Title
Self-directed speech alters visual processing

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

ISSN
1069-7977

Authors
Lupyan, Gary
Swingley, Daniel

Publication Date
2010

Peer reviewed

Self-directed speech alters visual processing

Gary Lupyan (lupyan@sas.upenn.edu)
Institute for Research in Cognitive Science, Center for Cognitive Neuroscience
University of Pennsylvania
Philadelphia, PA 19104 USA

Daniel Swingley (swingley@psych.upenn.edu)
Department of Psychology, Institute for Research in Cognitive Science
University of Pennsylvania
Philadelphia, PA 19104 USA

Abstract

A major part of learning a language is learning connections between spoken words and their referents in the world. An open question concerns the consequence this learning has for cognition and perception. According to the label feedback hypothesis (Lupyan, 2007), processing a verbal label can change ongoing perceptual processing, e.g., actually hearing "chair" compared to simply thinking about a chair temporarily makes the visual system a better chair detector. Here, we test whether engaging in a non-communicative verbal act—speaking to oneself—also affects visual processing. Participants searched for common objects, sometimes being asked to speak the target's name aloud. Speaking facilitated search, but only when there was a strong association between the name and the visual target. Speaking appeared to hurt performance when there was even a slight discrepancy between the name and the target. Together these results speak to the power of words to evoke associated visual information.

Introduction

Learning a language involves, among other things, learning to map words onto categories of objects in the environment. In addition to learning that chairs are good for sitting on, one learns that this class of objects has the name "chair." Clearly, this learning is critical for linguistic communication. But beyond communication, what consequences does naming things—hearing and producing verbal labels—have on perception and nonverbal cognition?

On one account language is a "transparent medium through which thoughts flow" (H. Gleitman, Fridlund, & Reisberg, 2004, p. 363). Therefore, words are mapped onto concepts, but do not affect them (e.g., L. Gleitman & Papafragou, 2005; Gopnik, 2001). Thus, while word-learning is significantly constrained by nonverbal cognition, nonverbal cognition is not significantly influenced by learning or using words (Snedeker & L. Gleitman, 2004).

The alternative is that words are not simply mapped on to concepts, but actually change them, affecting nonverbal cognition, and even perception. The idea that words can affect the concepts to which they refer is not new: William James, for example, remarked on the power of labels to make distinctions more concrete (James, 1890, p. 333), and it has been argued that words stabilize abstract ideas in working memory and make them available for inspection (Clark, 1997; Clark & A Karmiloff-Smith, 1993; Dennett, 1994; Goldstein, 1948; Rumelhart, Smolensky, McClelland, & Hinton, 1986; Vygotsky, 1962). This is not to say that different languages necessarily place strong constraints on the speaker's *ability* to entertain certain concepts. Rather, it is a claim that language richly interacts with putatively nonlinguistic processes such as visual perception.

Insofar as performance on putatively nonverbal tasks draws on language, interfering with language should interfere with performance on those tasks (Goldstein, 1948). Indeed, verbal interference impairs certain types of categorization in a way strikingly similar to impairments observed in aphasic patients (Lupyan, 2009). Interfering with language can also affect perception. A number of studies have shown that interfering with language impairs categorical color perception (e.g., Gilbert, Regier, Kay, & Ivry, 2006; Roberson & Davidoff, 2000; Roberson, Pak, & Hanley, 2008; Winawer et al., 2007), suggesting that language actively modulates visual processing.

An additional way to study affects of language on perception is by attempting to *increase* rather than decrease its putative effect. A surprising finding is that when asked to find a certain visual item among distractors actually hearing its name immediately prior to performing the search—even when the label is entirely redundant—improves speed *and efficiency* of searching for the named object (or searching among the named objects). For example, when participants search for the numeral 2 among 5's (for hundreds of trials), actually hearing the word "two" (or hearing "ignore fives") immediately prior to doing the search, improves search RTs and reduces search slopes (Lupyan, 2007a, 2008a). Indeed, hearing an object name can even make an otherwise invisible object visible (Lupyan & Spivey, 2008; under review).

One way to understand such findings is in terms of an interactive activation framework (Rumelhart & McClelland, 1982; Spivey, 2008) in which recognition involves the combination of bottom-up perceptual information, with higher-level top-down (conceptual) information. As one learns a verbal label, it becomes associated with features that are most diagnostic (or typical) of the named category. With such associations in place, hearing the label provides top-down activation of visual properties associated with the label. In effect, the object name makes an object a "better" object by augmenting the idiosyncratic perceptual features of a given object with features typical to the named category (Lupyan, 2007b, 2008b).

Aims and Hypotheses

In the present work, we investigate whether non-communicative (self-directed) speech can affect visual processing in the context of a visual search task. Does producing the name of a pre-defined target object enable subjects to find it faster? Participants were asked to find an object among distractors while speaking its name or not. We predicted that actually speaking the object's name would facilitate visual search—even though such speaking can be seen to constitute a form of distraction. We also predicted that the effect of speaking would be largest for items most strongly associated with the label, and speaking might actually be detrimental when searching for objects having weaker associations with the label, e.g., objects judged as being less typical of their categories.

Experiment 1

The participants' task was to find and click on a target object among 35 distractors, positioned randomly in a 6×6 grid on a computer screen (Figure 1). For half the trials, participants were asked to speak the name of the target as they searched for it.

Participants

Twelve University of Pennsylvania undergraduates participated for course credit.

Materials

The targets and distractors were drawn from a set of 260 colored images of common objects (Rossion & Pourtois, 2004). For the targets, we selected 20 images having 100% picture-name agreement, as assessed by Rossion and Pourtois (2004) (airplane, banana, barn, butterfly, cake, carrot, elephant, giraffe, chicken, ladder, lamp, leaf, truck, motorcycle, mouse, mushroom, rabbit, tie, umbrella, windmill).

Figure 1: A sample search trial from Exp. 1

For a given trial, any of the 259 non-target images could serve as distractors. Rossion and Pourtois provide a number of measures for these pictures, which we included for item-analyses. Most relevant to the present work are: RT to name the picture, familiarity, subjective visual complexity, and imagery-concordance. The latter measure was derived by presenting participants with a picture name (e.g., butterfly), asking them to form a mental image of the object, and then, on seeing the actual picture, providing a rating of imagery agreement. For the lexical items themselves, we obtained log frequency from the British National Corpus, word length in phonemes and syllables, actual age-of-acquisition (AoA) norms (Morrison, Chappell, & Ellis, 1997), and several measures from the MRC Psycholinguistic Database (www.psych.rl.ac.uk/): imageability, concreteness, and word familiarity.

Procedure

Each trial began with a prompt informing the participant what object they would need to find. The prompt also informed them whether they should repeat the object's name as they searched for it, or not. For example, immediately prior to a no-speaking trial, participants saw a prompt such as "Please search for a butterfly. Do not say anything as you search for the target" For a speaking trial, the second sentence was replaced by "Keep repeating this word continuously into the microphone until you find the target." The speech/no-speech trials were intermixed, as were the target identities. Participants completed 320 trials: 20 targets × speech condition (speaking vs. not speaking) × 8 blocks. A block included all target × speech condition combinations. Participants used a computer mouse to click on the target object.

Results and Discussion

Participants showed excellent compliance with the instruction to speak the name of the target on the label trials and to remain silent on the no-speaking trials. We focus on accuracy and median RTs to find the target as the main dependent measures. Comparisons between conditions were made using a mixed-effects ANCOVA with speech condition as a fixed effect, subject as a random effect, and block as a covariate. For reasons described in Thomas et al., (2009), separate tests were run to assess fixed factor main effects and those of the covariate × factor interaction.

Accuracy was extremely high, M=98.8%, revealing that (1) subjects had no trouble remembering which item they were supposed to find, and (2) the word cues were sufficiently informative to locate the correct object. Despite this very high accuracy, saying the object's name during search resulted in significantly higher accuracy, M=99.2% than not repeating the name, M=98.4%, $F(1,11)=12.19$, $p=.005$. Participants' accuracy increased over the course of the experiment, $F(1,11)=10.90$, $p=.001$, but there was no reliable speech-condition × accuracy interaction, $F(1,11)=1.49$, $p>.2$.

The analysis of median RTs included correct responses only. Unsurprisingly, participants' speed improved over the course of the experiment, $F(1,11)=22.85$, $p<.0005$. There

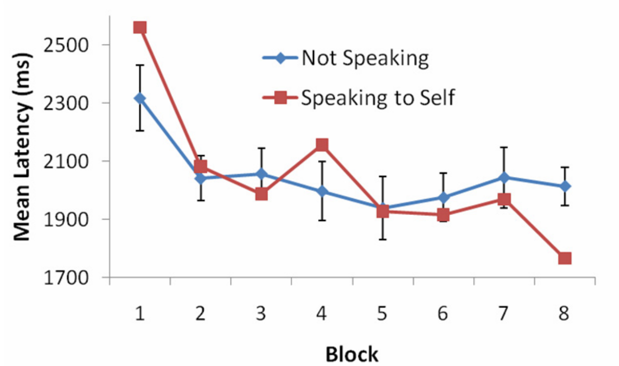

Figure 2: RTs in Exp. 1: Speaking significantly decreased RTs for the second half of the task. Error bars show ±1SE of the mean condition difference. Accuracy was significantly higher for the speaking condition throughout the task; see text.

was no main effect of the speech-condition on RTs, F<1, but there was a highly reliable speech-condition × block interaction, $F(1,11)=8.1$, $p=.004$. As shown in Figure 2, performance on the speech trials tended to be slower than on no-speech trials for the initial blocks, but this pattern reversed for the latter part of the experiment. Collapsing the last three blocks, participants were faster on speech trials than no-speech trials, $t(11)=2.91$, $p=.01$ (two-tailed). This finding suggests that although the target objects were very familiar, speaking the name decreased RTs only when participants had several opportunities to associate the picture name with the target picture, which presumably strengthened the picture-name association.

We next turn to the item analysis. A number of item factors predicted overall search performance. Search was faster, $r(18)=.55$, $p=.01$, and more accurate, $r(18)=-.54$, $p=.02$, for pictures that were visually simpler according to Rossion and Pourtois's (2004) norms. Search was faster, $r(18)=-.55$, $p=.01$, and slightly more accurate, $r(18)=.34$, $p=.15$ for pictures with higher imagery-concordance. Familiarity did not predict search times or accuracy. Lexical

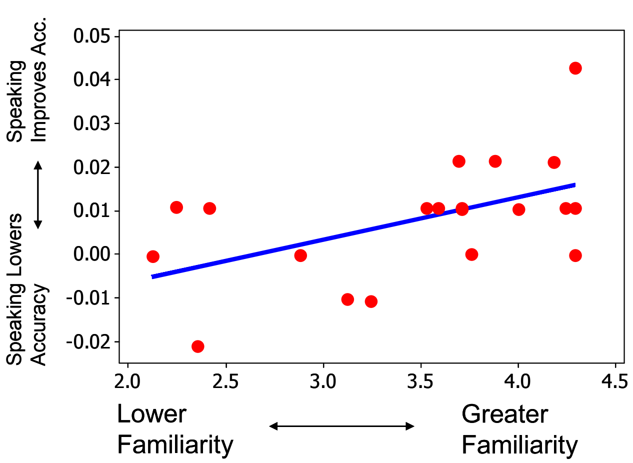

Figure 3: Relationship between item familiarity and effects of speaking on accuracy. Y-axis shows % correct when speaking - % correct when not speaking.

variables did not predict overall search performance, though there were marginal correlations of search times with word frequency, $r(18)=-.38$, $p=.10$, and of accuracy with age-of-acquisition (AoA) provided by adults, $r(18)=-40$, $p=.08$.

Finally, we examined which items were most affected by self-directed speech by subtracting performance on speech trials from performance on no-speech trials. Overall, speaking improved RTs most for the items which took, on average, the least time to find, $r(18)=-.57$, $p=.009$, and ones for which accuracy was, on average, the highest, $r(18)=.47$, $p=.037$. Recall that familiarity was not related to overall accuracy. However, separating accuracy into speech and no-speech trials revealed a very different pattern. Familiarity was unrelated to performance on no-speech trials, $p>.3$, but was highly correlated with performance on speaking trials, $r(18)=.55$, $p=.01$. The interaction was significant: speaking improved accuracy most for the more familiar items, $r(18)=.51$, $p=.02$ (Figure 3). Finally, RTs were improved marginally more for the items with the highest imagery-concordance, $r(18)=.39$, $p=.08$.

We also observed a relationship between AoA and self-directed speech. This relationship changed over the course of the experiment: for the first half of the task, AoA (both subjective and objective), correlated with the effect of speaking on search times, $r_{objective\ AoA}(28)=-.54$, $p=.02$, $r_{subjective\ AoA}=-.62$, $p=.003$: performance was impaired by saying words having higher AoA. By the second half of the task, these correlations disappeared entirely, $rs<.1$.

For interpretive ease, we performed a median split on the familiarity and imagery-concordance values. The label advantage ($RT_{without-speaking}-RT_{speaking}$) was larger for items having imagery-concordance scores above than below the median, $F(1,18)=6.32$, $p=.022$. Search items below the median were actually slowed by speaking, $t(10)=2.24$, $p=.049$ (two-tailed). The label advantage in accuracy trended in the same direction, being (marginally) larger for items with above-median familiarity ratings, $F(1,18)=4.19$, $p=.056$.

To summarize: speaking facilitated search for pictures judged in a separate norming study to be most familiar, and targets having the highest concordance between the actual image and the mental image formed by reading the name.

One way in which self-directed speech may help visual search is through verbal rehearsal: saying the name of the target might have helped participants remember what it was they were looking for. This account is not supported for two reasons. First, accuracy was extremely high, making it unlikely that difficulties in remembering the target played a significant role. Second, a memory-based account would predict that speech should help most for items that were most difficult to find. We found exactly the opposite pattern.

The item effects presented above place some constraints on the mechanisms by which labels affect visual search. One possibility is that saying the target name helps to find the target by activating and/or keeping active the visual features typical to that object (e.g., saying "cherry" makes it easier to attend to red things). Alternatively (or additionally), repeating a label helps to reject distractors. If speaking facilitated search only by improving rejection of distractors, one would not predict correlations between the magnitude of the speaking advantage and properties of the target. The presence of these correlations supports the hypothesis that speaking the target's name facilitates deployment of attention to the target item over and above seeing the printed name of the target.

The present results can be viewed as an extension of findings showing that hearing a label, even when it is entirely redundant, facilitates visual search, and this facilitation is greatest for the stimuli most strongly associated with the label (Lupyan, 2007a, 2007b, 2008a). When visual quality of the item is reduced, or the item is made more ambiguous, hearing a label can impair performance (Lupyan, 2007b). Thus, compared to just being told what to find, speaking a target name—just like hearing it—affects visual search.

Experiment 2

The goal of Experiment 2 was to test whether self-directed speech affects performance on a more difficult and ecologically valid "virtual shopping" task in which participants search for supermarket products in a visually complex display and were required to find several instances of a category.

Participants

Twenty-two University of Pennsylvania undergraduates (14 women) participated for course credit.

Materials

We photographed products on supermarket shelves in the Philadelphia area and selected 30 to serve as targets, e.g, apples, Pop-Tarts, Raisin Bran, Tylenol, Jell-O. For each product, we obtained three pictures depicting instances of the product in various sizes and orientations. Some pictures depicted multiple instances of the product, e.g., a shelf containing multiple cartons of orange juice. See Figure 4 for some examples.

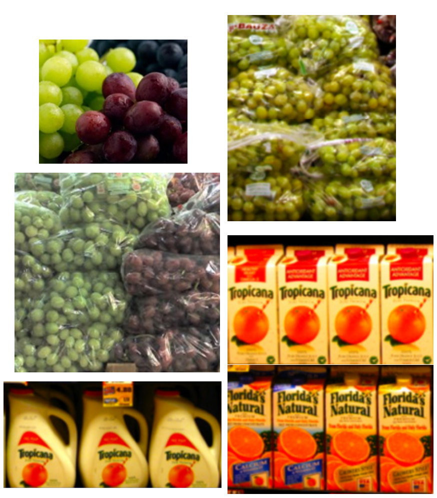

Figure 4: Samples of 2 search categories used in Exp. 2.

Procedure

As in Exp. 1, participants were instructed that they would need to search for various items while being asked to sometimes speak the items' names. Each trial included all three instances of the product and 13 distractors. Clicking on an object made it disappear, thus marking it as being selected. Once satisfied with their choices, participants clicked on a large "Done" button that signaled the end of the trial. To make the task more challenging, some of the distractors were categorically related to the target, e.g., whenever searching for "Diet Coke," some distractors were of other sodas, e.g., "Ginger Ale." There were a total of 240 trials (30 targets by × 8 blocks). Within each block, half the items were presented in a speech trial and half in a no-speech trial. Speech and no-speech trials alternated. Across the 8 blocks, each item was presented an equal number of times in speech trial and no-speech trials.

Prior to beginning the search task, participants rated each item on typicality ("How typical is this box of Cheerios relative to boxes of Cheerios in general?"), and visual quality ("How well does this picture depict a box of Cheerios?"). For each item category (i.e., all three images of Cheerios), participants rated its familiarity ("Overall, how familiar to you are the objects depicted in these pictures?") and visual similarity ("Considering only the visual appearance of these picture, how different are they from each other?"). In addition to providing us with item information, this task served to pre-expose participants to all the targets. We also obtained an imageability measure from a separate group of participants (N=28) who were shown the written product names, e.g., "Cheerios" and asked to rate how well they could visualize its appearance on a supermarket shelf.

Results and Discussion

The data were analyzed in the same way as in Exp. 1. Overall, participants were very accurate, averaging 1.5% false alarms and 97.7% hits (2.93 out of 3 targets). Overall performance (RTs, hits, and false alarms) correlated with all four item variables (visual similarity, visual quality, familiarity, and typicality). Correlation coefficients ranged from .35 to .65 (ps between .035 and <.0005). Items that were familiar, typical, of higher quality, and having least within-category similarity were found faster and with higher accuracy. Of course, the item variables were not all independent, e.g., familiar items and those of higher quality tended to be rated as more typical. The typicality and familiarity measures clustered together and were not independently predictive of performance (familiarity was the stronger of the two predictors). Within category visual similarity predicted performance independently of familiarity; multiple regression: $F(2,27)=9.15$, $p=.001$.

There was a reliable difference in hits between the two speech conditions: $M_{speech}=97.9\%$, $M_{no\text{-}speech}=99.1\%$, $F(1,21)=11.19$, $p=.003$. While speaking the product name, participants were more likely to miss one or more of the targets. As reported below, however, this effect was modulated strongly by the different targets in predictable ways. Speech-condition was not a reliable predictor of false-alarms, $F(1,21)<1$. There were no differences in total or per-click RTs between the speech and no-speech conditions, $F<1$. The speech-condition × block interaction was not reliable, $F<1$.

The item analyses in Exp. 1. suggested that effects of self-directed speech may be modulated by the relationship between the item and its name. Indeed, the cost in the hit rate incurred by speaking ($Hits_{no-speech}$-$Hits_{speech}$) was correlated with within-category similarity, $r(28)=-.34$, $p=.04$: the categories having the most dissimilar items incurred the highest cost when their names were repeated during search. The effect of self-directed speech ($RT_{no-speech}$-RT_{speech}) was also mediated by familiarity, $r(28)=-.51$, $p=.004$: labels tended to hurt performance for the less familiar items, but *improve* performance for the more familiar items (Figure 5). The label advantage also correlated positively with product imageability, $r(28)=.44$, $p=.01$. As an added confirmation of this finding, we divided the targets into those having characteristic colors (N=11), e.g., bananas, grapes, cheerios, raisin bran and those with weaker color associations, e.g., Jell-O, Pop-Tarts. The speaking advantage was greater for color-diagnostic items (for which speaking significantly improved RTs) than for non color-diagnostic items (for which speaking marginally increased RTs), $F(1,28)=7.35$, $p=.01$.

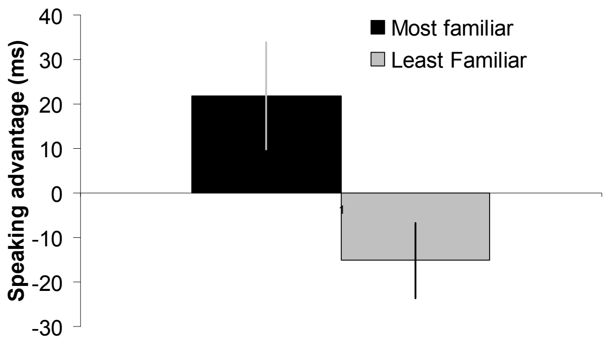

Figure 5: Speaking advantage (no-speech – speech trials) as a function of familiarity (median split). RTs were decreased by speaking the names of the more familiar items and increased by speaking the names of the least familiar items. Errors bars indicate 1±SE of the mean difference.

Exp. 2 revealed a striking gender difference in performance. Men had a significantly lower hit rate, $F(1,20)=5.02$, $p=.037$, and were significantly slower, $F(1,20)=6.37$, $p=.02$ to find the targets. The gender effect on RTs was substantial: men took on average 350 ms longer per trial. This effect was replicated in an item analysis, $F_2(1,29)=43.40$, $p<.0005$ (the *only* item on which men were faster than women was “Degree Deodorant”). There was a marginal gender × speech-condition interaction for hit rates, $F(1,20)=3.79$, $p=.066$: labels hurt performance slightly more for men than women. An examination of item ratings revealed that there were no gender differences in subjective ratings of familiarity, visual-quality, or visual-similarity, Fs<1, and only a marginal difference in typicality: women believed our items to be slightly more typical than did men, $F(1, 20)=2.66$, $p=.12$. In an effort to better understand the origin of this gender difference, we correlated the magnitude of the female advantage with various ratings of the stimuli. We observed a mildly reliable relationship between the magnitude of the female RT advantage and the measure of visual similarity: $r(26)=.38$, $p=.049$. The advantage was greatest for the most visually similar items (two items were excluded, as statistical outliers). There were no other reliable correlations.

Using a larger, more perceptually varied and true-to-life item set, the item analyses of Exp. 2 reinforced the conclusions of Exp. 1. As in Exp. 1, speaking aided search for the more familiar items. In contrast to Exp. 1, accuracy (hit rate) was actually decreased by speaking, though this decrease was limited to the items having low within-category similarity. This finding is consistent with the idea that speaking an object name activates a (proto)typical representation of the category. When the task requires finding items that diverge from this prototype (as when participants need to find visually heterogeneous items from the same category), speaking can impair performance.

General Discussion

Can language affect ongoing perceptual processing? A growing body of literature argues that it can. The present work is the first to examine effects of non-communicative (self-directed) speech on a visual task.

The findings show that speaking the name of the object that one is searching for improves search performance, provided that the object’s name is strongly associated with the visual depiction of the object.

The present results are somewhat less reliable than those of hearing labels on visual search (Lupyan, 2007a, 2008a). Subsequent work has shown that the effects of speech on visual processing have a characteristic timecourse, peaking about 0.5-1.5 seconds after the presentation of the label, and declining afterwards (Lupyan & Spivey, 2010, under review). In the present studies we did not have precise control over the timing of the label. Recordings of participants’ speech from the present work revealed a wide variability in the onset, speed, and duration of self-directed speech. Thus, more reliable effects may be obtained with finer control over speaking onset and rate.

Our results join work arguing for cognitive functions of self-directed speech. For example, even mild forms of articulatory suppression impair adults’ ability to switch from one task to another (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003; Miyake, Emerson, Padilla, & Ahn, 2004). The present results are consistent with Vygotsky’s claim that the function of self-directed speech extends far beyond verbal rehearsal (Carlson, 1997; Vygotsky, 1962)—itself a learned strategy (Flavell, Beach, & Chinsky, 1966).[1]

The present work comprises a first step in understanding effects of self-directed speech on visual processing. One unanswered question is whether effects of speaking on visual search arise from the act of production itself, or from

[1] It is worth noting that these articulatory suppression effects on putatively nonverbal task-switching were compelling enough for Baddeley to concur with Vygotsky’s claim (Baddeley et al., 2001, p. 655).

hearing one's speech. Although this distinction is of little practical importance—one almost always hears oneself speak—a full understanding of the mechanism by which speech and visual processing interact requires the two explanations to be teased apart. Despite these unknowns, the present results show that in the context of searching for a familiar object, knowing what an object is called is not the same as actually saying its name.

Acknowledgments

We thank Ali Shapiro, Joyce Shin, for their help with data collection and for assembling the stimulus materials.

References

Baddeley, A. D., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, *130*(4), 641-657. doi:doi:10.1037/0096-3445.130.4.641

Carlson, R. A. (1997). *Experienced Cognition* (1st ed.). Psychology Press.

Clark, A. (1997). *Being There: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clark, A., & Karmiloff-Smith, A. (1993). The Cognizer's Innards: A Psychological and Philosophical Perspective on the Development of Thought. *Mind & Language*, *8*(4), 487-519.

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

Emerson, M., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48*(1), 148-168.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous Verbal Rehearsal in a Memory Task as a Function of Age. *Child Development*, *37*(2), 283-299.

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, H., Fridlund, A., & Reisberg, D. (2004). *Psychology* (6th ed.). New York: Norton & Company.

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

Goldstein, K. (1948). *Language and language disturbances*. New York: Grune & Stratton.

Gopnik, A. (2001). Theories, language, and culture: Whorf without wincing. In *Language acquisition and conceptual development* (pp. 45-69). Cambridge, UK: Cambridge University Press.

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

Lupyan, G. (2007a). Reuniting categories, language, and perception. In D. McNamara & J. Trafton (Eds.), *Twenty-Ninth Annual Meeting of the Cognitive Science Society* (pp. 1247-1252). Austin, TX: Cognitive Science Society.

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

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

Lupyan, G. (2008b). 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. (2009). Extracommunicative Functions of Language: Verbal Interference Causes Selective Categorization Impairments. *Psychonomic Bulletin & Review*, *16*(4), 711-718. doi:10.3758/PBR.16.4.711

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*.

Miyake, A., Emerson, M., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, *115*(2-3), 123-142. doi:10.1016/j.actpsy.2003.12.004

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of Acquisition Norms for a Large Set of Object Names and Their Relation to Adult Estimates and Other Variables. *The Quarterly Journal of Experimental Psychology A*, *50*, 528-559. doi:10.1080/027249897392017

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition*, *28*(6), 977-986.

Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, *107*(2), 752-762. doi:10.1016/j.cognition.2007.09.001

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., & McClelland, J. (1982). An Interactive Activation Model of Context Effects in Letter Perception .2. the Contextual Enhancement Effect and Some Tests and Extensions of the Model. *Psychological Review*, *89*(1), 60-94.

Rumelhart, D., Smolensky, D., McClelland, J., & Hinton, G. (1986). Parallel Distributed Processing Models of Schemata and Sequential Thought Processes. In *Parallel Distributed Processing Vol II* (pp. 7-57). Cambridge, MA: MIT 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.

Spivey, M. (2008). *The Continuity of Mind*. Oxford University Press.

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research: JSLHR*, *52*(2), 336-358. doi:10.1044/1092-4388(2009/07-0144)

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

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.