two language groups in the baseline latencies.

As in the previous experiments, we started out with within-language analyses. As expected, participants in both language groups showed significantly longer response latencies for the taxonomically related cue/target pairs than for the taxonomically unrelated same_classifier pairs [Chinese: 344 ms vs. 327 ms, $F(1,\ 23) = 6.65$, $\eta_p^2 = .257$; Japanese: 334 ms vs. 308 ms, $F(1,\ 24) = 22.17$, $\eta_p^2 = .479$; both ps $< .01$] and for the control pairs [Chinese: 344 ms vs. 305 ms, $F(1,\ 23) = 41.06$, $\eta_p^2 = .639$; Japanese: 334 ms vs. 294 ms, $F(1,\ 24) = 94.69$, $\eta_p^2 = .793$; both ps $< .01$]. More importantly, speakers of both groups were slower to react to cue/target pairs from the same classifier category (Type 1–4) than to pairs from different classifier

Downloaded by [Keio University] at 05:51 26 July 2011

TABLE 11
Mean standardised-scores for each item pair type in each language in Experiment 5 (Phrase-picture priming)

Language	*N*	*Item pair type*	*Type 1 C-cls*	*Type 2 J-cls*	*Type 3 C/J-cls*	*Type 4 C/J-cls-tax*
Chinese	24	Same_classifier	−0.06	0.00	0.17	0.18
		Control	−0.23	−0.29	−0.14	−0.13
		Taxonomic	0.27	0.10	0.15	0.17
Japanese	25	Same_classifier	−0.08	−0.12	−0.05	0.28
		Control	−0.25	−0.23	−0.10	−0.24
		Taxonomic	0.30	0.18	0.23	0.17

Note: C-cls = same classifier only in Chinese; J-cls = same classifier only in Japanese; C/J-cls = same classifier in Chinese and Japanese; C/J-cls-tax = same classifier in Chinese and Japanese and taxonomically related; Same_classifier = item pair from the same classifier category of the respective type; Control = unrelated item pair of the respective type; Taxonomic = item pair from the same taxonomic category but from different classifier categories.

categories [Chinese: 333 ms vs. 312 ms, $F(1, 23) = 13.91$, $\eta_p^2 = .375$; Japanese: 316 ms vs. 301 ms, $F(1, 24) = 7.20$, $\eta_p^2 = .233$; both $ps < .05$].

To test whether the magnitude of the classifier effect was larger in speakers of a classifier language than in speakers of a nonclassifier language, we again conducted a set of Hierarchical Linear Models for the contrasts we were interested in, using the German data in Experiment 4 as a nonclassifier language reference. Parameter estimates and statistics are shown in Table 12.

Was there an amplified classifier similarity effect in Chinese speakers? We first tested (by presenting the classifier phrase as a cue) whether the explicit invocation of Chinese classifier categories caused a larger classifier priming effect in Chinese speakers than the bare noun cue in German speakers. As in the bare noun task, the magnitude of the classifier effect was not larger in Chinese speakers than in German speakers, as there was no significant Contrast (average of Type 1 and Type 3 same_classifier pairs vs. average of the Type 1 and Type 3 controls) × Language (Chinese vs. German) interaction (Chinese: 331ms vs. 304 ms, German (bare noun): 347 ms vs. 318 ms).

Was there an amplified classifier similarity effect in Japanese speakers? We also tested whether the classifier effect was stronger in Japanese speakers than in German speakers in this experiment. Again, we did not find an amplified classifier effect in Japanese speakers, as there was no significant Contrast (average of Type 2 and Type 3 same_classifier pairs

TABLE 12
Estimates for fixed effects and variance components in the Hierarchical Linear Models for Experiment 5 (Phrase-picture priming)

		Fixed effects		*Variance components*	
Contrasts		*Estimate (SE)*	T	*Estimate*	χ^2
Type 1 & $3_{(C\ vs.\ G)}$	Intercept	0.048 (0.073)	0.664	0.309	1261.12**
	Classifier × language	–0.026 (0.065)	–0.395	0.308	1249.09**
Type 1 & $2_{(J\ vs.\ G)}$	Intercept	–0.008 (0.088)	–0.098	0.450	1475.88**
	Classifier × language	0.072 (0.067)	1.077	0.442	1420.62**
Type $1_{(C\ vs.\ J)}$	Intercept	–0.043 (0.083)	–0.524	0.389	737.60**
	Classifier × language	0.064 (0.047)	1.361	0.394	733.30**
Type $4_{(C\ vs.\ G)}$	Intercept	0.230 (0.074)	3.077**	0.289	420.13**
	Classifier × language	–0.109 (0.116)	–0.943	0.287	414.30**
Type $4_{(J\ vs.\ G)}$	Intercept	0.204 (0.082)	2.466*	0.363	458.43**
	Classifier × language	–0.213 (0.125)	–1.696	0.357	446.11**

Note: Type 1 = same classifier in Chinese and Japanese; Type 2 = same classifier only in Japanese; Type 3 = same classifier only in Chinese; Type 4 = same classifier in Chinese and Japanese and taxonomically related; C = Chinese; J = Japanese; G = German. *$p < .05$; **$p < .01$.

Downloaded by [Keio University] at 05:51 26 July 2011

vs. average of Type 2 and Type 3 control pairs) and Language (Japanese vs. German) interaction (Japanese: 311 ms vs. 294 ms; German (bare noun): 345 ms vs. 322 ms).

Was there a difference between the magnitudes of the classifier similarity effect in Chinese as compared to Japanese speakers? Again there was no effect when the magnitudes of the contrast between the same_classifier pairs in Type 3 sets (in which the object pairs were drawn from the same classifier categories in both Chinese and Japanese) and the Type 3 control pairs were compared between Chinese and Japanese speakers (Chinese 338 ms vs. 303 ms; Japanese: 319 ms vs. 299 ms).

Was there a language-specific amplified classifier effect in Chinese and Japanese speakers when pairs were taxonomically related? In contrast to the bare noun case, no language-specific effect was found when response latencies for taxonomically related cue/target pairs from the same classifier category were contrasted with taxonomically related pairs from different classifier categories (Type 4 same_classifier pairs vs. Type 4 taxonomic pairs; Chinese: 352 ms vs. 334 ms; Japanese: 341 ms vs. 322 ms, German (bare noun): 359 ms vs. 359 ms).

In summary, the overall findings in the classifier phrase-picture matching task did not differ from findings in the (bare) word-picture matching task. The response latencies of all language groups reflected sensitivity to the similarity underlying classifier categories. But the magnitude of this effect did not differ between classifier and nonclassifier language groups, regardless of whether or not the classifier categories were explicitly invoked.

Discussion

In Experiments 4 and 5, we examined whether there was a language-specific amplified classifier effect in on-line processing, using a fast-speed picture-word matching paradigm in two contexts: a context in which the cue noun was presented without a classifier and a context in which the cue noun was presented with a classifier. In both cases, previous results showing that taxonomic relations between the cue and the target produced a larger delay in response (Saalbach & Imai, 2007; Yokosawa & Imai, 1997) were confirmed. It is important to note that when the two objects belonged to the same classifier category in Chinese or in Japanese, or in both languages, the response was significantly delayed in all of the three language groups, including German speakers. Thus, people seem to detect the similarity underlying classifier categories even in fast automatic processing. However, in contrast to Experiments 1 and 2, there was no language-specific amplified classifier effect in Chinese or in Japanese speakers, as compared to German

Downloaded by [Keio University] at 05:51 26 July 2011

speakers, in either of the two contexts (i.e., with classifiers unspecified or explicitly specified). We will discuss these results in the General Discussion.

GENERAL DISCUSSION

In this research, we examined whether the amplified classifier similarity effect shown in previous research (Saalbach & Imai, 2007) depends on the properties of the classifier language, with the aim of clarifying the conditions under which the effect emerges and, consequently, of gaining insight into how it arises.

An important finding is that the language-specific classifier effect found in previous research with Chinese speakers (Saalbach & Imai, 2007; see also Huettig et al., 2010; Gao & Malt, 2009; Zhang & Schmitt, 1998) cannot be generalised to another classifier language that has fairly comparable semantic structures. Consistent with the results of Saalbach and Imai (2007), we found that not only speakers of a classifier language but also speakers of a nonclassifier language judged objects from the same classifier categories (in either Chinese or Japanese, or in both) to be more similar and more likely to share the target property than pairs of objects that belonged to different classifier categories, and showed a delayed response although this classifier similarity effect was much smaller than the one observed in cases where objects were taxonomically related (but did not have classifier relations). Also consistent with the results of Saalbach and Imai (2007), Chinese speakers showed a language-specific amplified classifier similarity effect in the contexts of (nonspeeded) similarity judgements and inductive inference of a blank property. In contrast, no such language-specific amplified similarity effect was identified in Japanese speakers.

In summary, the whole pattern of results corroborates the conclusions that, contrary to the strong version of the Whorfian hypothesis, classifier systems do not provide speakers with the same way of organising concepts as taxonomic relations do. Classifier category membership may amplify similarity among objects due to the semantic features that underlie the classifier categories. However, unlike the similarity that is characteristic for taxonomic categories, where there is a cluster of interrelated features, the kind of similarity provided by classifier categories is mostly based on a single or, at most, a small number of perceptual or functional features. In this sense, classifier categories are, at best, weak categories.

Accounting for the amplified classifier similarity effect

The present research sheds light on how the amplified classifier similarity effect found in Chinese speakers arises. We replicated our previous finding

Downloaded by [Keio University] at 05:51 26 July 2011

that there is a relation, albeit a weak one, between classifiers and Chinese speakers' conceptual representation, but we also found that this effect cannot be automatically generalised to all classifier languages, at least not with the same degree of magnitude.

The strength of the classifier effect seems to interact with linguistic properties of the classifier system in a given classifier language. What, then, is the linguistic property that most strongly accounts for the amplified classifier similarity effect? Is it a semantic property or a structural one? As discussed above, this is an extremely difficult question given that there is a large variety of classifier categories whose semantic nature (e.g., category size, semantic transparency, and category coherency) is diverse within each language. In both Chinese and Japanese, there are very broad and incohesive categories, many members of which seem to have been included by way of metaphorical extensions from prototypical members (e.g., home runs being included in the *hon* category by association with a bat). Similarly, both languages have relatively tight and semantically cohesive categories that overlap with taxonomic categories. In both languages, broad and incohesive categories are more frequent than tight and cohesive categories. Beyond this, comparing the two languages with respect to semantics may not make much sense, as it would only amount to endless comparisons for each idiosyncratic classifier. In any case, in light of these points, we think the semantic natures of Chinese and Japanese classifier systems are largely comparable. We are therefore inclined to conclude that it is not very likely for the difference in the amplified classifier similarity effect to be due to the difference in the semantic structure of the Chinese and Japanese classifier systems, although we do acknowledge that in the current experiments, the semantic account cannot be definitely ruled out.

Based on linguistic analyses of the classifier systems in Chinese and Japanese in the literature (e.g., Denny, 1979; Downing, 1996; Erbaugh, 1986; Matsumoto, 1985) and on our own analyses of the Japanese-Chinese translation corpus described above, we thus think that the difference in the amplified classifier effect is due to structural rather than semantic properties. As discussed above, classifiers are used in more restricted contexts in Japanese than in Chinese and, as a consequence, appear much less frequently in text and discourse in the former than in the latter language. In harmony with this linguistic property, we found the amplified similarity effect for objects belonging to the same classifier category in Chinese speakers, but not in Japanese speakers. Thus, the effect seems to arise when the subtle inherent similarity between members of the same classifier category has been strengthened through repeated exposure to *indirect* associations with the same classifier. In other words, although it is unlikely for two nouns (e.g., "road" and "necktie", both members of the *tiao* classifier category) to

Downloaded by [Keio University] at 05:51 26 July 2011

Downloaded by [Keio University] at 05:51 26 July 2011

actually co-occur, speakers have, since childhood, been very frequently exposed to each of these nouns in association with the same classifier (e.g., *tiao*), and this association heightens the inherent semantic similarity that can be perceived even by speakers of a nonclassifier language. However, for this effect to occur, exposure to different-nouns/same-classifier associations has to be consistent and *very* frequent. In Japanese, classifiers may not appear frequently enough to build an amplified sensitivity to the similarity among objects from the same classifier category beyond the magnitude observed in speakers of a nonclassifier language.

This interpretation can be related to results from research on category learning which have shown that the frequency of speakers' exposure to a noun as a member of a category is crucial for the strength of the association among category members (e.g., Nosofsky, 1988). The fact that Japanese speakers' responses across the tasks were not different from those of German speakers, who had no exposure to classifiers, is worth noting. It suggests that the relation between the frequency of exposure to the noun-classifier association and the emergence of the amplified classifier similarity effect may not be linear: the amplified classifier similarity effect may only occur if classifier and noun co-occur frequently enough to exceed a certain threshold (as is the case in Chinese). The effect could in part be due to the fact that objects belonging to the same classifier category share an inherent similarity that even speakers of a nonclassifier language can detect. Developing an amplified sense of similarity on top of the inherent similarity might, then, require the presence of fairly strong associations, due to their frequent co-occurrence, of nouns and classifiers.

This reasoning differs from Vigliocco and colleagues' argument that a language-specific effect of grammatical categories arises from speakers' awareness, in the course of language acquisition, of an association between certain grammatical categories and the corresponding meanings (Vigliocco et al., 2005). These authors argue that the systematic correspondence of two grammatical categories of gender (masculine and feminine) and two semantic categories (male and female) for human terms (man, woman, uncle, and aunt, etc.) helps speakers to develop the implicit expectation that other sexuated entities (i.e., animals) belonging to the same grammatical categories should also be conceptually similar. However, when there are, as in German, three grammatical gender classes (masculine, feminine, and neuter) instead of two and, thus, no such direct and transparent correspondences between grammatical categories and conceptual representations, speakers do not develop the notion that entities belonging to the same grammatical gender category are conceptually similar. Vigliocco and colleagues therefore attributed the language-specific gender similarity effect to the semantic factor, i.e., to whether or not the initial mapping between grammatical categories and conceptual representation is transparent.

However, considering the large difference between the two grammatical categorisation systems, the difference between our explanation for the language-specific classifier effect and Vigliocco and colleagues' explanation for the language-specific gender effect is not surprising. With classifier systems, the correspondence between grammatical categories and semantics is generally much more visible than with gender grammar systems. Chinese and Japanese classifier systems are not likely to differ in this respect, as the semantics in both systems was transparent enough even for German speakers to detect the similarity among classifier category members. In the case of gender systems, however, transparency of the correspondence between grammatical and conceptual classes may be important for boosting the language-specific similarity effect, since different gender languages differ in this respect. Besides semantic transparency, classifier systems differ from gender grammar systems in how consistently nouns are marked with their grammatical class. In gender grammar systems, nouns' gender classes are very consistently marked, and a noun belongs to only one gender class. In contrast, in classifier grammar systems, nouns are not always accompanied by classifiers, and a noun is sometimes associated with more than one classifier. Under these circumstances, and given that classifiers are more frequently used in Chinese than in Japanese, the strength of the association between noun and classifier indeed seems to matter for boosting the amplified similarity effect.

Classifier relations and taxonomic relations

Does the amplified classifier similarity effect also apply to pairs which not only share classifier relations but also taxonomic relations? Our results suggest that this is not the case. The (subtle) effect of classifier relations seems to be superseded by the similarity that arises from taxonomic relations. In other words, when the object pairs already have taxonomic relations, the classifier relation will not add similarity on top of it.

In Experiment 1 (similarity judgement), however, we unexpectedly found a language-specific classifier similarity effect in Japanese speakers for pairs which shared both classifier and taxonomic relations (Type 4 same_classifier pairs). The effect, however, does not seem to be robust since no such effect was found in the inductive inference task with a blank property (Experiment 2a) which otherwise revealed an overall pattern similar to that of the similarity judgement task. This finding is thus difficult to interpret and needs to be further examined in future research.

Downloaded by [Keio University] at 05:51 26 July 2011

Absence of the amplified classifier similarity effect in online processing

What remains to be explained is why the amplified classifier similarity effect in Chinese speakers was not found in the fast-speed word-picture task—even when the classifier was explicitly specified with the cue noun. The null effect cannot be attributed to the insensitivity of the paradigm *per se*, as a strong effect (delay) was found when the cue was taxonomically related to the target. Also, a delay (while not as large as for the taxonomic items) was consistently found for the same_classifier pairs, as compared to the control pairs, in speakers of all three languages.

The fact that no language-specific amplified classifier effect occurred in Chinese speakers even in the presence of the classifier may appear to be at odds with the results of previous research. Zhou and colleagues (Zhou, Jiang, Zhang, Lou, & Ye, 2010) demonstrated that the violation of the classifier-noun agreement resulted in a large N400 effect and suggested that classifiers semantically constrain the subsequent noun in online processing. Furthermore, in their eye-movement study, Huettig and colleagues (Huettig et al., 2010) found shifts of eye-gaze to same-classifier objects only when the noun, presented with an auditory stimulus, was embedded in a classifier phrase within a sentence. These studies may thus suggest that classifiers automatically activate candidate nouns even before speakers actually hear the noun in online sentence processing. Automatic activation of the semantic features underlying classifiers may, however, occur only in the context of sentence processing, where participants continuously predict what comes next in the sentence. Our task did not involve sentence processing nor did it require participants to predict nouns from a classifier. Instead, it required them to automatically activate a target object from a given *cue noun*. A noun would activate many semantic features relevant to different conceptual relations, such as taxonomic or thematic relations, besides the features relevant to the classifier. As the same_classifier relation is much weaker than taxonomic or thematic relations (Saalbach & Imai, 2007), the activated features relevant to the classifier may have been superseded, given a very short processing time, by other semantic features relevant to more prominent conceptual relations.

Implications for language and thought

In Saalbach and Imai (2007), we highlighted the importance of contextualising the effect of given linguistic categories by comparing it to the effect of other kinds of conceptual relations in a range of different cognitive tasks. There, we revealed the nature and the magnitude of the cognitive effect that may have emerged due to the presence of a classifier grammar system. By

Downloaded by [Keio University] at 05:51 26 July 2011

comparing two classifier languages that differ in the frequency of classifier use, the present research sheds further light on the nature of the language-specific amplified classifier effect as well as on how this effect emerges. We replicated the language-specific amplified similarity effect in Chinese speakers for the same tasks and with the same magnitudes as in the previous research, but not in Japanese speakers.

These findings suggest that the relation between linguistic categories and thought (including both conceptual representation and cognitive processes) is not one-fold. If we find an effect in one cognitive task in a language, we should not automatically assume that this effect applies to other cognitive tasks, as well, nor should we assume that it can be generalised to other languages that have the same or comparable grammatical categories. The findings underscore the importance of examining the relation between linguistic categories and thought not in terms of *whether* there is one: rather, the conditions under which the effect arises should be specified, along with the magnitude and impact of the effect in a global picture of cognition and conceptual structures. This can only be achieved by testing the target effect in various task contexts and by comparing multiple languages that have the grammatical categories in question.

In fact, given the findings from the present research concerning the influence of classifier categories on overall object categories, it will be interesting to revisit previous work examining the influence of language on the construal of individuation (e.g., Imai & Gentner, 1997; Imai & Mazuka, 2007; Lucy, 1992; Mazuka & Friedman, 2000). These studies compared English with only one classifier language (Yucatec Mayan in Lucy's and Japanese in Imai's and Mazuka's studies). Further insights into the relation between language and the construal of individuation might be gained by extending these studies to include other classifier languages, using the same stimuli and procedure.

The fact that German participants judged objects belonging to the same classifier category to be more similar than unrelated objects should also be taken into account when studying the relation between language and cognition, as it supports the notion that grammatical categories are *motivated* (but of course not determined) *by* universally shared cognitive and perceptual experiences (e.g., Zubin & Köpcke, 1986; see also Malt et al., 2008 and Majid, Bowerman, van Staden, & Boster, 2007 for evidence that cross-linguistic diversity in lexical categories is also constrained by universally shared experiences). In this sense, our results are not incongruent with the view that grammatical categories reflect cognitive categories (e.g., Lakoff, 1987). Two aspects of our findings—that similarity underlying classifier categories can be detected by German speakers, on the one hand, and that similarity due to classifier relations is magnified by Chinese

Downloaded by [Keio University] at 05:51 26 July 2011

speakers, on the other hand—cogently suggest that the relation between language and thought is not unidirectional: linguistic categories reflect universally perceived commonalities in the world, but at the same time modify universally perceived similarities (see Imai & Mazuka, 2007; Majid, 2010; Regier, Kay, Gilbert, & Ivry, 2010; Roberson & Hanley, 2010, for a relevant discussion).

Manuscript received January 2010
Revised manuscript received January 2010
First published online month/year

Downloaded by [Keio University] at 05:51 26 July 2011

REFERENCES

Adams, K., & Conklin, N. F. (1973, April). *Toward a theory of natural classification.* Paper presented at the 9th Regional meeting of the Chicago Linguistic Society, Chicago.

Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun classification devices.* Oxford, UK: Oxford University Press.

Allan, K. (1977). Classifiers. *Language*, *53*(2), 285–311.

Beijing Center for Japanese Studies. (2003). *Chinese-Japanese parallel corpus.*

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think differently about time? *Cognition, 118*, 123–129.

Bowerman, M., & Levinson, S. C. (2001). *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press.

Craig, C. G. (1986). Jacaltec noun classifiers: A study in language and culture. In C. Craig (Ed.), *Noun classes and categorization* (pp. 263–293). Philadelphia, PA: John Benjamins.

Croft, W. (1990). *Typology and universals.* New York: Cambridge University Press.

Croft, W. (1994). Semantic universals in classifier systems. *Word*, *45*, 145–171.

Denny, J. P. (1979). Semantic analysis of selected Japanese numeral classifiers for units. *Linguistics*, *17*, 317–335.

Downing, P. (1996). *Numeral classifier systems: The case of Japanese*. Philadelphia, PA: John Benjamins.

Erbaugh, M. S. (1986). Taking stock: The development of Chinese noun classifiers historically and in young children. In C. Craig (Ed.), *Noun classes and categories* (pp. 399–436). Philadelphia, PA: John Benjamins.

Gao, Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes*, *24*, 1124–1179.

Gentner, D., & Goldin-Meadow, S. (2003). Whiter Whorf. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 3–14). Cambridge, MA: MIT Press.

Gomez-Imbert, E. (1996). When animals become "rounded" and "feminine": Conceptual categories and linguistic classification in a multilingual setting. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 438–469). Cambridge, UK: Cambridge University Press.

Grinevald, C. (2000). A morphosyntactic typology of classifiers. In G. Senft (Ed.), *Systems of nominal classification* (pp. 50–92). Cambridge, UK: Cambridge University Press.

Gumperz, J. J., & Levinson., S. C. (1996). *Rethinking linguistic relativity.* Cambridge, UK: Cambridge University Press.

Guo, X. Z. (2002). *Xiandai hanyu liangci yongfa cidian* [Modern Chinese classifier usage dictionary]. Beijing: Yuwen chubanshe.

Huettig, F., Chen, J., Bowerman, M., & Majid, A. (2010). 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*, 39–58.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, *98*(3), 377–389.

Iida, A. (2004). *Kasoe-kata no jiten* [Dictionary for counting things]. Tokyo, Japan: Shogakukan.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*(2), 169–200.

Imai, M., & Mazuka, R. (2007). Revisiting language universals and linguistic relativity: Language-relative construal of individuation constrained by universal ontology. *Cognitive Science*, *31*, 385–414.

Imai, M., Saalbach, H., & Stern, E. (2010). Are Chinese and German children taxonomic, thematic or shape biased?: Influence of classifiers and cultural contexts. *Frontiers in Cultural Psychology*, 1:194. doi: 10.3389/fpsyg.2010.00194

Kay, P., & Regier, T. (2006). Language, thought, and color: Recent developments. *Trends in Cognitive Sciences*, *10*(2), 51–54.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Lamarre, C. (2008). Gengo henka to kinougo [Language change and evolution of function words: The case of classifiers in Japanese and Chinese]. In T. Hasegawa, C. Lammer & T. Ito (Eds.), *Kokoro to kotoba: Shinka to ninchikagaku no apuroochi* [Evolutionary and cognitive approaches to language] (pp. 63–78). Tokyo: University of Tokyo Press.

Li, P., & Gleitman, L. R. (2002). Turning the tables: Spatial language and spatial reasoning. *Cognition*, *83*(3), 265–294.

Lucy, J. A. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. New York, NY: Cambridge University Press.

Majid, A. (2010). Words for parts of the body. In B. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 58–71). New York: Oxford University Press.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Science*, *8*, 108–114.

Majid, A., Bowerman, M., Van Staden, M., & Boster, J. S. (2007). The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics*, *18*(2), 133–152.

Malt, B. C., Gennari, S., Imai, M., Ameel, E., Tsuda, N., & Majid, A. (2008). Talking about walking: Biomechanics and the language of locomotion. *Psychological Science*, *19*(3), 232–240.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*, 230–262.

Malt, B., & Wolff, P. (2010). *Words and the mind: How words capture human experience*. New York: Oxford University Press.

Matsumoto, Y. (1985). Acquisition of some Japanese numeral classifiers: The search for convention. *Papers and Reports on Child Language Development*, *24*, 76–86.

Mazuka, R., & Friedman, R. S. (2000). Linguistic relativity in Japanese and English: Is language the primary determinant in object classification? *Journal of East Asian Linguistics*, *9*, 353–377.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Natsume, S. (1964). *Bocchan 'Master Daring'*. Tokyo: Kaisei-sha.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, *90*, 339–363.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 54–65.

Downloaded by [Keio University] at 05:51 26 July 2011

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Newbury Park, CA: Sage.

Regier, T., & Kay, P. (2009). Language, thought and color: Whorf was half right. *Trends in Cognitive Science*, *13*, 439–446.

Regier, T., Kay, P., Gilbert, A. L., & Ivry, R. B. (2010). Language and thought. Which side are you on, anyway? In B. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 165–182). New York: Oxford University Press.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. (2005). Colour categories in Himba: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*, 378–411.

Roberson, D., & Hanley, R. (2010). Relatively speaking. An account of the relationship between language and thought in the color domain. In B. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 183–198). New York: Oxford University Press.

Rowling, J. K. (1999). *Harry Potter and the chamber of secrets*. New York: Arthur A. Levine.

Saalbach, H., & Imai, M. (2007). The scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, *136*(3), 485–501.

Senft, G. (1996). *Classificatory particles in Kilivila*. New York: Oxford University Press.

Senft, G. (2000). What do we really know about nominal classification systems? In G. Senft (Ed.), *Systems of nominal classification* (pp. 11–50). Cambridge, UK: Cambridge University Press.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.

Vigliocco, G., Vinson, D., Indefrey, P., Levelt, W., & Hellwig, F. (2004). The interplay between meaning and syntax in language production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 483–497.

Vigliocco, G., Vinson, D., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General*, *134*(4), 501–520.

Yokosawa, K., & Imai, M. (1997, April). *Conceptual interference on object verification: Effect of superordinate vs. thematic associative relations.* Paper presented at the Meeting of the Asssociation for Research in Vision and Ophthalmology, Fort Lauderdale, FL.

Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, *4*(4), 375–385.

Zhou, X., Jiang, X., Ye, Z., Zhang, Y., Lou, K., & Zhan, W. (2010). Semantic integration processes at different levels of syntactic hierarchy during sentence comprehension: An ERP study. *Neuropsychologia*, *48*(6), 1551–1562.

Zubin, D., & Köpcke, K.-M. (1986). Gender and folk taxonomy: The indexical relation between grammatical gender and lexical categorization. In C. Craig (Ed.), *Noun classes and categorization* (pp. 139–180). Philadelphia, PA: Benjamins.

Downloaded by [Keio University] at 05:51 26 July 2011

APPENDIX 1: COMPLETE STIMULUS SET OF EXPERIMENTS 1–5

TABLE A1
Stimuli pairs of Type 1 (C-cls: same classifier in Chinese only)

	Classifier		*Item*		
Target	*Chinese*	*Japanese*	*Same Classifier*	*Control*	*Taxonomic*
Flower	*Duo*	*Hon*	Cloud	Cup	Tree
Comb	*Ba*	*Hon*	Key	Ticket	Hair dryer
Pistol	*Ba*	*Tyou*	Umbrella	TV	Cannon
Newspaper	*Zhang*	*Bu*	Bed	Tube	Book
Mountain	*Zuo*	*Ko*	Tower	Necklace	Hill
Piano	*Jia*	*Dai*	Ladder	Scarf	Violin
Plane	*Jia*	*Ki*	Swing	Chain	Boat
Necklace	*Tiao*	*Hon*	Blanket	Book	Rope
Chain	*Tiao*	*Hon*	Carp	Poster	Trumpet
Drum	*Mian*	*Ko*	Wall	Scissors	Ring
Scissors	*Ba*	*Hon*	Fan	TV	Cutter
Towel	*Tiao*	*Mai*	Eel	Potato	Handkerchief
Bell	*Zuo*	*Ko*	Building	Bike	Buzzer
Tent	*Ding*	*Tyou*	Hat	Table	Sleeping bag

Note: Same_classifier = item from the same classifier category as the target; Control = item unrelated to the target; Taxonomic = item from the same taxonomic category as the target but from different classifier categories; "Taxonomic" target-item pairs have only been used in Experiments 4 and 5.

Downloaded by [Keio University] at 05:51 26 July 2011

TABLE A2
Stimuli pairs of Type 2 (J-cls: same classifier in Japanese only)

	Classifier		*Item*		
Target	*Chinese*	*Japanese*	*Same Classifier*	*Control*	*Taxonomic*
Bicycle	*Liang*	*Dai*	Camera	Stick	Pedal boat
Bus	*Liang*	*Dai*	TV	Hat	Boat
Desk	*Zhang*	*Dai*	Calculator	Screw driver	Chair
Window	*Men*	*Mai*	Telephone card	Swing	Door
Shirt	*Jian*	*Mai*	Mirror	Teeth	Pants
Envelope	*Zhang*	*Mai*	Sheet	Cloud	Parcel
Skirt	*Tiao*	*Mai*	Ticket	Bed	Jacket
Fork	*Ba*	*Hon*	Matches	Ladder	Chopstick
Axe	*Ba*	*Hon*	Teeth	Bean	Power saw
Pencil	*Tiao*	*Hon*	Thread	CD	Typewriter

Note: Same_classifier = item from the same classifier category as the target; Control = item unrelated to the target; Taxonomic = item from the same taxonomic category as the target but from different classifier categories; "Taxonomic" target-item pairs have only been used in Experiments 4 and 5.

Downloaded by [Keio University] at 05:51 26 July 2011

TABLE A3
Stimuli pairs of Type 3 (C/J-cls: same classifier in Chinese and Japanese)

	Classifier		*Item*		
Target	*Chinese*	*Japanese*	*Same Classifier*	*Control*	*Taxonomic*
Pillar	*Gen*	*Hon*	Stick	Name card	Beam
Bone	*Gen*	*Hon*	Tube	Plaster	Muscle
Arrow	*Gen*	*Hon*	Cucumber	Table	Bullet
Comb	*Ba*	*Hon*	Knife	Calculator	Hair drier
Jump rope	*Gen*	*Hon*	Banana	Bike	Apple
Fridge	*Tai*	*Dai*	Computer	Knife	Vine cellar
Money bill	*Zhang*	*Mai*	CD	Matches	Coins
Map	*Zhang*	*Mai*	Name Card	Chain	Globe
Bean	*Li*	*Tsubu*	Sand	Key	Peppers
Sword	*Ba*	*Hon*	Umbrella	Stamp	Cannon

Note: Same_classifier = item from the same classifier category as the target; Control = item unrelated to the target; Taxonomic = item from the same taxonomic category as the target but from different classifier categories; "Taxonomic" target-item pairs have only been used in Experiments 4 and 5.

TABLE A4
Stimuli pairs of Type 4 (C/J-cls-tax) same classifier in Chinese and Japanese and same taxonomy)

	Classifier		*Item*		
Target	*Chinese*	*Japanese*	*Same Classifier*	*Control*	*Taxonomic*
Bed	*Zhang*	*Dai*	Table	Wire	Chair
Skirt	*Tiao*	*Mai*	Scarf	Stereo	Hat
Car	*Liang*	*Dai*	Bicycle	Table	Boat
Spoon	*Ba*	*Hon*	Knife	Street	Chopstick
Pliers	*Ba*	*Hon*	Skew driver	Scarf	Power drill
CD	*Zhang*	*Mai*	Floppy	Spoon	Tape
River	*Tiao*	*Hon*	Channel	Knife	Lake
Monitor	*Tai*	*Dai*	Printer	Pillow	Mouse
Banana	*Gen*	*Hon*	Long onion	Screw drive	Tomato
Carrot	*Gen*	*Hon*	Cucumber	Boat	Potato

Note: Same_classifier = item from the same classifier category as the target; Control = item unrelated to the target; Taxonomic = item from the same taxonomic category as the target but from different classifier categories; "Control" target-item pairs have only been used in Experiments 4 and 5.

Downloaded by [Keio University] at 05:51 26 July 2011