

On the Non-Uniformity of Asian Thinking (for Speaking): A Response to Masuda and Nisbett

Nigel Duffield & Yayoi Tajima
University of Sheffield & Keio University

1 Theoretical Background

We begin with an important clarification.¹ This paper is concerned with Cultural Relativity (CR) in a very precise sense, namely, with the putative effects of native culture on core aspects of basic cognition: visual perception and visual recall. We are *not* concerned with the possible impacts of culture—however this notion is defined—on higher level cognition: cross-cultural variation in social and ethical norms, belief attribution, notions of truth, *etc.* In short, this paper addresses questions of Cognitive, rather than Moral, Psychology.

The reason for stressing this is that there exist two distinct interpretations of Cultural Relativism: the first an objective one, to which it seems quite reasonable to subscribe; the second more subjective, more normative—and considerably more problematic. An informal Internet search reveals both positions: there are also many instances where the two are conflated. Consider first the objective definition in example (1), from *Wikipedia*:

- (1) Cultural relativism is the principle that an individual human's beliefs and activities should be understood in terms of his or her own culture. This principle was established as axiomatic in anthropological research by Franz Boas in the first few decades of the 20th century and later popularized by students. Boas himself did not use the term as such, but the term became common among anthropologists after Boas' death in 1942.

So defined, CR could be viewed as no more nor less than a guiding principle of classic Saussurean Structuralism: understanding a system, society or language in terms of its internal relations. This perspective is in stark opposition to the following presentation (from a website with an overt religious agenda):

- (2) Cultural Relativism: Truth Is Relative. Cultural relativism is the view that all beliefs are equally valid and that truth itself is relative, depending on the situation, environment, and individual. Those who hold to cultural relativism hold that all religious, ethical, aesthetic, and political beliefs are completely relative to the individual within a cultural identity. Relativism often includes moral relativism (ethics depend on a social construct), situational relativism (right or wrong is based on the particular situation), and cognitive relativism (truth itself has no objective standard). (<http://www.cultural.relativism.com/>)

¹ This article is based on research carried out by the second author as part of her 2008 MA thesis (Tajima 2008): the discussion and argument have been jointly constructed. A longer version of this paper is available for download at <http://www.ngduffield.staff.shef.ac.uk/index.html>. We should like to acknowledge a research project on *Culture and the Mind*, directed by Stephen Laurence (University of Sheffield), funded by the Arts and Humanities Research Council (UK): it was through participation in the workshops and reading groups associated with this project that the first author became aware of Richard Nisbett's work, and of the significance accorded to it by anthropologists, social psychologists and in the popular media.

The authors of this article represent CR as asserting that there are no absolute or objective standards, and thus implicitly that cross-cultural comparisons have no meaning: your version of truth is neither better nor worse than mine, it's simply different, and we can't argue about it. Needless to say, we do not subscribe to this view, though our reasons are very different from the article's authors, who also oppose it (on theological grounds).

If these two definitions (objective and normative) were kept distinct, matters would be a lot simpler. However, very commonly they are conflated. For example, the *Encarta* definition reproduced in (3) includes both, while Figure 1, from Rosado (1994)—another religious commentator—associates opponents of cultural relativism with closed-mindedness, cultural insensitivity and ethnocentrism. If nothing else, this figure shows just how laden the term has become.

- (3) Definition: judging cultures on their own terms: the principle that people should not judge the behavior of others using the standards of their own culture, and that each culture must be analyzed on its own terms.

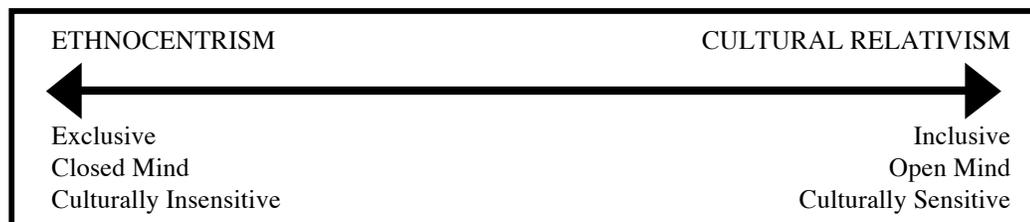


Figure 1. Representing Cultural Relativism (adapted from Rosado (1994))

This present article is exclusively concerned with the objective use of the term. The research questions we are concerned with relate to perception and memory, as follows:

- Do people in different communities literally see the world differently? Do they attend to different aspects of situations and events, and conceptualize those events in distinct ways?
- If the answer to these questions is affirmative, is this in virtue of their cultural or their linguistic environment?
- Finally, does either culture or language *determine* particular modes of thought, or simply *predispose* people to particular ways of thinking?

1.1 Nisbett's Cultural Dichotomy

The impetus for the present experiments is a body of work by Richard Nisbett and his co-workers, of which Nisbett (2003) is the most accessible to a popular audience. In this work, Nisbett asserts, as the subtitle of his 2003 book makes clear, that Asians and Westerners think—or at least “see the world”—differently: it is claimed that observed behavioral differences in visual recall tasks between various groups of Asian and Western participants are in large measure attributable to broad-range cultural differences that inform distinct ways of thinking in Asian and Western minds (Holistic *vs.* Analytic Thought). The quotes below are illustrative:²

² These lines attributed to Nisbett, have been widely reproduced in blogs and other commentaries across the Internet. The original source seems to be <http://www.wired.com/culture/lifestyle/news /2005/08/68626>.

Asians live in a more socially complicated world than we do...They have to pay more attention to others than we do. We are individualists. We can be bulls in a china shop, they can't afford it...The key thing in Chinese culture is harmony, Nisbett said, while in the West the key is finding ways to get things done, paying less attention to others.

These conclusions follow from a number of different empirical studies reported in Nisbett *et al* (2001). Here, we focus on only one of the key articles cited in that paper, namely, Masuda & Nisbett (2001, 2003). Our results thus speak to only one piece of the evidence for Nisbett's claim, a point we return to at the end. In the Masuda & Nisbett study, Japanese and American students were asked to observe and provide oral descriptions of underwater scenes ("animated vignettes"), and subsequently to recall whether they had seen particular features of a scene in following presentations in which either the background or the focal fish were substituted. It was found that Japanese participants tended to report significantly more—and to recall more accurately—on visual information available in the *Ground* of a given scene, whereas American participants tended to focus on *Figure* information at the expense of the Ground. American participants also invariably began their descriptions by mentioning the most salient object, whereas Japanese participants began by mentioning background elements. According to the interpretation of Masuda & Nisbett (2003), this contrast was explained by the fact that Asians 'think holistically'—all elements of the scene are equally important—whereas Americans 'think analytically' and thus tend to abstract away from the details of particular situations.

Westerners are inclined to attend to some focal object, analyzing its attributes and categorizing it in an effort to find out what rules govern its behavior. Rules used include formal logic. Causal attributions tend to focus exclusively on the object and are therefore often mistaken...East Asians are more likely to attend to a broad perceptual and conceptual field, noticing relationships and changes and grouping objects based on family resemblance rather than category membership. Causal attributions emphasize the context...(Masuda & Nisbett 2003: 11163).

These quotes make clear that, in Masuda & Nisbett's view, the explanation for these perceptual differences lies in culture. In our view, however, it is at least possible that a significant part of the explanation lies elsewhere, namely, in grammatical and discourse structure. Simply put, our contention is that many of the between-group differences observed in reporting and recall tasks in the Masuda & Nisbett study are attributable to the fact that Japanese people are Japanese speakers, and that Americans are speakers of English, rather than to any deeper cultural, social or environmental factors.

There is a way to tease apart these alternatives. As it turns out, Nisbett's putative cultural dichotomy is orthogonal to significant linguistic contrasts among Asian languages. Specifically, within the Asian group, there is a grammatical-typological split between HEAD-INITIAL languages such as Chinese, on one hand, vs. adjunct-initial/HEAD-FINAL languages like Japanese or Korean, on the other.³ If grammatical structure rather than cultural background plays a significant role in explaining the behavioral contrast between Japanese and English participants in Masuda & Nisbett's study, then native Japanese and native Chinese participants should diverge from each other in their results at least as much as from speakers of Western languages when the three groups are tested together. This is what our experiments examined.

1.2 An Alternative Approach: Weak Linguistic Relativity

Our research adopts a particular stance on the issue of Linguistic Relativity that may be termed Weak Linguistic Relativity (WLR). Roughly, this is the notion that language structure *predisposes* its speakers to particular conceptualizations of events and situations, and in ways that may influence attentional

³ See below for further discussion of the claim that Chinese is head-initial.

mechanisms. Such an hypothesis is opposed to two more radical positions. On the one hand, WLR is contrasted with Strong Universalism, which takes it that there is only one universal “Language of Thought”, and denies that language can penetrate cognition in any interesting way. Strong Universalism is commonly identified with Fodor (1975, 1983), also Chomsky (1975, and elsewhere). On the other hand, WLR is also opposed to Strong Linguistic Relativity, the idea that language constrains or ultimately determines thought.

The latter position is often identified with the classic work of Benjamin Whorf and Edward Sapir. In our view, however, such identification may be misplaced: it is at least arguable that the Sapir-Whorf hypothesis, like Cultural Relativity, has been misconstrued. Certainly, there is nothing in the one of the most famous quotes concerning the Sapir-Whorf hypothesis—from Sapir (1958)—that necessarily implies such a strong form of linguistic determinism:

Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society. It is quite an illusion to imagine that one adjusts to reality essentially without the use of language and that language is merely an incidental means of solving specific problems of communication or reflection. The fact of the matter is that the ‘real world’ is to a large extent unconsciously built upon the language habits of the group. No two languages are ever sufficiently similar to be considered as representing the same social reality. The worlds in which different societies live are distinct worlds, not merely the same world with different labels attached...We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation (Sapir 1958 [1929], p. 69.)

To say that “We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation” *can* be interpreted in a strong deterministic way, but it need not be: note especially the crucial use of the term “predispose” rather than “force” or “require”. Notice also the many hedges here: “to a large extent”, “very largely” and so forth. Whatever Sapir really intended, the position that we shall defend is markedly more tempered in tone.

1.2.1 ‘Thinking for Speaking’

The particular form of WLR we assume is the THINKING FOR SPEAKING hypothesis (TfS) proposed and advocated by Dan Slobin in a series of recent papers, including Slobin (1996, 2000, 2003). TfS embodies the claim that language structure affects conceptualization, at least at the moment of speaking. As Slobin (2003: 158) has it:

It has become clear to me that “we encounter the contents of the mind in a special way when they are being accessed for use” (Slobin 1987, 435). That is, there is a process of “thinking for speaking” in which cognition plays a dynamic role within the framework of linguistic expression:

The activity of thinking takes on a particular quality when it is employed in the activity of speaking. In the evanescent time frame of constructing utterances in discourse, one fits one’s thoughts into available linguistic forms. A particular utterance is never a direct reflection of “objective” or perceived reality or of an inevitable and universal mental representation of a situation. This is evident within any given language, because the same situation can be described in different ways; and it is evident across languages, because each language provides a limited set of options for the grammatical encoding of characteristics of objects and events. “Thinking for speaking” involves picking those characteristics that (a) fit some conceptualization of

the event, and (b) are readily encodable in the language. (Slobin 1987, 435)⁴

The paradigm case of a cross-linguistic difference in event construal is the encoding of the core semantic components of motion events, including PATH and MANNER OF MOTION. In research stemming from seminal work by Talmy (1975), also Talmy (1985, 2000), it has been repeatedly observed that languages may be broadly classified into two types—VERB-FRAMED *vs.* SATELLITE-FRAMED languages, according to how these two components are grammatically encoded: in verb-framed languages, such as Spanish and Japanese, PATH is obligatorily expressed as a component of the verb, while MANNER OF MOTION is (optionally) expressed as an adjunct phrase; by contrast, in predominantly satellite-framed languages, such as English or Dutch, MANNER OF MOTION is directly encoded on the verb, while PATH is expressed as a separate preposition (or particle). This contrast is illustrated in (4) and (5) below. The crucial point to observe about these examples is that the *linguistic typology is orthogonal to culture*: here, French and Japanese pattern together, in contrast to English or Dutch.⁵

- (4) a. Il traverse le fleuve en nageant. (French)
he cross-PRES the river in swim-GERUND
'He crosses the river, swimming.'
- b. 泳いで 川を 渡る (Japanese)
oyoi-de kawa-o wataru.
swim-by river-ACC cross-PRES
'(He) crosses the river, swimming.'
- (5) a. He is swimming across the river. (English)
- b. Hij zwemt de rivier over. (Dutch)
He swims the river over
'He swims over the river.'

The reason that this linguistic distinction is relevant is that it has been claimed to influence behavior in another area of cognition, namely, gesture: Özyürek & Kita (1999), Kita & Özyürek (2003), Kita et al (2007), for example, demonstrate that the divergent patterns of speakers' gestures accompanying linguistic descriptions are predicted by this typology; Van Hoof & Kellerman (2001) show that some L2 learners who acquire target language manner of motion verbs may nevertheless fail to acquire TL gestures, resulting in systematic asynchronicities between language (L2) and gesture (L1). These facts, taken with other evidence presented in Slobin (2003), suggest that structural properties of language are able to penetrate other areas of cognition, at least to this extent.

⁴ Slobin observes in the same article that even Pinker, a stark opponent of the Sapir-Whorf Hypothesis, endorses the intuition behind TFS :

'Whorf was surely wrong when he said that one's language determines how one conceptualizes reality in general. But he was probably correct in a much weaker sense: one's language does determine how one must conceptualize reality when one has to talk about it (Pinker 1989, 360)'

⁵ As has been noted elsewhere, Talmy's taxonomy is not unproblematic: For discussion of alternative morphosyntactic patterns, see Wienold (1995), Inagaki (2006) (for Japanese); Kim (1997) (for Korean); also Song & Levin (1998).

1.2.2 TfS effects beyond the time of utterance

A crucial question at this point is whether TfS effects extend beyond the time of utterance and thus is able to affect long-term memory (LTM) and visual recall. Some authors, including Pederson, Danziger, Senft, Levinson & Wilkins (1998) assume a positive answer—as indicated by the quote below; others, such as Levelt (1989), assume a negative one. Slobin’s own position appears more agnostic. This issue is an important one, in so far as it affects the predictions one might make for picture description and visual recall, respectively: all proponents of TfS assume that language structure affects linguistic descriptions, but only if it extends beyond the moment of speaking does one predict that language will have an effect on recall accuracy.

Far more than developing simple habituation, use of the linguistic system, we suggest, actually forces the speaker to make computations he or she might otherwise not make...That is, the linguistic system is far more than just an *available* pattern for creating internal representations: to learn to speak a language successfully *requires* speakers to develop an appropriate mental representation which is then available for non-linguistic purposes (Pederson, Danziger *et al.*, 1998: 586: [emphasis in original]).

1.3 The Head Parameter

The structural linguistic difference that we consider is at the heart of the results obtained in Masuda & Nisbett’s study is generally referred to in the linguistic literature as the Head Parameter, which refers to the position of a phrasal head to the left or right of its complement, within and across languages: see Travis (1984). As the examples in (6) and (7) show, Japanese is the mirror-image of English with respect to this parameter: all of the phrasal heads that appear to the left of their phrasal complements in English appear to the right in Japanese.⁶ In other words, English is a “head-initial” language, Japanese a “head-final” one. In each case, the relevant head is indicated in bold, the phrasal complement in italics:

- (6) a. John **broke** *the vase*.
b. John **should not have** *broken the vase*.
c. John said **that** he hadn’t broken the vase.
d. John danced **around** *the room in the palace*.
- (7) a. John-ga *kabin-o* **watta**.
John-NOM vase-ACC broke
b. John-wa *kabin-o waru* **bekide-wa nakatta**.
John-NOM vase-ACC break should-NOM not-have
c. John-wa *kabin-o watte nai* **to itta**.
John-nom vase-acc break not comp say-have
d. John-wa *kyuden-no hiroma-de* **odotta**.
John-nom palace-of room-in dance-past

⁶ Here, we ignore the fact that tense and agreement affixes appear as suffixes in finite verb forms in English: there is overwhelming theoretical and empirical evidence going back to Chomsky (1957) that supports the underlying head-initiality of the syntactic categories with which these morphemes are associated.

This contrast can be generalized to include adjuncts as well as complements. In non-generative approaches—and especially within the typological framework initiated by Greenberg (1978)—all head-modifier relations are potentially relevant to determining the head-initial or head-final status of phrases (see also Hawkins 1983, 1990). Thus, the position of attributive adjectives, relative clauses, possessor phrases, and subordinate adjunct clauses are also taken into account. By most of these measures also, Japanese is head-final, English head-initial (assuming DP-hypothesis).

The position of Chinese is slightly more problematic. In most mainstream generative analyses (for example, C.-T. J. Huang 1994), Chinese is treated as underlying head-final, but with a number of obligatory transformational operations that result in surface head-initiality for all but a few peripheral constructions, such as nominal complements (8a) and internally headed relative clauses (8b). Whatever the underlying facts may be—and Huang’s analysis is increasingly controversial, see X.-J. K. Huang (2008)—arguably what matters for language processing are the surface distributions, and in this regard Chinese is structurally much more like English than like Japanese, by any measure of constructional frequency (type or token): as shown in the examples in (9), from Huang (1994), verbs and prepositions canonically precede their complements:⁷

- (8) a. *yuyanxue de yanjiu*
 Linguistics de research
 ‘the study of linguistics’
- b. *ni zui xihuan de nei-ben shu mai-wan le.*
 You most like DE that-CL book sell-out PERF
 ‘The book that you like most has been sold out.’
- (9) a. Zhangsan **meiyou** *kanjian Lisi.*
 Zhangsan not-have see Lisi
 ‘Zhangsan did not see Lisi.’
- b. Zhangsan **zhidao** *Lisi bu chengshi.*
 Zhangsan know Lisi not honest
 ‘Zhangsan knows that Lisi is not honest.’
- b. Zhangsan zhu **zai** *Meiguo.*
 Zhangsan live at America
 ‘Zhangsan lives in the US.’
- d. Zhangsan fang-le yi-ben shu **zai** *zhuozi-shang.*
 Zhangsan put-perf one-cl book at table-top
 ‘Zhangsan put a book on the table.’

1.3.1 From Grammar to Processing...

This grammatical parameter is embedded in, and has implications for, processing models. Here, we assume that head-initial and head-final languages are subject to different default processing strategies (Nakayama (1999), Fodor & Inoue (1994): the many local indeterminacies occasioned by head-final

⁷ Ignoring derived environments, such as the BA-construction. See also McCawley (1992) for justification of the distinction between verbs and prepositions in Chinese.

structures, such as that involved in the Japanese sentence in (10b) means that Japanese cannot be processed serially, top-down, like English, but must instead be processed by a restricted delay model that postpones parsing commitments until the relevant head has been analysed, and arrives at a final parse, bottom-up, through a series of partial analyses: see, for example, Pritchett (1992).

- (10) a. John-ga kabin-o watta.
John-NOM vase-ACC broke
'John broke the vase.'
- b. John-ga [[kabin-o watta] kodomo]-o nagusameta.
John-NOM vase-ACC broke child-ACC consoled
'John consoled the child who had broken the vase.'

1.3.2 ...From Processing to Attentional Mechanisms

We hypothesize—along the lines of the Thinking for Speaking—that the top-down *vs.* bottom-up parsing mechanisms are reflected in the speakers' discourse strategies, and that this in turn may influence their attentional mechanisms. Japanese speakers build up phrase-structure from complements and/or modifiers to heads; in other words, from semantically and syntactically peripheral elements to the core elements in a bottom-up fashion. It plausibly follows from this that Japanese speakers are predisposed to plan and interpret discourse by placing peripheral elements ahead of the main point. In short, our hypothesis is that Japanese speakers are more likely to attend to contextual information primarily because the grammatical and discourse structure of the language requires speakers to mention contextual information ahead of focal information.

2 Experiments

We carried out three experiments—two were description and recall tasks directly modeled on Masuda & Nesbitt's experiments: these were preceded by an elicited production task. In each case, the general prediction was clear: if it is predominantly culture that determines response, Asian groups should pattern together, especially in the latter tasks; if however language structure better predicts response, the Asian response should be split, with the Chinese results patterning more closely with those of the English group.

The three experiments were presented to 120 participants (40 Chinese, 40 English, and 40 Japanese), tested in Japan and in the UK. All associated language materials were translated and presented to each group in their own language, by native-speaker experimenters.

2.1 Task 1: Cause-Effect Reporting Task

In the first task, participants were asked to describe four pictures extracted from popular children's story books, each of which involved a cause-effect relationship.⁸ In principle, the materials were such that all participants could either mention the effect ahead of the cause (approximately corresponding to Figure before Ground, Main Clause before Adjunct Clause) or the other way around.

Figure 2 below shows that the Japanese were the clear "odd group out" in this task, showing 100% reporting of Cause < Effect. This difference was borne out statistically: though all three two-way contrasts

⁸ The sources comprised illustrations from two Japanese story books *Emile* (Ungerer 1975) and *Shinkansen nozomi de kururinpa* (Yasui & Okamoto 1995); and two English books *The Gruffalo's Child* (Donaldson & Scheffier 2007) and *Noddy's Perfect Gift* (Blyton 2001).

(English-Chinese, English-Japanese, Japanese-Chinese) proved to be extremely significant—by Fisher’s Exact Test, all contrasts were reliable at less than $p < 0.0001$ —*the relative size of the effects varied considerably*. This is reflected in the χ^2 result for each comparison: English vs. Chinese $\chi^2 = 17.38$; English vs. Japanese $\chi^2 = 261.11$; Japanese vs. Chinese $\chi^2 = 140.79$. These results clearly show that Japanese is the exception in this task, as predicted by its grammatical typology.⁹

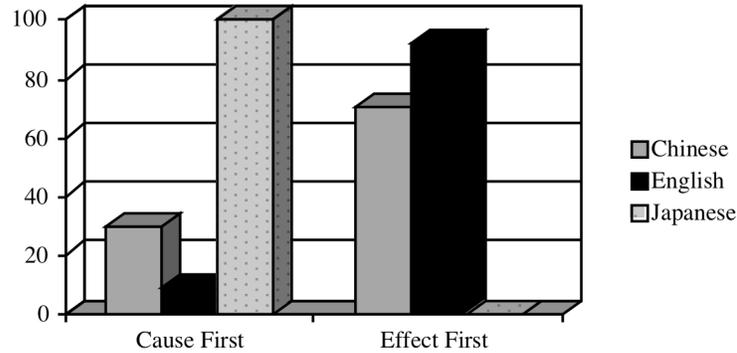


Figure 2. Task 1 Results: Cause-Effect Order (y-axis represents percentage of responses)

2.2 Picture Description Task

Next, we presented the same participants with a set of four colour photographs: a warm-up picture to familiarize them with the task—and for use in the subsequent recall task—and three photographs extracted from a Japanese travel publication (Sugimoto, 1998). Participants were asked to describe what they saw represented in the latter three photographs. These responses were then scored according to the number of central or peripheral elements mentioned.

The expectation, in light of Masuda & Nisbett’s findings, was that the Japanese participants should attend to and report consistently more peripheral items than central items: the question was whether the Chinese group would pattern with the English group, as predicted by WLR/TfS, or instead with the Japanese group, as CR would predict.



Figure 3. Photographs used in Task 2 (Original photographs full size/color)

⁹ As discussed in the full version of this paper, the Japanese participants also gave substantially fuller responses to each picture than did the Chinese participants, often mentioning additional background information—inferences about the ultimate rather than proximate causes of each event—that were not explicitly depicted in the pictures. See Tajima & Duffield, for details.

The participants' responses are charted in Figure 4 below. Notice that, as expected, the largest differences are observed in the mean number of *peripheral* items mentioned by each participant across groups. Here, the Japanese behave once again as predicted, with significantly higher mentions of peripheral items than the English group. Yet, as the statistics confirm,¹⁰ this facilitation for peripheral items is not shared by Chinese participants, who actually mention fewer peripheral elements than even the English group. Indeed, by all statistical measures, the Japanese and Chinese groups come out as more different from one other than either is from the English group. The “Asian preference” is thus effectively split.

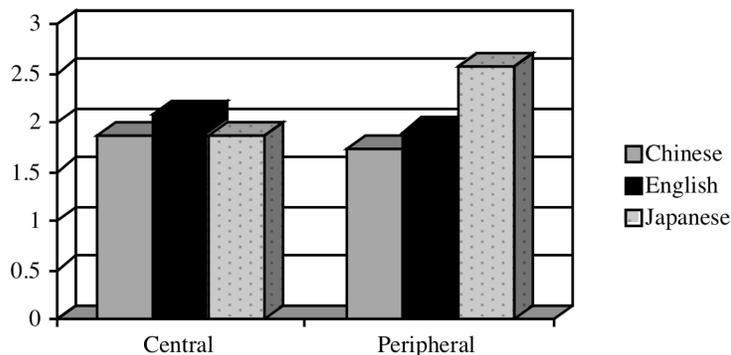


Figure 4: Task 2 Results. Y-axis represents mean no. of items mentioned (central vs. peripheral)

2.3 Task 3: Visual Recall

The final task, which tested participants' accuracy in identifying peripheral fragments, compares the Chinese and Japanese groups with respect to visual recall: thus, it investigates whether the effects of TfS extend beyond the time of utterance. If they do, then Japanese speakers were again predicted to show higher recall accuracy than Chinese or English speakers; if not, there should be no reliable differences between the various groups. Once more, if cultural rather than linguistic factors explain the variation, we should observe a pan-Asian advantage for recall over the English group.

The results of this identification task were somewhat more mixed than in the two previous experiments. Figure 5 shows the number of peripheral items recalled by each group overall, and excluding picture 2 (where the Japanese group did surprisingly badly).

The statistical analysis is especially important here. Overall, there was no main effect of Language in this experiment, its absence due to the Japanese group's performance on picture 2. However, if this picture is excluded from the analysis, then all of the predicted effects are observed to be significant: a reliable main effect of Language ($F(2, 79) = 3.465, p < .05$) is found, as well as post-hoc comparisons (Tukey HSD) showing significant differences between the scores for the Chinese and Japanese groups ($p < .05$), but no

¹⁰ A reliable main effect of Native Language ($F(2, 119) = 4.418, p < .05$) was observed, as well as an extremely significant interaction between Native Language and Position ($F(2, 119) = 12.63, p < .0001$). The source of this interaction is clearly suggested by the plot in Figure 4 below. Separate t-tests were run to examine between-language differences for central and peripheral items. The largest differences with respect to the numbers of peripheral items was observed between the Japanese participants ($mean = 7.72$) and Chinese participants ($mean = 5.2$), $t = -3.92, p < .0001$, respectively, and between the numbers of peripheral items for the Japanese participants (7.73) vs. English participants (5.67), $t = -3.17, p < .005$. There were no reliable differences between the English and the Chinese responses at any level of analysis.

difference for either of the other two pairs (English vs. Japanese, English vs. Chinese). Separate one-way ANOVAs for each picture also revealed that where there is a main effect of Language—in picture 4—it is the Japanese group that is the exception and that, in all cases, the only comparisons that reached or approached significance ($p < .1$) in post-hoc tests were those that involved Chinese vs. Japanese differences.

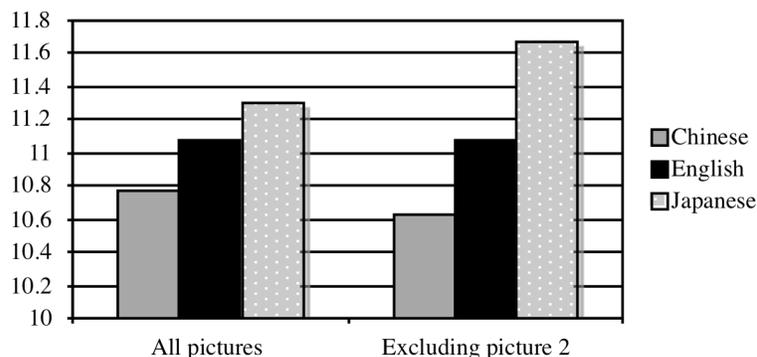


Figure 5. Visual Recall Task. Y-axis represents mean identification scores for each participant

However, even if one restricts attention to the overall results—there being no principled reason to exclude picture 2 on independent grounds—the results show clearly that there is no common Asian response in the identification task: instead, the Chinese participants pattern with the English group, separately from the Japanese.

3 Summary

In conclusion, the present results at the data level are entirely compatible with those obtained by Masuda & Nesbitt for their Japanese participants: we also tentatively endorse their concluding suggestion that “Japanese may simply see far more of the world than do Americans (Masuda & Nesbitt 2001: 933)”, with the caveat that this may only represent the typical case. Where we disagree, however is in the explanation of this difference: we argue that *if the suggestion is true*, then it is not primarily because they are Asians, but because they speak a head-final language, that Japanese speakers see—at least, *report*—more of the world.

It should also be stressed, as a reviewer of our work has pointed out, that this is not a “zero-sum game”. It is very plausible to believe that both cultural and linguistic factors play a role in determining inter-group differences of this kind, and we certainly do not discount the view that cultural traditions may influence modes of perception or categorization (though, it must be said, it is unclear to us how such cultural differences could be handed down or penetrate cognition *except through language*). All that is claimed here is that the group differences we observed are better explained in linguistic, than in cultural, terms: if it were otherwise, Chinese and Japanese participants should pattern together.

Notice finally that the alternative hypothesis presented here directly predicts other cross-linguistic groupings and contrasts within and across cultural spheres. For example, it predicts that Korean speakers, whose language is also head-final, should behave like Japanese participants, while Vietnamese and Thai speakers should pattern with the Chinese participants. This is something that we are currently testing. If these further predictions are borne out, and if effects of language structure continue to outweigh effects of culture for speakers of other languages, then our results can be said to constitute a significant challenge to

this part of the evidence base for cultural relativism, as well as further endorsement of the explanatory potential of the Thinking for Speaking Hypothesis.

References

- Blyton, E. (2001). *Noddy's perfect gift*. London: Collins.
- Boroditsky, L. (2003). Linguistic relativity. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. London: Macmillan Press. 917-921.
- Cheng, L. L.-S. (1997). *On the typology of WH-Questions*. New York: Garland Publishing.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht, Holland; Cinnaminson, N.J.: Foris Publications.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Donaldson, J., & Scheffier, A. (2007). *The Gruffalo's Child*. London: Campbell Books.
- Greenberg, J. (1978). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.), *Universals of Language*. Cambridge, MA: MIT Press.
- Hawkins, J. (1990). A parsing theory of word order universals. *LI*, 21, 223-260.
- Huang, C.-T. J. (1994). More on Chinese word order and parametric theory. In B. Lust, M. Suñer & J. Whitman (Eds.), *Syntactic Theory and First Language Acquisition: Crosslinguistic Perspectives: Vol. 1, Heads, projections and learnability*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Huang, X.-Y. K. (2008). Initialness of sentence-final particles in Mandarin Chinese: An antisymmetry movement approach. *ms*. National Tsing Hua University, Taiwan.
- Kim, Y.-J. (1997). Verb lexicalization patterns in Korean - with focus on motion conflation in complex verb constructions. In H. Min Sohn & J. Halg (Eds.), *Japanese/Korean linguistics, Volume 6*. Stanford: CSLI Publications.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Lucy, J. (1997). Linguistic relativity. *Annual Review of Anthropology* 26, 291-312.
- Masuda, T., & Nisbett, R. E. (2001). Attending holistically vs. analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology* 81, 922-934.
- Nakayama, M. (1999). Sentence processing. In N. Tsujimura (Ed.), *Handbook of Japanese linguistics*. Oxford: Blackwell. 398-424.
- Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review* 108, 291-310.
- Pederson, E., Danziger, E., Levinson, S., Kita, S., Senft, G., & Wilkins, D. (1998). Semantic typology and spatial conceptualization. *Language* 74, 557-589.
- Pritchett, B. (1992). *Grammatical competence and parsing performance*. Chicago: Chicago University Press.
- Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical Constructions: Their Form and Meaning*. Oxford: Oxford University Press. 195-219.
- Slobin, D. I. (1997). Mind, code, and text. In J. Bybee, J. Haiman & S. A. Thompson (eds.), *Essays on Language Function and Language Type*. Amsterdam: John Benjamins Publishing Company. 437-467.
- Slobin, D. I. (2000). Verbalized events – a dynamic approach to linguistic relativity and determinism. In S. Niemeyer & R. Dinsen (Eds.), *Evidence for Linguistic Relativity*. Amsterdam: John Benjamins Publishing Company.
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press. 157-192.
- Song, G., & Levin, B. (1998). *A compositional approach to cross-linguistic differences in motion expressions*. Paper given at the 72nd Annual Meeting of the LSA, New York.
- Sugimoto, Y. (1998). *Otona no Bali*. Tokyo: Shueisha.
- Talmy, L. (1975). Semantics and syntax of motion. In J. P. Kimball (Ed.), *Syntax and Semantics* Vol. 4. New York: Academic Press. 181-238.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description. Vol. 3: Grammatical categories and the lexicon*. New York: Cambridge University Press. 57-149.

- Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring* (Vol. 2). Cambridge, MA: MIT Press.
- Travis, L. (1984). *Parameters and effects of word order variation*. Unpublished PhD dissertation, MIT.
- Ungerer, T. (1975). *Emile*. Tokyo: Bunka Publishing Bureau.
- Wienold, G. (1995). Lexical and conceptual structures in expressions for movement and space: With reference to Japanese, Korean, Thai, and Indonesian as compared to English and German. In U. Egli, P. Pause, C. Schwarze, A. von Stechow & G. Wienold (Eds.), *Lexical knowledge in the organization of language*. Amsterdam: John Benjamins Publishing Company. 301-340.
- Yasui, S., & Okamoto, S. (1995). *Shinkansen nozomi de kururinpa*. Tokyo: Kyouikugageki.