LINGUISTIC AND CONCEPTUAL PROCESSES IN THE COMPREHENSION OF SPATIAL ANTONYMS

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY

LISA FRIEDENBERG

1975

G9(?)...

ABSTRACT

LINGUISTIC AND CONCEPTUAL PROCESSES IN THE COMPREHENSION OF SPATIAL ANTONYMS

By

Lisa Friedenberg

This study compared three pairs of antonyms in a picture-sentence matching task. The three word pairs refer to the vertical perceptual dimension, and should reflect the asymmetry of that dimension in reaction time. Although all three pairs showed an intra-pair difference (referred to as the lexical marking effect), grammatical class influenced the size of that difference and the sensitivity of that difference to other experimental variables (e.g., sentence negation, the correspondence of the sentence to a picture, etc.). The differences between these word pairs strongly affected the fit of a model of picture sentence matching to the data, indicating that grammatical class is an important variable in the study of intra-pair differences in antonyms.

Approved: Gary M. Olson

LINGUISTIC AND CONCEPTUAL PROCESSES IN THE COMPREHENSION OF SPATIAL ANTONYMS

By

Lisa Friedenberg

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1975

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The purpose of this experiment is to compare three sets of antonyms in a picture-sentence matching task. The three pairs of terms refer to the same perceptual dimension, but differ according to grammatical class. H. Clark (1973) proposed that intra-pair differences in spatial terms are direct reflections of the characteristics of the perceptual dimensions to which they refer. According to this theory, the characteristics of man's perceptual space have been "mapped" onto corresponding lexical items. An asymmetrical perceptual dimension should result in asymmetries in the use and comprehension of the referential terms. It follows that the asymmetries observed between members of pairs of spatial terms should be similar, if all pairs refer to the same perceptual dimension. Pairs of spatial terms referring to the same perceptual dimension should behave in an analogous manner when applied to the same task. The present study examines the plausibility of this inference by comparing the adjectives <u>high-low</u>, the prepositions <u>above-below</u>, and the verbs <u>rising-falling</u>.

This concept of intra-pair differences in lexical items redefines the "lexical marking effect" in spatial terms. Lexical marking is present whenever one member of a pair of antonymous terms is linguistically and psychologically more complex than its counterpart. When first studied with adjectives (H. Clark 1969a, 1969b), lexical marking was defined as the presence of an additional linguistic feature in the representation of one pair member. This additional feature increases the psychological complexity of that item, observable as longer latency to response (H. Clark 1969a) and higher error rates (H. Clark 1969b) to problems using that item.

The simpler adjective, the "unmarked" one, may be used in two ways: nominally, to name the dimension in question; or contrastively, to locate an object on that dimension. The more complex adjective, the "marked" one, may only be used contrastively. Using good-bad as an example, "good" is unmarked with respect to "bad." Consider the question: "How good is Dick?". There is no a priori assumption that "Dick" is "good" or "bad." The question may be answered by either adjective. This constitutes nominal use. However, using the marked adjective "bad" in a question presupposes that "Dick" is "bad," and located somewhere at the lower portion of that dimension. The question "How bad is Dick?" could not be answered by the unmarked term "good" without some additional knowledge on the part of the respondent. Even in question form, the marked adjective does not have nominal use. Statements about the "good-ness" or "bad-ness" of an object constitute contrastive use, and the locating of that object somewhere on that scale. Linguistically, this distinction is expressed by the presence of two feature representations for unmarked adjectives, with only one feature representation for marked adjectives: "good" in the nominal sense being (+evaluative(polar)); and "good" and "bad" in the contrastive sense being (+evaluative(+polar)) and (+evaluative(-polar)) respectively (H. Clark 1969a). As shown by this example, the nominal use of an unmarked adjective involves the construction of a representation with one less feature than the contrastive use of either adjective. The differential performance or asymmetry within such pairs was at first related to the storage and retrieval of the nominal representation of unmarked adjectives (H. Clark 1969a).

The development of this theory led to a flurry of research with pairs of antonyms in many grammatical classes. Additional studies with adjectives confirmed the above findings (H. Clark and Card 1969, H. Clark 1972). Developmental research revealed that comprehension of unmarked adjectives was superior at earlier ages (E. Clark 1971a), as was the tendency to use them in descriptions (Donaldson and Wales 1970). Investigation of the temporal adverbs before-after showed similar trends (E. Clark 1971b). In comparing the acquisition of lexically marked word pairs in two semantic fields, dimensional terms like "good-bad" and spatio-temporal terms like "above-below" and "before-after," the latter were acquired at an earlier age (E. Clark 1972).

Since the marking effect emerges earlier in spatio-temporal terms, an investigation of the derivation of this effect was undertaken (H. Clark 1973). An attempt was made to differentiate the source of lexical marking in spatio-temporal terms from the effect in dimensional terms. Spatial terms have direct correspondence to perceptual space, making them more salient than other types of terms. Temporal terms are in turn derived from these, i.e., the concept of time as a "spatial metaphor" (H. Clark 1973). Dimensional terms, however, comprise a special subset of terms that refers to abstract dimensions. These dimensions are only analogies to spatial dimensions. Taking "good-bad" again as an example, people often think of a continuum relating these two words, with one word at either end of that continuum. Typically, the unmarked dimensional term is thought of as at the "top" of that dimension. This line of thought is derived from analogy to perceptual space (H. Clark 1973).

What about spatial and temporal terms leads to a lexical marking effect? According to H. Calrk (1973) lexical marking in spatial terms is the expression of asymmetries in our perceptual space. Using the ground level as a reference plane, all of our perceptual apparatus is located above this plane, giving that direction salience. This is reflected in language as the presence of marking in terms referring to the vertical dimension, with the item referring to the lower part of that dimension always being marked (e.g. below, lower, etc.). Using a reference plane that divides the body into front and back sections, we again find perceptual asymmetry, in that all of the perceptual apparatus is located on the front section. This is expressed in language as the presence of marking in word pairs referring to the front-back vertical dimension (e.g. front-back, forwards-backwards, push-pull, etc.). The last reference plane considered divides people into right-left sections, with the perceptual apparatus evenly divided. It would appear that lexical marking should not occur in word pairs referring to this dimension. However, there are additional considerations that may induce asymmetry in this plane. Most people have a preferred hand, commonly the right, which could lead to a preference for that direction. Many societal conventions emphasize the "right" direction (e.g. reading, writing, driving). These considerations do lead to a somewhat consistent asymmetry in this dimension, and the presence of a marking distinction between the terms left-right (Olson and Laxar 1973).

The theory of perceptual mapping enabled researchers to explain the existence of marking in terms not showing the distinction between nominal and contrastive senses. In addition, it provided further justification for the differences in reaction time and error rate observed in such pairs. But by emphasizing the role of perceptual correspondence, certain distinctions related to grammatical class have been overlooked. The semantic field of spatiotemporal terms includes word pairs of many grammatical classes. If grammatical class is a relevant variable in the study of intra-pair differences, certain characteristics of a particular grammatical class could influence the expression of lexical marking within a semantic field.

To examine this question, this study compares three pairs of terms referring to the vertical perceptual dimension in an identical task. All refer to the same asymmetrical vertical dimension; all should show lexical marking effects. Several predictions about the role of grammatical class in determining the size of this effect can be made. First, since the terms high-low are adjectives, there could be a significantly larger marking effect within this pair than within either of the other two pairs. This larger intra-pair difference would be a direct reflection of the importance of nominal and contrastive senses in comprehension, a distinction present only in adjectives. Second, since the pair rising-falling are verbs, which denote action, there may be a smaller lexical marking effect within this pair. This could be related to the absence of a static, easily measurable difference between the locations indicated by these words. Finally, although all three pairs describe locations on the vertical dimension, syntactic considerations make certain ones more directly applicable to describing those relationships. This could result in modifications of the comprehension process for certain word pairs. There could be significant differences in reaction time to identical problems as a function of grammatical class. This difference could in turn affect the fit of a processing model developed for one pair when applied to the others. Since a model of sentence processing in picture-sentence matching for the pair above-below has already been developed (H. Clark and Chase 1972), the fit of this model to data on above-below, high-low and rising-falling can be tested. If the comprehension processes in picture-sentence matching vary according to grammatical class, then even the perceptual correspondence of these three word pairs cannot lead to a similar fit of this model in all cases.

The Model

The model of picture-sentence matching developed by Clark and Chase (1972) involves four processes: encoding the picture, encoding the sentence, comparing the two codes, and executing a response. It was originally applied to data on affirmative and negative sentences using above-below. A picture of a plus and a star in a vertical relationship, and a sentence using "above" or "below" were presented to subjects as slides. Subjects were instructed to view the picture, view the sentence, compare the two, and decide whether or not the sentence accurately described the picture. Reaction time (RT) is described by an additive factors model of processing. RT is viewed as the sum of a series of additive factors which increase RT from a baseline amount as task complexity is increased (Sternberg 1969, 1971). This view of RT differs from earlier attempts to partition RT into a series of processing stages. The stage models use a subtractive method to determine stage durations (cf: Donder's subtractive method, Sternberg 1969, 1971). This approach necessitates the inclusion of an entire new stage to account for increased task complexity. Instead, the additive factors model examines the relationship between factors influencing stage durations. RT is seen as the sum of a series of non-interacting, additive factors--not stages. This allows for estimation of mean RT from the sum of the mean RT components, and estimation of RT variance from the sum of the variance of the RT components. When an interaction between two factors occurs, the two factors are influencing a stage in common, weakening the assumption of additivity (Sternberg 1969, 1971).

An additional assumption about higher order cumulants permits the derivation of the entire RT distribution from the distributions of the RT components (Sternberg 1969). It is possible through this approach to partition RT in picture-sentence matching into a baseline amount, and a series of factors that relate directly to increased task complexity. For example, encoding a negative sentence is more difficult than encoding a positive one (Clark and Chase 1972; Trabasso 1974; Trabasso, Rollins and Shaughnessy 1971). It is not necessary to include another processing stage to account for this difference. The effect can be handled by the factor of "negation," which is here influencing the stage of stimulus encoding.

The model developed for above-below uses a total of five factors to estimate mean RT in picture-sentence matching. Each factor influences a separate encoding or comparison process. This allows for the prediction of additivity, since no factors are assumed to influence any stages in common. The authors examined several other models, and found all unable to account for the data.[1]

The first process in picture-sentence matching is encoding the picture. Because subjects are aware of the comparison to be made, it is assumed that the pictures are

---

[1] The alternative models examined and rejected by Clark and Chase (1972) included: visual imagery models, conversion models, and picture negation models. For a complete description of these models, refer to that article.

encoded in linguistic form. The picture $\begin{matrix}+\\ *\end{matrix}$ could be encoded as either (plus above star) or as (star below plus). Clark and Chase (1972) present several reasons why the first code appears more probable. First, although both codes describe the same physical situation, the former describes the position of the plus relative to the reference point of the star, while the latter describes the situation in reverse. English typically describes relationships on the vertical dimension from a reference point at the bottom of the dimension. Clark (1973) relates this to the location of the perceptual apparatus, with the ground level as the reference plane, and direction defined upwards from that level. The dimension of "height" is typically described as "direction upwards," not "direction downwards." The use of "above" in the picture code is consistent with the common English description of such a situation.

Second, in developing the model, the authors assumed that the "above" code was used in all cases of picture encoding. The model was subsequently able to account for 97.5% of the variance in the data (Clark and Chase 1972). This indicates that the use of that code was highly consistent with the performance of subjects in the task.

Third, in an additional study reported in the same manuscript, the authors compared the performance of subjects when specifically instructed to attend to either the top of the picture, the bottom of the picture, or the entire picture. If we assume that the instructions influence the preference for code construction, the fact that no differences in encoding time resulted indicates that neither code is more difficult to construct. As the model quite accurately predicts subjects' performance when the "above" code is assumed to be used, this assumption seems justified.

The second encoding process, encoding the sentence, necessitates a series of encoding stages and influential factors. The main clause of the sentence is encoded first. The presence of <u>lexical marking</u> is a factor that may increase encoding time. The construction of a code for the sentence "The star is below the plus" should exceed the time taken to encode the sentence "The star is above the plus" by a factor attributable to lexical marking.[2] The second stage of sentence encoding deals with embedding of the main clause, as in the case of <u>negative</u> sentences. Here, the subject must deal with the embedding of the clause "the star is above the plus" in the clause "It is false that . . . ." The addition of this embedding, yielding a code of the form (false(star above plus)), will

---

[2]While encoding a picture in a marked term evidently does not affect picture encoding time, encoding a sentence with a marked term in it does. This illustrates the point that the marking effect is inherent in the actual word, not the location it indicates.

increase RT by a factor due to negation (Clark and Chase 1972).

After the two codes have been created, it is necessary to compare them. According to Clark (1969a), two codes can be compared only for congruence (i.e. identity). If there are any mismatches in the forms (subject-object agreement) of these codes, some manipulation is needed before comparison can continue. The goal of this process is to obtain enough information about the underlying meaning of the picture and sentence to judge them as equivalent or not. The model assumes that subjects keep track of their response preference via a "truth index" that changes from "true" to "false" as various comparisons are made. The initial value of this index is set at "true," since it is assumed that subjects are prepared to make a "true" judgement unless there is evidence to the contrary (Clark and Chase 1972).[3]

Three stages are involved in the comparison process, with three factors capable of increasing RT. Subjects begin by comparing the inner or "embedded" strings of the codes. If the <u>subjects</u> of these codes do not match, a

---

[3]The concept of a "truth index," whose initial value is set at "true," has been employed by other researchers as well (cf: Trabasso 1974--cited in Clark and Chase 1972).

transformation is made on the sentence code. In comparing the picture code (star above plus) to the sentence code (plus above star), the latter would be transformed to the form (star below plus), with a corresponding increase in RT. The next comparison made is of the relational term of the embedded strings. Since the forms of the codes are already congruent, a mismatch of relational terms will not result in transforming either code. Instead, a mismatch in this stage will lead a subject to change his/her response preference from "true" to "false," increasing RT again. The final comparison stage involves the outer or "embedding" string of the sentence code. These embedding strings are only present with negative sentences. The lack of an embedding string in the picture code will result in a mismatch, a change in truth index value, and an increase in RT.

The entire processing model is diagrammed in Figure 1, with the specific factors involved in each condition listed in Table 1. In the simplest situation, where an unmarked term is used in the sentence, and there are no mismatches in the comparison stages, the total RT is represented by t, the baseline time. This time represents an estimate of the amount of time needed to complete all processing stages when no additional factors affect those stages. When the marked item is used, sentence encoding is increased

Sentence Type	STAGE I: Encode Picture	STAGE II: Encode Relat. Term → Marking / Encode Embedding Str. → Negation		STAGE III: Compare Subjects → Mismatch	STAGE III: Compare Preps → Falsification	STAGE III: Compare Embedding Str. → Negation	STAGE IV: Output Response
ATU	(A above B)	(A above B)					T
ATM	(A above B)	*(B below A)	→	°(A above B)			T
AFU	(A above B)	(B above A)	→	°(A below B)	¤T → F		T
AFM	(A above B)	*(A below B)			¤T → F		T
NTU	(A above B)	((B above A) false)+	→	°((A below B) false)	¤T → F	•F → T	T
NTM	(A above B)	*((A below B) false)+			¤T → F	•F → T	T
NFU	(A above B)	((A above B) false)+				•T → F	T
NFM	(A above B)	*((B below A) false)+	→	°((A above B) false)		•T → F	T

* add "a"
\+ add "b"
° add "e"
¤ add "f"
• add "d"

Figure 1.--Model of Picture-Sentence Matching.

TABLE 1.--Latency Components.

Sentence Type			Sentence	Latency Components
Affirmative	true	above	A is above B.	t
		below	B is below A.	t + a + e
	false	above	B is above A.	t + e + f
		below	A is below B.	t + a + f
Negative	true	above	B is not above A.	t + (b+d) + e + f
		below	A is not below B.	t + (b+d) + a + f
	false	above	A is not above B.	t + (b+d)
		below	B is not below A.	t + (b+d) + a + e

Source: Clark and Chase Model.

Note: In the present design, the adjectives high-low and the verbs rising-falling will be substituted in these sentence types.

by a, the lexical marking effect. When a negative sentence is used, sentence encoding time is increased by b, and comparison time by d. Negation as a unitary variable is represented by b+d, since both factors occur in the processing of all negative sentences.[4] The presence of a subject mismatch, and the transformation of the sentence code to obtain congruence, is represented by e. A mismatch of the

[4]The ability of the additive factors model to represent negation as a unitary variable, despite its occurrence in different processing stages, illustrates another aspect of its superiority over stage models.

relational terms in the embedded strings, yielding a change in truth index value, is *f*, falsification time. The present study found the fit of this model for data on *above-below* to be quite good, but found certain consistent deviations from model predictions with the adjectives *high-low* and the verbs *rising-falling*.

METHOD

A modified version of the Clark and Chase paradigm (1972) was used. The picture and sentence slides were presented in succession, to isolate the time for sentence encoding and comparing the codes.

Subjects were taken from the Michigan State University introductory psychology classes. A total of forty-six subjects were employed. Each received two credits for his/her participation. Instructions explained the purpose of the experiment as an attempt to study the effect of using different words on people's ability to match verbal and visual descriptions. It was emphasized that the study was concerned both with the speed of their decisions and the accuracy of those decisions.

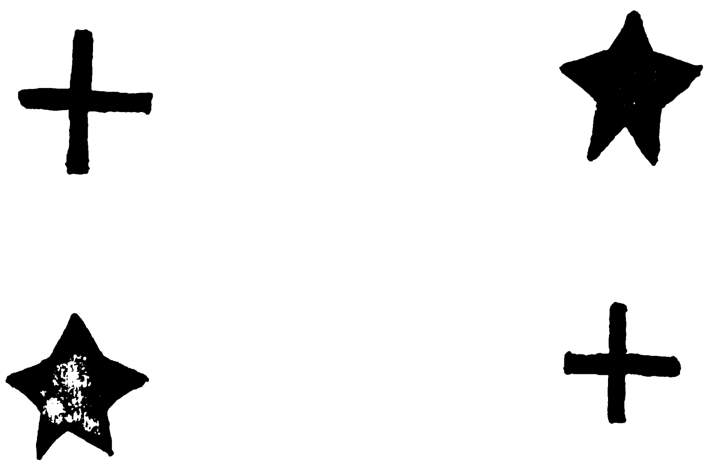
The stimuli consisted of twelve sentences and two pictures, listed in Table 2. The twelve test phrases were introduced to the subjects prior to beginning the experiment, with visual displays showing "true" and "false" matches. Subjects were further informed that each test phrase would appear an equal number of times in the "true" and "false" conditions. All sentences were randomized within five slide trays, and the trays themselves presented

TABLE 2.--Stimuli.

Sentences:

	is (not) above	
	is (not) below	
The star/ The plus	is (not) higher than	the plus/ the star.
	is (not) lower than	
	is (not) rising away from	
	is (not falling away from	

Pictures:



in random order. Each of the twelve test phrases were presented four times in the "true" condition and four times in the "false" condition, for a total of ninety-six responses per subject. The four presentations in each condition were two of the form "The star.......the plus" and two of the form "The plus.......the star." A tray of practice slides was used to familiarize the subjects with the equipment. Feedback on correct answers and RTs were given for the practice slides only.

The subject began a trial by pressing a button that released the picture slide. This slide was available for study for eight seconds. After a 1.5 second interval, during which a blank slide was presented, the sentence slide was released automatically. Exposure of this slide activated a timer, which counted RT in milli-seconds until a response was made. Executing a response terminated the presentation of the sentence slide, releasing another blank slide. Responses were made by depressing one of two telegraph keys, labelled "true" and "false." Half of the subjects had the "true" key on the left, half had the "true" key on the right. The experimenter then recorded RT in milliseconds, and the answer as correct or incorrect. When the experimenter indicated that the answer was recorded, the subject was free to begin the next pair whenever ready. After each slide tray, there was a brief rest period (2-3 minutes), after which testing was resumed. Testing generally lasted forty minutes. After completing all slide trays, subjects were asked to answer certain questions about the strategies they had used during the experiment on a mimeographed sheet. The experimenter then answered any questions, and the subject was dismissed.

RESULTS

Mean RT was used to determine the parameter estimates and the fit of the model. Four RT scores were averaged to obtain a mean RT value for each subject in each of the eight conditions for each grammatical class: two of the form "The star.......the plus," two of the form "The plus.......the star." This yielded an eight by forty-six matrix for each grammatical class. (The eight conditions will be abbrieviated as follows: affirmative true unmarked -ATU, affirmative false unmarked-AFU, negative true unmarked -NTU, negative false unmarked-NFU, affirmative true marked -ATM, affirmative false marked-AFM, negative true marked -NTM, negative false marked-NFM.)

The estimation of parameter values involved a least squares analysis. This necessitated first the collapsing of these matrices by averaging over the forty-six subjects to obtain a mean observed RT for each condition. The observed mean RTs were then categorized according to the specific parameters characterizing each one. (This breakdown has already been presented in Table 1.) The principle of least squares analysis is to obtain a set of simultaneous equations, one equation for each unknown value, that can be solved to yield parameter estimates. These estimates, when summed together to obtain a predicted RT for each condition, should yield a value that is as close as possible to the obtained value, within the confines of the model. This minimizes the error of estimation, or the deviation of the obtained data from the predictions of the model.

The obtained parameter estimates are presented in Table 3, along with the observed and predicted mean RT for each condition. In all cases, b+d (negation) was the largest parameter, followed by a (lexical marking). This pattern conforms to the one obtained by Clark and Chase (1972). The root mean squared deviation (RMSD), a measure of the deviation of the obtained data from the predictions of the model, varied according to grammatical class.

An analysis of variance was performed on the mean data. The results are presented in Table 4. Although this analysis does not directly test the significance of the parameters involved in picture-sentence matching, it allows for inferences about the size of these parameters. Parameter a, lexical marking, is reflected by the main effect of marking, significant in all classes, $p < .001$. Negation represents parameter b+d, also significant in all classes, $p < .001$.

Parameters e and f are best reflected by interactive comparisons in the analysis of variance. Falsification time,

TABLE 3.--Parameter Estimates.

Parameter	Clark and Chase	above-below	high-low	rising-falling
t	1793	1394	1380	1694
a	128	138	240	134
b+d	504	428	473	387
e	212	115	134	141
f	91	118	30	79

Parameters	Clark and Chase	above-below	high-low	rising-falling
t	obs: 1783	1395	1338	1677
	pre: 1798	1394	1380	1694
t + a + e	obs: 2139	1648	1710	1935
	pre: 2122	1647	1754	1969
t + e + f	obs: 2130	1640	1627	1981
	pre: 2103	1627	1544	1914
t + a + f	obs: 2077	1637	1654	1891
	pre: 2008	1650	1650	1907
t + (b+d) + e + f	obs: 2614	2043	1935	2234
	pre: 2659	2055	2017	2301
t + (b+d) + a + f	obs: 2499	2092	2120	2310
	pre: 2564	2078	2123	2294
t + (b+d)	obs: 2354	1821	1895	2098
	pre: 2349	1822	1853	2081
t + (b+d) + a + e	obs: 2678	2075	2272	2391
	pre: 2678	2075	2227	2356
RMSD		9	51	39

TABLE 4.--Analysis of Variance.

Source	df	high-low	above-below	rising-falling
I By Grammatical Class				
True-False	1,45	8.22 +	.01	3.14
Marked-Unmarked	1,45	72.54 *	28.35 *	16.18 *
Negative-Affirm	1,45	184.30 *	105.51 *	100.63 *
T x N	1,45	1.39 #	19.19 *	11.90 *
T x M	1,45	2.57	.55	.83
M x N	1,45	2.90	.46	5.44 +
T x M x N	1,45	17.17 *	7.24 *	13.89 *

Source	df	combined F
II Combined Classes		
True-False	1,45	6.23 +
Marked-Unmarked	1,45	73.32 *
Negative-Affirm	1,45	208.88 *
Grammatical Class	2,90	71.41 *
T x N	1,45	21.94 *
M x N	1,45	5.63 +
M x G	2,90	5.44 *
T x M x N	1,45	19.74 *

Key: * = significant effect, $p < .05$.

+ = significant effect, $p < .05$, violates model predictions

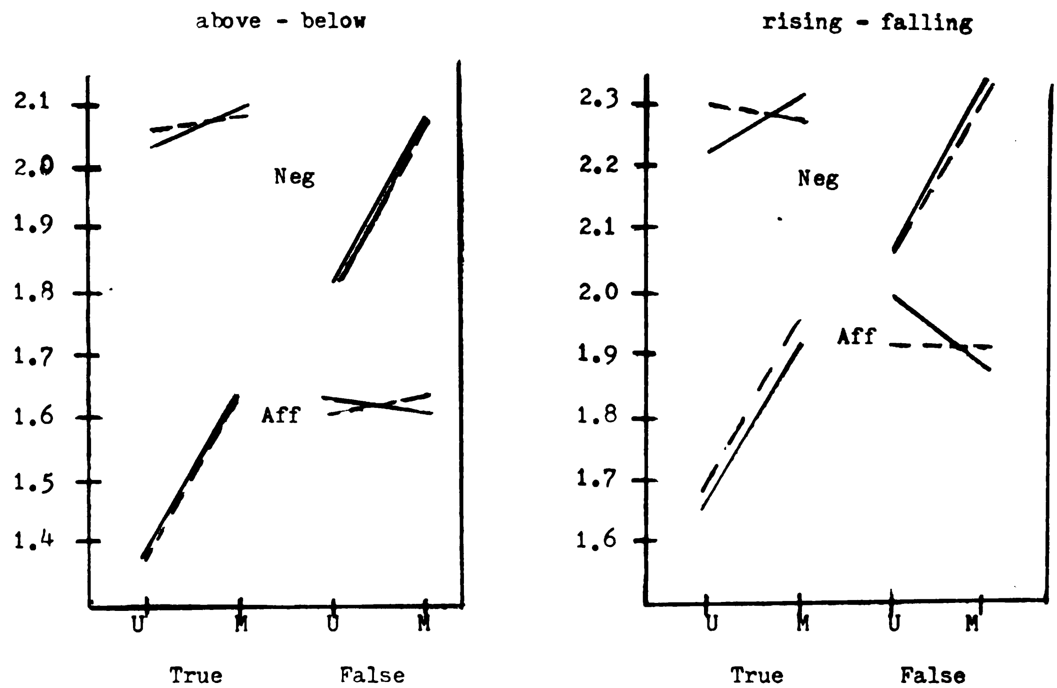
= non-significant effect, violates model predictions.

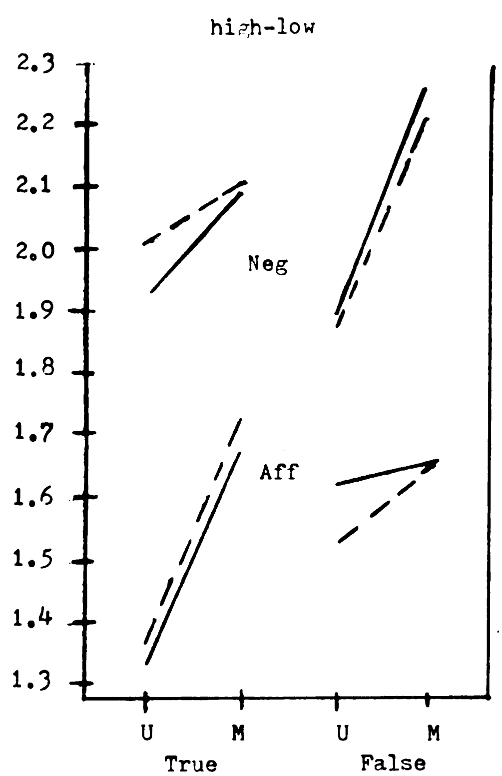
f, is characteristic of affirmative false and negative true sentences. It does not occur in affirmative true and negative false sentences. It depends, then, on the level of two variables: true-false and negative-affirmative. The T x N interaction was significant for both *above-below* and *rising-falling*, but did not reach significance for *high-low*. This locates one source of deviation from the model's prediction for this adjective pair. In addition, it provides an excellent example of one of the principles of additive factors models. When the RMSD of a particular model exceeds the value of one factor in that model, that factor value must be regarded as untenable (Sternberg 1969, 1971). In the case of *high-low*, the RMSD is 51 msec, while parameter *f* is 30 msec.

Parameter *e*, subject mismatch time, occurs in ATM sentences, AFU sentences, NTU sentences, and NFM sentences. It does not occur in the complementary forms: AFM, ATU, NFU, and NTM. It is dependent on the levels of three variables: marking, true-false, and negative-affirmative. It can best be examined by the T x M x N interaction, significant in all classes, $p < .05$. Graphs of these interactions are presented in Figure 2.

The analysis of variance located certain deviations from the model's predictions for *high-low* and *rising-falling*. High-low contained two deviations, one being the failure of parameter *f* to reach significance. The other

Figure 2: T x M x N Interactions





——— observed ----- predicted

Sentence Type	Parameters
ATU	t
ATM	t + a + e
AFU	t + e + f
AFM	t + a + f
NTU	t + (b+d) + e + f
NTM	t + (b+d) + a + f
NFU	t + (b+d)
NFM	t + (b+d) + a + e

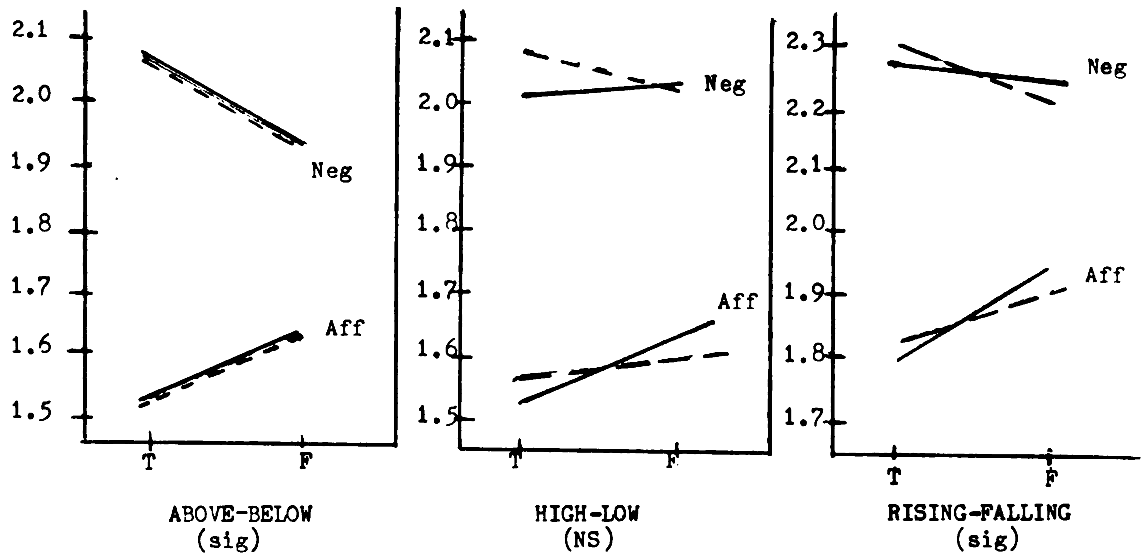
was a significant main effect for true-false, $p < .05$. False RTs averaged 86 msec longer than true RTs. This effect is in violation of the model's predictions since the identical number of each parameter is used in the total true and false conditions (averaged over marked-unmarked and affirmative-negative). Rising-falling contained one deviation from the predictions of the model. A marking by negation (M x N) interaction was significant, $p < .05$. The T x N and M x N interactions in each grammatical class are graphed in Figure 3.

An analysis of variance was performed on the combined data to determine the role of grammatical class in picture-sentence matching. All predicted main effects and interactions (i.e. those involved in parameter tests) reached significance. The main effect of true-false reached significance, $p < .05$, due to its appearance in the high-low data. The M x N interaction also reached significance, due to its appearance in the data on rising-falling. (The results of this analysis are presented in Table 4, along with the analysis by class.)

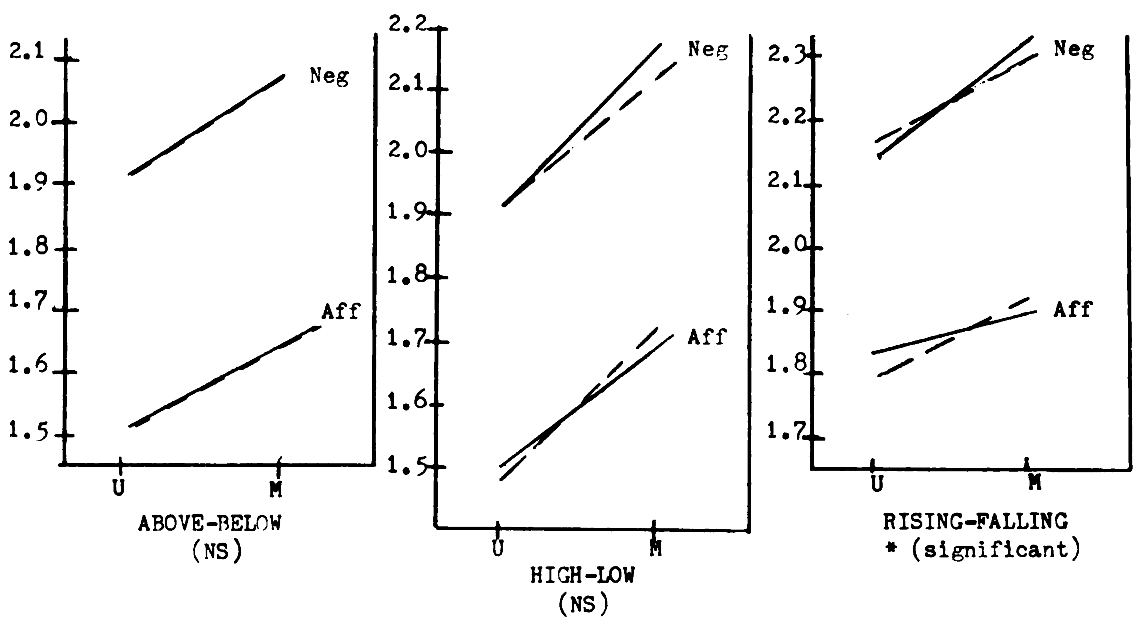
Two effects related to grammatical class were significant. First, the main effect of grammatical class was significant, $p < .001$. This illustrates that the mean RT values varied significantly between the three word pairs, when averaged across subjects. In addition, it provides an indication that the fit of the model could vary

Figure 3: T x N and M x N Interactions

T x N Interactions



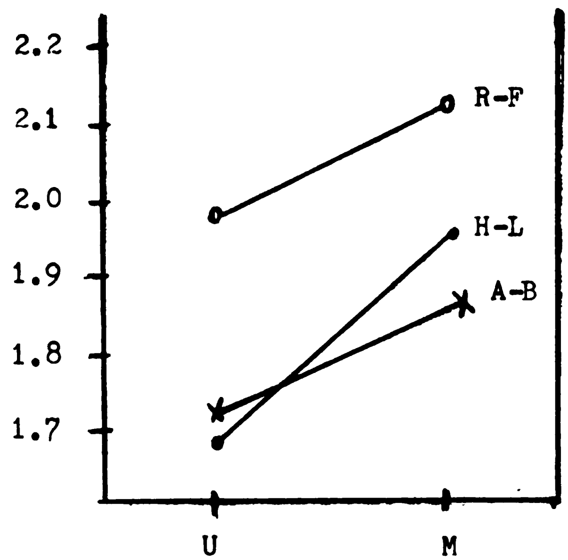
M x N Interactions



significantly according to grammatical class. Second, the M x G interaction was significant, $p < .02$. This effect shows that the difference in size of parameter a (lexical marking) as a function of grammatical class is significant. Consulting the table of parameter values, this effect is due to the large difference between high-low (a = 240 msec), and the almost equivalent size of parameter a is above-below and rising-falling (138 and 134 msec respectively). This interaction is graphed in Figure 4.

Error rates were generally low throughout the study. The average error rate was 3.87%. The breakdown of error data is presented in Table 5, along with the results of an analysis of variance performed on the error data. In both a by class and combined analysis of variance, the only significant effect was negation, $p < .001$. This finding parallels earlier findings about the increased complexity of negative sentences (Clark and Chase 1972; Trabasso, Rollins and Shaughnessy 197). The errors did not follow any systematic pattern, indicating that no speed/accuracy tradeoff occurred.

Figure 4: M x G Interaction



	High-Low	Above-Below	Rising-Falling
$(M - U)_{obs}$	.240	.138	.134
$(M - U)_{pre}$	.240	.138	.134
	●——●	✕——✕	o——●

TABLE 5.--Error Data.

	ATU	AFU	ATM	AFM	NTU	NFU	NTM	NFM		
				I Percentage Data						
high-low	1	2	4	4	9	6	9	6	41	2.78%
above-below	2	3	2	2	12	4	16	13	54	3.66%
rising-falling	8	2	18	3	6	18	12	19	76	5.16%
Total	11	7	14	9	27	28	37	38	171	
Percent	1.9%	1.2%	2.5%	1.6%	4.8%	5%	6.7%	6.8%		3.87%

Source	df	above-below	high-low	rising-falling
		II Analysis of Variance by Grammatical Class		
True-False	1,45	.688	.727	1.585
Marked-Unmarked	1,45	.395	4.07	.557
Negative-Affirm	1,45	5.99 *	20.76 *	14.36 *

Note: All interactions were non-significant.

Source	df	combined F ratio
		III Analysis of Variance on Combined Classes
True-False	1,45	.552
Marked-Unmarked	1,45	3.788
Negative-Affirm	1,45	49.106 *
Grammatical Class	2,90	.099

Note: All interactions were non-significant

DISCUSSION

This study attempted to integrate research in two areas: the nature of intra-pair differences, and the processes involved in picture-sentence matching. The rationale for this approach was to determine whether or not the model for picture-sentence matching developed by Clark and Chase (1972) for the terms _above-below_ could account for analogous processing of other pairs of terms referring to the same perceptual dimension. If words used to refer to perceptual dimensions reflect the characteristics of those dimensions, the factors influencing the processing of those terms should function in a similar manner. This is apparently not the case.

The fit of the Clark and Chase model for the data on _above-below_ was excellent. The largest deviations from the model's predictions occurred in the negative true unmarked (NTU) and affirmative false unmarked (AFU): 13 msec in each (non-significant). All parameters reached significance, and there were no unpredicted effects. This is strong evidence in favor of using such a model to describe sentence processing in this task.

However, when grammatical class was introduced as a variable, differences between the three pairs of terms were evident. These differences took several forms. First, the size of the intra-pair difference varied according to grammatical class. The adjectives high-low showed a significantly larger marking effect than the prepositions above-below or the verbs rising-falling. The predicted difference in size of marking effect between the verbs and the other two classes did not materialize. Instead, the size of the marking effect in the verbs studied varied significantly according to the nature of the sentence. The significant M x N interaction reflected a difference in the size of the marking effect when the sentence was affirmative or negative. The marked-unmarked difference between rising-falling was 100 msec longer when the sentences were negative. Despite the fact that the size of parameter a for this pair was 4 msec less than in the pair above-below, the marking effect in these verbs functions differently.

To what can we attribute the differences in lexical marking effect amongst these pairs? In the case of the adjectives high-low, the larger marking effect is a direct reflection of a linguistic distinction characteristic of the class of adjective pairs. This distinction has been explained previously as the presence of an additional feature in the representation of the marked adjective (H. Clark 1969a). The representation of the term "low" must always include information about the dimension being used and the polarity of the object in question on that dimension (contrastive use). The adjective "high" need only include the feature of dimensionality in its representation (nominal use). This difference affects the ease of code construction and interpretation with adjectives. The necessity of including two features in the representation of "low" increases significantly the amount of time needed to construct that code. The absence of this distinction in the other word pairs studied leads to a smaller lexical marking effect in those pairs. The estimated size of parameter _a_ for _high-low_ (240 msec) is close to the original marked-unmarked difference reported by Clark (1969a) for the dimensional adjectives _good-bad_: 170 msec in affirmative problems, 340 msec in negative problems, yielding an average marked-unmarked difference of 255 msec.

This raises an interesting question of semantics. Are the adjectives _high-low_ part of the semantic field of dimensional terms, or spatio-temporal terms? They are commonly used to refer to perceptual space. Subjects did not have trouble applying them to a perceptual dimension. (The predicted baseline RT for the adjectives _high-low_ was actually shorter than the baseline time for the prepositions _above-below_.) _Above-below_ have consistently been classified as spatio-temporal terms. However, E. Clark (1972) classified _high-low_ as dimensional terms in her study of the acquisition of lexically marked word pairs in two semantic fields. The fact that high-low are usually applied to perceptual space seems to make their inclusion in the class of spatio-temporal terms relevant. In light of this, the observed parallelism in size of lexical marking effect follows grammatical class boundaries more closely than semantic field boundaries. It is obvious that certain linguistic considerations enter into the expression of perceptual asymmetry in these terms.

In the case of the verbs rising-falling, the marking effect is under the control of other variables. The presence of sentence negation significantly increases the size of the effect. The significance of the M x N interaction weakens the assumption of additivity of parameters a and b+d, since these factors must be influencing a common processing stage (Sternberg 1969, 1971). Perhaps the size of the marking effect is dependent on negation because of certain linguistic properties of verbs. Consider the sentence "The star is not rising away from the plus." There are several interpretations of his sentence: the star can be stationery, the star can be rising towards the plus[5]

---

[5]The difference in reference points indicated by the phrases "rising towards" and "rising away from" is interesting. While the direction is consistent (upwards, an "unmarked" direction), one employs a reference point at the bottom of the vertical dimension, the other a point at the top. This reversal of reference point could have an unusual effect on lexical marking.

(meaning that it would be located under the plus), or the star could be falling away from the plus (again meaning that it would be located under the plus). Since subjects were cued to interpret "not rising" as "falling," they were able to correctly match the sentences and pictures. However, in normal language use, "not rising" is not necessarily the equivalent of "falling," and vice versa. Verbs, being action terms, do not function the same way as other grammatical classes when negated. The presence of a larger intra-pair difference when these words are negated indicates that certain linguistic characteristics of verbs could influence the expression of perceptual asymmetry. Yet again we face the problem of semantics. Are the verbs <u>rising-falling</u> spatial terms? They always refer to relationships on the asymmetrical vertical dimension, making it feasible to include them in the class of spatial terms. But they are not often used in describing relationships on this dimension unless movement is indicated.

The difference in lexical marking effects amongst these word pairs implicate several additional research studies to clarify the situation. A comparison of adjective pairs would yield additional information about the role of grammatical considerations in lexical marking. If possible, such an analysis should include adjectives referring to the asymmetrical vertical dimension, the asymmetrical horizontal dimension, and those referring to abstract conceptual dimensions. According to the findings of this study, regardless of the dimension to which these word pairs refer, the size of the marking effect should be consistently larger in adjectives than in word pairs of other grammatical classes. A comparison of various verb pairs is also warranted. It has not been conclusively determined that the fluctuation in size of marking effect in the verbs pairs studied is due to characteristics of the class of verbs. It is possible that this effect was the result of the experimental procedure (using verbs to describe a static situation) or the stimulus selection procedure (the fluctuation in marking being a characteristic of only this verb pair). By comparing other verb pairs referring to the asymmetrical vertical and horizontal dimensions (e.g. push-pull, lead-trail), a more accurate characterization of the marking effect in verbs can be obtained. Such a comparison should involve applying verbs to movement situations (e.g. using streamers to indicate that the objects in question are moving) to control for the difficulty in applying verbs to static situations.

The second expression of the importance of grammatical class occurred in the deviations from the model predictions for high-low and rising-falling. For the adjectives studied, these deviations related to the failure of parameter *f* to reach significance, and the significant difference between true and false RTs (false RTs averaged 86 msec longer than true RTs). The deviations for the verbs *rising-falling* related to the significant M x N interaction. However, if we move to examine how this deviations affect the fit of the Clark and Chase model, a remarkable parallelism exists. As stated earlier, the largest deviations for the *above-below* data were in the NTU and AFU conditions (13 msec, non-significant). Referring to the graph in Figure 2, the largest deviations in the other two word pairs occur in the same conditions. In this graph, each data point corresponds to a predicted or observed mean RT, with all means accounted for. The large deviations in the adjectives and verbs are: for *high-low*, 82 and 83 msec for NTU and AFU respectively; for *rising-falling*, 67 msec in both NTU and AFU conditions. Even allowing for the 13 msec deviations in the above-below "ideal cases," these deviations are still quite large. It appears that the model is insensitive to a parameter of importance in both of these cases.

The question becomes, then: is this parameter an artifact of the experimental procedure, or indicative of processing differences amongst these three word pairs? The sentences were randomized for presentation. Subjects had no way of predicting which stimulus word would occur after any given picture slide. According to subjects' own reports, there was a strong preference for coding the picture in a linguistic form using an unmarked term (59%). Of these subjects, 27% reported using "above" in their picture code, 22% reported using "on top of," and 10% used a variety of other unmarked terms. Since the code constructed for the sentence is determined by the stimulus word used in that sentence, it is possible that some form of "translation strategy" is needed before the codes can be compared for congruence. If "above" were primarily used in the picture code, this additional process would only be required for the adjectives and verbs in the study. If "on top of" were used, although a translation mechanism would be needed in all cases, it is possible that there is a more obvious correspondence between "above" and "on top of" than in the case of "high" or "rising." This would decrease the value of that parameter in the <u>above-below</u> sentences. It is possible that something akin to a translation process is responsible for the poorer model fit in the other two word pairs. If so, this is an artifact of the experimental procedure. This question can be tested directly by isolating the three word pairs for presentation, and performing the identical experiment.

If the model fit does not improve in such a study, the translation strategy is no longer a viable explanation for the poorer model fit. The difference between the three word pairs could then be related to qualitative differences amongst them. Perhaps such differences necessitate a reformulation of the comprehension processes in picture-sentence matching. The failure of parameter *f* to reach significance, the significant main effect for true-false, and the presence of an M x N interaction all weaken the assumption that identical processing stages are used in the comprehension of these three pairs of terms.

This question also implicates several additional research projects. First, a comparison of three other pairs of terms referring to a different perceptual dimension is needed to determine whether or not these findings are replicable with other word pairs. The stimulus words should be isolated to prevent the interference of a translation process in any of the conditions. If the poorer model fit for adjectives and verbs persists, a new series of processing stages are needed to describe picture-sentence matching in these two classes. This reformulation of comprehension processes could only necessitate a reordering of the stages in comprehension, so that different parameters would characterize each sentence type. Conversely, this alteration could involve the combining of existing stages (e.g., negation and marking occurring in the same stage for verbs, and represented by one parameter), or adding new stages (e.g. separating the transformation of the sentence code when comparing for congruence from the change in truth index value for adjectives, represented by two parameters).

Second, the reasons for poorer model fit with adjectives and verbs could be examined developmentally. Previous research has shown that spatio-temporal terms are acquired earlier than abstract dimensional terms (E. Clark 1971b). It is believed that the direct correspondence of spatial terms to perceptual space makes spatial terms easier for the child to understand. Temporal terms are similar to spatial terms because temporal concepts have been derived through analogy to spatial concepts (H. Clark 1973). However, certain mechanisms in language acquisition operate in the learning of word pairs in both semantic fields. Children first acquire global characteristics of antonymous word pairs (e.g. the dimension to which they refer) and later acquire additional features to differentiate the pair members from each other and from similar pairs. E. Clark (1971a) found that when acquiring those first global characteristics, children use and understand unmarked words first. The features acquired later are those needed to comprehend the marked term. This often results in overextension of unmarked terms, since young children do not understand the polarity feature (e. g. using the unmarked term "big" as the opposite of "small," "little," "short," "young," etc.). This is often followed by a stage in which children use marked and unmarked words interchangeably (e.g. using "more" and "less" as meaning the same thing--"different"). When both dimensionality and polarity are understood, children exhibit a lexical marking effect consistent with the form shown by adults. Recent research has uncovered a series of non-linguistic strategies which result in identical performance to the marking effect (E. Clark 1974). The situation has thus become more complicated, since it is difficult to determine whether linguistic or non-linguistic strategies are responsible for the intra-pair differences observed in children.

A better understanding of the role of grammatical class in determining intra-pair differences in children could help clarify the acquisition of lexically marked word pairs. The three original word pairs could be compared in a placement task, where children are required to manipulate objects to construct displays of sentences they hear, and error rate is used as the dependent measure. An increase in the number of correct responses is expected with age. More importantly, an increase in the ratio of marked to unmarked word errors should occur as children acquire the features needed to differentiate these words. Since all three pairs refer to observable perceptual space, children should be able to understand these sentences fairly early (3-4 years--cf: E. Clark 1971a). However, differences among the error rates to the three word pairs could indicate how directly these words refer to perceptual space. In addition, the pair with the larger lexical marking effect (high-low) should show a greater ratio of marked-unmarked word errors, since the marking effect is much stronger within this pair. Since the prepositions above-below seem to be least affected by additional linguistic considerations, comprehension of this pair could be superior at earlier ages. In such a study, an examination of possible alternative non-linguistic strategies is important, to insure that the results are reflective of linguistic processes.

It is further possible to perform a reaction time experiment with children, to make as many generalizations to the present research as possible. Additional information about the characteristics of reaction time studies with children is needed before such a design is formulated. In either an error rate or reaction time study, the use of terms referring to another perceptual dimension (e.g. the asymmetrical horizontal dimension) would strengthen any conclusions about the modification of perceptual characteristics by linguistic considerations.

REFERENCES

REFERENCES

Clark, E. What's in a word? On the child's acquisition of semantic in his first language. Conference on Developmental Psychology, 1971 (a).

Clark, E. On the acquisition of the meaning of before and after, JVLVB, 1971, 10, 266-275, (b).

Clark, E. On the child's acquisition of antonyms in two semantic fields, JVLVB, 1972.

Clark, E. Non-linguistic strategies and the acquisition of word meanings, Cognition, 1974, 2, 161-180.

Clark, H. Linguistic processes in deductive reasoning, Psyc Rev, 1969, 76, 387-404, (a).

Clark, H. The influence of language in solving three term series problems, JEP, 1969, 82, 205-215 (b).

Clark, H. Difficulties people have in answering the question "Where is it?", JVLVB, 1972, 12, 256-277.

Clark, H. Space, time, semantics and the child, Presented at the Conference on Developmental Psycholinguistics, SUNY at Buffalo, 1971.

Clark and Chase, On the process of comparing sentences to pictures, Cog Psyc, 1972, 2, 513-521.

Olson, G. and Laxar, K. Asymmetries in processing the terms "right" and "left," JEP, 1973, 100, 284-290.

Sternberg, S. The discovery of processing stages: extensions of Donder's method, in W. G. Koster (Ed) Attention and Performance II, 1969, 276-315.

Sternberg, S. Decomposing mental processes with reaction time data, Presented at the Annual Meeting of the Midwestern Psychological Association, Detroit, 1971.

Trabasso, T. Mental operations in language comprehension, in J. B. Carroll and R. O. Freddle (Eds) Language Comprehension and the Acquisition of Knowledge, Washington: V. H. Winston & Sons, 1974.

Trabasso, Rollins and Shaugnessy, Storage and verification stages in processing concepts, Cog Psyc, 1971, 2, 239-289.