The semantics of locative information in pictures and mental images[1]

Marcel Adam Just

and

Patricia A. Carpenter

DEPARTMENT
of
PSYCHOLOGY

Carnegie-Mellon University

In press, British Journal of Psychology

CIP WORKING PAPER #280
February 1975

The semantics of locative information in pictures and mental images[1]

Marcel Adam Just

and

Patricia A. Carpenter

Psychology Department

Carnegie-Mellon University

Pittsburgh, Pennsylvania 15213

Abstract

Three experiments examined how people compare sentences about spatial location to pictures and images. Previous investigations have found that people are faster at judging relative location when the description contains the word <u>above</u> or <u>right</u> than when it contains the word <u>below</u> or <u>left</u>. Experiment I showed that this asymmetry persisted when the words were replaced by arrows, indicating that the effect is not specific to particular lexical items. Experiment II showed the asymmetry persisted even when the response latency did not include the time to encode the description, indicating that the asymmetry does not lie in the description-encoding stage. Finally, Experiment III investigated how people compare sentences to information from a previously memorized picture. In this situation, the usual asymmetry was not present. The three studies suggest that the asymmetry arises from the way descriptions influence the encoding of perceptual events. The results also showed that the information encoded in a mental representation of a picture is ordered, such that certain features can be accessed more quickly than others. However, the same features are equally quickly accessed in a picture that is physically present.

Judging the relative location of objects in a visual display is easier when the display is described in terms of the word above than when it is described in terms of below (Seymour, 1969; Chase & Clark, 1971). For example, people are faster to confirm that the word above correctly describes an object in an upper location than confirming that below correctly describes a lower location. Similarly, Olson and Laxar (1973) have found that verification of descriptions with the word right is faster than verification of left. Performance in these word (or sentence)-picture verification tasks has been characterized in terms of four independent stages whose durations are additive (Chase & Clark, 1971): (1) reading and encoding the verbal description, (2) scanning and encoding the picture, (3) comparing the representation of the description with the representation of the picture and (4) executing the result of the comparison stage (true or false) as an overt response. The processing asymmetries between above and below (and right and left) have been ascribed to various stages, such as the reading and encoding of the description (Chase & Clark, 1971; Olson & Laxar, 1973), the scanning and encoding of the picture (Seymour, 1969), or the effect of the first three stages on response availability (Seymour, 1974). The present paper will demonstrate that the asymmetry between above and below (and right and left) is not specifically linguistic, but results from a more general asymmetrical conception of spatial dimensions. The experimental results are used to develop a model for the processing of relative spatial location which accounts for the asymmetry.

A secondary purpose of this paper is to compare the way information is retrieved from a picture to the way it is retrieved from a mental representation of that picture. The study of locatives (such as above and below) inherently involves the spatial distribution of information, one of the defining properties of pictures and images. The comparison of information retrieval from pictures and mental images will arise in Experiments II and III. The two experiments are identical except that in II, sentences are verified as true or false with respect to a picture, while in III the same sentences are verified with respect to a mental image of that same picture. A cognitive theory which copes with imagery must be able to specify the properties of an image--the kind of information that is stored, the format of the information, and the kinds of mental operations that can be applied to the information. Perhaps most importantly, the theory must specify how these properties are different for imaginal and non-imaginal representations. The comparison of performance between Experiments II and III provides some of the preliminary data about the relative accessibility of information in pictures and mental images.

There have been several kinds of theories to explain the asymmetries like those for above and below in word-picture comparison tasks. One of these theories has attributed the difference to processes in the word-encoding stage. Chase and Clark (1971) and Olson and Laxar (1973) argued that the words below and left take longer to encode than above and right, respectively. The proposal was that below and left refer to the negative pole of their underlying dimension, and this negativity takes extra time to encode.

The strongest evidence in favor of the word-encoding explanation of the asymmetry comes from a verification task which was identical to the above-below study, except that the words above and below were replaced with the corresponding non-linguistic symbols (↑) and (↓) (Chase & Clark, 1971). An example of the displays is shown in Table I. This procedure eliminated the word-encoding process, and the above-below (or ↑↓) latency difference disappeared. Similarly, Olson and Laxar (1973) eliminated the right-left difference when the words were replaced by the symbols (→) and (←). The conclusion drawn in these two papers was that the advantage of above and right had been due to the word encoding process.

It is possible that subjects in these "arrow" experiments could have

Insert Table I about here

responded on the basis of perceptual properties of each display. For example, the decision rule could have been "Respond true if the arrowhead is close to the dot; otherwise respond false". Subjects may not have compared the "meaning" of the arrow to the representation of the location of the dot. Rather, they might have responded on the basis of global perceptual features. If subjects were basing their response on such perceptual properties, then the entire process is very dissimilar to the usual word-picture comparison task. The use of these particular displays may have eliminated both the word-encoding stage and the subsequent process that compares the two representations. While these "arrow" experiments led to the conclusion that the asymmetry lay in the word-encoding stage, the alternative interpretation of the experiments leaves that conclusion in doubt.

Experiment I

The purpose of Experiment I was to investigate whether there was any latency advantage in favor of a non-linguistic symbol that denotes an upper location. Like Chase and Clark (1971), we used arrows to denote above and below. However, the arrow in the Chase and Clark experiment was strongly asymmetrical, with the direction of the asymmetry defining its meaning. This structural property of the arrow itself, rather than its meaning, may have been a perceptual cue that people used to make their decisions. To diminish this type of cue, we used arrowheads as shown in the lower part of Table I, paired with a consonant and a vowel, one letter above the other. Our intention was to force subjects to encode the direction of the arrow, encode the location of the vowel, and then compare the two locations. The subjects' task was to decide whether the direction of the arrow (up or down) was the same as the location of the vowel.

The hypothesis was that people would have to explicitly encode the direction in which the arrowheads were pointing, and compare this direction to the encoded location of the vowel. If the usual advantage of the word above is in some process other than word-encoding, that is, either in picture encoding or in the comparison process, then the upward-pointing arrows should enjoy a similar advantage. If the results show this advantage, then it would follow that the marking advantage of above over below may not lie in the word-encoding stage.

Method

The experiment was a timed verification task. The subject first examined an arrow that pointed up or down, as shown in Table I. Then, he examined an accompanying display of a consonant and a vowel to determine whether the vowel was above or below the consonant. If the direction of the arrow corresponded to the location of the vowel, the subject responded YES; if it didn't, he responded NO.

Stimuli. The letters in the display were 9° of visual angle to the right of arrow. The letter arrays were formed by randomly pairing one of the vowels AEIOU with one of the consonants BHLRZ. There were five examplars of each condition, composed of five different vowel-consonant pairs. There were four experimental conditions: arrow pointing up or down, and vowel above or below the consonant. Thus, there were 20 different stimuli.

Procedure. The subject initiated a trial by pressing a microswitch.

Half a second later, the display appeared in the tachistoscope. The subject was instructed to look first at the arrow and then at the letters. The subject responded either YES or NO by pressing one of two response buttons. The assignment of dominant hand to response button was balanced across subjects.

Subjects were instructed as to which stimuli were "matching" and which "mismatching", but the words above and below were never used. Prior to the beginning of the experiment, the subject was given 8 practice trials, as well as 4 lead-in trials before each block of trials. Feedback as to the correctness of responses was given only during practice trials. The subject then went through 8 blocks of 20 test trials. The order of trials within blocks was random and different for each subject. The subjects were 12 students who were paid for their participation.

Results

The analysis showed that latencies were shorter when the arrow pointed up than when it pointed down, as shown in Table II. The data were collapsed over the five or fewer correct responses for each condition. An analysis of variance examined the effects of the upward or downward pointing arrows, the matching or mismatching displays, and the three blocks of practice. The analysis showed that the advantage of upward-pointing arrows was 56 msec (S.E. = 17 msec), $F(1,11) = 10.32$, $p < .01$. The results also showed that matching (YES) trials were faster by 98 msec, $F(1,11) = 15.21$, $p < .01$. Also, subjects responded faster with increased practice, $F(7,77) = 5.68$, $p < .01$, although practice did not interact with the other effects. None of the interactions were significant. The overall error rate was a low 2.1%.

Insert Table II about here

The results of Experiment I show that people are generally faster when they process the concept of upwardness than when they process the concept of downwardness. The words above and below were not used in the experiment. Thus, the advantage of upwardness is not specific to lexical items like above. Rather, it may take less time to encode a symbol denoting upwardness than to encode one that denotes downwardness.

This interpretation is only slightly constrained by the fact that the task involved the position of a vowel relative to a consonant. Vowel localization is a new task that may have characteristics of its own. For example, Posner and Mitchell (1967) found that decisions that two letters belong to the same category (either vowels or consonants) were faster when the letters were vowels. If a vowel effect does happen to interact with location, then it does so in a way that preserves the upwardness advantage. The presence of this advantage, in the absence of verbal symbols such as above or below, fulfills the major purpose of Experiment I.

The current result suggests that the advantage of upwardness does not lie in the encoding of words per se. However, the results do not indicate at what stage the advantage does occur. It could be occuring in the initial encoding of the meaning of the arrows, the encoding of the subsequent consonant-vowel display, the comparison process, or a response process. Experiment II examined whether the advantage persists when the duration of the first stage is eliminated from the response latency.

Experiment II

The purpose of Experiment II was to show that the advantage of the word above was not in the initial encoding stage. The experiment uses a methodology that removes the word-encoding process from the response latency, without altering the nature of the sentence-picture comparison task. Very simply, the person heard and encoded a sentence, indicated that he had comprehended it, and 500 msec later a picture appeared. The person decided whether the sentence was true or false of the picture. Response latency was timed only from the onset of the picture. Thus, the latency should not include the time to hear and encode the sentence. If in this paradigm the latencies are shorter for above than for below and shorter for right than for left, it would indicate that the asymmetry between the terms is due to a stage of processing later than word encoding.

Experiment II also examined the way information is retrieved from a picture. People verified sentences about a simple picture containing 4 consonants, each at the vertex of a rectangle. Even this simple display allows us to determine whether information from certain parts of the display is more quickly accessed than from other parts. For example, suppose the picture is $\begin{smallmatrix} W & X \\ Y & Z \end{smallmatrix}$. The test sentence can ask if W is to the left of X, or if Y is to the left of Z. Both these statements require the response YES, and both contain the same preposition, left. They differ only in whether they refer to the top or bottom part of the picture. If the statements concerning whether W is left or right of X are verified faster than the questions concerning Y and Z, the inference can be made that the top of the display is more accessible than the bottom part. Similarly, questions can interrogate whether W is above or below Y and whether X is above or below Z. If the above-below statements involving W and Y are verified faster, it would suggest that the left side of the picture is more accessible than the right. Thus, this procedure will show whether one part of the picture is more accessible than another without confounding the meaning of the sentence with its referent.

Method

The experiment was a verification task in which the subject listened to a sentence like W is below Y. Then, he was timed while he looked at a picture and decided whether the sentence was true or false of the picture.

Stimuli. The pictures consisted of four letters arranged to form the vertices of a rectangle or square. There were four possible figures, with dimensions (in cm) 8 x 8, 8 x 3, 3 x 8, and 3 x 3. They were viewed in a tachistoscope at a distance of 58.1 cm so that 8 and 3 cm subtended visual angles of approximately 8° and 3°, respectively. The letters in the figures were four consonants randomly chosen from 16 possible consonants. A different quadruplet of consonants was assigned to each shape. The position of the four consonants varied from trial to trial, with the constraint that in all variations, a pair of letters was always on the same axis. For example, in all pictures W was either above or below Y, and W was either to the left of or right of X. This constraint allows four possible arrangements of consonants for each shape.

The sentences were of three types: X is above/below Y, X is to the right/left of Y, and X is diagonally opposite Y. The diagonally opposite sentences were included to discourage people from treating the consonants as two ordered pairs. The importance of this will become apparent in the next experiment. Moreover, a sentence with above or below could refer to the letters on the left of the figure or on the right. Analogously, a sentence with left or right could refer to letters on the top or on the bottom of the figure. Finally, a sentence with diagonally opposite could refer to either of the two diagonals, and could name the letters from upper to lower, or vice versa. Thus, there were twelve true sentences associated with each shape. In addition, twelve false sentences were constructed by reversing the position of the two letters in the true sentence. For example, if it was true that X is above Y, then the false counterpart was Y is above X. Sentences with diagonally opposite were falsified by citing two of the letters that were not diagonally related.

Procedure. On each trial, the experimenter read a sentence to the subject, like X is above Y. The subject pushed a button as soon as he had understood the sentence. Half a second later, the picture was presented. The subject was timed from the onset of the picture until he responded true or false. Therefore, the response latency did not include the time to hear and represent the sentence. The S indicated his response by pushing one of two response buttons. The assignment of dominant hand to response was balanced across subjects.

Prior to the test trials, the subject was given 24 practice trials on a test figure, 6 cm x 3 cm, composed of four vowels. Then, he was given two blocks of 24 test trials where all the sentences were followed by pictures of a particular shape, for example, only 8 x 8 cm squares, all containing the same four consonants, but not always in the same arrangement. Then, the next two blocks of trials dealt with pictures of a different shape, and so on, until the subject had gone through each of the four shapes. Thus, there was a total of 192 trials. The order of the 24 sentences was randomized within and between blocks. Also, each subject went through the four shapes in a different order. The subjects were 12 right-handed college students who were paid for their participation.

Results

The data for the correct responses were averaged over the two blocks of trials for each shape. The latencies for questions about vertical relations (*above* or *below*) were analyzed separately from statements about horizontal relations (*left* and *right*). The latencies for questions about diagonal relations were not analyzed because they were irrelevant to the theoretical issues of interest here.

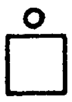
Vertical relations. The latencies for the correct responses were submitted to an analysis of variance, whose factors were the four shapes, the two truth values, the marked (*below*) and unmarked (*above*) adjective, and the position of the interrogated items (i.e., the left or right side of the rectangle). Only one term in the analysis of variance was significant, and that was the term of major interest. Sentences with *above* were verified an average of 56 msec (S.E. = 23 msec) faster than sentences with *below*, $F(1,11) = 6.12$, $p < .05$. These results are shown in Table III.

Insert Table III about here

Horizontal relations. Sentences with *right* were verified an average of 95 msec (S.E. = 41 msec) faster than sentences with *left*, $F(1,11) = 5.41$, $p < .05$. Aside from this term, three interaction terms reached significance. The *right-left* difference was larger for the two wider shapes, $F(3,33) = 4.40$, $p < .05$. There was a three-way interaction between true-false, *right-left*, and top-bottom positions, $F(1,11) = 7.24$, $p < .05$. There was also a four way interaction between the three factors just mentioned and the four shapes, $F(3,33) = 4.93$, $p < .05$.

One important point to notice is that the position of the interrogated item had no main effect on latencies. It did not matter whether the items were on the left or on the right (for *above* or *below* sentences), and whether they were on the top or bottom (for *left* or *right* sentences). So it appears that the various locations of the picture are equally quickly accessible.

Discussion

The size of the marking effect in this experiment, that is the advantage of *above* over *below*, and *right* over *left*, is similar to those found in experiments that included the time for word-encoding. This experiment attempted to eliminate the word-encoding stage by instructing subjects to initiate the picture presentation after they "understood" the verbally-presented sentence. And beyond that, the subjects were given an additional half-second before the picture appeared. Thus, subjects were probably given sufficient time to encode the sentence. Even so, this experiment resulted in a 56 msec marking effect for *above-below* and 95 msec for *right-left*. We can compare this effect to that found in other experiments. Clark and Chase (1972) had people verify sentences like *The star is(n't) above/below the plus* and found a marking effect of 93 msec. In experiments where words (*above/below*) were compared to the location of a circle in a display like , the marking effect was 75 msec (Chase & Clark, 1971) and 53 msec (Seymour, 1969). Similarly, the 95 msec preference for the term *right* in the current experiment is very similar to the 94 msec preference found by Olson and Laxar (1973) in their word-picture comparison task. Thus, even though the latencies in this experiment exclude sentence reading and encoding time, the marking effect is still present.

Moreover, it is of the same magnitude as the marking effects found in experiments like those cited above, that do include reading and representation time. Thus the advantage of above and right does not seem to lie in the word-encoding stage.

Experiment III

Comprehension and verification of sentences entails the comparison of information from the sentence to information from another source. The latencies necessarily include the time to scan and encode the picture. But the comprehension and verification process can be assessed in the absence of picture scanning and encoding in cases where the information from the sentence is compared to information already stored in long-term memory (Just, 1974). In Experiment III, people became very familiar with a simple picture (very similar to the ones used in Experiment II) and were then timed as they decided whether statements about the memorized picture were true or false. Thus, the task contains all the stages of the verification process, except that the picture scanning and encoding stage is replaced by the stage that retrieves information about the picture from long-term memory. This experiment should indicate the similarities and differences between retrieving information from a picture and retrieving information from a memorial representation of the picture.

Method

The experiment was a verification task in which subjects were timed as they read a sentence like X is above Y and decided whether it was true or false of a picture they had previously studied. The design of the experiment was similar to that of Experiment II. The main difference was that the arrangement of the quadruplet of consonants in the picture of a given shape was kept constant so that subjects could memorize the picture and respond to the sentence on the basis of their memory.

Stimuli. The pictures were of the same size and shape as in Experiment II, and presented at the same visual angle. However, the four consonants assigned to a given shape were kept in a single arrangement. The sentences were exactly the same as in Experiment II, but presented visually.

Procedure. The subject was given a test figure and asked to form a mental image of it since he would have to verify statements about the figure. The S studied each figure until he was certain he had memorized it. The purpose of including statements about diagonal relations was to encourage imaginal coding. The order of the 24 sentences was randomized for each block and subject. After each block of 24 trials the subject was asked to recall the test figure by drawing it exactly as he had seen it. There were two blocks of trials for each shape. The subject memorized and verified sentences about one shape at a time. The order in which the four shapes were examined was different for each S.

The subject initiated a trial by closing a microswitch. Half a second later, a sentence was presented and the subject was timed until he responded whether it was true or false of the memorized picture.

The Ss were 12 undergraduates who participated as part of an Introductory Psychology requirement. Eleven of the subjects were right-handed.

Results

Vertical relations. The results were treated exactly as in Experiment II. Of the four factors and their interaction terms, there were only two significant effects. People answered faster (by about 404 msec) when they were asked about letters on the left side of the figure than when they were asked about the right side, $F(1,11) = 19.28$, $p < .01$, as shown in Table IV. This left-right difference cannot be due to the comprehension of the sentences, since the sentences were the same and, moreover, referred to vertical rather than horizontal relations.

Insert Table IV about here

The second significant effect was an interaction between above-below and true-false, $F(1,11) = 37.80$, $p < .01$. Above was verified faster when it was true, while below was verified faster when it was false.

Horizontal relations. The data were treated analogously to the vertical relations, and the analysis yielded precisely analogous results. Statements about letters on the top were verified about 355 msec faster than statements about the botton, $F(1,11) = 12.37$, $p < .01$. The left-right true-false interaction was also present, $F(1,11) = 13.75$, $p < .01$. Sentences with the preposition left were verified faster when they were true while sentences with the preposition right were faster when false. Even though there were many conditions in this experiment, the results can be characterized in terms of two factors, the interaction term and the position of the interrogated items.

The difference between above and below was not significant, $F(1,11) = 3.39$. Only 8 of the 12 subjects were faster on above, but one of them by over 1000 msec. The difference between left and right (left being slightly faster) was also not significant, $F(1,11) = 1.16$. These results can be compared to those of Experiment II. In Experiment II, the advantage of above over below was 56 msec, with a S.E. of 23 msec, while the advantage of right over left was 95 msec with a S.E. of 41 msec. In Experiment III, which had the same sentences, the same number of subjects, and the same number of observations, the advantage of above was 192 msec, but was highly variable, with a S.E. of 105 msec. Furthermore, there was no advantage of right; in fact, left became faster by 125 msec, but this difference was highly variable, with a S.E. of 116 msec. So the sentence-picture experiment, which does not include sentence-encoding time, shows the usual small but reliable advantage of above and right. The experiment that referred to pictures in long-term memory, where latencies did include sentence-encoding time, showed large and unstable effects.

The response latencies were more variable and about three times longer in Experiment III than in Experiment II. It is possible that some of this extra duration and variability in Experiment III is due to a visualization conflict of the kind reported by Brooks (1967). Reading a sentence may interfere with maintaining or scanning the representation of the array of letters. After the subject read the sentence, it may have been necessary for him to retrieve the information about the array from long-term memory, and this retrieval process might produce the top- down and left-right positional effects that were obtained. Interference of this sort could potentially introduce the kind of variability and longer latencies that were observed, and these effects could have masked a small marking effect.

The overall error rate was 3.6%, similar to the 3.8% in Experiment II. The data from the one left-handed subject was much like everyone else's.

Discussion

The results demonstrated that the small but reliable advantage of *above* found in previous studies becomes highly variable, and for *left-right*, reverses direction when people retrieve information about pictures from long-term memory. The absence of a reliable advantage suggests that perhaps the effect is due primarily to picture encoding, the stage that was eliminated with the present procedure.

Both of the major results of Experiment III can be explained by assuming that subjects represented the information from the picture in a particular order: from top to bottom and from left to right. One major result was that sentences with *above* were verified faster when true, but those with *below* were faster when false. What these two faster conditions have in common is that the first consonant in the probe sentence is the upper one in the picture. Since this condition was faster it may be inferred that the consonants in the picture were represented from top to bottom. For example, suppose the left-hand column in the picture $\begin{matrix}W & X\\ Y & Z\end{matrix}$ were encoded as the proposition *W is above Y*. The coding matches the order of the consonants in the true sentence *W is above Y*, as well as the false sentence W is below Y. And latencies to both were short, presumably because the order of the consonants in the sentence matched the represented order from the picture. However, subjects took much longer when the order of the consonants in the sentence did not match the top-to-bottom order in the picture. They were equally slow for the true sentence Y is below W and for the false sentence Y is above W. One possibility is that when the order of the consonants in the sentence did not match the top-to-bottom order, subjects recoded their representation of the picture or the sentence and then recompared them. A similar explanation accounts for the true-false x left-right interaction if it is assummed that a horizontal row was encoded from left to right (e.g., W to the left of X). Thus, a top-to-bottom, left-to-right representation of the picture can account for the obtained interactions.

This form of representation can also account for the result that the above-below sentences were verified faster when they referred to the consonants in the left-hand column. This is entirely consistent with the hypothesis that in the picture representation, the proposition about the left-hand column precedes the proposition about the right-hand column. The search through these propositions could start with the first, and terminate with the proposition containing the interrogated consonants. Then it would follow that test sentences that refer to the left side of the picture would be verified faster. Similarly, the left-right sentences were verified faster if they referred to the upper row of consonants than if they referred to the lower row. This is again consistent with the proposal that the proposition about the top of the picture precedes the proposition about the bottom. So, a top-to-bottom, left-to-right representation of the consonants in the picture accounts for the position effect as well as for the true-false x unmarked-marked interaction described above.

The above-below asymmetry seems to be more general than the right-left asymmetry. Above has an advantage over below in two ways. Experiment II showed that when a description with above precedes a picture, the picture may be encoded faster than if the description is in terms of below. In addition, Experiment III showed that people spontaneously code the information in a picture so that it is congruent with the true description containing above. Therefore, the advantage of above is consistent across the two tasks. However, the right-left asymmetry is not as consistent. Experiment II showed that when a description with right precedes a picture, then the picture is encoded faster than if the description is in terms of left. However, Experiment III showed that people spontaneously tended to code the information from the picture so that it is congruent with the true description containing left. This demonstrates that the left-right asymmetry is more task specific and it may even be culture specific.

The comparison between Experiments II and III reveals an interesting difference between verifying a sentence against a picture and verifying the same sentence against a long-term representation of that picture. Experiment III shows that different parts of the memorial representation of a picture are differentially accessible and the relative accessibility seems to be determined by the order of encoding. This is in marked contrast to the results for retrieval of information from an actual picture. Experiment II showed that different parts of the picture, left or right, top or bottom, were equally accessible. In this respect at least, mental representations of pictures are not like real pictures.

Even though the task encouraged the use of imagery (by asking about horizontal, vertical, and diagonal relations), and subjects were instructed to form images, and reported doing so, there is no independent evidence that they did use imagery. In fact, the results indicate that subjects used propositional representations of the simple pictures in performing this task.

General Discussion

The concept of lexical marking has its beginnings in the field of linguistics. Bierwisch (1967) pointed out the consistent asymmetry between antonyms that refer to spatial dimensions, e.g., high-low. One of the adjectives, in this case high, is linguistically more basic, or unmarked, while the other is more derived or marked. The unmarked term always refers to extent, (for example, high, tall, deep), while the marked term refers to lack of extent (for example, low, short, shallow). One distinguishing property is that unmarked adjectives can be used to ask neutral questions (e.g., How high is that flag?, How tall is that man?, How deep is that pool?). These questions presuppose nothing about the relative extent of the object while questions with the marked adjective presuppose lack of extent (e.g., How low is that flag?, How short is that man?, How shallow is that pool?). In addition, the name of the adjectival dimension is sometimes morphologically related to the unmarked term (e.g., height, depth), but it is never morphologically related to the marked term. Bierwisch suggested that these properties of adjectives might be universal to all languages.

A linguistic universal like the marking of spatial adjectives may be based on powerful conceptual and environmental universals (Clark, Carpenter & Just, 1973). Therefore, lexical marking should be manifested in a number of psychological processes. In fact, unmarked adjectives are processed faster and with fewer errors than marked ones in many tasks, such as deductive reasoning (Clark, 1969), sentence-picture verification (Just & Carpenter, 1971; Clark, et al., 1973), and sentence memory (Carpenter, 1974). In the sentence-picture verification task, the marking effect probably lies in one of the four stages that contribute to the total latency: encoding the word or sentence, encoding the picture, comparing the two representations, and outputting the result of the comparison process in a response.

The locus of the marking effect

The proposal that we advance attributes the advantage of above and right to the picture-encoding stage. In a situation where a sentence precedes a picture, the preposition of the sentence determines the nature of the encoding of the picture. The words above and right engender a more natural, canonical and therefore faster encoding. Below and left cause the picture to be encoded in a way that is not canonical and consumes more time. This encoding could take longer for one of two reasons. One possibility is that all relations are first encoded in a canonical way corresponding to the unmarked form. But if the preceding description contains a marked word, then the representation of the picture is recoded into the format that matches the marked form. Thus, there would be an extra time-consuming recoding process involved in arriving at a picture representation that followed a marked description. The other possibility is that if the preceding verbal description is marked, then the picture is directly encoded in a format corresponding to the marked description, but it may take longer to derive this encoding. In either case, our proposal specifies that the picture-encoding process takes longer when a marked, rather than an unmarked, verbal description precedes the picture.

This model can account for the results of the current experiments, as well as a number of previous studies. Our sentence-picture verification task (Experiment II) showed the usual marking effect, although sentence encoding time was not included in the response latency. This is consistent with the proposal that the marking effect is located at the picture-encoding stage. In Experiment III, where the picture had been previously encoded (in fact, memorized), then the marking effect was abolished for right-left, and made unreliable for above-below. The main difference between Experiments II and III was that III didn't include a picture-encoding stage. Otherwise, the two experiments had the same pictures, sentences, and responses. Since the marking effect was gone, or at least highly variable, it is reasonable to conclude that the effect is in the picture-encoding stage.

Although we have proposed that the marking effect occurs at the picture-encoding stage, a more general formulation would be that the marking effect occurs whenever a linguistically marked description determines the encoding of information from a second source, like a picture, semantic memory, or an auditory display. This generalization seems sound because marking effects are found in tasks that do not involve pictures. For example, there are marking effects when the word above or below is verified with respect to the frequency of a test tone relative to a reference tone (Harvey, 1973). Also, marking effects occur when sentences are verified with respect to one's knowledge of the world, e.g., Giraffes are taller than goats (Carpenter, 1974). Thus, the effect of a marked description is to increase the time it takes to encode information from any second domain when the description governing that encoding is marked.

This generalization also accounts for the difference between semantic memory experiments that show a marking effect and Experiment III, where sentences were verified with respect to previously memorized pictures. In the semantic memory experiment cited above, with sentences like Giraffes are taller than goats, the information about the relative heights was probably computed after the sentence was read. Thus, the sentence could influence the way the information was retrieved and coded. And, there was a reliable marking effect. By contrast, in Experiment III the relative locations of the consonants in the picture were encoded before the sentence appeared. Thus, the sentence had no opportunity to influence the encoding of the picture and there was no reliable marking effect. The marking effect occurs only when a description can influence the coding of its referent.

The influence of a description on the encoding of a subsequent picture has been investigated by monitoring subjects' eye fixations during a sentence-picture verification task (Carpenter & Just, 1972). Subjects were given sentences with unmarked quantifiers like A majority (or large proportion) of the dots are red or marked quantifiers like A minority (or small proportion)..., and a picture of a large and a small subset of dots. Following the word large proportion or majority, people fixated primarily the larger subset, while following small proportion or minority they fixated primarily the smaller subset. So marked and unmarked words made people look at different parts of the same picture in that study. This difference in locus of fixation can be interpreted as a difference in picture encoding. And the verification latencies in this task support the hypothesis that the encoding following a marked description takes extra time. Sentences with unmarked quantifiers were verified significantly faster than sentences with marked quantifiers, (Just & Carpenter, 1971).

Two recent studies have shown that it is possible to control the picture encoding stage (and hence the marking difference) by means of explicit instructions on how to code the picture. Using stimuli like those in the eye-movement experiment described above, Just and Carpenter (1971) instructed subjects on whether to attend to the large or small subset of dots in the picture that preceded the sentence. When they were explicitly told to code the color of the larger subset (and in a condition without explicit coding instructions) quantifiers that referred to the large subset (e.g., a majority, a large proporiton) were verified faster. However, when subjects were explicitly instructed to encode the picture in terms of the color of the smaller subset, then the sentences that referred to the small subset were verified faster. Uniformly imposing an unnatural encoding of the picture eliminated the usual marking effect.

In a similar study, Clark and Chase (1972) asked people to verify sentences like *The star is above the plus* with respect to previously presented pictures like $\overset{*}{+}$. When subjects were not instructed on how to encode the pictures, or when explicitly instructed to attend to the top of the picture, or to both parts of it, the advantage of *above* was between 93 and 117 msec. However, when subjects were instructed to attend to the bottom, the advantage dropped to 30 msec. In fact, for the true affirmative sentences, *above* was 33 msec slower than *below*. These results are completely consistent with our proposal that *below* usually causes an unnatural and time-consuming coding of the picture. When people are explicitly instructed to use this unnatural encoding all the time, then the advantage of *above* diminishes or disappears.

A different class of explanation of the marking effect is offered by response availability models (Seymour, 1973; 1974), which attribute variation in response latencies to the ease or difficulty of translating from a particular word-picture display to a particular response (e.g., YES or NO). Positive features in the representation of the word-picture display reduce the threshold for selection of a YES response, which is selected by sampling semantic features until a threshold is exceeded (Seymour, 1973). The two studies that experimentally deal with response selection and allocation unfortunately provide contradictory results. Seymour (1974, Experiment II) asked people to judge whether the position of a dot relative to a face was correctly described by the word *above* or *below*. When the face was inverted (i.e., mouth at the top, eyes at the bottom), the advantage of *above* disappeared, consistent with the response availability notion. However, in an analogous study where subjects judged whether the word *right* (or *left*) correctly described the position of a dot relative to a square *from the perspective of a person facing the subject*, the advantage of *right* over *left* persisted (Olson & Laxar, 1973). Both experiments changed the usual mapping from the physical display to the response, such that people respond positively (YES) when there is a mismatch between the word and the physical position of the dot. The presence of the mismatch (a negative feature, according to Seymour, 1974) should have affected response availability similarly in the two studies. The inconsistency between the two results leaves the response availability model in some doubt.

Summary

We have proposed that when a linguistic description contains a lexically marked term, then the process of encoding subsequent information from the referential domain takes longer. This view has direct implications for what makes a sentence or word "hard" or "easy" to process. According to this view, there are no hard or easy words or sentences per se. Rather, there are hard or easy correspondences between descriptions and their referents (Carpenter & Just, 1975). In general, the unmarked terms more closely correspond to the way we normally encode and conceptualize the environment.

Bierwisch, M. Some semantic universals of German adjectivals. Foundations of Language, 1967, 3, 1-36.

Brooks, L. The suppression of visualization by reading. Quarterly Journal of Experimental Psychology, 1967, 19, 289-299.

Carpenter, P. A. On the comprehension, storage and retrieval of comparative sentences. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 401-411.

Carpenter, P. A., & Just, M. A. Semantic control of eye movements during picture scanning in a sentence-picture verification task. Perception and Psychophysics, 1972, 12, 61-64.

Carpenter, P. A., & Just, M. A. Sentence comprehension: An information processing model of verification. Psychological Review, 1975, in press.

Chase, W. G., & Clark, H. H. Semantics in the perception of verticality. British Journal of Psychology, 1971, 62, 311-326.

Clark, H. H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404.

Clark, H. H., Carpenter, P. A., & Just, M. A. On the meeting of semantics and perception. In W. G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.

Harvey, N. The semantics of high-low continua. British Journal of Psychology, 1973, 64, 51-54.

Just, M. A. Comprehending quantified sentences: The relation between sentence-picture and semantic memory verification. Cognitive Psychology, 1974, 6, 216-236.

Just, M. A., & Carpenter, P. A. Comprehension of negation with quantification. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 244-253.

Olson, G. M., & Laxar, K. Asymmetries in processing the terms "right" and "left". Journal of Experimental Psychology, 1973, 100, 284-290.

Posner, M. I., & Mitchell, R. F. Chronometric analysis of classification. Psychological Review, 1967, 74, 392-409.

Seymour, P. H. K. Asymmetries in judgments of verticality. Journal of Experimental Psychology, 1974, 102, 447-455.

Seymour, P. H. K. Judgements of verticality and response availability. Bulletin of the Psychonomic Society, 1973, 1, 196-198.

Seymour, P. H. K. Response latencies in judgements of spatial location. British Journal of Psychology, 1969, 60, 31-39.

Footnotes

[1]The order of authors is arbitrary. The paper represents a collaborative research effort. The research was supported in part by the National Institute of Education, grant NIE-G-74-0016 and by the United States Public Health Service, National Institute of Mental Health, Grant MH-07722. We thank Joe Christie for his aid in collecting data for Experiments I and II. Requests for reprints should be sent to the Psychology Department, Carnegie-Mellon University, Pittsburgh, PA. 15213, U.S.A.

Table I

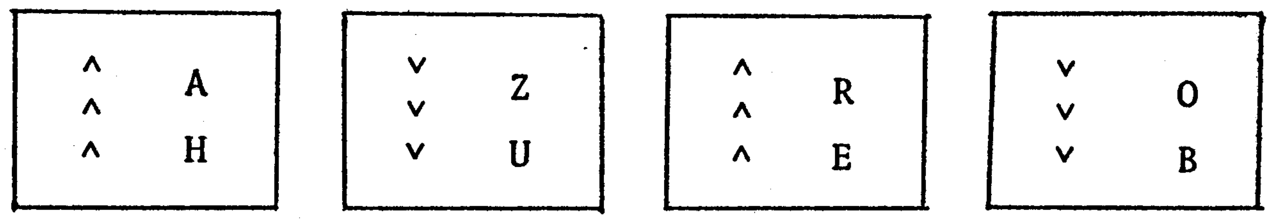
Schematic diagram of displays

used by Chase and Clark (1971)

Arrow:	'above'	'below'	'above'	'below'
Response:	yes	yes	no	no

Schematic diagram of displays

used in Experiment I



Arrow:	'above'	'below'	'above'	'below'
Response:	yes	yes	no	no

Table II

Response latencies in msec (% error)

for Experiment I

	Stimulus Description	
Response	Arrow pointing upward	Arrow pointing downward
Yes	1111 (2.1%)	1187 (2.5%)
No	1230 (2.7%)	1265 (1.0%)
Mean	1170	1226

Table III

Response latencies in msec (% error)

for Experiment II

	Stimulus Sentence			
	Above	Below	Right	Left
Response	(Letters on the right side)		(Letters on the top)	
True	1106 (1.0)	1127 (2.1)	1233 (5.2)	1210 (7.3)
False	1123 (2.1)	1122 (1.0)	1168 (2.1)	1439 (9.4)
	(Letters on the left side)		(Letters on the bottom)	
True	1052 (2.1)	1187 (5.2)	1181 (4.2)	1340 (7.3)
False	1122 (0.0)	1192 (0.0)	1278 (7.3)	1250 (5.2)
Mean	1101	1157	1214	1309

Table IV

Response latencies in msec (% error)

for Experiment III

	Stimulus Sentence			
	Above	Below	Right	Left
Response	(Letters on the right side)		(Letters on the top)	
True	3095 (3.1)	3916 (6.3)	3964 (4.2)	3346 (3.1)
False	3649 (0.0)	3357 (1.1)	3660 (3.1)	4223 (4.2)
	(Letters on the left side)		(Letters on the bottom)	
True	2662 (3.1)	3490 (3.1)	4460 (9.4)	3592 (5.2)
False	3416 (0.0)	2835 (1.1)	4068 (2.1)	4492 (8.3)
Mean	3206	3400	4038	3913