The Magic Spell of Language: Linguistic Categories and Their Perceptual Consequences

Diederik A. Stapel
Tilburg University

Gün R. Semin
Free University Amsterdam

Language is a tool that directs attention to different aspects of reality. Using participants from the same linguistic community, the authors demonstrate in 4 studies that metasemantic features of linguistic categories influence basic perceptual processes. More specifically, the hypothesis that abstract versus concrete language leads to a more global versus local perceptual focus was supported across 4 experiments, in which participants used (Experiment 1) or were primed either supraliminally (Experiments 2 and 3) or subliminally (Experiment 4) with abstract (adjectives) or concrete (verbs) terms. Participants were shown to display a global versus specific perceptual focus (Experiments 1 and 4), more versus less inclusiveness of categorization (Experiments 2 and 3), and incorporation of more rather than less contextual information (Experiment 3). The implications of this new perspective toward the language–perception interface are discussed in the context of the general linguistic relativity debate.

Keywords: linguistic relativity, language, perception, priming, linguistic categorization model

Linguistic relativity is the reverse of the view that human cognition constrains the form of language. Relativity is the view that the cognitive processes of a human being—perception, memory, inference, deduction—vary with the structural characteristics—lexicon, morphology, syntax—of the language he speaks. Of course, both can be true, but in different domains of language and cognition. It has proved to be more difficult, however, to find convincing examples of language affecting cognition than of cognition affecting language, but then it is very difficult to invent a really good experiment on linguistic relativity. (Brown, 1986, p. 482)

The relationship between language and cognition is a longstanding puzzle in Western intellectual history (Boas, 1949; von Humboldt, 1843; Sapir, 1951; Whorf, 1957). Whereas, as Brown (1986) noted, there is sufficient and reliable evidence that cognition constrains language, the evidence on whether language constrains cognition is more contentious, especially when it concerns the impact of language on basic, lower-level cognitive processes, such as attention, memory, and perception. This premise of “linguistic relativity” (Whorf, 1957) has been the subject of numerous theoretical and empirical studies, and there is still considerable debate as to whether it should be abandoned or retained (see Gentner & Goldin Meadow, 2003; Gumperz & Levinson, 1996).

It is not surprising that a typical empirical approach to testing the linguistic relativity hypothesis is to compare two different linguistic communities and examine whether or not the same categorical domain (e.g., color) is linguistically represented in the same way between these linguistic communities. Insightful examples occur when two cultures show linguistic differences in categorization (e.g., naming color), because one can then examine whether such differences affect nonlinguistic processes in a directly implicated cognitive domain (e.g., perception of color, memory for color). Thus, typical research questions are as follows: Do cultural differences in color coding influence the actual perception of color (e.g., Özgen, 2004)? Do cultural differences in gender marking influence gender-related memory (e.g., Stahlberg, Szcsényi, & Braun, 2001)? Do cultural differences in the types of spatial metaphors that people use influence their concept of time (e.g., Boroditsky, 2001)? (For more research questions, see Gentner & Goldin Meadow, 2003, and Gumperz & Levinson, 1996.)

Although comparative cultural analyses of the domain-specific impact of language on cognition have yielded a wealth of insights, they have by no means provided a crystalline resolution to the controversy about whether and how language may shape cognition. As Boroditsky (2003) noted in a review of the empirical literature on linguistic relativity, “definitely answering the ‘does language shape thought’ question has proven to be a very difficult task. Some studies have claimed evidence for the affirmative . . . , while others report evidence to the contrary” (p. 917). In other words, even today the decades-old lamentations of Brown (1986, p. 493) that “we are still without any convincing evidence that language structure affects cognition” ring more than true.

The current research is designed to cast a different light on the linguistic relativity debate. The four studies presented here are predicated upon a general assumption, namely that the chief function of language as a tool is to channel the direction of attention...
(Semin, 2001; 2000b). Given this functional take on language, we present a fresh perspective on the linguistic relativity by investigating how language channels generic (rather than domain-specific) cognitive processes by examining this function within the same language (rather than between languages). Thus, our approach is experimental, rather than correlational or quasi-experimental, and our focus is on generic or metasemantic effects of language (the influence of generic predicate categories on perception), rather than on domain-specific effects (e.g., specific categories, such as time or color). We investigated how—within the same linguistic community—particular linguistic devices (i.e., predicate categories, such as interpersonal verbs and adjectives) influence people’s perception of objects and events. Specifically, in four experiments, we tested the hypothesis that the use or cognitive activation of global predicates (e.g., adjectives such as “aggressive”) or concrete predicates (e.g., verbs such as “punch”) channel people’s focus mainly on global or detailed features in their environments.

Thus, we argue that even though comparative analyses of domain-specific language–cognition effects can be helpful in understanding the influence (or lack thereof) of language upon cognition, neither a comparative nor a domain-specific approach is necessary to unravel the language–cognition interface. One can conceptualize and examine the effects of language upon cognition very well within the same language that provides its speakers with different means of representing the same thing (see also Kay, 1996; Semin, 2001; and Stahlberg et al., 2001). Furthermore, language may have a more generic influence upon cognition, rather than a domain-specific one (e.g., color labels, color categorization). These are the two pillars of the research reported here with which we hope to find new empirical, experiment-based answers to the intriguing question of whether and how language may shape cognition.

**Linguistic Diversity, Cognitive Diversity**

The languages of the world are a testament to diversity, with the different ways they permit people to represent their physical, psychological, and social environments. In Turkish, for instance, a suffix on the verb explicitly specifies whether speakers themselves have witnessed an event or are recounting it from hearsay. Similarly, there are differences with respect to the types of lexical categories that are available to describe persons in English and Chinese (e.g., Hoffman, Lau, & Johnson, 1986). In the same way, there are striking differences in the manner in which spatial locations are described (Majid, Bowerman, Kita, Haun, & Levinson, 2004), with some languages relying on relative spatial terms (e.g., left, right) and others relying upon absolute spatial terms (e.g., east, west). Likewise, there are considerable variations in the availability of basic color terms across a diversity of linguistic communities (e.g., Berlin & Kay, 1969).

This list of dramatic differences between languages can be extended considerably with respect to languages that differ in how gender is marked (e.g., English vs. Turkish) and languages that differ in how emotional states are represented (e.g., Dutch vs. Hindustani Surinamese; Semin, Görts, Nandram, & Semin-Goossens, 2002). It seems logical to assume that, given the enormous diversity the languages of the world display in the ways they permit people to cognitively represent their physical, psychologi- cal, and social environments, these languages are also likely to channel people’s actual perceptions differentially to their physical, psychological, and social environments. Because language and cognition are so intrinsically related (people typically need words to express what they think), it is a small step to link linguistic diversity to cognitive diversity. And of course, it is precisely this notion that sparked comparative analyses of how linguistic differences between cultures may be related to differences in nonlinguistic, cognitive processes (Sapir, 1951; Whorf, 1957).

As we noted before, there is no doubt that comparative research on the language–cognition interface has provided fascinating insights that would otherwise not be possible. However, although it cannot be denied that the question of whether different linguistic devices can have differential effects on cognition is of considerable interest, comparative, cross-cultural analyses of differences between languages retain some problematic features that are not always easy to surmount. As a number of researchers have argued, one of the core problems of such studies is translation commensurability (see Gumpertz & Levinson, 1996; Ji, Zhang, & Nisbett, 2004). To know exactly, for example, to what extent two different languages carve the color spectrum at different places or have a different concept of time, one needs translations. The lenity or strictness of such translations is then an important determinant of whether or not languages are perceived as different (e.g., Brown, 1958, 1986). Similarly, comparability of populations used in such cultural studies is also a problem: What other characteristics vary with the linguistic difference? This, in conjunction with the translation commensurability issue, remains a potential source of weakness in comparative research.

**A New Look at the Linguistic Relativity Debate: Generic Effects of Metasemantic Categories Within Language**

For the current studies, we adopted a new approach to the investigation of the language–cognition interface. Our main question is whether linguistic devices within the same language (rather than between languages) influence people’s generic (rather than domain-specific) perception of objects in a systematic manner. We think that this new approach may avoid the possible pitfalls because of the incommensurability of translations and samples between linguistic communities.

In the current studies, we investigated how the generic property of abstractness versus concreteness of specific linguistic devices may influence people’s perception of objects accordingly. Specifically, we tested the hypothesis that the use or cognitive activation of abstract predicates (i.e., adjectives) is more likely to lead to a focus on global features of a stimulus, whereas the use of concrete predicates (i.e., verbs) is more likely to lead to a focus on local, detailed features of this stimulus.

The research question we address here thus relies on what could be called a language-use-and-function approach (Semin, 2001). In contrast to most work on the language–cognition interface, which takes a symbolic representational perspective on language, we advocate a functional perspective. In our perspective, language is not merely a set of abstract rules that are virtual and outside of time (Riceour, 1955). Rather, language is functional. Language is for seeing, thinking, and doing. Taking this perspective introduces new ways of looking at language, such as focusing on its attention-driving function, a focus that is unlikely to arise in an abstract
representational perspective on language. Thus, although our approach is derived from considerations about language use, it is not an investigation of language in use, per se. It is an investigation of language and its effects from a functional perspective. We posit that language is a tool that is used to give public shape to people’s goals, motives, or intentions, and it thereby directs attention to different aspects of reality. Accordingly, different linguistic devices serve different perceptual functions.

It is worthwhile to note that this approach and our emphasis on how language may direct people’s attention and shape their perceptual focus is at the heart of Whorf’s (1957) original linguistic relativity hypothesis: “Users of markedly different grammars are pointed by their grammars toward different types of observations and different evaluations of external similar acts of observation, and hence are not equivalent as observers, but must arrive at somewhat different views of the world” (p. 221; italics added). Notably, what may hold true across languages, we argue, should also hold true within a language.

The Current Studies

In the current studies, we used the linguistic category model (LCM; Semin, 2000a, 2000b; Semin & Fiedler, 1988, 1991) as a conceptual framework to investigate our hypothesis that different linguistic devices within a language may have generic, metasemantic effects on cognition and, more specifically, that abstract predicates will induce a global perceptual focus, whereas concrete predicates are likely to induce a local perceptual focus.

The LCM is a model of interpersonal language that provides the means to investigate the type of linguistic devices that are used to represent social events. In this model, adjectives and interpersonal verbs are ordered on a dimension from concrete to abstract. On the most concrete end of the LCM continuum are action verbs; these terms describe a single, observable event and preserve perceptual features of the event (e.g., “A punches B”). The most abstract category in the LCM is adjectives; these are properties that generalize across specific events and objects and describe only the subject (e.g., “A is aggressive”). The properties described by adjectives show low contextual dependence; the use of adjectives is governed by abstract, semantic relations and not by the contingencies of contextual factors. The opposite is true for action verbs. Action verbs refer to contextual and situated features of an event. If it is the case that concrete terms such as verbs of action are used predominantly in situated contexts and refer to the specific details of a social event, then their obvious function—aside from providing a semantic representation of the event—is to draw attention to the situated, local features of the event. For instance, “Jack punched David” or “Jack helped David” draws attention to the specific act. In contrast, adjectives draw attention to global features that are extracted from the event: “Jack is aggressive” or “Jack is helpful.” It is the generic implication of the classification advanced by the LCM that introduces a novel look at the language–cognition interface. The question is thus “Do these categories have a generic attention-channeling effect that affects the perception of a stimulus environment?”

In our view, the LCM provides us with an excellent tool to study whether the generic features of specific linguistic categories (e.g., adjectives vs. action verbs) induce different perceptual foci. The LCM affords this possibility because it has been shown that abstractness and concreteness are generic to the entire predicate class (e.g., Semin, 2000b; Semin & Fiedler, 1988, 1991). Thus, in contrast to more conventional and domain-specific linguistic categories, such as color, the LCM is not domain-specific. Moreover, the difference between conventional approaches to meaning (e.g., semantics) and meaning of the concrete–abstract dimension is that the inferential properties identified by the LCM are not specific to particular semantic domains. Thus, because the inferential properties apply across semantic domains, one can refer to the inferential properties identified by the LCM as metasemantic (Semin, 2000b). It is precisely because of this feature of the LCM—because of the fact that the LCM categorizes linguistic devices in a way that goes beyond their direct, semantic meaning—that the LCM can serve as a useful tool to test the hypothesis that the use or cognitive activation of certain categories of linguistic devices (i.e., abstract vs. concrete predicates) may have generic, metasemantic effects that go beyond the specific, descriptive meaning that is directly implicated by these devices.

Overview

In sum then, we posit that language is a tool that directs attention to specific aspects of the physical, psychological, and social environment. As an illustration of this, we hypothesize that different predicates are likely to direct attention to different features of an object. It is important to note, however, that we are not addressing questions concerning surface meaning–perception relationships, such as “Does the availability of gender markers influence gender-related memory?” (e.g., Stahlberg et al., 2001) or “Does the availability of color categories influence color perception?” (e.g., Ozgen, 2004). Our research question is more general and focuses on the metasemantic, rather than on the semantic, effects of language. Specifically, we hypothesize that concrete terms (e.g., action verbs) are more likely to direct attention to the local properties and details of an object, whereas abstract terms (e.g., adjectives) are more likely to draw attention to the global properties of an object.

We tested this hypothesis in four experiments. All these experiments used an unrelated-tasks paradigm: In a first task participants used or were exposed to abstract predicates (adjectives) and concrete predicates (action verbs) and then, in an ostensibly unrelated second task, they were asked to complete one or more dependent measures. To test the strength of our hypothesis and to increase the conceptual generalizability of our experimental method, we used a variety of independent and dependent variables. We studied the impact of abstract adjectives versus specific action verbs by putting participants in situations in which they were likely to use the linguistic categories (Experiment 1) but also by subtly priming these categories supraliminally (Experiments 2 and 3) and subliminally (Experiment 4). Furthermore, we studied the impact of language on cognition by using diverse dependent measures, self-report (Experiments 1 and 4), Kimchi and Palmer’s (1982) perceptual, global–specific focus task (Experiments 1 and 4), Isen and Daubman’s (1984) categorical inclusiveness task (Experiment 2), and Kitayama, Duffy, Kawamura, and Larsen’s (2003) framed-line test (FLT; Experiment 3).
Experiment 1: See What You Say

In Experiment 1, we put to a first test our hypothesis that the use of adjectives versus action verbs influences the basic level of perceptual focus of the users of these predicates. Participants first watched short films with chess pieces moving in ways that invited anthropomorphic interpretations and were given the instruction to describe either the personalities or the behaviors of the chess pieces. Research by Norenzayan and Schwarz (1999) showed that even the subtlest of cues, namely a research letterhead that read “Institute for Social Research,” induced more situational explanations, whereas a letterhead that read “Institute for Personality Research” resulted in more dispositional explanations (see also Stapel & Koomen, 1996). Previous LCM research (see Semin, 2000b, 2001) has also shown that subtle instructions reliably instigate the use of adjectives (“She is aggressive,” “He is sad”) versus verbs (“She hits him,” “He moves away from him”).

To measure the impact of these linguistic categories on participants’ perceptual focus, we used a task that was designed by Kimchi and Palmer (1982) to measure differences in a perceptual, global–specific focus task (see also Gasper & Clore, 2002). On each trial of this task, participants were asked to indicate which of two geometric comparisons was more similar to a target figure. Each figure could be viewed from either a global or a specific, local perspective. The hypothesis was that participants who had described the personality of chess pieces would match the figures more on the basis of their global details than would participants who had described the behavior of the chess pieces.

Method

Twenty-four undergraduate students were randomly assigned to a personality or a behavior condition. After arrival in the laboratory, participants were placed in individual cubicles and were told that they would participate in a number of unrelated tasks. First, they were shown two short (less than 1 min) films with chess pieces moving in ways that invited anthropomorphic interpretations (to see the films, go to www.stapel.socialpsychology.nl/magicmovies.html). These films are modern versions of the famous Heider–Simmel film (Heider & Simmel, 1944) designed to study the activation of anthropomorphic descriptions when watching moving geometric figures.

Participants in the personality condition were instructed to “tell us what you see. This is a film about the personality of chess pieces.” Participants in the behavior condition were instructed to “tell us what you see. This is a film about the behavior of chess pieces.” Previous studies have shown that such instructions effectively induce the use of adjectives and action verbs, respectively (see Semin, 2000b, 2001).

Next, participants were given the 24-trial perceptual, global–specific focus task, modeled after Kimchi and Palmer (1982) and Gasper and Clore (2002). On each trial of this task, participants had to indicate which of two geometric comparisons was more similar to a target figure. Each figure could be viewed from either a global or a specific, local perspective. Each object was either a triangle (specific forms) or triangles (global forms). Participants indicated whether each geometric figure was more similar to a group of objects that matched its local, specific components. Participants also answered the following question on a scale ranging from 1 (local) to 9 (global): “When you did the shape task, to what extent did you focus on local matches (e.g., a square of triangles goes with a triangle of triangles) or global (e.g., a square of triangles goes with a square of squares) matches?”

Finally, on completion of these tasks and questions, participants were carefully debriefed about the goal and purpose of the experiment, following the funneled debriefing procedure for priming experiments, as advocated by Bargh and Chartrand (2000). None of the participants spontaneously indicated suspicion of the actual goal of the study. After debriefing, participants were thanked and dismissed.

Results and Discussion

The number of times that participants matched the shapes on the basis of their global forms, rather than on specific, local details, was calculated (see Table 1). An analysis of variance (ANOVA) showed that, as predicted, participants in the personality condition were more likely to use the global form as a basis for matching objects ($M = 14.25, SD = 3.39$) than were participants in the behavior condition ($M = 11.08, SD = 3.40$), $F(1, 22) = 5.23, p < .05, \mu^2 = .19$. Participants’ self-reports showed a similar pattern. Personality participants reported basing their choices more on global forms ($M = 7.33, SD = 1.16$) than did behavior participants ($M = 5.08, SD = 1.73$), $F(1, 22) = 14.04, p < .01, \mu^2 = .39$. Analyses showed that the partial correlation (controlling for experimental condition) for these two dependent measures was high ($r = .70, p < .01$).

These results provide strong support for the hypothesis that the impact of language on cognition can be studied and be experimentally and empirically demonstrated within one language when one focuses on the impact of metasemantic categories (such as abstract adjectives vs. concrete action verbs) on generic cognitive process (such as the level of perceptual focus).

---

1 In all experiments reported here, respondents were native Dutch undergraduate students who received partial course credit for participation. For detailed descriptions of the materials used in these experiments, please contact Diederik A. Stapel.

2 This high correlation between implicit (perceptual) and explicit (self-report) measures may be somewhat puzzling to some, because there exists a host of theorized differences between implicit and explicit measures (for a detailed discussion, see Stapel & Blanton, 2004, Experiment 5). Assuming researchers have valid implicit and explicit measures, the interesting question is determining when implicit and explicit judgments will share empirical and conceptual overlap and when they will not. One answer is the following: Correlations should be high when people view their responses as true, uncontaminated, unbiased expressions of their evaluations. Correlations should be low when people do think that their explicit responses are biased and are motivated and able to correct for these biases. We think that in the current study (as well as in the other studies reported in this article) the high correlations occurred because our participants were unaware of the influence of language on the perceptual-focus task. They not only failed to realize that they had been exposed to linguistic categories that might have influenced their implicit responses (e.g., matching task), but they also failed to realize that they had been exposed to stimuli that might have influenced their explicit responses (for further information, see Stapel & Blanton, 2004).
Table 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Film instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Personality</td>
</tr>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Global matches</td>
<td>14.25</td>
</tr>
<tr>
<td>Self-report</td>
<td>7.33</td>
</tr>
</tbody>
</table>

Note. On the self-report measure, scale range is from 1 (local) to 9 (global). Thus, for both measures, higher numbers indicate a more global perceptual focus.

Experiment 2: Categorize What Is Said

In Experiment 1, we showed how instructions (“personality” vs. “behavior”) that are known to lead to the active utterance or production of adjectives versus action verbs influenced the level of perceptual focus. Although Experiment 1 was an important first step, to test our hypothesis more effectively, we need to investigate directly how certain linguistic devices affect cognition and perception. In Experiment 2, we thus investigate whether similar effects can occur when language is processed, rather than produced. The question is then whether the subtle activation of certain linguistic categories could be sufficient to shape cognition. We addressed this question by assessing whether supraliminal priming of adjectives versus action verbs affected the inclusiveness or globality of categorization. Participants first unscrambled a number of word jumbles into meaningful short sentences (see Srull & Wyer, 1979; Stapel, Koomen, & Zeelenberg, 1998). These sentences contained either action verbs or adjectives. Then, in an ostensibly unrelated task, participants completed a measure of inclusiveness of categorization (see Isen & Daubman, 1984; Smith & Trope, 2006), indicating to what degree weak or atypical exemplars (e.g., camel) were good members of a given category (e.g., vehicle). We made the following hypothesis: If priming with adjectives leads to more global thinking than does priming with action verbs, then adjective-primed participants will be more inclusive in their categorizations than will verb-primed participants.

Method

Thirty undergraduate students were randomly assigned to an adjective or an action-verb priming condition. The experiment was part of a general testing session in which participants were tested in individual cubicles and received a number of unrelated questionnaires. Filler tasks preceded the priming and categorization tasks. This made it unlikely that participants would be able to guess the true purpose of the experiment. When participants were finished, the questionnaires were collected, and, as in Experiment 1, participants were probed carefully for awareness of the relation between the priming and categorization tasks. No participant showed suspicion of a relation between the experimental tasks or stated that their ratings in the categorization task were influenced by the priming task.

For the priming task, we used the scrambled-sentence test. This test consisted of 40 scrambled four- or five-word groups (e.g., “hits table he him”). Participants’ task was to reorganize the word groups into meaningful sentences, using only three or four words from each group (see Stapel et al., 1998). Fifteen word groups were fillers and contained neutral information (e.g., “empty vacation glass is the”). In the adjective conditions, the remaining 25 word groups contained adjectives (e.g., aggressive, friendly, humble). In the action-verb condition, these word groups contained action verbs (e.g., punch, help, swim). The adjective and action-verb priming sentences were pilot tested (n = 30) to show no differences (F < 1) in the positive or negative affect they elicited in participants (as measured by Watson, Clark, & Tellegen’s, 1988, Positive and Negative Affect Scale). The sentences were also pilot tested (n = 28) to make sure there were no differences (F < 1) in the perceived evaluative extremity (“How positive or negative is the behavior/trait described in this sentence”).

After participants finished the priming task, they completed the categorization task, modeled after the work of Isen and Daubman (1984) and Smith and Trope (2006). They were told to rate 10 items on a scale ranging from 1 (definitely does not belong to the category) to 9 (definitely does belong to the category). For each item, participants saw the general category (e.g., vehicle), the specific item (e.g., camel), and the rating scale. The items included two strong, three moderate, and five weak exemplars, based on Rosch’s (1975) norms and pretesting.

When they had completed the categorization task, participants were given a one-item mood measure that has been proven to be sensitive to mood changes in previous research (e.g., Stapel & Koomen, 2006; Stapel, Koomen, & Ruys, 2002). Participants were also asked to indicate their current mood by responding on a scale ranging from 1 (negative or positive do you feel at this moment?) to 9 (positive or negative do you feel at this moment?) to the question “How positive or negative is the behavior/trait described in this sentence?”

Results and Discussion

First, we performed an ANOVA on the inclusiveness scores, treating exemplar strength (strong, moderate, weak) as a within-subjects factor and condition (adjective, action verb) as a between-subjects factor. This revealed a main effect of condition, F(1, 24) = 18.20, p < .01, μ² = .39; a main effect of exemplar strength, F(2, 24) = 83.67, p < .01, μ² = .75; and an Exemplar Strength × Condition interaction, F(2, 24) = 4.41, p < .05, μ² = .14. As expected (see Table 2), participants in the adjective condition were more inclusive than participants in the action-verb condition, and this effect was strong for weak exemplars (M adjective = 5.25, SD = 1.82; Maction verb = 2.91, SD = 1.34; F(1, 24) = 16.15, p < .01, μ² = .37), moderate for moderate exemplars (M adjective = 6.24, SD = 1.16; Maction verb = 4.56, SD = 1.74; F(1, 24) = 9.83, p < .01, μ² = .26), and weak for strong exemplars (M adjective = 8.17, SD = 1.06; Maction verb = 7.57, SD = 0.78; F(1, 24) = 3.12, p < .10, μ² = .10).

Next, we calculated an average inclusiveness score over the 10 category-inclusion items to obtain one comprehensive dependent measure. An ANOVA showed that, as predicted, participants in
the adjective condition were more inclusive ($M_{\text{adjective}} = 6.13$, $SD = 1.25$) than were participants in the action-verb condition ($M_{\text{action verb}} = 4.33$, $SD = 0.98$), $F(1, 28) = 19.40$, $p < .01$, $\mu_r = .41$.

This finding provides an unequivocal answer to the question that inspired this experiment: “Can the subtle activation of certain linguistic categories be sufficient to shape cognition?” Yes. It thus provides further support for our hypothesis concerning the impact of metasemantic linguistic devices on generic cognitive process. Compared with Experiment 1, it does so assessing the impact of the mere activation of these devices (rather than their active use) on a cognitive, high-level categorization measure (rather than a perceptual, low-level attentional-focus measure).

Experiment 3: See What Is Said

In Experiment 3, we aimed to replicate and extend the supraliminal priming logic of Experiment 2 by adding a control priming condition in which neither adjectives nor action verbs were primed and by adding a different, less “wordy” dependent measure. Specifically, in this experiment, we used the FLT (Kitayama et al., 2003), a test that can be used to assess the ability to attend to or ignore global, context-insensitive versus specific, context-sensitive information on a basic, perceptual level.

In our adapted version of this test (see Stapel & Van der Zee, 2006, for further details), an FLT trial looks like this: Participants are presented with a square frame, within which is printed a vertical line. Next, participants are presented with another (smaller or larger) square frame and are asked to draw a line that is identical to the first line in absolute length. Thus, in this task, participants have to ignore both the first frame (when assessing the length of the line) and the second frame (when reproducing the line). Thus, they will do well (exactly copy the framed line) when they are independent, noncontextual, less situational, less localized perceivers.

Hence, if our hypothesis concerning the impact of metasemantic linguistic categories (adjectives vs. action verbs) on generic perceptual processes (global vs. local perceptual focus) is correct, then adjective-primed people should perform better on our FLT than action-verb-primed (dependent) people.

Method

Forty-seven undergraduate students were randomly assigned to an adjective, an action-verb, or a control priming condition. The experimental procedure was similar to the one used in Experiment 2: The experiment was part of a general testing session in which participants received a number of unrelated questionnaires. As in Experiments 1 and 2, an awareness and suspicion check revealed that none of the participants thought priming had influenced their performance on the dependent measure.

The priming task was identical to the one used in Experiment 2. However, in addition to the adjective and action-verb conditions, a control condition was included in which all 45 word groups contained neutral or filler information (e.g., “empty vacation glass is the”).

After they had completed the priming task, participants answered the mood measure that was also included in Experiment 2. Even administered immediately after the priming task, this mood measure showed no impact of the priming manipulation ($F < 1$). Next, participants performed a six-trial FLT, modeled after the work of Kitayama et al. (2003) and Stapel and Van der Zee (2006). In each of these trials, participants were first shown a square frame within which a vertical line was printed. Then, they were shown a second square frame that was either larger or smaller than the first frame. The task was to draw a line in the second frame that was the same absolute length as the line in the first frame. The lines drawn by each participant were measured, and the absolute differences between these lines and the correct lines were calculated. Averaging these differences resulted in a mean error score (in mm).

When they had completed the FLT, participants were given the categorization task that was also used in Experiment 2.

Results and Discussion

We measured the FLT lines drawn by each participant and the differences between these lines and the correct lines (see Table 3). Participants’ errors were in the expected direction (too small when the second frame was smaller, too large when the second frame was larger; see Stapel & Van der Zee, 2006). Averaging these differences resulted in a mean error score (in mm). An ANOVA showed that priming affected this score, $F(2, 44) = 25.28$, $p < .01$, $\mu_r = .53$. As predicted, for adjective participants, FLT errors were smaller ($M = 4.40$, $SD = 0.51$) than for action-verb participants ($M = 6.20$, $SD = 0.86$). The scores of control participants were in

Table 3

<table>
<thead>
<tr>
<th>Measure</th>
<th>Adjectives</th>
<th>Control</th>
<th>Action verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Frame-line test</td>
<td>4.40</td>
<td>0.51</td>
<td>5.00</td>
</tr>
<tr>
<td>Category inclusiveness</td>
<td>6.33</td>
<td>0.82</td>
<td>5.71</td>
</tr>
</tbody>
</table>

Note. On the framed-line-test measure, higher numbers indicate a more localized, contextualized focus. On the category-inclusiveness measure, scale range is from 1 (definitely does not belong to the category) to 9 (definitely does belong to the category). Thus, on this measure, higher numbers indicate more inclusive categorizations.

Table 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>Supraliminal priming</th>
<th>Adjectives</th>
<th>Action verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Category inclusiveness</td>
<td>6.23</td>
<td>3.39</td>
<td>4.33</td>
</tr>
</tbody>
</table>

Note. On the category-inclusiveness measure, scale range is from 1 (definitely does not belong to the category) to 9 (definitely does belong to the category). Thus, higher numbers indicate more inclusive categorizations.
between these two extremes ($M = 5.00, SD = 0.71$; for all individual comparisons, $p < .05$).

Next, as in Experiment 2, we performed an ANOVA on the inclusiveness ratings, treating exemplar strength (strong, moderate, weak) as a within-subject factor and condition (adjective, action verb, control) as a between-subjects factor. This revealed a main effect of condition, $F(2, 39) = 23.53, p < .01, \mu^2 = .52$; a main effect of exemplar strength, $F(2, 39) = 264.49, p < .01, \mu^2 = .86$; and an Exemplar Strength \times Condition interaction, $F(4, 39) = 11.24, p < .01, \mu^2 = .19$. As expected, participants in the adjective condition were more inclusive than participants in the action verb condition (and scores of control participants were between these two extremes). And as expected, this effect was strong for weak exemplars ($M_{\text{adjective}} = 5.37, SD = 1.35; M_{\text{action verb}} = 3.07, SD = 0.64; M_{\text{control}} = 4.54, SD = 1.03; F(2, 39) = 18.73, p < .01, \mu^2 = .46$), moderate for moderate exemplars ($M_{\text{adjective}} = 6.67, SD = 0.88; M_{\text{action verb}} = 5.42, SD = 0.80; M_{\text{control}} = 6.06, SD = 0.36; F(2, 39) = 11.65, p < .01, \mu^2 = .35$), and weak for strong exemplars ($M_{\text{adjective}} = 8.20, SD = 1.05; M_{\text{action verb}} = 7.63, SD = 0.48; M_{\text{control}} = 8.09, SD = 0.48; F(2, 39) = 2.70, p < .10, \mu^2 = .11$). As in Experiment 2, we also calculated an average inclusiveness score over the 10 category-inclusion items. Again, an ANOVA showed that priming affected this score, $F(2, 44) = 25.75, p < .01, \mu^2 = .54$. As predicted, participants in the adjective condition were more inclusive ($M_{\text{adjective}} = 6.33, SD = 0.82$) than were participants in the action-verb condition ($M_{\text{action verb}} = 4.67, SD = 0.49$). The scores of control participants were in between these two extremes ($M_{\text{control}} = 5.71, SD = 0.60$; for all individual comparisons, $p < .05$). Analyses showed that the partial correlation (controlling for experimental condition) for the two dependent measures (FLT and categorization) was high, $r = .62 (p < .01)$, after reverse coding the FLT scores such that higher scores denote a more global perceptual focus.

These findings further strengthen the support for our hypothesis by showing the impact of abstract versus specific word categories on a relatively low-level perceptual measure (the FLT) as well as on a relatively high-level cognitive measure (the inclusiveness-of-categorization measure).

**Experiment 4: See What You Do (Not) See**

In Experiment 4, we push the envelope and put our “(metasemantic) language affects (generic) perception” hypothesis to a final test by assessing the impact of subliminally (rather than supraliminally) primed linguistic action verbs versus adjectives on the global–specific focus task that we used in Experiment 1. The question is therefore whether the subtle, unconscious activation of certain linguistic categories could be sufficient to shape conscious perceptual focus.

**Method**

Fifty undergraduate students were randomly assigned to an adjective, an action-verb, or a control priming condition.

Upon arrival in the laboratory, participants were seated in front of a computer. First, participants performed a parafoveal vigilance task, in which words were presented outside of awareness. After having completed the vigilance task, participants were thanked for their participation and given the global–specific focus questionnaire (see Experiment 1). Next, participants received a prime recognition task and a funnel debriefing procedure.

Priming stimuli were taken from the list of adjectives and action verbs used in Experiments 2 and 3. The priming task was modeled after the paraphrase priming task Stapel et al. (2002) used previously. This priming task has been used reliably to prime information without awareness (see Bargh & Chartrand, 2000; Stapel et al., 2002). Once participants were seated in front of their computers, the experimenter explained the vigilance task. The experimenter then instructed participants to place their index fingers on two keys of the keyboard and to press the left key, labeled “L,” if a flash appeared on the left side of the screen and the right key, labeled “R,” if a flash appeared on the right side of the screen. A fixation point consisting of one X was presented continually in the center of the screen. Participants were given 10 practice trials to become familiar with the procedure. After answering any questions from the participants, the experimenter began the 60 experimental trials of the vigilance task. The words that were presented in the 10 practice trials and in 20 of the experimental trials were neutral object words (e.g., table, chair). In the remaining 40 experimental trials, either an adjective (e.g., aggressive, friendly, humble) or an action verb (e.g., punch, help, swim) was presented. Words were presented for 80 ms and were immediately followed by a 120-ms mask.

Previous subliminal priming studies have shown that the paradigm employed here provides sufficient safeguards to prevent participants from becoming aware of the priming stimuli (see Bargh & Chartrand, 2000; Stapel & Koomen, 2006; Stapel et al., 2002). However, to ensure that participants were not aware of the priming stimuli, we used an extensive funnel debriefing procedure in which participants were asked increasingly specific questions about the study.

All participants reported that they had seen flashes, but none of the participants could recall any specific word. Furthermore, when confronted with a list of words that consisted of primed words and nonprimed words, participants’ guesses of whether or not a word was primed did not exceed chance, nor did they differ between conditions ($F$s $< 1$). Finally, no participants thought the vigilance and the global-specific tasks were related. Thus, we can conclude that we were successful in presenting our priming stimuli outside of awareness (see Bargh & Chartrand, 2000).

**Results and Discussion**

The number of times that participants matched the shapes on the basis of their global form rather than on the basis of local details was calculated (see Table 4). An ANOVA showed that priming affected this measure, $F(2, 47) = 20.83, p < .01, \mu^2 = .47$. As predicted, participants in the adjective condition were more likely to use the global form as a basis for matching objects ($M = 14.71, SD = 2.02$) than were participants in the action-verb condition ($M = 11.24, SD = 1.52$). The scores of control participants were in between these two extremes ($M = 12.19, SD = 1.17$; for all individual comparisons, $p < .05$).

Participants’ self-reports showed a similar pattern. An ANOVA showed that priming also affected this measure, $F(2, 47) = 5.93, p < .01, \mu^2 = .20$. As predicted, adjective-primed participants in the adjective condition reported basing their choices more on global forms ($M = 6.29, SD = 1.11$) than did verb-primed par-
participants ($M = 4.71$, $SD = 1.57$). The scores of control participants were in between these two extremes ($M = 5.69$, $SD = 1.35$; individual comparisons, $p < .05$, except control-verb, $p < .10$). Analyses showed that the partial correlation (controlling for experimental condition) for the two dependent measures was high, $r = .69$ ($p < .01$).

The findings of this final experiment show that even subtle, unconscious activation of linguistic categories can be sufficient to shape conscious perceptual focus. Our participants had a more global focus on reality when they were subliminally primed with adjectives than when they were subliminally primed with action verbs. Thus, even words one does not perceive can affect perception, such that what one sees (or, better, does not see) influences how one sees.

General Discussion

Language Shapes Cognition

Language shapes cognition. At least that is a common assumption, not only among students of the language–cognition interface, but also among the general public at large (see Gentner & Goldin Meadow, 2003; Griffin & Ross, 1991; Gumpertz & Levinson, 1996; Stapel & Marx, in press; Stapel & Spears, 1996). And of course, public relations consultants and political propagandists have long recognized the power of language to shape public perception and policy. That is why in political debates or times of international conflict, we see propagandists on both sides struggling to control the semantics of what is supposedly going on. The underlying assumption is then that people’s cognitive representation of reality may be shaped by the ways in which it is coded linguistically. This belief in the power of language has been sufficient to shape conscious perceptual focus. Our participants had a more global focus on reality when they were subliminally primed with adjectives than when they were subliminally primed with action verbs. Thus, even words one does not perceive can affect perception, such that what one sees (or, better, does not see) influences how one sees.

Given this list of dramatic examples of how language is and can be used to shape people’s cognitive representation of (social) reality, it is probably not surprising that in the scientific community, there is hardly any debate over these content-specific, higher-level effects of language on cognition. As numerous studies in psycholinguistics and social psychology suggest, linguistic labels help to give meaning to and disambiguate our (intrinsically ambiguous) world. Linguistic labels help people make sense of their complex environments by encouraging them to go beyond the information given and make inferences and form associations consistent with those labels (e.g., to go from the label of freedom fighters to images of virtuous, self-sacrificing patriots; or to go from the label of terrorists to images of cruel, relentless psychopaths). In fact, the principle that the content of language may affect the content of cognition and behavior is at the core of one of the hallmarks of modern social psychology: semantic priming effects. Social-psychological priming studies (for reviews, see Bargh, 2006; Greenwald & Banaji, 1995; Higgins, 1996) show how the semantic content of language affects the content of specific thoughts and associations (as tapped by cognitive-activation measures), evaluations (as tapped by judgment tasks), and behaviors (as tapped by observational measures).

But Does Language Shape Basic Perception?

However, the debate over the language–perception interface is not about whether or not language may exert effects on relatively content-specific, higher-level cognitive processes. That is, the question is not whether or not, say, priming stereotype labels (“Blacks“) activate specific semantic associations (e.g., “aggressive”) that accordingly result in corresponding evaluations (e.g., rating Tyrone as aggressive) or behaviors (e.g., shouting at someone). The debate over the impact of language on cognition is about the question whether language may exert effects on relatively basic, content-unspecific, lower-level processes, such as attention, memory, and perception. Although there is reliable evidence showing the impact of language on the content of cognition, when it concerns the impact of language on basic cognitive processes, the evidence has been relatively lacking (see Boroditsky, 2003; Brown, 1986; Gentner & Goldin Meadow, 2003; Gumpertz & Levinson, 1996).

One of the main reasons why evidence for the impact of language on basic cognitive processes, such as attention and perception, is far from conclusive is perhaps that most investigations of this issue have used comparative, cross-cultural analyses of domain-specific language-on-cognition effects. Thus, students of linguistic relativity have compared and contrasted different cultures to find out, for example, whether cultural differences in color coding influence the perception of color, or whether cultural differences in time coding influence the perception of time (e.g., Boroditsky, 2001; Özgen, 2004).

Such cross-cultural analyses of domain-specific language-on-cognition effects will always be somewhat inconclusive (see Brown, 1986; Gumpertz & Levinson, 1996; Ji et al., 2004; Pinker, 1994), especially when it concerns the question of whether language use can have causal effects on basic cognitive processes, such as attention and perception. Such analyses are problematic (a) because of the possible confounding of cultural and linguistic differences and (b) because of the problem of cross-translation
between languages. This latter problem is especially an issue when investigating whether linguistic differences in categorization (e.g., naming color) are related to cognitive differences in a semantically directly implicated cognitive domain (e.g., seeing color).

The New Look

The current studies were designed to avoid these defining features of many previous studies of linguistic relativity (cross-cultural comparisons and domain specificity) by offering a new look at the way language influences cognition. The present research thus focused on how—within the same language (rather than between languages)—the use or cognitive activation of different metasemantic linguistic categories may influence people’s generic (rather than domain-specific) perceptions of objects in a systematic manner. The results of these studies provide strong support for our hypothesis. Across four experiments, using various experimental techniques to expose people to different linguistic categories (e.g., supraliminal and subliminal priming) and using various tasks, each of which is known to tap generic (abstract/global/inclusive vs. concrete/local/exclusive) categorization or perceptual-focus effects, we reliably showed that metasemantic cues can influence basic perceptual processes. Specifically, we consistently showed that abstract predicates (i.e., adjectives) lead to a global focus, whereas concrete predicates (i.e., verbs) lead to a local focus. Of course, each of the measures we used in these studies focused on a slightly different aspect of the variable we are interested in. But as the correlations between these measures suggest, each of them was originally designed to tap the categorization and perception effects of a global, inclusive focus versus a local or specific focus. Thus, the four studies reported here open a new way of looking at the interface between language and perception by emphasizing the attention-driving function of language. In developing this argument in our introduction, we emphasized the advantages of investigating the influence of language upon cognition within a single linguistic community, compared with comparative cultural studies on the subject. Now is the time to look back and see if the implications of current findings can have a bearing upon comparative studies. Thus, the question is whether the current studies can provide a bridge between cultural differences in preferential language use and perception.

Research reported by Semin et al. (2002) has shown that there is a preference for concrete language use in interdependent cultures relative to independent cultures. Maass, Karasawa, Politi, and Suga (2006) showed similar findings. In their studies, (independent) Italians were shown to rely more on adjectives, whereas (interdependent) Japanese used more action verbs in person description and memory. The question that arises in the context of our studies is whether such preferences for abstract versus concrete linguistic devices are likely to give rise to systematic differences in the way different groups of people (e.g., Japanese vs. Italians) perceive stimulus objects. If the results of our third experiment were taken as an instance from which to generalize, then one would expect systematic differences between, say, the way Japanese versus Italian participants make errors in the FLT. Indeed, this is precisely what Kitayama et al. (2003) have shown. Interdependent individuals (Japanese participants) were more sensitive to contextual information, compared with more independent individuals (North American participants). Although quite speculative, this suggests that cultures that are more likely to use concrete language are also more likely to attend to contextual (local) features of a stimulus, relative to cultures that use a more abstract language.

The current studies may thus be seen as providing the missing “corner” in the culture-language-perception triangle, as providing the scaffolding for specific relationships between cultural differences and the attention-driving feature of language. The link that is suggested by these comparisons is that:

1. Generic features of language can be shown to drive attention to different features of a stimulus environment, and

2. Cultural differences in the habitual use of the very same generic properties of language give rise to differences in the way the stimulus environment is perceived.

Although the current research was designed to investigate the interface between language and perception from a generalized perspective, namely investigating the perceptual consequences of generic linguistic categories, it was also designed to constitute a first, if speculative, step in grounding the possible implications of this approach for the interface between language (and cultural differences in habitual use of language) and perception. The distinctive flavor of such an approach is to uncover the general features of the link between language as an attention-driving tool and perception within the same language, thus avoiding potential pitfalls that beset comparative research. It is important to note that the perceptual tasks we used in the current studies not only were especially suited to examine the hypotheses under investigation, but also have been used independently in investigations that have addressed cultural differences (e.g., Kitayama et al., 2003). Moreover, researchers who have addressed differences in linguistic strategies at the cultural level (e.g., Maass et al., 2006; Semin et al., 2002) have relied on the very same feature of language both conceptually and operationally, namely the LCM (Semin & Fiedler, 1988, 1991). We feel that one of the added values of the current research enterprise is that it provides glue and cement that may help to bring these convergent and independent research strategies together. The current research program opens possibilities of investigating the language-perception interface on a substantially broader surface than has been done so far. It effectively introduces a multilevel research program, which has the potential of integrative theory construction.

Conclusion

The current studies attest to the idea that by adopting a new, experiment-based approach to the study of linguistic relativity that focuses on the impact of metasemantic (rather than semantic) linguistic categories and on generic (rather than specific) perceptual processes within (rather than between) languages and cultures, it is possible to provide reliable, empirical evidence for the core of Whorf’s (1957) linguistic-relativity hypothesis: Linguistic categories point people to different types of observations. Language is a tool that directs attention to different aspects of reality.
References


