Pictures and Words: Priming and Category Effects in Object Processing

Karima Kahlaoui*, Thierry Baccino, Yves Joanette and Marie-Noëlle Magnié
One central issue in cognitive psychology concerns how the human brain stores and accesses conceptual representations of objects in response to both pictures and words. Two general classes of models have been proposed. Some researchers have argued for a multiple semantic system model, asserting that the visual and verbal input modalities have separate conceptual representations (e.g., Paivio, 1986, 1991; Shallice, 1988). For example, Paivio’s dual coding theory posits the “imagen” system for picture processing and the “logogen” system for word processing. These systems can communicate with another but operate differently and have their own organization and processing parameters. Others have argued for a single amodal semantic system model, claiming that all processing routes converge on a single set of conceptual representations, common to both modalities (Caramazza, Hillis, Rapp & Romani, 1990). These two types of accounts make different predictions about the pattern of results that should be obtained when one compares picture and word processing and looks for interactions between them. The multiple semantic system model predicts that picture and word processing should be processed differently, and that interactions across modalities should be delayed and weak compared to within-modality interactions. In contrast, the single amodal semantic system model predicts that pictures and words will be processed similarly and there should be no differences in facilitation or interference between within and across-modality interactions. Behavioral studies in both normal and neuropsychological participants have produced a large amount of data on this topic (Job & Tenconi, 2002; Lloyd-Jones & Humphreys, 1997; Seifert, 1997; Snodgrass & McCullough, 1986; Vitkovitch, Cooper-Pye & Leadbetter, 2006). However, most findings can be interpreted as supporting either model, and thus no consensus has been reached.

Numerous studies in cognitive psychology literature have also shown that object processing varies as a function of semantic category. Studies of category-specific effects
have become increasingly important in discovering the organization of semantic memory (e.g., Warrington & McCarthy, 1994). This field of research originated in neuropsychological studies with brain-damaged patients. Most reports of category-specific deficits describe patients with impaired recognition of biological (e.g., animals, fruits) relative to man made objects (e.g., furniture, tools), while the opposite pattern is reported much less frequently (for reviews, see Capitani, Laiacona, Mahon & Caramazza, 2003; Laws, 2005). While some accounts of category-specific deficits have focused on explanations pertaining to semantics (Warrington & McCarthy, 1987), an increasing number of accounts have emphasized that category-specific deficits for biological objects may arise following damage to pre-semantic stages of visual object recognition (Gerlach, Law & Paulson, 2004; Humphreys & Forde, 2001; Humphreys, Riddoch & Quinlan, 1988; Lloyd-Jones & Humphreys, 1997; Moore & Price, 1999). According to the visual similarity hypothesis, biological objects have similar shapes and are therefore perceptually not differentiated from each other. For that reason, biological objects are structurally similar. In contrast, man made objects have distinct shapes and, for that reason, they are considered to be structurally dissimilar. In addition, man made objects, which can be manipulated, are recognized on the basis of functional semantic knowledge. Indeed, some authors have shown that the action plays an important role in their recognition; thus, they have argued that sensorimotor experiences play a critical role in man made object processing (Chainay & Humphreys, 2002; Magnié, Besson, Poncet & Dolisi, 2003; Magnié, Ferreira, Giusiano & Poncet, 1999). Thus, it is assumed that the visual identification of structurally similar objects is more difficult than the identification of structurally dissimilar objects. Consistent with this hypothesis, it has been found that control participants make more errors and have longer reaction times when identifying biological as opposed to man made objects (Gerlach, 2001). However, the most influential studies that have been presented as evidence for category-specific effects have employed pictures. Few studies have demonstrated that category-specific deficits have implications for modality-specific organization, and their results are contradictory. Whereas some authors found that category-specific effects may only be present in the picture processing (Farah, McMullen & Meyer, 1991), others demonstrated that they can be also selectively observed in the word modality (McCarthy & Warrington, 1988). Thus, there is some evidence indicating that performance on object semantic processing differs as a function of both category-specificity and presentation modality, but the precise mechanisms underlying this phenomenon remain unknown. Another important point concerns the semantic context. Indeed, most studies of the organization of semantic memory have been conducted in a semantic context (e.g., using a semantic priming paradigm). It is not known how these factors interact in the absence of any semantic context.

In this article, we investigate object semantic processing as a function of category when objects are presented in isolation as pictures or words (Experiment 1), and when they are preceded by a prime (Experiment 2). In order to further describe the similarities and differences between both modalities, the processing of pictures was compared with that of words in a reality decision task (Kroll & Potter, 1984). The reality decision task is an original task that mixes object and lexical decisions; it is very informative for investigating both picture and word processing. In addition, the task requires different levels of abstraction as a function of how the meaningless stimuli are constructed. In a recent investigation, Magnié et al. (2003) proposed differentiating between two kinds of meaningless stimuli: chimeric objects (i.e., stimuli constructed out of parts of two real objects of the same semantic category), and non-objects (i.e., stimuli created by mixing up
the line drawings of real objects with the constraint that they were not reminiscent of any actual objects). These authors point out that chimeric objects and non-objects can be considered similar to the pseudo-words and non-words, respectively, used in a lexical decision task.

This study had three specific goals: (i) To investigate whether the presentation modality may affect different semantic categories in different ways. According to the visual similarity hypothesis, we predicted slower decisions for pictures of biological stimuli compared with man made objects, due to the greater perceptual overlap of the former and, hence, the subjects' less efficient access to stored structural information. For words, the predictions are less clear. However, because words are all quite visually similar, there should not be any differential similarity effect; this effect should be found in both experiments. (ii) To determine the influence of semantic priming on both picture and word processing as a function of semantic category. According to the multiple semantic system model, we predicted that different semantic priming would be found for pictures than for words, and this effect should be greater within than across modalities. In addition, we hypothesized greater priming for biological targets than for man made stimuli presented in the picture modality. (iii) To examine the processing of pseudo-items (i.e., chimeric objects vs. pseudo-words) and non-items (i.e., non-objects vs. non-words) in a reality decision task. For Experiment 1, we predicted that pseudo-items would be the most complex stimuli to process as they are created from real stimuli (i.e., real words or objects), and thus they should have the longest RTs. For experiment 2, we hypothesized that pseudo-items would be sensitive to semantic priming since partial clues may permit subjects to access semantic information (Damasio, 1990), whereas non-items would not, since they are truly meaningless. In order to allow a straightforward interpretation, the same real objects were used in both pictures and words, and the same participants performed both experiments. The presentation order of experiments was counterbalanced between subjects.

Eighty undergraduate students at the University of Nice-Sophia Antipolis, ranging in age from 18 to 35 years old (mean: 21.81 ± 3.15) participated in this study. They were all French native speakers and reported normal or corrected-to-normal vision. Participants were tested in two separate sessions, with a week between them.

The material consisted of 144 pictures and 144 letter strings: half of them were meaningful and the other half meaningless. Meaningful picture stimuli consisted of black-and-white line drawings of 72 common objects. Half of the stimuli belonged to biological categories, and the other half to man made categories. Pictures were selected from the Snodgrass and Vanderwart (1980) standardized set, modified by Cycowicz, Friedman, Rothstein, and Snodgrass (1997), and were matched with respect to concept familiarity, visual complexity and name agreement (F < 1 in all cases). Meaningless picture stimuli comprised 36 chimeric objects and 36 non-objects selected from the Magnié et al. (2003) standardized set. Chimeric objects were made up of two halves of objects, and non-objects were constructed by mixing up the lines of real objects. Examples of meaningless pictures are presented in Figure 1. Real french words corresponded to the usual names of objects presented in the pictures. Pseudo-words were constructed according to the orthographic and phonological rules of French and non-words were strings of consonants. All linguistic stimuli were matched for length, and real words also for lexical frequency (F < 1, in all cases).
Stimuli were presented in the center of the screen with a 70-Hz refresh rate. Each trial sequence consisted of a central fixation cross displayed for 200 ms followed by the stimulus for 400 ms. Participants were instructed to decide, as quick and accurately as possible, whether the stimulus represented a picture of a real object or a French word by pressing one of two keys on a response box. The hand used for each response was counterbalanced across participants. Response times (RTs) were recorded from the target onset until the participant’s manual response. The intertrial interval lasted 1,500 ms after the response was given. Stimuli were pseudo-randomly presented; with the constraint that no more than three stimuli of the same presentation modality, stimulus type or semantic category were presented in succession. Twenty-four practice trials were given at the beginning of the experimental session. Participants were assigned to one of two lists constructed to counterbalance stimuli and conditions.

Figure 1. Sample meaningless pictures used in Experiments 1 and 2

Mean RTs of correct responses for stimuli from each condition are presented in Figure 2. In all cases, trials with RTs shorter than 200 ms or longer than 1,500 ms were defined as outliers and filtered out of the data set. The percentage of data categorized as outliers did not exceed 2% overall in either of the experiments. Analyses of variance (ANOVAs; Statistica 2000) were conducted with Modality (picture vs. word) and Stimulus (biological object vs. man made object vs. chimeric object/pseudo-word vs. non-object/non-word) as within-participant factors. In all analyses, the level of statistical significance was fixed at \( p < .05 \). The mean RT was 708 ms, and the mean error rate (ER) was 5.36%. Interaction effects were analyzed using planned comparisons.

There was a significant main effect of Modality [RTs: \( F(1,78) = 66.52; p < .001; \) ER: \( F(1,78) = 66.32; p < .001 \)], with longer RTs and more errors for pictures than for words. The main effect of Stimulus was significant [RTs: \( F(3,234) = 197.34; p < .001; \) ER: \( F(3,234) = 100.06; p < .001 \)], with longer RTs and more errors for chimeric objects/pseudo-words than for biological/man made objects or non-objects/non-words (\( p < .001 \), in all cases). In addition, there was a significant Modality x Stimulus interaction [RTs: \( F(3,234) = 22.93; p < .001; \) ER: \( F(3,234) = 2.97; p < .05 \)]. In the picture modality, there were longer RTs and more errors for...
biological than for man made objects ($p < .0001$). In the word modality, more errors were produced for biological than for man made objects ($p < .05$), but no significant difference in RTs was found. Non-objects were associated with longer RTs and more errors than non-words ($p < .001$, in both cases). Participants produced more errors for chimeric objects than for pseudo-words ($p < .001$), but no significant difference was found for RTs.

Figure 2

![Mean reaction times in milliseconds (ms) in Experiment 1.](image)

11 Experiment 1 resulted in longer RTs and more errors for pictures than for words, arguing for the superiority of words in a reality decision task. This result is inconsistent with the classical picture superiority reported in numerous studies comparing picture and word processing (Amrhein, McDaniel, & Waddill, 2002; Seifert, 1997; Stenbert, Radeborg, & Hedman, 1995). However, it is important to note that this advantage for pictures is usually reported when the experimental task requires access to the semantic system (Lloyd-Jones & Humphreys, 1997; Paivio, 1986, 1991; Seifert, 1997; Snodgrass & McCullough, 1986). In addition, it is generally considered that, just as lexical decisions focus more on lexical than on semantic properties, object decisions are based mainly on structural rather than semantic knowledge (Chertkow, Bub, & Caplan, 1992; Gerlach, Law, Gade, & Paulson, 1999; Lloyd-Jones & Humphreys, 1997; Sheridan & Humphreys, 1993). Thus, this word superiority effect supports the hypothesis that semantic access is not needed in a task such as a mixed reality decision that can be performed entirely at the lexical or structural level.

12 Another important result is related to category effects. Our data demonstrated longer RTs and more errors for biological than man made objects in the picture modality. This finding is consistent with the hypothesis that visual similarity between pictures and stored perceptual neighbors slows down processing (Lloyd-Jones & Humphreys, 1997). Because of the structural similarity of biological objects, a larger range of related representations would be activated and would compete in visual long-term memory during stimulus selection (Humphreys & Forde, 2001). Moreover, it is important to emphasize that this category effect was not found in the word modality, which supports
Among the meaningless stimuli, chimeric objects were associated with longer RTs and more errors than pictures of objects or non-objects. Similarly, pseudo-words were associated with longer RTs and more errors than words or non-words. These findings are in line with previous studies performed in a single modality (i.e., lexical decision vs. object decision; Kroll & Potter, 1984; Magnié et al., 2003; Spetch et al., 2003). This pattern may be related to the effort required to process ambiguous stimuli. No difference was found in RTs between chimeric objects and pseudo-words. Nevertheless, whereas chimeric objects and pseudo-words have been proposed to be equivalent (Magnié et al., 2003), chimeric objects were associated with a higher error rate than pseudo-words, in accordance with their different visual complexity, which is higher for pictures than for words. To further track object semantic processing, it would be interesting to consider the influence of semantic priming on both picture and word processing as a function of object category. That was the purpose of Experiment 2.

One hundred and forty-four stimuli were used to construct 576 prime-target pairs. Primes were always meaningful; half were pictures and half words. Targets were meaningful in half of the trials, with 144 targets that were semantically related to the prime (e.g., apple-pear), and 144 targets that were unrelated (e.g., apple-hammer). In the other half of the trials, targets were meaningless, including 144 pseudo-objects (72 chimeric objects and 72 pseudo-words) and 144 non-objects (72 non-objects and 72 non-words). Each prime stimulus appeared four times, twice in the picture modality and twice in the word modality. Each target stimulus appeared twice, once in each modality. The combinations of prime-target pairs as a function of presentation modality were: Picture prime-Picture target (PP); Picture prime-Word target (PW); Word prime-Picture target (WP); and Word prime-Word target (WW). All conditions appeared an equal number of times for each participant. This design was chosen to show whether the presentation of a semantically related prime facilitated performance in this task regardless of the prime-target modality (Bajo, 1988; Lloyd-Jones & Humphreys, 1997). A control study was conducted to evaluate the degree of prime-target semantic relatedness by having a new group of 80 undergraduate students rate the pairs on a scale ranging from 1 (unrelated) to 5 (very related).

The procedure was similar to that of Experiment 1 but modified by using a priming paradigm. For each trial, a central fixation cross was displayed for 200 ms. Then the prime appeared for 400 ms followed by a 1,000-ms interval during which the screen was black. The target then appeared for 400 ms. Participants were instructed to decide whether or not the target corresponded to a picture of a real object or to a French word. Thirty-two practice trials were presented at the beginning of the experimental session. Participants were assigned to one of eight lists constructed to counterbalance stimuli and conditions.

Mean RTs of correct responses for stimuli pairs from each condition are presented in Figure 3. ANOVAs were conducted with Modality pair (PP vs. PW vs. WP vs. WW) and Target (related vs. unrelated vs. chimeric object/pseudo-word vs. non-object/non-word) as within-participant factors. The mean RT was 737 ms, and the mean ER was 5.93%. Interaction effects were analyzed using planned comparisons.

There was a significant main effect of Modality [RTs: F(3,216) = 26.83; p < .001; ER: F(3,216) = 52.71; p < .001], with longer RTs and more errors for picture than for word targets (p <
.001, in both cases; i.e., RTs to PP/WP conditions were longer than for WW/PW conditions). For RT analyses, there was no significant difference between the PP/WP and PW/WW conditions, or between the PP/WW and PW/PP conditions. The main effect of Target was significant [RTs: $F(3,216) = 286.34; p < .001$; ER: $F(3,216) = 110.69; p < .001$], with longer RTs and more errors for chimeric objects/pseudo-words than for unrelated/related conditions ($p < .001$, in all cases) and non-objects/non-words ($p < .001$, in all cases). Participants produced more errors for unrelated objects than for related objects ($p < .001$, cf. Figure 4) but no significant difference was found for RT analyses.

There was a significant Modality x Target interaction [RTs: $F(9,648) = 43.02; p < .001$; ER: $F(9,648) = 4.08; p < .001$]. Pseudo-words triggered longer RTs than chimeric objects independently of the prime modality ($p < .01$ for prime picture; $p < .001$ for prime word). In contrast, there were more errors for chimeric objects than for pseudo-words but only for picture primes ($p < .05$). The non-object and non-word comparison showed significant differences, with longer RTs for non-objects than for non-words, independently of the prime modality ($p < .001$, for both picture and word primes). There were also more errors for non-objects than for non-words for picture primes ($p < .001$) but no significant difference was found for word primes.

A further analysis was conducted to investigate the biological/man made dichotomy. There was a significant effect of Modality [RTs: $F(3,216) = 47.16; p < .001$; ER: $F(3,216) = 67.07; p < .001$], with longer RTs and more errors for picture target (PP/WP) conditions than for word target (PW/WW) conditions ($p < .001$, in all cases). The main effect of Category was significant [RTs: $F(1,72) = 72.71; p < .001$; ER: $F(1,72) = 66.58; p < .001$], with longer RTs and more errors for biological than man made objects. There was a significant Modality x Target interaction [TRs: $F(3,216) = 12.36; p < .001$; ER: $F(3,216) = 12.99; p < .001$]. Planned comparisons showed longer RTs for biological than for man made targets presented in the picture modality than in the word modality, independently of the prime modality ($p < .001$ in all cases). There were more errors for biological targets than for man made targets for PP, PW and WP conditions ($p < .001$, in all cases).
As in Experiment 1, longer RTs and more errors were produced for picture than for word targets, whatever the prime modality, confirming the superiority of words in a mixed reality decision task with a semantic priming paradigm. However, no difference was found between within- and across-modality conditions. We originally demonstrated that the critical modality in a reality decision is the target modality, regardless of the prime modality. Thus, our data revealed differences in how pictures and words representing the same semantic concept are processed. These differences between the two modalities seem to be consistent with the multiple semantic system model. Moreover, the question that...
remains is whether these differences are truly modality-specific or are instead driven by another factor. Indeed, one potentially important, and frequently overlooked, factor is related to the physical characteristics of pictures and words, that is, the difference between large pictures and small words. Whereas some studies have shown similar processing for both pictures and words when stimulus size, spatial frequency and featural line width are controlled (Theios & Amrhein, 1989a, 1989b), other studies have presented contrary results (Seifert, 1997). However, these studies used different tasks and stimuli and it is difficult to compare them. It would be interesting to replicate our experiment (i.e., using the same stimuli and experimental task) to directly address this hypothesis.

Concerning priming effects, our study showed no differences in RTs between related and unrelated conditions. These results are inconsistent with previous studies using either lexical or object decision tasks with semantic priming, in which targets were usually processed faster when they were preceded by a semantically related prime than when the prime was neutral or unrelated (Kroll & Potter, 1984; Neely, 1991). Our data might be related to differences in the instructions given to participants. In previous studies, subjects were usually informed that two kinds of stimulus pairs would be presented: semantically related or unrelated. In contrast, in our experiment, no information was given to participants concerning the possible relationships between prime and target stimuli. Indeed, we chose not to encourage participants to use different strategies to process stimuli pairs. Bajo (1988) demonstrated that semantic priming depended on both the task and the participant’s strategies. When the task required semantic processing of pictures and words, priming occurred across all modality combinations. However, when the task did not require participants to evaluate the relationship between primes and targets, a semantic priming effect was not found. More recently, Friederici, Opitz, and von Cramon (2000) showed, in an fMRI study, that task demands modify the activation pattern during word recognition. Thus, the null effect of semantic priming observed on RTs in our experiment could be explained by the fact that the task does not encourage semantic relatedness judgments. However, it should be noted that a higher error rate was found for unrelated pairs than for related pairs. Since the reality decision does not explicitly require deep processing of stimulus meaning, electrophysiological studies are needed to further track semantic priming effects by comparing evoked responses to related and unrelated pairs of objects as a function of the instructions given to participants. These follow-up studies are currently under way.

As for category effects, our results showed longer RTs and more errors for biological than man made targets in both modalities. In contrast to Experiment 1, the category effects were not limited to the picture modality, suggesting that this result could be related to the prime presentation and thus arguing in favor of the hypothesis that there is implicit access to semantic representations. Moreover, the data from Experiment 2 favor the functional interpretation of the biological/man made dichotomy. Indeed, the superiority effect found for pictures of man made objects presented in isolation in Experiment 1, without any difference for words, supports the hypothesis based on the structural overlap of biological objects. In contrast, the superiority effect for man made objects found in both modalities in the second experiment argues that there is an advantage at the semantic level when stimuli are presented in a semantic priming paradigm.

Concerning the meaningless stimuli, Experiments 1 and 2 had different patterns for pseudo-items whereas the same pattern is reported for non-items in both experiments. In contrast to Experiment 1, chimeric objects preceded by a prime were associated with
shorter RTs than pseudo-words. This different pattern suggests that the semantic priming provides an advantage for picture targets since chimeric objects may activate semantic representations via the halves of the two real objects of which they are made, whereas pseudo-words did not appear to do so. Our data provide additional evidence of the usefulness of these kinds of meaningless pictures since they are processed in different ways. Taken together, our data are consistent with the account that object processing is modulated by task demands (Lloyd-Jones & Humphreys, 1997; Snodgrass, 1984). Thus, the results of our study suggest that the reality decision task might be performed only at a structural level but can also involve implicit semantic processing, depending on the experimental task.

In conclusion, our data revealed that the word superiority in reality decisions occurs independently of semantic priming, and that the critical modality is the modality of the target, whatever the prime modality. These findings are in line with both structural and functional interpretations of the biological/man made dichotomy proposed in the literature. Moreover, our study demonstrates that these interpretations might be considered complementary regarding the superiority of man made objects and they might be involved at different levels in object processing. In addition, the data showed the importance of distinguishing between two kinds of meaningless stimuli. Consequently, our results showed that the mixed decision task only requires lexical/structural processing when stimuli are presented in isolation and might involve implicit semantic access when participants perform the task in a semantic paradigm but are unaware of the relatedness conditions.

25This research forms part of a doctoral dissertation by the first author, and was supported by grants from the French Ministry of National Education, Research and Technology (MENRT). We would like to thank Sylvain Denis and Régis Desneulin for computer programming and Sylvane Faure for her helpful comments and suggestions.

BIBLIOGRAPHY


NOTES

* Corresponding author. Tel.: +514-340-3540; fax: +514-340-3548.
E-mail: kahlaoui@unice.fr (K. Kahlaoui).
ABSTRACTS

This study investigated the influence of semantic priming on object processing as a function of both presentation modality and object category in a reality decision task. Participants performed a mixed decision (i.e., object and lexical decisions) on picture and word stimuli presented in isolation (Experiment 1) and in a semantic priming paradigm (Experiment 2). The results showed longer RTs and more errors for picture targets than for word targets, in both experiments. Category effects were also demonstrated: biological objects were associated with longer RTs and more errors than man made objects, only for pictures in Experiment 1 but in both modalities in Experiment 2. Thus, our data reveal a word superiority effect in reality decisions, independently of semantic priming, and provide additional evidence favoring the biological/man made dichotomy. Finally, our data show that the mixed decision task only requires lexical/structural processing when stimuli are presented in isolation and may involve implicit semantic access when participants perform the task as part of a semantic paradigm.

INDEX

Keywords: pictures, words, reality decision, semantic priming, category effect

AUTHORS

KARIMA KAHLAOUI*
Centre de Recherche, Institut Universitaire de Gériatrie de Montréal, Canada . E-mail: kahlaoui@unice.fr

THIERRY BACCINO
Laboratoire de Psychologie Expérimentale et Quantitative, Université de Nice-Sophia Antipolis, France

YVES JOANETTE
Centre de Recherche, Institut Universitaire de Gériatrie de Montréal, Canada & Ecole d'Orthophonie et d'Audiologie de Montréal, Université de Montréal, Canada

MARIE-NOÉLE MAGNIÉ
Laboratoire de Physiologie, Université de Nice-Sophia Antipolis, France