

Visual effects in picture and word categorization

LORELLA LOTTO, REMO JOB, and RINO RUMIATI
University of Padua, Padua, Italy

Two experiments are reported in which participants categorized stimuli as belonging or not belonging to the category of fruits. Blocks of pictures and words were used, with items referring to exemplars having either high or low intercategory visual similarity and/or semantic relatedness. For both pictures and words, response time was longer in the semantically related conditions than in the unrelated condition. Furthermore, there was a strong effect of visual similarity for pictures but not for words when semantic relatedness was held constant: Participants took longer to classify pictures of fruits when these were mixed with visually similar vegetables than when they were mixed with visually dissimilar vegetables. Reducing the stimulus visibility by adding a dot pattern had an additive effect for words but an interactive effect for pictures. The results are explained in terms of a unique locus for category decisions about pictures and words.

Both picture and word categorization can be thought of as a two-step process requiring first a visual discrimination and then a semantic decision about category membership. However, the visual information initially made available by pictures and words is quite different, and this may bring about differences in processing (Potter & Faulconer, 1975). For words, visual-orthographic information is instrumental for gaining access to the mental lexicon, and semantic information is available only afterward. For pictures, visual information can be directly used to access semantic information, and, as a result of the interaction between categorization demands and perceptually available information (Goldstone, 1994; Oliva & Schyns, 1997), may even be diagnostic of category membership. Thus, given the word *crow*, we may infer semantic properties—for example, that it is a bird—only after its meaning has been retrieved, but given the corresponding picture, such information may already be available from visual cues.

Therefore, semantic factors should affect both pictures and words analogously, but visual factors should have different effects on categorizing pictures and words.

Semantic effects in categorization have been studied extensively, with both word and picture stimuli. One such effect is the so-called semantic relatedness effect. *Semantic relatedness*, or semantic distance, as it is sometimes called, is defined as the number of properties two concepts have in common (Collins & Loftus, 1975) and has been operationally defined as how semantically close or far two elements are rated (Rips, Shoben, & Smith, 1973)

or the number of levels or links separating one element from another in a semantic net (Rosch, 1975). Semantic relatedness affects both positive trials (e.g., deciding that *a canary is a bird* is faster than deciding that *a canary is an animal*) and negative trials (deciding that *a canary is a mammal* takes longer than deciding that *a canary is a car*) (McCloskey & Glucksberg, 1979; Rips et al., 1973). Guenther and Klatzky (1977) have shown that when people are required to decide whether a pair of items belongs to a given category, both picture and word categorization is affected by semantic relatedness, even though the effect tends to be more pronounced for pictures than for words.

Another semantic effect shown to play a role in categorization is degree of category membership (Rosch, 1975). This notion refers to the different degrees of goodness-of-exemplar shown by category members. Some are considered central members since they are more prototypical in that they share many typical features with the other elements of the category, whereas others members are peripheral in that they tend to have idiosyncratic features. It has been shown that central members are more easily processed than peripheral members in a number of cognitive tasks (e.g., deciding that *a canary is a bird* takes less time than deciding that *a penguin is a bird*). With respect to the categorization of pictures and words, Rosch (1975) performed an experiment in which people had to decide whether pairs of items were drawn from the same category or not. Each pair was preceded by a prime, either a category name or the word *blank*. The results showed that priming with the category name facilitated responses to central members of the category but slowed down response times (RTs) to peripheral members. This was true for both picture and word stimuli.

The role of concepts' visual similarity in categorization has been less studied, but some evidence is available on this issue. First, context-sensitive models (see, e.g., Markman & Gentner, 1996; Medin & Shaffer, 1978; Nosofsky, 1986) suggest that classification decisions are affected by

We thank Franca Agnoli, James Antes, Alistair Ewins, Geoffrey Loftus, Joan Gay Snodgrass, and two anonymous reviewers for their helpful comments on an earlier version of the paper, and Eraldo Nicotra for his advice on the statistical analyses. This research was supported by grants from MURST (40%). Correspondence should be addressed to R. Job, Dipartimento di Psicologia dello Sviluppo e della Socializzazione, via Venezia 8, 35131 Padova, Italy (e-mail: remojob@psico.unipd.it).

the degree of similarity among the items to be classified.¹ Similarity, according to this view, rather than being an invariant relationship, changes as a function of the experimental context (e.g., the actual set of exemplars the participants are presented with). Indeed, Nosofsky (1989) has shown that systematic changes in the structure of the multidimensional space in which concepts are represented occur as a function of the dimensions selectively attended to. With reference to simple perceptual classifications, Nosofsky and Palmeri (1997) made the explicit prediction that rapid class decisions should be made to exemplars from one of two alternative categories, whereas slower decisions should be made to items similar to exemplars from both categories.

More directly relevant to our study are the results reported by Snodgrass and McCullough (1986). These authors have shown that the degree of visual similarity among concepts affects categorization time of pictures but not of words. In their study, participants were presented with lists of pictures and lists of words to be categorized. Within each list, elements were drawn from two categories that were either visually similar (i.e., fruits and vegetables) or visually dissimilar (i.e., fruits and animals). Their results showed that pictures were categorized faster in the visually dissimilar than in the visually similar condition but that word categorization times did not differ in the two conditions. Using the same task, Job, Rumiati, and Lotto (1992) have shown that word stimuli are also affected, albeit less strongly, by the visual similarity of their referents.

How can these effects of visual similarity in categorization be accounted for? The explanation offered by Snodgrass and McCullough (1986) rests on the assumption that, under some conditions, word and picture categorization is mediated by different routes. In particular, words are always categorized through access to the semantic system, which provides information about category membership. Pictures are categorized by the same procedure (i.e., through the semantic system) only when the elements of the contrasting categories tend to be similar to each other. Instead, when the pictures depict elements that are quite dissimilar from each other, a nonsemantic route is available based on the evaluation of the degree of visual similarity between a picture and a comparison perceptual prototype.

An alternative interpretation can be given on the basis of Caramazza, Hillis, Rapp, and Romani's (1990; see also Rapp, Hillis, & Caramazza, 1993) proposal. Let us consider the asymmetry in picture and word categorization. According to Caramazza et al., the locus for the category membership decision for pictures and words is the same—that is, the semantic system—and the semantic constraints affecting the categorization of pictures and words are analogous. But pictures and words differ with respect to the relationship between stimulus features and meaning. For words, the relationship between orthography and mean-

ing is arbitrary, whereas for pictures the features available from perception are closely linked to the functional characteristics of the concepts they represent, so that these features can activate their corresponding semantic predicates. This multiple activation would facilitate the identification of the concept as a whole, and would be responsible for faster responses to pictures than to words—the pattern generally obtained.

In the present study, the categorization of pictures and words was studied by manipulating the degree of visual similarity and the degree of semantic relatedness. There are a number of reasons for taking this approach. First, in the studies so far conducted, there is a potential confounding between visual similarity and semantic relatedness. For example, in Snodgrass and McCullough's (1986) study, the categories of fruits/vegetables versus fruits/animals were used as visually similar and visually dissimilar categories, respectively, but, as pointed out by Job et al. (1992), according to any semantic theory, fruits and vegetables are more closely semantically related than are fruits and animals. Therefore, the effects ascribed to visual similarity might instead be due, at least in part, to semantic relatedness. Second, it is important to detect individually the role of these two variables. So, even though it is impossible to manipulate factorially visual similarity and semantic relatedness (see Carr, McCauley, Sperber, & Parmelee, 1982; Rosch, 1978), it may be productive to manipulate visual similarity while holding semantic relatedness constant.

The rationale underlying the present study is as follows. Let us consider the target category fruits and the two contrasting categories, vegetables and weapons. If we select randomly an element from the target category and an element from one of the contrasting categories, the probability that the two elements are visually similar to each other is higher for the fruits/vegetables condition than for the fruits/weapons condition. This follows from the fact that the elements from the former two categories tend to share parts (Humphreys, Riddoch, & Quinlan, 1988) and/or global shape (Lotto, 1988). However, it is possible to select different subsets of fruits and vegetables so as to vary systematically the similarity relationship among elements from the two categories. For example, one subset may be composed of roundish fruits and vegetables (e.g., apple and potato), and another subset may be composed of roundish fruits and elongated vegetables (e.g., apple and celery). In this case, the semantic relationship between the elements remains constant, but their visual similarity varies.

For the control condition, the fruits/weapons categories may be used because they differ both semantically and visually. In Experiment 1, the fruits selected were visually similar to one another and were contrasted once with a set of vegetables visually similar to them and once with a set of vegetables visually different from them. In Experiment 2, the same items were presented with a super-

imposed dot pattern, with the aim of investigating the relationship between stimulus form and visual similarity in the early stage of processing.

The predictions for Experiment 1 were that if participants use a visual strategy to discriminate between the two contrasting categories, they should be facilitated in the condition of low between-category similarity relative to the condition of high between-category similarity. However, both conditions should be more difficult than the unrelated condition, in which stimuli are visually as well as semantically dissimilar. In addition, if the visual-structural information available from pictures is used directly in the categorization process, words should not be affected by the manipulation of their referents' visual similarity, as pictures would be.

The degree of category membership was also manipulated in order to contrast some predictions derived from Caramazza et al.'s (1990) and Snodgrass and McCullough's (1986) models. If a common locus for both word and picture categorization is postulated, as the Caramazza et al. model does, then the degree of category membership should be expected to affect picture and word categorization analogously. However, if pictures and words are categorized through different procedures in the visually dissimilar condition, as proposed by Snodgrass and McCullough, we should expect differential effects of category membership for pictures and words in the fruits/weapons condition.

PRELIMINARY STUDY

Collecting Visual Similarity Ratings

In order to obtain a measure of concepts' visual similarity, we asked participants to rate pairs of concepts on a 5-point scale. Since the critical experimental comparisons are those involving fruits and vegetables, our aim was to select vegetables with either a high or a low degree of similarity to fruits. Therefore, the pairs to be rated consisted of either two fruits or a fruit and a vegetable. Ratings were collected for both picture pairs (see examples in Figure 1) and the corresponding word pairs to obtain a measure as free as possible from the influence of the format of the stimuli (i.e., pictorial vs. verbal).

Method

Materials and Procedure. A set of 20 fruits and 16 vegetables was selected.² Using elements from the Snodgrass and Vanderwart (1980) set of elements drawn in a similar fashion, pairs of pictures were constructed. Each pair was obtained by associating each exemplar of fruits with (1) any of the remaining exemplars of the same category and (2) any of the exemplars of the category of vegetables, for a total of 510 pairs. No vegetable/vegetable pairs were created since such a condition was not part of the experimental design. From the 510 pairs, four booklets were constructed, two composed of 127 picture pairs and two composed of 128 picture pairs; each participant was presented with only one of the booklets. On each page of the booklet, a 5-point scale appeared below a picture pair, and participants were asked to rate how visually similar the elements of the pair were (1 = *low visual similarity*, 5 = *high visual similarity*).

A similar procedure was used to obtain ratings using word pairs. The same 510 pairs of concepts were presented to a different group of participants, but in this case the words, rather than the pictures, corresponding to the two concepts appeared above the 5-point scale. The instructions provided to the participants stressed that they were to judge the visual similarity of the words' referents, not the similarity of the words themselves.

Participants. A total of 160 Italian-speaking participants took part in the study, half in the picture pairs condition and half in the word pairs condition. In both conditions, there were 20 participants for each of the four booklets.

Results

On the basis of the mean rating score obtained for each pair, three groups of items were constructed (see the Appendix): (1) a group of eight fruits with high rated within-category similarity (fruits), (2) a group of eight vegetables with high rated visual similarity to the elements of the fruits group (similar vegetables), and (3) a group of eight vegetables with low rated similarity to the elements of the fruits group (dissimilar vegetables). The mean ratings of visual similarity are reported in Table 1. Participants judged the fruits group to be more visually similar to the similar vegetables group than to the dissimilar vegetables group [picture pairs, $t(126) = 18.88$, $p < .001$; word pairs, $t(126) = 7.91$, $p < .001$].

The groups of items so identified constituted, along with eight elements drawn from the category of weapons, the experimental materials used in the following experiments.³ In each group, half of the elements were central members and half were peripheral, according to the

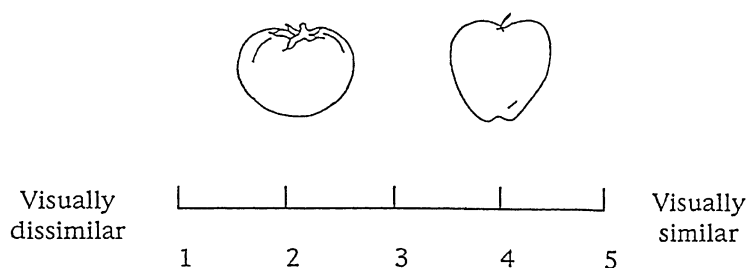


Figure 1. An example of a picture stimulus pair used in the preliminary study.

Table 1
Mean Rated Intra- and Intercategory Visual Similarity

Sets	Picture	Word
Fruits (intracategory similarity)	3.50	2.50
Vegetables		
Similar (intercategory similarity)	2.91	1.97
Dissimilar (intercategory similarity)	1.11	1.37

norms reported by Job (1981). In addition, the names of the items were matched as closely as possible for length and frequency (Bortolini, Tagliavini, & Zampolli, 1974).

EXPERIMENT 1

Method

Participants. Twenty-four Italian-speaking undergraduates at the University of Padua were employed.

Materials and Apparatus. Three conditions were created by manipulating the visual similarity and the semantic relatedness of the categories considered. The first condition involved items that were semantically unrelated and visually dissimilar (unrelated condition). The second condition involved items that were semantically related but were visually dissimilar (low-similarity condition). The third condition involved items that were both semantically related and visually similar (high-similarity condition).

In the unrelated condition, the list of elements to be categorized consisted of the eight items of the fruits group and the eight items of the weapons group. In the low-similarity condition, the list of elements to be categorized was formed by the eight items of the fruits and the eight items of the dissimilar vegetables groups. Finally, in the high-similarity condition, the list of elements to be categorized was formed by the eight items of the fruits group and the eight items of the similar vegetables group.

Participants were tested in a $2 \times 3 \times 2$ within-subjects design. The independent variables were form of stimuli (pictures vs. words), type of relationship between categories (unrelated vs. low similarity vs. high similarity), and degree of typicality (central vs. peripheral).

Procedure. We required participants to categorize each item as belonging to one of the two categories appearing in each condition. Before each condition, we told participants what two categories they would receive and the form of presentation (picture or word). Half of the participants were instructed to press the right response key if the stimulus belonged to the target category (i.e., fruits) and the left response key if the stimulus belonged to the nontarget category. The response key order was reversed for the other half of the participants. Participants were encouraged to respond as quickly as possible. Response type (correct/incorrect) and response latencies to the nearest millisecond were recorded for each stimulus.

Pictures and words were presented in different blocks. The order of presentation of the two blocks and, within block, of the three conditions, was counterbalanced across participants. The order of stimulus presentation was determined randomly for each condition, with the constraint that no more than three stimuli from the same category could appear sequentially. Once the random order of stimuli for each condition was established, it remained the same for every participant. In each condition, the items were presented twice.

The stimuli appeared on a screen controlled by a Macintosh computer. Pictures subtended a visual angle of about 3.5° horizontally $\times$ 3° vertically, and the pictures' names subtended a visual angle of 0.6° per letter horizontally $\times$ 0.7° vertically. The participant was seated at a table in a sound-attenuated room separate from the experimenter. The stimulus duration, interstimulus intervals, intertrial

interval, and visual signal duration preceding the appearance of each item were controlled automatically. Before the presentation of the stimulus, a fixation point appeared on the center of the screen for 500 msec followed by a 500-msec lighted blank field. Each stimulus disappeared as soon as the response was given or after 1.5 sec had elapsed. It was followed by a 1-sec lighted blank field that served as the intertrial interval. The onset of the stimulus started a timer. Pressure on either of the two response keys stopped the timer, and reaction time was automatically recorded.

Practice items consisted of 10 pictures and the corresponding words. Half of the stimuli were selected from the category of birds, and the other half were selected from the category of mammals. The 10 practice pictures and the 10 corresponding practice words were placed at the beginning of the picture block and the word block, respectively.

To familiarize participants with the stimuli, and to ensure that their categorization of items agreed with that of the experimenters, at the beginning of the session participants in the picture-first condition were presented the 32-picture experimental list and participants in the word-first condition were presented the 32-word experimental list. Participants were required to categorize each item by saying the category label aloud. At the beginning of the second block, participants were presented with the items in the complementary form. Again, participants were required to provide the items' category label.

Results and Discussion

Mean correct RTs and percentage of error rates to fruits in the six conditions are presented in Table 2. Error rates were low (total, 2.76%; to fruits, 1.50%) and were consistent with the pattern of RTs, ruling out a speed-accuracy tradeoff.

Two aspects of the results should be singled out. First, for both pictures and words, there was an increase in response time in the conditions in which fruits were compared with vegetables relative to the condition in which they were compared with weapons. However, as expected, the increase for words (45 msec) was about half of that for pictures (92 msec). Second, there was a strong effect of visual similarity for pictures but not for words when semantic relatedness was held constant: When pictures of fruits were to be discriminated from visually sim-

Table 2
Mean Response Times (RTs) in Milliseconds and Percent Errors (%) to the Items of the Category "Fruits" in Experiment 1

Conditions	Picture		Word		RT difference
	RT	%	RT	%	
Unrelated					
Central	555	.02	606	.00	-51
Peripheral	592	.07	639	.18	-47
<i>M</i>	574	.05	622	.09	
Low similarity					
Central	618	.02	622	.02	-4
Peripheral	715	.09	713	.19	+2
<i>M</i>	666	.06	667	.10	
High similarity					
Central	670	.11	615	.02	+55
Peripheral	763	.39	706	.39	+57
<i>M</i>	716	.25	660	.20	

ilar vegetables, participants were 50 msec slower to classify them than when these pictures were to be discriminated from visually dissimilar vegetables. RTs to the corresponding words were 7 msec faster in the visually similar condition than in the visually dissimilar condition. These conclusions were confirmed by an analysis of variance (ANOVA) with both participants and materials as random factors (Clark, 1973), carried out on the correct RT data of fruits. The effects of very long or very short latencies were reduced by establishing a cutoff point equal to 2.0 *SD* from a participant's mean and replacing outlying data with those values. Outlying values amounted to 2.08% of all data.

The factors considered in the ANOVA were form of the stimulus (picture vs. word), type of relationship between categories (unrelated vs. low similarity vs. high similarity), and degree of typicality of elements (central vs. peripheral). The ANOVA by participants was a $2 \times 3 \times 2$ within-subjects design, and the ANOVA by items was a $2 (2 \times 3)$ design with degree of typicality as a between-subjects factor. Post hoc analyses were conducted using Newman-Keuls. In all analyses, the criterion for statistical significance was $p < .05$.

RTs in the unrelated condition were faster than those in either the low- or the high-similarity conditions, which did not differ from each other [$F'_{\min}(2,51) = 22.24$, $MS_e = 8,062.46$].⁴ RTs to central members were faster than those to peripheral members [$F'_{\min}(1,8) = 9.94$, $MS_e = 4,856.01$]. Also, two first-order interactions reached significance. The interaction of stimulus form with type of relationship [$F'_{\min}(2,49) = 8.29$, $MS_e = 4,969.89$] was due to the fact that for pictures, each condition differed significantly from the others (with RTs in the unrelated condition being the fastest and RTs in the high-similarity condition being the slowest). On the contrary, for words, RTs in the unrelated condition were again the fastest, but there was no difference between the low- and high-similarity conditions. It is worth noting that in the unrelated condition, pictures elicited faster responses than words, while in the high-similarity condition, the opposite pattern was obtained. In the low-similarity condition, there was no difference between pictures and words.

The interaction between type of relationship and degree of typicality [$F'_{\min}(2,46) = 7.48$, $MS_e = 2,074.31$] was due to the fact that typicality played a role mainly in the low- and high-similarity conditions: RTs to central members were 94 and 92 msec faster in these two conditions, respectively; RT to central members was only 36 msec in the unrelated condition. Critically, there was no indication of a three-way interaction (whose implications will be examined in the General Discussion). No other main effect or interaction reached significance.

The pattern of results clearly suggests that we were successful in obtaining a purely visual effect differentially affecting pictures and words. Whereas for both types of stimuli, RTs in the fruits/weapons condition were the

fastest, RTs in the similar fruits/vegetables condition were longer than those in the dissimilar fruits/vegetables condition only for picture stimuli.

The pattern obtained cannot be easily explained by Snodgrass and McCullough's (1986) proposal, which postulates—in addition to the semantic route—a parallel, nonsemantic route for the categorization of pictures but not for words. Such a route is used only when there is a high degree of visual dissimilarity between the categories involved. Two aspects of our data render this proposal inadequate. First, the advantage of the fruits/weapons over both the similar and the dissimilar fruits/vegetables conditions was also present for the word data. So, it would be more economical to have a single explanation for the semantic distance effect for both types of stimuli rather than having two separate explanations, one for pictures and one for words. Second, it is not clear how Snodgrass and McCullough's proposal can explain the different pattern of the high- and low-similarity conditions found in Experiment 1 for pictures and words, since it assumes that for visually similar categories, the decision is based on information available in the semantic system. But such a system is common to pictures and words, and therefore the same pattern should be obtained for both types of stimuli.

Also, the Caramazza et al. (1990) model, although able to account for more data, is not fully supported at the empirical level. Thus, given the present functional architecture, the model cannot explain why in the semantically related, low-similarity condition of Experiment 1, the picture superiority effect disappeared. According to the model, pictures should still have enjoyed direct links between visual features and semantic predicates, and thus should have been responded to faster than the corresponding words.

On the basis of our results, we propose that the categorization process is a function of the semantic relatedness between (elements of the) categories as well as of the visual similarity between (elements of the) categories. The closer two categories are semantically, the more difficult it is to discriminate between them; analogously, the more similar two categories are visually, the more difficult it is to discriminate between them. Further, we postulate that the categorization process is always mediated by the semantic system, and that the information initially available for the membership decision process is different for pictures and words: both visual and semantic for the former, and only semantic for the latter. The data supporting this view are (1) the parallel pattern for pictures and words in the fruits/weapons condition yielding fast RTs, (2) the analogous effect of the degree of typicality for both pictures and words, and (3) the effects of visual similarity limited to the picture stimuli.

Within this interpretation, we attribute the faster responses in the unrelated condition—common to both pictures and words—to the semantic relatedness between the contrasting categories: Since there was no overlap

among semantic features of the two categories, no in-depth processing was needed to emit the response. For words, this is the only cause of fast responses in the unrelated condition, but for pictures, visual dissimilarity is an additional, facilitating factor. There is also another locus for this factor: We attribute the slow responses in the visually similar condition—affecting only pictures—to the need to discriminate among visual features that are shared by elements of the two categories in order to minimize false alarms (e.g., deciding that a potato is a fruit).

If so, making the visual analysis of the stimuli more difficult should have different effects for pictures and words. Experiment 2 was designed to test this hypothesis. With respect to the pattern obtained in Experiment 1, we hypothesized that stimulus degradation would produce a general increase in RTs, but we predicted such an effect to be additive for words and nonadditive for pictures. This means that for words, we expected the same increase in RTs in all the three conditions, whereas for pictures we expected the increase to be progressively larger from the unrelated condition to the low-similarity condition to the high-similarity condition.

EXPERIMENT 2

Method

Participants. Participants were 48 Italian-speaking undergraduates at the University of Padua.

Materials, Design, and Procedure. The same three conditions from Experiment 1 were used. The first condition involved items that were semantically unrelated and visually dissimilar (unrelated condition), the second condition involved items that were semantically related but were visually dissimilar (low-similarity condition), and the third condition involved items that were both semantically related and visually similar (high-similarity condition). The only difference between the present experiment and Experiment 1 was that half of the participants performed the task in exactly the same way as in Experiment 1, and half were tested in a condition in which both word and picture stimuli were degraded by a dot pattern that appeared together with the stimulus. An example of stimuli in the latter condition is provided in Figure 2.

Participants were tested in a $2 \times (2 \times 3)$ design. The independent variables were condition (intact vs. degraded stimuli), form of stimuli (pictures vs. words), and type of relationship between cate-

gories (unrelated vs. low similarity vs. high similarity). The apparatus and procedure were similar to those of Experiment 1. Participants were again asked to categorize each picture or word as belonging to one of the two categories appearing in each condition.

Results and Discussion

Mean correct RTs and percentage of error rates to fruits in the different conditions are presented in Figure 3a, where the confidence intervals for between-subjects data (Loftus & Masson, 1994) are also reported.⁵ Error rates were low (total 3.37%; to fruits 1.78%) and were consistent with the pattern of RTs, ruling out a speed–accuracy tradeoff.

The pattern of results in the intact stimuli condition was analogous to that found in Experiment 1. For pictures, RTs in the unrelated condition were the fastest, those in the low-similarity condition were intermediate, differing significantly from the other two, and those in the high-similarity condition were the slowest. On the contrary, for words, RTs in the unrelated condition were again the fastest, but the difference between the low- and high-similarity conditions was not significant. RTs to pictures were 57 msec faster than those to words in the unrelated condition, 2 msec slower than those to words in the low-similarity condition, and 26 msec slower than those to words in the high-similarity condition.

In the degraded stimuli condition, the overall pattern was quite different. For words, the increase in RT in the low-similarity versus the unrelated condition, and in the low-similarity versus high-similarity condition, was almost the same as in the intact stimuli condition (51 vs. 64 msec, and 11 vs. 11 msec, respectively). For pictures, instead, the increase in RT in the low-similarity condition relative to that in the unrelated condition was larger than in the corresponding conditions with intact stimuli (161 and 123 msec, respectively). Analogously, the increase in RT in the low-similarity condition relative to that in the high-similarity condition was larger than in the intact stimuli condition (78 and 35 msec, respectively).

These conclusions were confirmed by an ANOVA with both subjects and materials as random factors carried out on the correct RT data of fruits. The effects of very long

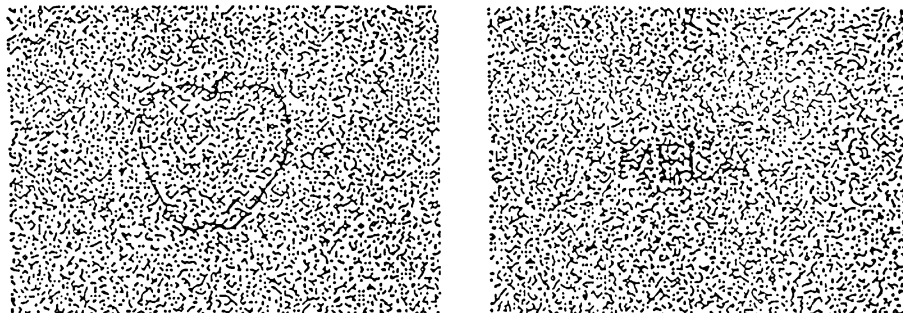


Figure 2. An example of the picture and word stimuli with a superimposed dot pattern used in Experiment 2.

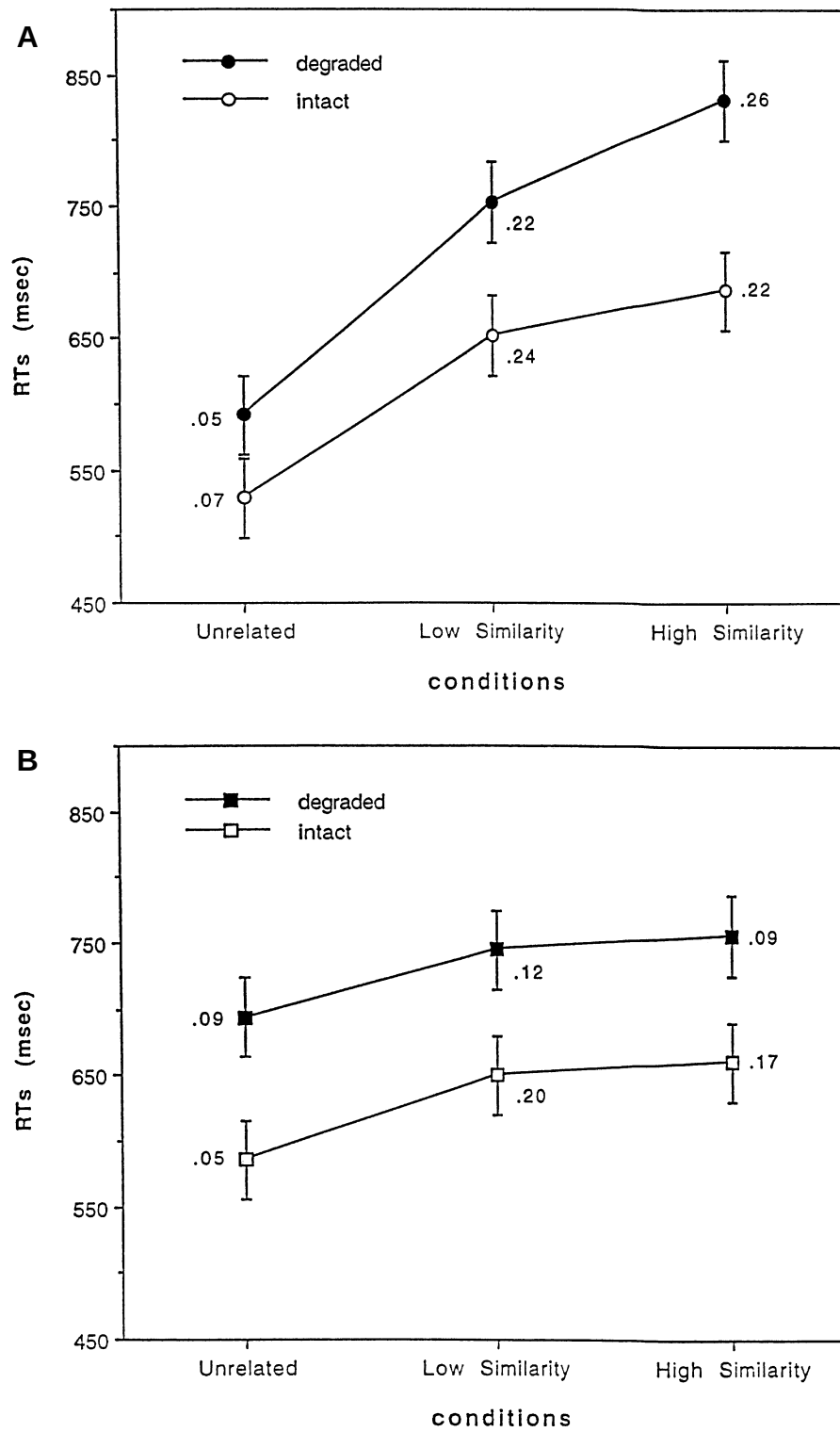


Figure 3. Response times (RTs) to pictures (A) and to words (B) in the different experimental conditions of Experiment 2. Segments indicate confidence intervals. Percentages refer to errors.

or very short latencies were reduced by establishing a cutoff point equal to 2.0 *SD* from a participant's mean and replacing outlying data with those values. Outlying values amounted to 2.29% of all data.

The factors considered in the ANOVA were condition (intact vs degraded stimuli), form of the stimulus (picture vs. word), and type of relationship between categories (unrelated vs. low similarity vs. high similarity). The ANOVA by subjects was a $2 \times (2 \times 3)$ design, with condition as a between-subjects factor, and the ANOVA by items was a $2 \times 2 \times 3$ within-subjects design. Post hoc analyses were conducted using Newman-Keuls. In all analyses, the criterion for statistical significance was $p < .05$.

RT in the intact condition (628 msec) was faster than that in the degraded condition [728 msec, $F'_{\min}(1,50) = 12.66$, $MS_e = 54,579.69$]. RT in the unrelated condition (600 msec) was faster than that in the low-similarity condition (700 msec), which in turn was faster than that in the high-similarity condition [734 msec; $F'_{\min}(2,37) = 28.23$, $MS_e = 6,640.76$]. The interaction of stimulus form with type of relationship was also significant [$F'_{\min}(2,71) = 12.11$, $MS_e = 5,497.58$] due to the fact that for pictures, RT in each condition differed significantly from that in the others (with RT in unrelated condition being the fastest and that in the high-similarity condition being the slowest). On the contrary, for words, RT in the unrelated condition was again the fastest, but there was no difference between RTs in the low- and high-similarity conditions. This interaction was qualified by the three-way interaction among condition, stimulus form, and type of relationship, significant in the analysis by item [$F(2,14) = 10.259$, $MS_e = 496.83$] and marginally so in the analysis by subject [$F(2,92) = 2.437$, $MS_e = 5,497.58$]: Whereas the RT difference to word stimuli in the intact and degraded condition was constant, the RT difference to picture stimuli increased steadily from the unrelated, to the low-similarity, to the high-similarity condition.

This pattern of results shows that words and pictures were affected differently by the dot pattern used as mask: RTs for the words showed a slowdown of about 100 msec in each of the three experimental conditions (108, 95, and 95 msec in the unrelated, low-similarity, and high-similarity conditions, respectively), while for pictures, such a slowdown rose markedly from the 63 msec in the unrelated conditions to the 101 msec in the low-similarity condition, to the 144 msec in the high-similarity conditions. Tetraedic contrasts between pairs of means showed that for words, the difference between the intact and the degraded stimuli in the three experimental conditions was constant (53.87, 47.46, and 47.10 in the unrelated, low-similarity, and high-similarity conditions, respectively), whereas for pictures, it was not (31.48, 50.52, and 71.92, in the unrelated, low-similarity, and high-similarity conditions, respectively). This confirms that the visual similarity of the concepts' referents affects picture categorization but not word categorization.

GENERAL DISCUSSION

The main results obtained in this study can be summarized as follows: In the categorization task we used, there were effects of semantic and visual similarity for pictures but only semantic effects for words. The visual similarity effects were such that the more visually similar (the elements of) the categories were, the longer the categorization time was. The magnitude of these effects was increased by degradation, with a different pattern for pictures and for words: Whereas for words, degradation had a constant effect over the three experimental conditions (unrelated, low similarity, high similarity), for pictures it had a small effect in the unrelated condition, a stronger effect in the low-similarity condition, and an even stronger effect in the high-similarity condition.

The semantic effects—which were analogous for both pictures and words—were of two types. First, there was a semantic distance effect, so that categorization was more difficult when the contrasting categories were semantically related (fruits/vegetables) than when they were semantically distant (fruits/weapons). Second, the degree of category membership affected categorization time, with RTs to central members being faster than those to peripheral members. Furthermore, the degree of category membership had a very strong effect in both the high- and the low-similarity condition but a weak effect in the unrelated condition, and this pattern held true for both pictures and words.

This pattern of results cannot be fully explained by either Snodgrass and McCullough's (1986) or Caramazza et al.'s (1990) models. The fact that the RT for the fruits/weapons comparison was also the fastest for word stimuli shows that its advantage is not due to visual factors only, as assumed by Snodgrass and McCullough: Words—unlike pictures—provide no visual-structural cues for inferring category membership. But this finding is easily explained if we consider that there are semantic as well as visual differences among the categories considered: Fruits and vegetables are semantically related whereas fruits and weapons are not. Thus, it seems plausible to suggest that the pattern obtained for pictures reflects a joint effect of visual and semantic factors, whereas that for words reflects the effect of semantic relatedness alone. Evidence congruent with this conclusion is the finding that the RT advantage of the fruits/weapons condition over the fruits/vegetables conditions was larger for pictures than for words in both experiments.

The proper role of visual similarity in categorization is further specified by comparing the high- and low-similarity conditions. These two conditions were constructed by manipulating the visual similarity among different sets of fruits and vegetables, holding semantic relatedness constant. When the visual similarity between fruits and vegetables was high, there was an increase in response time relative to the condition in which between-category visual similarity was low. This pattern held for

pictures but not for words, for which no effect of visual similarity emerged. Since the manipulation of visual similarity in the present experiments maintained categorical relatedness constant, the effect obtained is genuinely visual and cannot be attributed to semantic factors. This conclusion is strengthened by the lack of a visual similarity effect in the comparable word conditions.

Many features of the present data seem compatible with findings of recent studies investigating the role of similarity in speeded classification. The model proposed by Nosofsky and colleagues (Nosofsky, 1986; Nosofsky & Palmeri, 1997) predicts effects of both within- and between-category similarity, and therefore our results with picture stimuli are interpretable in a direct way within such a model; however, since the model does not explicitly deal with the lexical representation of concepts, its viability in accounting for our results is rather limited. Therefore, we will discuss the results with reference to the Snodgrass and McCullough (1986) and the Caramazza et al. (1990) models.

Snodgrass and McCullough's (1986) model cannot account for the difference between the high- and low-similarity conditions, since it assumes that for visually similar categories the decision is based on information available in the semantic system. But such a system is common to pictures and words, and therefore the same pattern should be obtained for both types of stimuli. In addition, the model runs into problems in explaining the analogous effects of the degree of category membership on picture and word categorization found in Experiment 1. Overall, RTs to central members were faster than those to peripheral members, but the size of the effect was more pronounced in the two semantically related conditions than in the unrelated condition, and this pattern held for both pictures and words. If word and picture categorization was indeed mediated by different routes in the unrelated condition, then it is not clear why the degree of membership affected the two processes in a similar way.

As for the Caramazza et al.'s (1990) model, given the present functional architecture, it cannot explain why in the semantically related low-similarity condition, the picture superiority effect disappeared. According to the model, pictures should still enjoy direct links between visual-structural features and semantic predicates, and RTs should thus be faster than those to the corresponding words.

A proposal that takes into account aspects of both Snodgrass and McCullough's (1986) and Caramazza et al.'s (1990) models and that accounts for the pattern obtained here, is that the semantic system is accessed by both an input lexicon—a system for the representation of orthographic units (see Morton & Patterson, 1980)—and a structural description system. The latter is thought of as a presemantic, modality-specific (in this case, visual) representation system for concepts (Jackendoff, 1987; Marr & Nishihara, 1978; Riddoch, Humphreys,

Coltheart, & Funnell, 1988; Sartori & Job, 1988; Schacter, Cooper, & Delaney, 1990). This system is called upon when objects are visually presented, but also when different object-oriented tasks requiring visual information to be processed are performed (see Sartori, Job, & Coltheart, 1993). The structural description feeds into the semantic system, just as the input lexicon does. However, unlike the input lexicon, the structural description is sensitive to the visual similarity of objects, and of their perceptual attributes, in that attributes shared by many concepts will tend to activate, to different degrees, the concepts themselves. Therefore, when the categorization task requires discrimination between visually similar objects, several representations of similar concepts will all be active in the structural description system. These will in turn activate, to different degrees, the corresponding entries in the semantic system, and the competition between the possible outputs will be quite high. In the visually dissimilar condition, instead, shared visual attributes will be fewer, and so the number of activated similar concepts competing for response will be smaller.

In addition to explaining the effect of visual similarity, this proposal also accounts for the differences between pictures and words. Since the structural description, but not the input lexicons, is sensitive to the visual similarity of objects and their perceptual attributes, pictures in the visually similar condition need more processing to enable discrimination among activated entries than do words in the corresponding condition, leading to a reversal of the picture superiority effect.

REFERENCES

- BOCCARDI, M., & CAPPA, S. F. (1997). Valori normativi di produzione categoriale per la lingua italiana [Category production norms for Italian]. *Giornale Italiano di Psicologia*, **24**, 425-436.
- BORTOLINI, U., TAGLIAVINI, C., & ZAMPOLLI, A. (1974). *Lessico di frequenza della lingua italiana* [Frequency norms for Italian]. Milan: Garzanti.
- CARAMAZZA, A., HILLIS, A. E., RAPP, B. C., & ROMANI, C. (1990). The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology*, **7**, 161-189.
- CARR, T. H., MCCAULEY, C., SPERBER, R. D., & PARMELEE, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 757-777.
- CLARK, H. H. (1973). The language as a fixed-effect fallacy. A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior*, **12**, 335-359.
- COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, **82**, 407-428.
- GOLDSTONE, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, **123**, 178-200.
- GUENTHER, R. K., & KLATZKY, R. L. (1977). Semantic classification of pictures and words. *Journal of Experimental Psychology: Human Learning & Memory*, **3**, 498-514.
- HUMPHREYS, G. W., RIDDOCH, M. J., & QUINLAN, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, **5**, 67-103.

- JACKENDOFF, R. (1987). On beyond zebra: The relation of linguistic and visual information. *Cognition*, **26**, 89-114.
- JOB, R. (1981). *Giudizi di appartenenza categoriale per 547 concetti della lingua italiana* [Category membership ratings for 547 concepts of the Italian language] (Rep. No. 50). Padua: University of Padua, Institute of Psychology.
- JOB, R., RUMIATI, R., & LOTTO, L. (1992). The picture superiority effect in categorization: Visual or semantic? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1019-1028.
- LOFTUS, G. R., & MASSON, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, **1**, 476-490.
- LOTTO, L. (1988). *Il ruolo dell'informazione visiva nel processo di categorizzazione* [The role of visual information in the categorization process]. Unpublished doctoral dissertation, University of Padua.
- MARKMAN, A. B., & GENTNER, D. (1996). Commonality and differences in similarity comparisons. *Memory & Cognition*, **24**, 235-249.
- MARR, D., & NISHIHARA, H. K. (1978). Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society: Series B*, **200**, 269-294.
- MCCLOSKEY, M. E., & GLUCKSBERG, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, **11**, 1-37.
- MEDIN, D. L., & SHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.
- MORTON, J., & PATTERSON, K. (1980). A new attempt at an interpretation, or an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 91-118). London: Routledge & Kegan Paul.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.
- NOSOFSKY, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics*, **45**, 279-290.
- NOSOFSKY, R. M., & PALMERI, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, **104**, 266-300.
- OLIVA, A., & SCHYNS, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, **34**, 72-107.
- POTTER, M. C., & FAULCONER, B. A. (1975). Time to understand pictures and words. *Nature*, **253**, 437-438.
- RAPP, B. C., HILLIS, A. E., & CARAMAZZA, A. (1993). The role of representations in cognitive theory: More on multiple semantics and the agnosias. *Cognitive Neuropsychology*, **10**, 235-249.
- RIDDOCH, J., HUMPHREYS, G. W., COLTHEART, M., & FUNNELL, E. (1988). Semantic system or systems? Neuropsychological evidence re-examined. *Cognitive Neuropsychology*, **5**, 3-23.
- RIPS, L. J., SHOEN, E. J., & SMITH, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning & Verbal Behaviour*, **12**, 1-20.
- ROSCH, E. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology: General*, **104**, 192-233.
- ROSCH, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.
- SARTORI, G., & JOB, R. (1988). The oyster with four legs: A neuropsychological study on the interaction of visual and semantic information. *Cognitive Neuropsychology*, **5**, 105-132.
- SARTORI, G., JOB, R., & COLTHEART, M. (1993). The organization of object knowledge: Evidence from neuropsychology. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 451-465). Cambridge: MIT Press.
- SCHACTER, D. L., COOPER, L. A., & DELANEY, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General*, **119**, 5-24.
- SNODGRASS, J. G., & MCCULLOUGH, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 147-154.
- SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

NOTES

1. In these studies, similarity, although not exclusively visual, is defined also with reference to the visual domain, and may thus be relevant for the issue under discussion.
2. More fruits were used at this stage because it was crucial to be able to select the best possible elements for the target category.
3. The distinction between fruits and vegetables follows people's judgments rather than botanical classification, and thus may vary from culture to culture. In this research, they were selected for, and tested with, Italian people (Boccardi & Cappa, 1997).
4. The MS_e used is that of the analysis by participants.
5. A preliminary analysis showed that the degree of typicality had the same effects as in Experiment 1 and did not interact with either degradation or stimulus form. Therefore, in order to simplify data presentation and analysis, this factor is not considered any further.

(Continued on next page)

APPENDIX
Italian Names and English Translations of the
Stimuli Used in the Experiments

Fruits			
<u>Mela</u>		Apple	
<u>Albicocca</u>		Apricot	
<u>Arancia</u>		Orange	
<u>Pesca</u>		Peach	
Castagna		Chestnut	
Limone		Lemon	
Melone		Melon	
Melograno		Pomegranate	
Similar Vegetables		Dissimilar Vegetables	
<u>Carciofo</u>	Artichoke	<u>Asparago</u>	Asparagus
<u>Lattuga</u>	Lettuce	<u>Broccoli</u>	Broccoli
<u>Peperone</u>	Pepper	<u>Carota</u>	Carrot
<u>Pomodoro</u>	Tomato	<u>Sedano</u>	Celery
Aaglio	Garlic	Mais	Corn
Cipolla	Onion	Cetriolo	Cucumber
Zucca	Pumpkin	Porro	Leek
Rapa	Turnip	Fungo	Mushroom
Weapons			
<u>Cannone</u>		Cannon	
<u>Pugnale</u>		Dagger	
<u>Mitraglia</u>		Machine-gun	
<u>Pistola</u>		Pistol	
Ascia		Axe	
Mazza		Club	
Mortaio		Mortar	
Forbici		Scissors	

Note—Central items are underlined.

(Manuscript received August 7, 1997;
revision accepted for publication June 9, 1998.)