# A Cognitive Study of the Semantic Memory activated by Pictures and by Words.

**Martine Cornuéjols (mc@limsi.fr)**
University of Paris XI - Orsay / CNRS-LIMSI
Human Cognition group. Bat 508, B. P. 133 -
91403 Orsay Cedex - France

**Jean-Pierre Rossi (rossi@limsi.fr)**
University of Paris XI - Orsay / CNRS-LIMSI
Human Cognition group. Bat 508, B. P. 133 -
91403 Orsay Cedex - France

**Abstract**

Here, the fundamental question is the existence of a mental lexicon for pictures, as it exists one for words. And so, what is the structure of the semantic memory and its relationships with access modalities ?
In the first experiment we determine if a picture or a word can activate the same semantic structure by using the semantic priming paradigm. The classical effect of facilitation was found for word priming, but not for picture as prime. An hypothesis is that the semantic network activated by pictures is different from that activated by words. In the second experiment, we demonstrate that the semantic association structure for pictures and for words is different. So on the basis of these data, we replayed the experiment I. In this case, pictures as words could induce semantic priming of a word. And so we discuss the issue of multiple vs unitary semantic storage.

## Introduction

Analysis of semantic representation is central to the study of language comprehension. Semantic memory is initially described as a network (Quillian, 1967) constituted by nodes interconnected by links. Nodes are generally words and the semantic network is described as an organization of verbal entities. So, semantic networks are essentially verbal networks. This brings up the question of the semantic processing of pictures. Both kinds of models were designed to account for picture processing : in one the semantic memory is a common structure for words and pictures; in the other, for each category of input (words, pictures...) there is a specific corresponding semantic network. The aim of this article is to analyze the assumption of a common semantic representation.

The thesis of a common semantic network is supported by studies of picture/word interference tasks (Stroop effect) which demonstrated the influence of semantic activation in the processes of color naming (Potter & Faulconer, 1975; Snodgrass, 1980; Glaser & Glaser, 1989; La Heij, Happel & Mulder, 1990; Smith & Magee, 1980). In fact, interpretation of the interference effect is difficult because interference can result from the propagation of activation either inside a common network or between two networks. In the analysis of interference effects, the hypothesis of two semantic networks must be distinguished from the hypothesis of independence of the networks. To sum up, relationships between the two networks can generate interferences. The discussion on the number of semantic files requires comparison of word identification models with the picture naming model. Indeed, only a precise description of these two models can allow an explanation to be made about the origin of interference effects. This comparison can be facilitated by the fact that some picture naming models reproduce reading models (Levelt and al., 1991).

In general, the study of semantic networks is carried out using the priming technique. The experimental paradigm has been labeled as the single-word semantic paradigm (Neely, 1991). A trial consists of two events: first, a semantic context is provided by the presentation of a single stimulus (more often a word) called the prime, to which no overt response is required. Second, this prime is followed by a presentation of a single string of letters, called the target. In nearly all experiments, subjects have been required either to make a word/nonword decision about the target (the lexical decision task). In the lexical decision task, decision latency decreases when the prime and the target are semantically associated. This facilitation was examined in some detail by Neely (1976). Semantic priming effects were initially interpreted in terms of a spread of activation in a semantic network to nearby nodes for related words (Collins & Loftus, 1975). In particular

conditions (very short prime-target stimulus onset asynchronies (SOAs)), this spread of activation might be considered as an automatic phenomenon. In this condition the priming paradigm can be used to study the organisation of the semantic network.

## Experiments

The goal of the research presented here is to test the validity of the hypothesis of a common semantic structure, independent of the input. In the first experiment, we will determine if a picture, a word presented auditorily or the same word presented visually can activate the same semantic structure. Results from the first research led us to construct tables of association for pictures (experiment 2). The third experiment will allow us to specify if there exists a specific semantic system for pictures.

### 1 - First experiment

The aim of the first experiment is to examine the role of the nature of input in the activation of the associative network. For the lexical decision task, most experiments have used words as stimuli. Words could be associated by a graphemical, a phonological (using homophones), a syntactical, or a semantical relation. The purpose of our experiment is to determine if the presentation of a picture or an auditorily presented word (prime) facilitates the lexical decision on the written word (target).

***Method*** : *Subjects.* A group of eighteen people, made up of graduate students and researchers in computer science and in cognitive science at the University of Paris-Sud, took part in this experiment. They were native French speakers who had either normal or corrected to normal vision and normal hearing. Participation was voluntary.

***Materials and design.*** The stimuli were 240 pairs of items where the prime was either a written word (i.e. a visually presented prime word), or a heard word (i. e., an auditorily presented prime word), or a picture ( colored drawings). The target was always a written word. Each prime evoked an object, an animal or a vegetable, etc. Words were controlled concerning their length (4-9 letters), and the word frequency was medium. The pictures were selected from the Snodgrass and Vanderwart (1980) norms and colored to resemble the natural colors of the objects considered. These were drawings of familiar objects and animals or vegetables, with a moderate amount of detail, all easy to identify. The auditory material was composed by heard words, from the list of written words, recorded without intentional specific intonation. The auditory prime words were recorded via a Sound Blaster card on a PC.

The experiment was divided into three parts, each corresponding to one of the three categories of prime (auditorily presented word, visually presented word. and picture). Each part was made up of 40 associated couples (prime and target semantically associated), 40 couples of controls (with neutral prime), 40 couples of distractors (with non-word as target), and 40 couples of distractor's controls (with neutral prime and non-word as target). The associated couples were selected on the basis of the French table of word association (Lieury, Iff, & Duris, 1976). The target is the first word spontaneously associated to the prime by more than 50 % of the subjects.

For the picture condition, the neutral prime was a shapeless picture, for the written-word condition a number X equal to the number of letters in the associated word target, and for the auditory condition a continuous sound. There was an equivalent proportion of associated and control conditions. Instructions and stimuli were presented using a PC, either on the center of the screen or via an acoustic headset. The computer was equipped to permit millisecond accuracy.

***Procedure.*** Subjects were tested individually. All subjects went through the three conditions of the experiment and therefore also went through the three modalities of presentation. Effects of order and ranks were counterbalanced. The three conditions of priming were contrasted with the control condition (with neutral prime). Subjects made a decision task and respond ed "yes" or "no" by means of pushbuttons. Right-handed subjects responded " yes " with the right index finger, left-handed subjects with the left index finger.

A trial consisted of the presentation of the prime, an ISI and the presentation of the target. The picture prime appeared for 40 ms, the written-word prime for 60 ms. The identification threshold for pictures and written words was under 10 ms, weaker for the pictures than for the written words. The target remained in appearance until the subject had made a decision, but after 1500 ms. the message " it is too late " appeared . The SOA was 90 ms (as is supposed to be a condition for automatic activation).

Each session began by practice trials. Trial order was randomized. Reaction times were measured.

**Results and discussion:**

Mean values for lexical decision latencies of positive answers (exact decision for target words) are presented in Table 1. The number of errors is negligible, i.e. always inferior to 10%.

**Table1** : Mean values for lexical decision latencies (msec.)

| Prime | Picture | |
|---|---|---|
| Cond. | Associated | Control |
| m | 591.12 | 597.94 |
| $\sigma$ | 27.18 | 22.63 |

| Prime | Written-Word | |
|---|---|---|
| Cond. | Associated | Control |
| m | 571,66 | 599,86 |
| $\sigma$ | 15,81 | 29,76 |

| Prime | Heard-Word | |
|---|---|---|
| Cond. | Associated | Control |
| m | 585,87 | 605,36 |
| $\sigma$ | 17,22 | 23,35 |

For each situation (picture, heard word and written word) we have compared the associated condition with the control condition. Significant effects were found in the situation where the prime is a written word ($t(17) = 4.810$, $p = .0002$) and in the situation where the prime is a heard word ($t(17) = 3.451$, $p = .0031$); in contrast, when the prime is a picture there was no difference ($t(17) = 1.092$) between results for associated and control conditions. An ANOVA indicated that there were no differences ($F<1$) between control conditions: for prime pictures mean latency was 597,94, for prime heard-word, 605,36 and for prime written-word, 599,86.

When the prime is a word, the classic effect (Neely, 1991) of facilitation was found: the lexical decision is shorter when the prime and the target are semantically linked. This result usually found with written words was also found when the prime was a word presented acoustically. Thus, visual and acoustic words activate the same semantic system. In contrast, the effect of facilitation was not found when the prime was a picture. An explanation of our data is that there exists two semantic networks, one for pictures and one for verbal inputs.

## 2. Experiment 2 : Associations table

In the first experiment, the associated couples were selected from the "table of verbal associations" (Lieury and al., 1976). If the semantic network is different for words than for pictures, the table of associations will also be different. So, to test the hypothesis of a dual semantic network it is necessary to define a " table of picture associations ". To compare the picture prime condition with the verbal prime condition a new " word association table " has been constructed with the same group of subjects.

***Method*** : *Subjects*

Thirty graduate students at the University of Paris-Sud volunteered for this experiment. They were all native French speakers. None had previously seen the set of pictures.

*Materials and design* : The stimuli were 284 pictures or corresponding words. They are the same as those used in the first experiment. The colored pictures were selected from the Snodgrass and Vanderwart norms (1980). These were drawings of familiar objects and animals or vegetables and were easily labeled. 284 words corresponding to pictures were used to construct the word association table. The words were 4 to 9 letters in length and their frequency was medium.

***Design and procedure*** : Stimuli were presented on a screen for 10 s. They were numbered. At the beginning of the experiment subjects received an 11-page notebook containing the numbers of the stimuli (from 1 to 284). Subjects were instructed to write the first response associated with the stimulus in front of its number. They had the duration of the presentation (10 s) to respond to each stimulus. For the word association table they wrote

the first word evoked by the word stimulus, and for the picture association table they wrote the word that named the first image evoked by picture stimulus. This last task was very difficult because, as we have tested in a previous experiment, the first answer evoked by the subjects is the name of the picture and then the second word evoked is the word associated with the first word naming, which brings us back to the verbal associations table. Thus, in this experiment we had to encourage them to perform the task of generating an image associated with the picture presented. To avoid name response, subjects were trained to activate a mental image and to name it (one word) or to describe it by few words.

Subjects who took part in the construction of the two tables (word association and picture association) were tested in two sessions : one session with picture stimuli and one session with word stimuli. Half of the subjects began with picture stimuli and were tested two months later with word stimuli and half of the subjects began with word stimuli and were tested two months later with picture stimuli. The experiment was run collectively by group of 15 to 20 subjects.

**Results**

The first analysis involves calculating the dispersal of the answers for each stimulus. For word stimuli, 59.27 % of responses are identical. Thus, more than half the subjects gave the same association response when the stimulus was a word. For picture stimuli, the percentage is slightly weaker at 60.26 %. The difference between these two percentages is not significant. Despite the difficulty of the task, when the stimulus is a picture, associative responses are no more dispersed than when it is a word. Thus, the picture association table is as valid as the word association table.

Comparison between the two tables was completed by the study of the rate of recovery, i.e. the ratio of cases where the majority's response is the same in the two association tables (word or image). This rate of recovery is 36.26 %. Thus, the rate of divergence is 63.74 %.. This result indicates that associations activated by pictures are different from those activated by words.

The most important result of this analysis is the fact that the presentation of the picture is most likely to evoke the situation in which the stimulus is habitually encountered. It is the case for 43.15% of the responses. That leads to the thought that the associative structures corresponding to pictures and to words are different. This hypothesis is evaluated in the third experiment.

## 3. Experiment 3

The hypothesis of the dual semantic network is tested in this experiment. The experiment 1 is replayed without the acoustic condition. In the situation where the prime is a picture, associated couples are chosen by using the associative picture table.

***Method*** **:** *Subjects.* There were 16 volunteer subjects, all graduate students in computer science or in cognitive science at the University of Paris-Sud. They were native French speakers who had either normal or corrected to normal vision and normal hearing.

***Materials and procedure.*** The method was identical to that used in Exp. I with respect to instructions and presentation rates. In the case where the prime was a picture, the target of the associated couple was the first associated word given by a minimum of 50 % of subjects in the test of association.

***Results and discussion***

Mean values for lexical decision latencies of positive answers (exact decision for target words) are presented in Table 1. As in experiment 1, the number of errors is always inferior to 10%.

**Table 2** : Mean values for lexical decision latencies

| Prime | Picture | | Written-word | |
|---|---|---|---|---|
| Cond. | Associated | Control | Associated | Control |
| m | 553.71 | 575.44 | 565.32 | 581.96 |
| $\sigma$ | 21.70 | 24.84 | 24.86 | 19.43 |

Comparison between associated and control conditions indicated significant effect for written-word ($t(15) = -2.298$, $p = .0354$) and for picture ($t(15) = -3.132$, $p = .0064$). Difference between control conditions was not significant. So, the presentation of a picture activates word naming associated with an image. Semantic activation with picture prime is similar to semantic activation with word prime when the target is a word.

## General discussion

The hypothesis of a single semantic network would be strengthened if phonological encoding preceded semantic activation. Phonological encoding would generate a single output for both picture and word. The hypothesis of phonological encoding before semantic activation is not supported by studies of picture identification. Levelt and al. (1991), Dell and O'Seaghdha (1991) and Vikovitch, Humphreys and Lloyd-Jones (1993) have proved that semantic processing precedes phonological encoding of the picture. This conclusion is confirmed by Experiment 2 and 3 results: were this not the case, and the naming of picture preceded its semantic process, words and pictures would have the same associated item. In fact, results from experiments 1 and 3 as well as the study of the table of figure association indicate that the presentation of a picture does not evoke the same associated item as the presentation of a word. In the majority of cases, the presentation of a picture activates the situation where the object is generally encountered (bread - bakery) while the reading of a word evokes another object (bread - butter). These results do not involve two semantic networks; rather they indicate that the part of the semantic network that is activated varies with categories of input. This interpretation of these results brings up the question of the format of the semantic network since it is necessary to imagine a semantic network that could be activated by pictures, or by visual or auditory presentation of words.

### References


Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13-42.

Holcomb, P. J., & Anderson, J. E. (1993) : Cross-modal semantic priming: A time course analysis using event-related brain potentials. *Language and Cognitive Processes, 8 (4),* 379-411.

La Heij, W., Happel, B., & Mulder, M. (1990). Components of Stroop-like interference in word reading. *Acta Psychologica, 73,* 115-129.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., Havinga, J.(1991). Normal and deviant lexical processing : Reply to Dell and O'Seaghdha. *Psychological Review, 98 (4)*, 615-618.

Lieury, A., Iff, M., Duris, P. (1976). Normes d'associations verbales. Laboratoire de psychologie expérimentale et comparée associé au CNRS.

Meyer, D., & Schvaneveldt , R. W. (1971). Facilitation in recognizing pairs of words : Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Neely, J. H. (1991). Semantic priming effects in visual words recognition : A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds), *Basic processes in reading : visual word recognition* (pp. 264-336). Hillsdale, NJ : Erlbaum.

Snodgrass, J. G. , & Vanderwart, M. (1980). A standadized set of 260 pictures : Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology : Human Learning and Memory, 6*, 174-215.