The evocative power of words: Activation of visual information by verbal and nonverbal means.

Gary Lupyan (lupyan@sas.upenn.edu), Sharon L. Thompson-Schill (sschill@psych.upenn.edu)
Institute for Research in Cognitive Science, Center for Cognitive Neuroscience
University of Pennsylvania
Philadelphia, PA 19104 USA

Abstract

A major part of learning a language is learning to map spoken words onto objects in the environment. An open question concerns the consequence this learning has for cognition and perception. We show that hearing common words (e.g., dog) activates visual information more than equally informative non-linguistic information (e.g., a dog bark). The main results show that (1) pictures were verified more quickly after hearing a word than after hearing a nonverbal sound, even after hundreds of trials of practice. (2) Verbal labels activated visual information more effectively than nonverbal sounds as tested by a simple visual discrimination task that required minimal semantic processing. (3) The advantage of the verbal modality did not arise simply due to greater familiarity of verbal labels: when experience with novel labels and sounds was equated, verbal labels continued to activate the associated visual information more reliably than the equally well-learned nonverbal sounds. These results inform the understanding of how human cognition is shaped by language and hint at effects that different patterns of naming can have on individuals' conceptual structure.

Introduction

Two hallmarks of human development are the development of conceptual categories—learning that things with feathers tend to fly, that animals possessing certain features are dogs, and that foods of a certain color and shape are edible (Carey, 1987; Keil, 1992; Rogers & McClelland, 2004), and learning *names* for those categories. The latter achievement is unique to humans. While many have commented on the transformative power of names (Clark, 1998; Dennett, 1994; Harnad, 2005; James, 1890; Vygotsky, 1962), it is only recently that the interplay between verbal labels and concepts has become a subject of rigorous empirical study.

The learning of categories is, in principle, separable from the learning of language. A child can have a conceptual category of "dog" without having a verbal label associated with the category. However, in practice the two processes are intimately linked. Not only does conceptual development shape linguistic development (Snedeker & Gleitman, 2004), but linguistic development—specifically learning words—impacts conceptual development (Casasola, 2005; Lupyan, Rakison, & McClelland, 2007; Gentner & Goldin-Meadow, 2003; Spelke, 2003; Spelke & Tsivkin, 2001; Waxman & Markow, 1995; Yoshida & Smith, 2005). The effects of words on nonverbal cognition only begin at word-learning. The learned associations between words and their referents appear to continue to influence cognitive processes such as visual recognition memory (e.g., Lupyan, 2008a) and even visual processing (Gilbert, Regier, Kay, & Ivry, 2006; Lupyan, 2008b; Winawer et al., 2007). For example, hearing a verbal label such as "chair" facilitates the visual processing of the named category compared to trials on which participants know the relevant object category but do not actually hear its name (Lupyan, 2007, 2008b; Lupyan & Spivey, 2010). Hearing a label can even make an invisible object, visible (Lupyan & Spivey, 2008). One way to think about such results is that processing a verbal label preactivates the sensory and higher-level representations of objects denoted by the label—over and above activation caused by just thinking about the object category.

The present work addresses the question of how special words are in evoking visual information. Is a highly familiar concept accessed equivalently through verbal and nonverbal means with words being a merely convenient way to activate conceptual information? Or, do words evoke conceptual representations in a special way? We focus here on visual representations and compare the power of verbal and nonverbal cues to evoke visual information of both familiar and novel categories.

It has been long known that a response to a visual stimulus can be altered by a cue presented prior to the target stimulus. These cues can be nonverbal (Egly, Driver, & Rafal, 1994; Eriksen & Hoffman, 1972; Posner, Snyder, & Davidson, 1980) as well as verbal. For example, verbal cues in the form of words like "left" and "right" produce automatic shifts of attention just as reliably as nonverbal cues such as directional arrows even when the words are entirely nonpredictive of the target's location (e.g., Hommel, Pratt, Colzato, & Godijn, 2001). Words related to motion, e.g., "float," have been shown to affect visual motion processing (Meteyard, Bahrami, & Vigliocco, 2007). Several studies have also shown visual object processing can be altered by verbal cues (Puri & Wojciulik, 2008; Vickery, King, & Jiang, 2005). Such effects of cues on visual processing have been linked to increases in category-specific cortical activity. For example, after seeing the word "face," participants are not only better at making a gender judgment of faces embedded in visual noise, but this enhanced discrimination correlates with activity in the fusiform face area (Esterman & Yantis, 2009). These experiments have typically used verbal labels as cues, as language is a natural way to convey information about objects. It is at present unknown whether a verbal label should be thought of as merely a convenient method of cuing—it is a primary function of language to convey information not presently in view—or whether there is something special in the way language activates visual information. In other words, is the type of visual activation produced by hearing the word "cow" somehow special or can it be achieved by nonverbal cues similarly associated with the concept of cows, e.g., a mooing sound. Although both "cow" and the sound of a cow mooing are associated with cows, only the former is treated (in the normal course of things) as referring to a cow.

We present six experiments comparing the powers of verbal and nonverbal cues to evoke visual information. Experiments 1a-1c contrast verbal and nonverbal cues in a series of picture-verification tasks. Experiments 2a-2b contrast verbal and nonverbal cues in a visual discrimination task that requires minimal semantic processing. Experiment 3 tests whether the verbal advantage arises due to participants' greater familiarity with the verbal cues or whether the verbal advantage in evoking visual information is due specifically to the referential status of words.

Experiments 1a-1b

Experiments 1a-b comprised picture verification tasks in which participants heard an auditory cue (a label or a nonverbal sound), and then saw a matching or mismatching picture. If verbal labels activate visual information more reliably than do nonverbal cues, participants should be able to respond more quickly after hearing a label than a nonverbal sound.

Participants

A total of 116 University of Pennsylvania undergraduates volunteered in the experiments in exchange for course credit: 18 in Exp. 1a, 15 in Exp. 1b, 20 in Exp. 1c, 18 in Exp. 2a, 25 in Exp. 2b, and 20 in Exp. 3.

Materials

We selected 10 objects that were easily nameable and that had characteristic sounds (cat, car, dog, frog, gun, motorcycle, rooster, train, cow, whistle). Each category was instantiated by 5 images: a normed color drawing (Rossion & Pourtois, 2004), 3 photographs obtained from online image collections, and 1 "cartoon" image (e.g., a drawing of a cartoon dog). We used several instances of each category to introduce some visual heterogeneity. Spoken labels comprised basic-level names (listed above). Nonverbal sounds were obtained from online environmental sound libraries and judged to be unambiguously related to the target categories through piloting. All sounds were volume and length-normalized.

Procedure

On each trial participants heard a label or nonverbal sound followed by a picture, which, with equal probability, either matched the cue or did not. In the latter case, the picture was randomly selected from among the non-matching category images. Participants responded by pressing a "match" or "does not match" key on a keyboard. Immediately following their response, auditory feedback in the form of a buzz or bleep indicated whether the response was correct. Exps. 1a and 1b differed in one respect: in Exp. 1a the delay between cue offset and picture onset was 400 ms. In Exp. 1b this was increased to 1 s—a common delay used in verification tasks (Stadthagen-Gonzalez, et. al.2009). The rationale for this long delay is that it gives plenty of time for the word or sound to be encoded thoroughly by the time the picture appears. Thus, the verification RTs will be largely determined by the time it takes to recognize the picture rather than reflecting residual processing of the label or sound cue.

All factors were within-subjects and each participants completed 400 verification trials: 10 categories × 5 category exemplars × 2 levels of congruence × 2 cue-types (sound vs. label) × 2 repeats.

Results and Discussion

The data were analyzed using a mixed-effects ANOVA with all factors as within-subject effects. Only correct RTs were included. RTs less than 200 ms or greater than 1500 ms were excluded (1.9% of all trials). An analysis of RTs revealed a highly reliable validity advantage, M_{valid}=552 ms, $M_{invalid}$=600 ms, $F(1,18)=35.72$, $p<.0005$ and a strong advantage for label trials, M_{label}=563 ms, M_{sound}=588 ms, $F(1,18)=24.77$, $p<.0005$ (Figure 1A). This advantage was also observed in accuracy, M_{label}=96.2%, M_{sound}=95.2%, $F(1,18)=6.38$, $p=.02$. There were no reliable cue-type × validity interactions.

Experiment 1b likewise revealed a validity advantage for RTs, $F(1,14)=20.80$, $p<.0005$, and a strong label advantage, M_{label}=583 ms, M_{sound}=620 ms, $F(1,14)=26.80$, $p<.0005$ (Figure 1B). The label advantage was also observed in accuracy, M_{label}=97.8%, M_{sound}=96.0%, $F(1,14)=13.11$, $p=.003$. There was no significant prime-type × item interaction, $F<1$. A replication of Exp. 1b with a 1.5 s delay yielded virtually identical results.

It is conceivable that the advantage of labels is short-lived, owing its existence to the initial unfamiliarity of the sound cues. If so, the advantage should vanish or be diminished with practice. We divided each participant's data into four equal blocks and ran an ANCOVA with block as a covariate. Although participants became faster, and more accurate over time (Fs >10), there were no hints of an interaction between block and cue-type for either RT or accuracy in either experiment, Fs<1.

Experiments 1a-1b show that hearing a verbal label compared to a nonverbal sound affords a quicker identification of a subsequent picture most likely by pre-activating visual information associated with the label allowing for quicker and more accurate acceptance of a congruent picture and a quicker rejection of an incongruent picture.

Experiment 1c

The results from Exps. 1a-1b suggest that labels may play a special role in evoking visual representations owing to their referential nature (see below for discussion). Alternatively, participants may have simply been more familiar with verbal labels than the sounds we used. This latter account predicts that, in a verification context, participants should, on seeing an image, be faster to activate a label than its nonverbal sound. Experiment 1c tested this possibility by reversing the order of the label/sound and picture. Participants now saw a picture first and had to judge a subsequently presented auditory label or nonverbal sound as either matching the picture or not. A finding of a continued advantage of labels would support the familiarity account (but would not necessarily contradict the reference-based account). A disappearance of the label advantage would provide evidence against the familiarity-based account.

Materials and Procedure

Materials were identical to Experiments 1a-1b. The procedure was identical except for the reversal of cue and target identities. On each trial, participants saw a picture for 1 s. One additional second after it disappeared, a verbal label or nonverbal sound was played and the participants task was, as quickly as possible, to press the appropriate key indicating whether the sound matched the picture (valid trial) or not (invalid trial). Participants could start responding at any time after the onset of the target label or sound, although responses generally occurred after the offset of the label or sound. Accuracy feedback was provided immediately after the response.

Results and Discussion

The data were analyzed identically to Exps. 1a-1b. There was a significant validity advantage in RTs, $F(1,19)=17.45$, p=.001: $M_{\text{valid}}=575$ ms, $M_{\text{invalid}}=614$ms. There was no significant difference between label and sound trials, $F(1,19)=2.62$, $p=.12$, with a trend for slower responses times to label trials than to sound trials, $M_{\text{label}}=502$ ms, $M_{\text{sound}}=587$ ms. An analysis of accuracy also failed to find a difference between label and sound cues, $M_{label}=94.6\%$, $M_{\text{sound}}=94.9\%$, $F<1$, further demonstrating that the nonverbal sounds were as recognizable as the labels. Comparing Exps. 1b and 1c revealed a highly reliable experiment × cue-condition interaction, $F(1,33)=24.19$, $p<.0005$.

If the label advantage observed in experiments 1a-1b was a simple consequence of participants' greater familiarity with labels, it was expected that a label advantage would be observed in the present study because viewing the picture would activate the stronger associate—the label—more quickly than the weaker associate—the nonverbal sound. However, that is not what we observed. Rather, the label advantage appears to be asymmetric, occurring when visual information is to be activated by a label cue, but not when a label needs to be activated by a visual cue. An alternative explanation is that lexical items are more complex than environmental sounds and thus require additional processing time. On this account, however, it is unclear why, if labels required greater processing time, we found a reliable verification advantage in Exps. 1a-1b.

Experiments 2a-2b

A limitation of Experiments 1a-1c is that the response requires the participants to semantically classify the image. It is thus unclear whether the label advantage derives from a faster activation of associated visual information (which facilitates subsequent recognition) or if it arises from faster activation of a semantic category itself. To tease apart these accounts, we use a task with a response that depends on visual processing, but only minimally dependent on semantic processing: discriminating an upright image from an upside-down one. The task is similar to one used by Puri and Wojciulik (2008) to examine effects of general and specific cues on visual processing.

Materials

The verbal and nonverbal sounds were identical to Experiments 1a-1c. In addition, a non-informative cue was created consisting of white noise of the same length and volume as the other auditory cues. For the pictures, only the standardized and normed instances of each category were used (Rossion & Pourtois, 2004).

Procedure

On each trial, participants saw two pictures for 200 ms. presented simultaneously to the left and right of a fixation cross. These pictures were identical except one was upside-down (flipped about the x-axis). Participants' task was simply to indicate which side of the screen contained the upright picture by pressing the 'z' key with their left index finger if it was the picture on the left, and the '/' key with their right index finger if it was the picture on the right. It was stressed that it did not matter what object was shown in the picture. The pictures were preceded by an auditory cue. The trials were evenly divided into label cues, sound cues, and uninformative noise cues. The label and sound cues validly cued the upcoming picture on 80% of the trials. On the remaining 20% the cue was invalid, for example, participants would hear "cow" (or hear a mooing sound) but then see a car. This allowed us to measure both the advantage of a valid cue relative to a noise cue (are people faster to locate the upright cow after hearing "cow"/a moo sound?) and the cost of an invalid cue relative to a noise cue baseline (are people slower to identify the upright cow after hearing "car"/a car-starting sound?) And, critically, we can compare these benefits and costs for label and sound cues.

Exps. 2a and 2b differed in only one respect: in Exp. 2a the delay between the offset of the cue and onset of the pictures was 400 ms. In Exp. 2b it was lengthened to 1 s to determine whether the results observed in Exp. 2a were due to insufficient time to process the nonverbal sound. There were 20 practice and 300 real trials.

Results and Discussion

RTs were analyzed by mixed-effects ANOVAs followed by directed t-tests. The first analysis included validity and cue-type (sound vs. label) as fixed factors (validity is undefined for noise cue trials) and subject as a random factor. Results are shown in Figure 1. We found a highly reliable effect of validity, with valid trials being reliably faster than invalid trials, $F(1,17)=39.72$, $p<0005$. There was a significant validity × cue-type interaction with label cues showing a larger cuing effect than sound cues, $F(1,17)=8.23$, $p=.011$. Relative to the no-cue baseline, valid sound cues improved performance by a significant amount, $t(17)=2.84$, $p=.03$). Label cues also improved performance, $t(17)=5.01$, $p<.0005$, and this improvement was significantly greater than the improvement due to sounds, $t(17)=2.93$, $p=.009$. Conversely, relative to the no-cue baseline, invalid label cues significantly slowed responses, $t(17)=4.38$, $p<.0005$; sounds cues did not, $t(17)=1.19$, $p>.2$. There was a significant difference between the cost of invalid labels and the cost of invalid sounds, $t(17)=2.12$, $p=.048$. Accuracy was very high (M=97.8%) and did not vary between conditions.

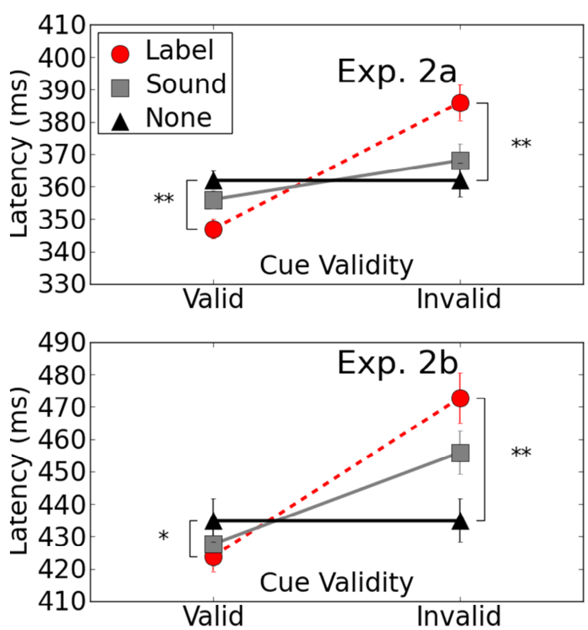

Figure 1: Results of Exps 2a-2b. Error bars show ±1 SE of the difference between noise cues and the condition closest to its mean. The mean of the noise cue trials is plotted twice for ease of comparison.

Did the label advantage result from a lack of time to process the sound cue? This was unlikely given the results of Exp. 1c, but nevertheless, we conducted a replication of Exp. 2a with a longer (1 s) delay between cue offset and picture onset. As shown in Figure 2b, valid labels helped relative to baseline, $t(24)=2.45$, $p=.022$, while sounds did not, $t(24)=1.13$, $p>.2$, though the interaction was not significant. Invalid sounds now hurt performance relative to baseline (although not as much as labels). In sum, labels continued to function as more effective cues than sounds.

With a longer time to process the cue, the nonverbal cues start to act more like verbal cues, quite possibly because participants may explicitly label the nonverbal sounds.

Experiment 3

The studies thus far examined effects of words/sounds on visual processing of objects with which participants have had extensive prior experience. We had no way of knowing whether and what types of differences in experience may have produced the label advantage observed in Exps 1-2. The label advantage is unlikely to be a product of a simple familiarity difference between labels and sounds (see Exp. 1c), but it is possible that labels have a greater power to evoke visual information because they have been more frequently encountered in the context of the visual referent. In Exp. 3, we exerted complete control over by training different groups of participants to associate either novel labels or nonverbal sounds with novel stimuli. A finding of a label advantage in this context would lend support to the idea that words have a special power to evoke visual information.

Materials

The learning set consisted of 6 novel 3D objects (Figure 2). There were 3 variants of each object to increase visual heterogeneity. These variants involved changes in viewpoint and slight changes in feature configuration. Each category was paired with a novel label (shonk, whelph, scaif, crelch, foove, and streil). Each of these nonce words was designed to have approximately equal bigram and trigram statistics and similar real-world lexical neighborhoods. We also created 6 nonverbal sounds: one for each category. These were created by modifying and combining environmental and animal sounds to create 6 sounds that were not readily nameable, as judged by pilot testing.

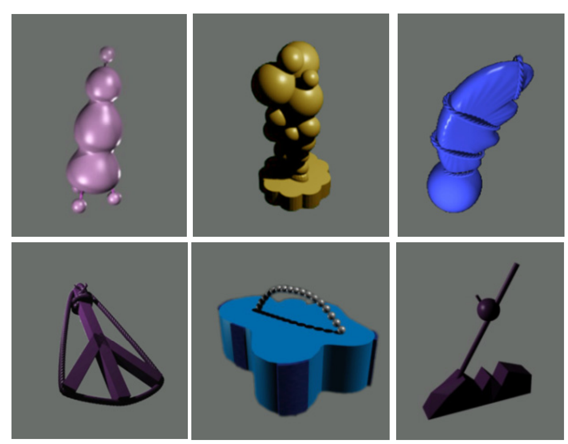

Figure 2: Materials used in the learning task for Exp 3.

Procedure

Participants were randomly assigned into *label* and *sound* groups. There were 3 parts to the experiment presented in immediate succession. In the first part, participants passively viewed 12 trials during which all three exemplars of each category were presented together with a recording, e.g., "These are all shonks" (for the label condition), or These all make the sound___" (for the sound condition).

Part 2 consisted of a verification task. Participants saw two exemplars from different categories followed by a prompt, e.g., “Which one’s the streil?” or “Which one makes the sound___”) and had to select whether the left or right stimulus matched. There were 180 training trials.

The last part was a replication of Experiment 2b with the novel stimuli. That is, participants judged whether the left or right picture was upright (i.e., in the familiar orientation) after hearing a sound or label cue (now without a sentential context). The images were presented for 200 ms after a 1 s delay which was timed to the offset of the auditory cue.

Results and Discussion

Participants were remarkably adept at learning the 6 categories. After Part I—just two exposures to each category—participants could correctly perform the 2AFC task of Part II with ~95% accuracy. The label group was slightly less accurate and slower than the sound group, *p*s=.08, and there were no reliable condition × block interactions. By block 5 both groups were performing at 99%, demonstrating that learning names for novel categories is no more or less difficult than learning what sounds they make.

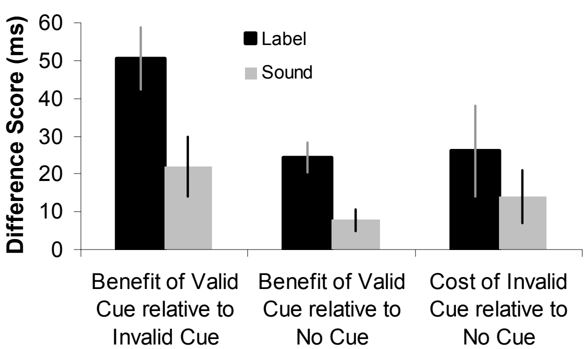

Figure 3: Cuing effects in Experiment 3. Left: $RT_{invalid}$-RT_{valid}. Middle: $RT_{no\text{-}cue}$-RT_{valid}. Right: $RT_{invalid}$-$RT_{no\text{-}cue}$. Error bars show ±SE of the mean difference score.

The critical part of the experiment was subsequent orientation judgment task. Having ruled out entirely differences in familiarity and association strength between labels and sounds, would labels continue to evoke visual activations in a more robust way than sounds? Indeed, that is what we found. As shown in Figure 3, there was a significant validity advantage, $F(1,18)=49.55$, $p<.0005$, but this advantage was significantly larger for label than sound cues, $F(1,14)=6.14$, $p=.023$. The valid cue also benefited RTs relative to the uninformative noise cue, $F(1,18)=38.34$, $p<.0005$, and this benefit was larger for the label than sound trials, $F(1,18)=10.73$, $p=.004$. Finally, there was a significant cost of hearing an invalid cue relative to no cue, $F(1,18)=8.08$, $p=.011$, but this cost was not reliably different for the two groups, $F<1$. An identical pattern of results was found when we used proportions instead of RT differences.

The two groups did not differ in overall response times, M_{label}=433 ms, M_{sound}=389 ms, $F(1,18)=1.70$, $p>.2$, or accuracy, M_{label} = 96.4%, M_{sound} = 95.7%, $F<1$.

In Experiment 3 we had complete control over participants’ exposure to the pictorial stimuli, labels, and sounds. We could thereby ensure that they were equally familiar with the labels and nonverbal sounds. Participants were equally proficient in learning to associate the novel categories with labels or sounds. After only about 10 minutes of training, hearing a label or sound activated the corresponding visual form, as revealed by an RT advantage on valid trials and an RT cost on invalid trials. This in itself is quite remarkable. Critically for our thesis, the label cues were more reliable in activating the corresponding visual form than the sound cues, confirming that even when familiarity and experience with verbal and nonverbal associates is fully equated, verbal cues activate visual information more reliably than nonverbal cues.

General Discussion

Humans learn an elaborate system of sounds (or gestures in case of sign language) that refer, in a largely arbitrary way, to objects, actions, and relations. Beyond enabling linguistic communication, does the acquisition and use of the system confer certain cognitive and perceptual abilities?

In this work, we have investigated whether information communicated verbally (through words denoting concrete objects) and nonverbally (through sounds associated with those objects) activates visual information in the same way. We found that it does not. Cuing categories by using words is more effective than cuing them using nonverbal cues. Verbal cues, more than nonverbal cues appear to preactivate a visual representation of the cued category, helping when the cue is valid and hurting performance when the cue is invalid. This phenomenon is robust, being observed in virtually every subject. A number of control experiments rule out the possibility that this effect is due to different levels of familiarity with verbal versus nonverbal cues.

These findings contradict the popular view that language simply activates nonverbal concepts (Gleitman & Papafragou, 2005; e.g., Li, Dunham, & Carey, 2009; Snedeker & Gleitman, 2004) because presumably such concepts should have been activated in the same way by equally well-learned nonverbal information (as in Exp. 3), but they were not. The finding that representations of very familiar categories (e.g., dogs, cats, and cars) can be evoked more reliably by labels than by sounds, even a full second after cue offset hints at the powerful effects of language on visual activation.

How do words come to have such evocative powers? We believe it is unlikely that there is innately privileged access to vision from the verbal modality (indeed, it is unclear what an innate verbal modality would entail). Rather, the special status of words may derive from accumulated experience of treating them in a referential way (Waxman, 1999), although what exactly this entails vis-à-vis a neural mechanism remains unknown. The present results show that verbal labels serve as powerful cues (Elman, 2004; Rumelhart, 1979), invoking associated concepts and percepts in a unique way, even when the concept in question is a highly familiar one such as [dog]. The finding that after only ~10 minutes of experience, labels affect representations of new concepts (Exp. 3), hints that long-term differences in linguistic experience can have significant effects on the ease of activating specific mental states. Rather than being simple constituent feature of the concept with which it is associated, a name appears to offer a particularly efficient route to the activation of visual and perhaps other information. Although the verbal activation of conceptual information can be deemed a human universal—perhaps the defining feature of language—the present results hint that the substantial differences in lexicalization patterns between languages may translate to cross-linguistic differences in how particular mental states can be evoked.

Acknowledgments

This work was supported by an IGERT training grant to G.L. and NIH R01DC009209 and R01MH70850 to S.T-S. We thank Nina Hsu for designing the stimuli used in Exp. 3, and thank Joyce Shin, Ali Shapiro, and Arber Tasimi for help with data collection.

References

Carey, S. (1987). *Conceptual Change in Childhood* (First paperback edition.). The MIT Press.

Casasola, M. (2005). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology*, *41*(1), 183-192. doi:10.1037/0012-1649.41.1.188

Clark, A. (1998). Magic Words: How Language Augments Human Computation. In *Language and Thought: Interdisciplinary themes* (pp. 162-183). Cambridge University Press.

Dennett, D. (1994). The Role of Language in Intelligence. In *What is Intelligence? The Darwin College Lectures*. Cambridge University Press.

Egly, R., Driver, J., & Rafal, R. (1994). Shifting Visual-Attention Between Objects and Locations - Evidence from Normal and Parietal Lesion Subjects. *Journal of Experimental Psychology-General*, *123*(2), 161-177.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, *8*(7), 301-306. doi:10.1016/j.tics.2004.05.003

Eriksen, C., & Hoffman, J. (1972). Temporal and Spatial Characteristics of Selective Encoding from Visual Displays. *Perception & Psychophysics*, *12*(2B), 201-&.

Esterman, M., & Yantis, S. (2009). Perceptual Expectation Evokes Category-Selective Cortical Activity. *Cereb. Cortex*, bhp188. doi:10.1093/cercor/bhp188

Gentner, D., & Goldin-Meadow, S. (2003). *Language in Mind: Advances in the Study of Language and Thought*. Cambridge, MA.: MIT Press.

Gilbert, A., Regier, T., Kay, P., & Ivry, R. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 489-494.

Gleitman, L., & Papafragou, A. (2005). Language and thought. In *Cambridge Handbook of thinking and Reasoning* (pp. 633-661). Cambridge: Cambridge University Press.

Harnad, S. (2005). Cognition is categorization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (pp. 20-45). Elsevier.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science*, *12*(5), 360-365.

James, W. (1890). *Principles of Psychology. Vol. 1*. New York: Holt.

Keil, F. C. (1992). *Concepts, Kinds, and Cognitive Development*. The MIT Press.

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology*, *58*(4), 487-524. doi:10.1016/j.cogpsych.2008.12.001

Lupyan, G. (2007). *The Label Feedback Hypothesis: Linguistic Influences on Visual Processing*. PhD. Thesis. Carnegie Mellon University.

Lupyan, G. (2008a). From chair to "chair:" A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General*, *137*(2), 348-369.

Lupyan, G. (2008b). The Conceptual Grouping effect: Categories matter (and named categories matter more). *Cognition*, *108*, 566-577.

Lupyan, G., Rakison, D., & McClelland, J. (2007). Language is not just for talking: labels facilitate learning of novel categories. *Psychological Science*, *18*(12), 1077-1082.

Lupyan, G., & Spivey, M. (2008). Now You See It, Now You Don't: Verbal but not visual cues facilitate visual object detection. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 963-968). Austin, TX.

Lupyan, G., & Spivey, M. (2010). Redundant spoken labels facilitate perception of multiple items. *under review*.

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs - Language affects low-level visual perception. *Psychological Science*, *18*(11), 1007-1013.

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology-General*, *109*(2), 160-174.

Puri, A., & Wojciulik, E. (2008). Expectation both helps and hinders object perception. *Vision Research*, *48*(4), 589-597.

Rogers, T., & McClelland, J. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, MA: Bradford Book.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217-236.

Rumelhart, D. (1979). Some problems with the notion that words have literal meanings. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 71-82). Cambridge University Press.

Snedeker, J., & Gleitman, L. (2004). Why is it hard to label our concepts? In D. G. Hall & S. R. Waxman (Eds.), *Weaving a Lexicon* (illustrated edition., pp. 257-294). The MIT Press.

Spelke, E. (2003). What Makes Us Smart? Core knowledge and natural language. In *Language in Mind: Advances in the Study of Language and Thought* (pp. 277 -311). Cambridge, MA.: MIT Press.

Spelke, E., & Tsivkin, S. (2001). Initial knowledge and conceptual change: Space and number. In *Language acquisition and conceptual development* (pp. 475-511). Cambridge, UK: Cambridge University Press.

Stadthagen-Gonzalez, H., Damian, M. F., Pérez, M. A., Bowers, J. S., & Marín, J. (2009). Name–picture verification as a control measure for object naming: A task analysis and norms for a large set of pictures. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1581. doi:10.1080/17470210802511139

Vickery, T. J., King, L., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), 81-92. doi:10:1167/5.1.8

Vygotsky, L. (1962). *Thought and Language*. Cambridge, MA: MIT Press.

Waxman, S. (1999). The dubbing ceremony revisited: Object naming and categorization in infancy and early childhood. In *Folkbiology* (pp. 233 -284). Cambridge, MA: MIT Press.

Waxman, S., & Markow, D. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*(3), 257-302.

Winawer, J., Witthoft, N., Frank, M., Wu, L., Wade, A., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 7780-7785.

Yoshida, H., & Smith, L. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, *16*(2), 90-95.