AN EVALUATION OF THE APPLICABILITY OF ADDITIVE FACTOR MODELS TO SENTENCE-SENTENCE VERIFICATION TASKS

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY

ELIZABETH ANNE MAIER

1977

ABSTRACT

AN EVALUATION OF THE APPLICABILITY OF ADDITIVE FACTOR MODELS

TO SENTENCE-SENTENCE VERIFICATION TASKS

By

Elizabeth Anne Maier

Three experiments were conducted investigating rules of processing in sentence-sentence verification tasks. In each experiment, the content of the sentences referred either to the midpoint of the dimension of time, that is, On Time, or the endpoint, Late.

Experiment 1 examined subjects' verification of both explicitly conveyed information (Assertions) and implicit information (Inferences) over three testing sessions using a within-subjects design. Subjects heard a set of two sentences. The task required them to indicate whether the two sentences conveyed the same meaning responding as quickly and as accurately as possible.

In general, Inference conditions, conditions containing an On Time or a negative in Sentence 2, or in which an inner string mismatch occurred, were most difficult for the subjects to process. A special analysis suggested that the effects of On Time and the Inference condition were restricted to the encoding phase of processing, whereas the effects of negation were seen in both encoding and comparison operations. No evidence was obtained to support the idea that in more complex situations subjects adopt conversion operations in processing.

Experiment 2 investigated subjects' verification of explicitly relayed information (Assertions) in the auditory modality. As in Experiment 1, conditions in which an inner string mismatch occurred, or in which an On $Time_2$ or $Negative_2$ appeared were difficult for the subjects to process. The role of Sentence 1 was more important in this experiment, as indicated by the Term 1 main effect and the Term 1 x Negation 1 x Term 2 interaction. The latter result suggested that when a negative occurred in Store 1, the subjects encoded both affirmative and negative representations of the sentence. Which representation was chosen in comparing Sentence 1 and Sentence 2 was dependent upon what had been encoded in Sentence 2.

Experiment 3 was designed to determine whether differences in processing occur between the auditory and visual modalities. As in Experiment 2, only the verification of Assertions was examined. On $Time_2$, $Negative_2$, and inner string mismatch conditions were more time consuming for the subjects to process than their counterparts. The role of Sentence 1 was more pronounced in this study in that both the Term 1 and Negation 1 main effects were significant. RT orderings suggested that after the subjects encoded Sentences 1 and 2, they went back and re-encoded Sentence 1 before commencing comparison operations.

The results of all three experiments were interpreted in terms of additive factor models. The three models constructed adequately represented the obtained results in each case. Although some problems arose within the encoding stages of processing in the model of Experiment 2, it was the conclusion of the study that additive factor models which make distinctions between the encoding and comparison stages of processing are appropriate frameworks in which to interpret the results of sentence-sentence verification tasks.

AN EVALUATION OF THE APPLICABILITY OF ADDITIVE FACTOR MODELS TO SENTENCE-SENTENCE VERIFICATION TASKS

By

Elizabeth Anne Maier

A THESIS

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

MASTER OF ARTS

Department of Psychology

1977

To

C. M.

ACKNOWLEDGMENTS

I would like to thank Rose Zacks for her concern, guidance and assistance; and Lauren Harris, Gordon Wood, and Gary Olson for their thoughtful comments on this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Early investigations of negation (Gough, 1965, 1966; Slobin, 1966; Wason, 1959, 1961; Wason & Jones, 1963) uniformly reported that negative sentences took longer to comprehend than affirmative sentences whether the sentences were true or not. More recent investigations of the comprehension of negation have had as their aim the discovery of specific underlying components of true-false verification processes, components which would adequately represent and accurately predict information processing (Carpenter & Just, 1975; Chase & Clark, 1972; Clark, 1970, 1974; Clark & Chase, 1972; Trabasso, Rollins & Shaughnessy, 1971). In such experiments the tasks are generally quite simple: subjects are shown sentence-picture stimuli simultaneously, and are instructed to specify by pressing a true or false button as quickly as possible whether the illustration correctly represents the sentence. From such a task, Clark and Chase (1972) have constructed a four-stage model of negation characterizing sentence-picture comprehension based upon Sternberg's (1969a,b) additive factor approach to mental operations.

The four-stage sentence-first model assumes that during Stage 1 the subject will encode sentences in a linguistic representation which preserves all meaningful aspects of the original sentence. For example, sentence (1) is presumed to be encoded as representation (2). The

(1) The star is above the plus.

(2) (star above plus)

internal representation of an affirmative is assumed to involve a constant amount of processing time. Its negative counterpart, sentence (3), would be represented as (4). The actual representation of the negative

(3) The star is not above the plus.

(4) false (star above plus)

as a falsehood indicator is adopted from Wason's (1965) proposition that the negative naturally functions to deny a proposition, as in (5). Based upon previous studies of negation comprehension diffi-

(5) It is false that the star is above the plus.

culties (Gough, 1965, 1966; Slobin, 1966; Wason, 1959, 1961), Clark and Chase (1972) posited that the encoding of a negative would be a more difficult operation that the encoding of an affirmative, requiring an additional amount of processing time, time <u>b</u>.

Just as time measurement is sensitive to variations in encoding time between affirmatives and full explicit negatives (Clark, 1970, 1974), Clark and Chase (1972) further hypothesized that less explicit forms of negation (Clark, 1969, 1970, 1974) also would increase encoding time. In the Clark and Chase (1972) paradigm, <u>below</u> was considered to be less neutral than its contrary <u>above</u>. The encoding of sentence (6) as (7) was predicted to exceed that of sentence (1) as (2) by the

(6) The star is below the plus.

(7) (star below plus)

constant <u>a</u> amount of time. Similarly, since the model proposed by Clark and Chase (1972) was an additive factor model, the encoding of sentence (8) as (9) was expected to exceed (1) encoding by (<u>a</u> + <u>b</u>) time.

(8) The star is not below the plus.

(9) false (star below plus)

During Stage 2, the subject focuses his attention on the picture stimulus, presumably encoding its meaning in a representation like that used in Stage 1 sentence encoding. Clark and Chase's (1972) principle of congruence asserts that for any comparison of the meanings of sentence and picture to be made, sentence and picture must be encoded in similar formats. A final assumption underlying the second stage of processing is that the encoded picture representation employs the same proposition appearing in the previous sentence. Thus, if sentence (8) appeared with picture stimulus (10), (10) would be encoded as (11);

(10) *
+

(11) (plus below star)

while the same picture (10) would be encoded as (12) if paired with

(12) (star above plus)

sentence (6). An interesting feature of Stage 2 is that a picture is *never* assumed to be encoded in a negative format. Thus, the possibility of two explicit full negatives co-existing in the same comparison operation is consistently averted.

Finally, the Clark and Chase model makes no prediction of differential time encoding of pictures. It assumes that picture encoding will take a constant amount of time regardless of the proposition; picture encoding time was a component of the base processing time, $\underline{t_o}$.

The major processing to be accomplished during the verification task occurs during Stage 3, the comparison stage. Assuming common representations of picture and sentence encodings, the model proposes that the processor will first examine the innermost propositional strings seeking identity and subsequently check the identity of the outer strings.

The purpose of Stage 3 is to determine the validity of the sentence picture pair. Hence, for each mismatch the subject will convert the value of the truth index of the pair to its polar opposite. The model assumes that the value of the truth index is set at true at the onset of processing. A single mismatch, of either inner or outer strings, would convert the truth index from true to false, whereas concurrent inner and outer string mismatches would result in two reversals of the truth index value, yielding a true response. Each change in the value is expected to be accompanied by an increase in processing time, times $\underline{c}$ and $\underline{d}$, for inner and outer string mismatches, respectively.

The outcome of Stage 3 is a true or false response input to Stage 4, in which the subject executes the response by pressing the appropriate response button. Clark and Chase did not elaborate further on this stage. It is assumed, however, that response execution time (time taken to mobilize the hand and press the appropriate button) adds a constant amount of time to the base reaction time (RT), $\underline{t}_o$.

Experimental manipulations by Clark and Chase (1972) using above and below provided strong support for the model. Parameters were estimated by means of a least squares analysis to be: Below time, $\underline{a}$, 93 msec., negation time, ($\underline{b}$ + $\underline{d}$), 685 msec., and falsification time, $\underline{c}$, 187 msec. Collectively, the five parameters accounted for 99.8% of the total between conditions variance, and each parameter proved to be highly significant ($\underline{p}$ < .05). No interactions were statistically significant, lending further credence to the independence of stages notion.

Clark (1970, 1974; Clark & Chase, 1972) referred to this conceptualization as the "true" model of negation, being "true" is so far as the linguistic representations of sentence and picture were assumed to preserve all semantic and syntactic features of the actual stimuli. Another model of negation has been proposed (Clark, 1970), the conversion model, accounting for those situations in which the negative is converted to its affirmative equivalent prior to the onset of Stage 3 comparison operations (Trabasso et al., 1971; Wason & Jones, 1963; Young & Chase, 1971). Clark (1970, 1974) felt that the conversion model was subordinate to the true model conceptualization, given that by the conversion of a proposition, some of the information contained in the original representation was forfeited. Furthermore, in order for conversion operations to be executed, either the use of contradictory content words or an experimental design strictly defining binary operations is required. For example, the conversion of sentence (13) to (14) is true

(13) The plus is not above the star.

(14) The plus is below the star.

only of a particular spatial configuration (10). Conversion becomes possible, however, when the subject is informed that only binary situations will be under consideration. Use of a conversion strategy was found (Clark, 1970) to yield a pattern of RT's similar to the sentence construction subjects of Wason (1961): True Affirmatives (TA) < False Affirmatives (FA) < True Negatives (TN) < False Negatives (FN).

Both the true and converstion models of negation have proven to be powerful instruments in the measurement of psychological processes. The former model, using five parameters, accounted for 99.8% of the total between conditions variance (Clark & Chase, 1972); the latter, using four parameters (t_o, *a*, *c*, and *k*, a conversion parameter) was responsible for 95% of the between conditions variance (Clark, 1974). Clark (1974) also successfully interpreted earlier studies of negation (Gough, 1965, 1966; Green, 1970a,b; Slobin, 1966; Trabasso et al., 1971; Wason, 1959, 1961; Wason & Jones, 1963) according to either the true or conversion models.

Thorough examination and evaluation of the true and conversion models of negation reveal serious experimental shortcomings, however. Simplistic tasks have been used to demonstrate principles of understanding which are theoretically applicable to any information processing task. With the exception of Just and Clark's (1972) work with presuppositions and implications, validation of this model in more complex situations is lacking.

Furthermore, the restriction to simplistic paradigms overlooks cases in which negatives occur in both stimulus inputs (i.e., Stage 1, Stage 2, or Stages 1 and 2). Neither model would predict a significant effect on processing in the case of the double negative, but no investigation has tested this premise. Finally, virtually all studies have been investigations of negation using solely the visual modality. Serious inattention to the auditory modality has resulted, although most verbal information is heard rather than seen (e.g., read). Processing limitations may exist within the auditory modality prompting the subject to use a different comprehension strategy from those used for visually presented information.

In an attempt to resolve some of these issues, a study was conducted of the effects of negation (Stage 1, Stage 2, and Stages 1 and 2) on the verification of Assertions and Inferences. Assertions were defined as simple declarative statements about an individual's activity on a particular day; Inferences referred to the individual's usual pattern of behavior, derived from information in the sentence. Sentences were restricted to the dimension of time (On Time vs. Late).

Pilot Study

Subjects. Twenty-five male and female undergraduates participated as a requirement in an introductory course at Michigan State University.

Stimuli. Two hundred fifty-six sentences, recorded by a female speaker, were presented in two sentence test sequences. Sentence 1 constructed a context for an individual's activity, ending with a clause indicating what the particular actor did today, i.e., whether he was On Time or Not On Time. Sentence 2 was a simple statement containing the proposition Late or On Time; it also varied as to whether is was an Assertion or an Inference, and whether it was a negative or an affirmative sentence. Half the sentences represented each factor. The subjects received eight examples of each of the 16 possible conditions (see Table 1.1 for a complete list of the possible sentence combinations).

Apparatus. Stimuli, recorded on a Teac stereo tape recorder, were presented through stereo headphones. Concurrent with the cessation of speech a click (inaudible to the subject) automatically started Hunter times recording subjects's RT's. A three button panel was located directly in front of the subject: a central button for comprehension responses, one button for true responses, the other for false. Depression of any button stopped the Hunter timers and triggered the presentation of the next stimulus.

Procedure. The subjects were tested individually. Each was intructed to listen carefully to the stimulus presentations. Following

TABLE 1.1

Examples of Experimental Conditions Presented in the Pilot Study

---

S_1: Mr. Jones rides the train downtown to work everyday; surprisingly, he was . . .

. . . on time today. . . . not on time today.

S_2: A. Mr. Jones was on time today.

B. Mr. Jones was late today.

C. Mr. Jones was not on time today.

D. Mr. Jones was not late today.

E. Mr. Jones is usually on time.

F. Mr. Jones is usually late.

G. Mr. Jones is not usually on time.

H. Mr. Jones is not usually late.

Sentence 1, they were required to depress a button marked "C" (Comprehension) as quickly as possible indicating that they had understood the sentence. Sentence 2 immediately followed. A button marked "T" (True) or "F" (False) was to be depressed indicating the truth value of the match. The position of the T and F buttons was randomly varied between subjects. Both accuracy and speed were emphasized in decision making. The intertrial interval between each test comparison was five seconds. The subjects received eight practice trials before the experiment began. If more than two errors occurred, the subjects were informed of their inaccuracy in an attempt to improve their performance. At the completion of 128 sentence comparison trials, the subjects were asked to describe what strategies, if any, they had used during the experiment. The experimenter recorded all of the subjects' comprehension and verification RT's.

Results

Mean RT data per treatment for 25 subjects is contained in Table 1.2. Four factors were manipulated in the experiment: Sentence 1 Negation (presence vs. absence), Sentence 2 Negation (presence vs. absence), Sentence 2 Term (Late vs. On Time), and Sentence 2 Statement (Assertion vs. Inference). A 2 x 2 x 2 x 2 within subjects analysis of variance was performed on mean RT data. All main effects as well as all two-way interactions involving the Stage 2 Term factor were significant, where $p < .05$. In general, therefore, comparisons involving Late were more time consuming than On Time verifications; likewise, negatives as well as Inferences involved more processing time than affirmatives and Assertions, respectively. The overall error rate was 8.25%.

Discussion

If the true model of negation were applied to the present paradigm, the RT's (adding only an i parameter for the encoding of the Inference in Stage 2) represented by the parameter components in Table 1.3a would be predicted to occur. Similar extension of the conversion model would produce the parameter groupings appearing in Table 1.3b. Comparison of these predictions with the findings in Table 1.2 shows that neither model can account for the subjects' processing.

While it is difficult to conceptualize all treatments in terms of True and False Affirmatives and True and False Negatives because of the presence of the false double negative condition, we still can examine these features in those conditions most comparable to the original

TABLE 1.2

Mean RT in Msec. Per Condition, Estimated and Observed

Conditions	Projected Parameters	Estimated	Observed
On Time- On Time	t_o	284	310
On Time- Late	t_o+c	653	626
On Time- Not On Time-	t_o+b+d	488	477
On Time- Not Late	t_o+b+c+d	857	900
Not On Time- On Time	t_o+d	428	373
Not On Time- Late	t_o+k	597	652
Not On Time- Not On Time	t_o+b	345	329
Not On Time- Not Late	t_o+b+2k+c	1338	1321
Not Usually On Time- Usually On Time	t_o+i+d	608	585
Not Usually On Time- Usually Late	t_o+i+n	776	678
Not Usually On Time- Not Usually On Time	t_o+i+n	821	796
Not Usually On Time- Not Usually Late	t_o+i+n+k+d	1277	1261
Usually On Time- Usually On Time	t_o+i	464	529
Usually On Time- Usually Late	t_o+i+k+d	920	976
Usually On Time- Not Usually On Time	t_o+i+n+d	965	969
Usually On Time Not Usually Late	t_o+i+n+k	1133	1170

TABLE 1.3

Condition Predictions According to the True and Conversion Models of Negation (Clark, 1970; Clark & Chase, 1972)

Conditions	True (3a)	Conversion (3b)
On Time- On Time	t_o	t_o
On Time- Late	t_o+c	t_o+c
On Time- Not On Time	t_o+b+d	t_o+k+c
On Time- Not Late	$t_o+b+c+d$	t_o+k
Not On Time- On Time	t_o+d	t_o+k+c
Not On Time- Late	t_o+c+d	t_o+k
Not On Time- Not On Time	t_o+b	t_o+2k
Not On Time- Not Late	t_o+b+c	t_o+2k+c
Not Usually On Time- Usually On Time	t_o+i+d	$t_o+i+k+c$
Not Usually On Time- Usually Late	$t_o+i+c+d$	t_o+i+k
Not Usually On Time- Not Usually On Time	t_o+i+b	t_o+i+2k
Not Usually On Time- Not Usually Late	$t_o+i+b+c$	$t_o+i+2k+c$
Usually On Time- Usually On Time	t_o+i	t_o+i
Usually On Time- Usually Late	t_o+i+c	t_o+i+c
Usually On Time- Not Usually On Time	$t_o+i+b+d$	$t_o+i+k+c$
Usually On Time- Not Usually Late	$t_o+i+b+c+d$	t_o+i+k

Clark and Chase (1972) paradigm: On Time_1-On Time_2; On Time_1-Late_2; On Time_1-Not On Time_2; and On Time_1-Not Late_2. Recall that corresponding to the true model of negation, TA < FA < FN < TN, or that On Time_1-On Time_2 < On Time_1-Late_2 < On Time_1-Not On Time_2 < On Time_1-Not Late_2. The conversion model predicts the inversion of False and True Negatives. Examination of Table 1.2 again reveals that neither model accounts for the results found that TA < FN < FA < TN. This pattern of results was also found in the comparable more complex Inference conditions in which Usually On Time_1-Usually On Time_2 < Usually On Time_1-Not Usually On Time_2 < Usually On Time_1-Usually Late_2 < Usually On Time_1-Not Usually Late_2. Furthermore, neither model would predict the false double negative Assertion as the most difficult condition, which, in fact, was the finding.

The inversion of FN and FA was a surprising result. It appears that inner string mismatches were more difficult to resolve than outer string mismatches, contrary to earlier findings (Clark, 1970, 1974; Clark & Chase, 1972). This could result, in part, from the higher order processing required by the task (the verification of Assertions and Inferences) or from the different demands required of the subject in an auditory task.

The interactions yielded by the 2 x 2 x 2 x 2 within-subject analysis of variance were also unexpected. Additive factor models, as outlined by Sternberg (1969a,b), should produce no interactions, as each stage of processing is assumed to be independent of all others. Hence, any model generated using the factors suggested by this design would be an inaccurate reflection of subjects' processing. The findings suggest that an additive factor approach is inappropriate in the present paradigm.

Certain aspects of the data indicated, however, that subjects were operating according to some type of routine. For example, in less complicated conditions, why should a FN treatment consistently be easier than a FA? Of greater interest is the fact that neither model is able to account for the extreme difficulty of the false double negative Assertion condition. Furthermore, no differences in mean RT patterning were found between subjects reporting using some sort of strategy and those not. Whatever subjects were doing, they all seemed to be acting in the same manner.

In an attempt to explain these issues and characterize the results of the pilot study according to a stage theory of information processing, it was hypothesized that the data reflect a mixture of the true and conversion models. A four-stage, seven-parameter additive factor model, abstracting features from both accounts of negation has been projected to account for the subjects' performance (see Table 1.2 for treatment parameters and estimates; Figure 1.1 for a representation of the model itself).

The Model

It is proposed that in Stage 1, Sentence 1 is encoded in some propositional format, similar to that suggested by Clark (1970; Clark & Chase, 1972). Encoding is expected to take a fixed amount of time; however, since this time was recorded separately, and since an underlying premise of any additive factor model is independence of stages, Stage 1 is not expected to interfere with future processing or encoding times.

It became clear to subjects shortly after the onset of testing that they would be required to verify information directly encoded in Stage 1, or an abstraction (Inference). Thus, a further assumption of Stage 1

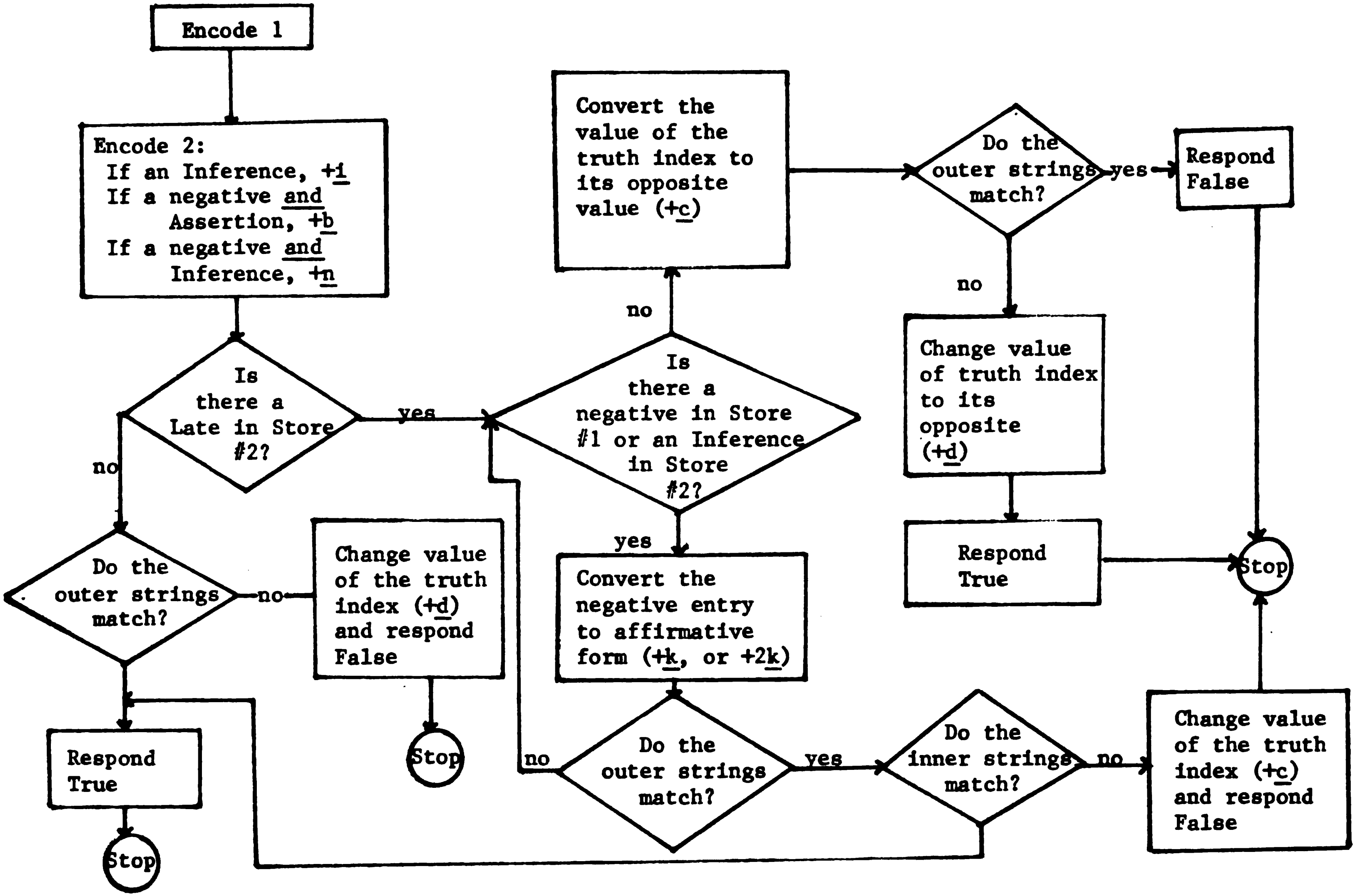

Figure 1.1. Process Model for Sentence-Sentence Verification

is that upon encoding Sentence 1, the comparable Inference will also be encoded.

During Stage 2, the processor encodes Sentence 2. As in the Clark and Chase (1972) model, the encoding of a negative is expected to involve a given amount of time longer than an affirmative. It was apparent from the data, however, that the encoding of a negative in the presence of an Inference took considerably longer than in the presence of an Assertion. This was probably because of differences in their scope (Clark, 1970, 1974; Klima, 1964). Consider the following statements (15a,b). The latter negates a larger, more complex proposition which

(15) a. Mr. Jones was not on time today.
false (on time)

b. Mr. Jones is not usually on time.
false (usually (on time))

should in fact require more time. It was thus proposed that parameter *b* represents the additional time needed to encode a negative in an Assertion, and *n* be the negation encoding time in an Inference.

The nature of Sentence 2 also varied as to whether it was an Assertion or an Inference. Since an Assertion was assumed to be cognitively less complex than an Inference, dealing with information directly relayed in Sentence 1, the encoding of an Inference was predicted to require additional processing time, time *i*.

In Stage 2, Clark and Chase (1972) hypothesized that *below* would take *a* time longer to encode than *above*. This extra time was interpreted as being the result of the inherent negativity of *below* along the vertical dimension. Implicit negation or lexical marking was a principle introduced by Clark (1969; Clark & Chase, 1972) which accounted for the asymmetrical features of certain dimensional contraries, e.g., high-low, above-below, and deep-shallow. Applied to the present study, it would be reasoned that should marking effects be found, Late would be marked with respect to On Time. Along a dimension of timeliness, On Time is the focal point, that is, the most neutral point along that dimension. Proceeding outward in either direction, one becomes more Early or more Late, acquiring a negative sense. Since Clark (1969) posited that lexically unmarked items would be stored in and retrieved from memory more rapidly than marked items, one would predict that On Time would be verified more rapidly than Late.

Although the analysis of variance reported a significant main effect for Late, which could reflect Late's being inherently negative with respect to On Time, it cannot be determined from the present design whether the effect stems from lexical marking or to simply a change in the stimulus. Hence, no marking parameter has been included.

The occurrence of Late in Stage 2, however, does play a significant role in the processing of infomration in the present experiment. Upon encoding Late, it appears that the subject will search his internal representations for the presence of either a negative qualifier in Representation 1, or an Inference in Representation 2. If either or both of these conditions is found, the subject will automatically convert the negative representation to its affirmative equivalent. Since there are two negatives in the false double negative Assertion treatment, two conversions would be predicted. Each conversion is expected to involve an additional time, k. At this point, all encodings and elaborations are presumed complete.

There are two important features to note about this particular substage of processing. First, this stage is a critical pivot point on which all other processing hinges. A "no" response to question (16)

(16) "Is there a Late in Representation 2?"

results in a simple outer string identity check, at most involving $\underline{d}$ (outer string mismatch) additional time. A "yes" response, however, will lead either to a semantic truth value change (+$\underline{c}$) or at least one conversion (+$\underline{k}$), two in the case of the false double negative Assertion condition (+2$\underline{k}$), with the possibility of inner (+$\underline{c}$) or outer (+$\underline{d}$) string mismatches also occurring.

Following the "yes" route at this initial decision point would account for the difficulty of the false double negative Assertion Condition. It is the only condition in which two conversion operations may occur. In addition, the outcome of the two conversion executions is an inner string mismatch, incurring additional time $\underline{c}$. Hence, the presence of these three operations (2$\underline{k}$+$\underline{c}$), each quite timely, within one treatment greatly increases the processing load of the subject.

The processor is now able to commence comparison operations in Stage 3 in order to determine the truth value of the paired representions, as in Clark and Chase (1972). If the subject followed the "no" route at the initial decision point (16), all that remains for him to do is a check of the identity of the outer embedding strings. At most, this will take more time, $\underline{d}$, if a mismatch is detected.

Subjects following the "yes" route must undergo a more complex set of operations. If no Inference in Store 2 or negative in Store 1 is found at decision point (17), in Stage 3 the subject will convert the

(17) "Is there a negative in Store 1 or an Inference in Store 2?"

value of the truth index to its opposite, an operation involving time $\underline{c}$, (18). The subject will then proceed to check the identity of the outer

(18) (on time)
(late) $+c$

string entries; a mismatch and change of the truth index to its opposite will involve time d (19). If, however, a conversion operation

(19) (on time)
false (late) $+(c+d)$

has occurred, as a result of an Inference in Store 2 or a negative in Store 1, the subject, finding identity of the outer strings (all negative entries are eliminated by model specifications), will recheck the identity of the inner strings. Finding identity, he will respond true, incurring no additional time. It is only in the case of the false double negative Assertion (20) that a mismatch will be found, incur-

(20) false (on time)
false (late) $+c$

ring additional time c.

Stage 3 operations aid in the explanation of why the FN treatment is less difficult than the FA. All FN's traverse the "no" path. Only one operation is required (outer string identity check) before the subject is able to leave the system. FA's, however, require the subject to follow the "yes" route. Before changing the value of the truth index to its opposite and checking the identity of the outer strings, the subject must conduct a search of his internal representations. This extra operation would account for the increased difficulty of the FA's.

The outcome of Stage 3 is then input to Stage 4 where a response, true or false, is executed. Any time remaining has been allotted to parameter t_o, reflecting base processing time common to all treatments.

A least squares analysis of mean RT's yielded the following parameter estimates: t_o, 284 msec.; b, 60 msec.; n, 357 msec.; i, 180 msec.; k, 312 msec.; c, 368 msec.; and d, 143 msec. The root mean squared deviation (RMSD) was calculated to be 43 msec. Since the RMSD was lower than the smallest parameter, _b_ (See Sternberg, 1969a,b), the model proposed was interpreted as being a reliable description of subjects' processing. In addition, related measures _t_-tests of each parameter all proved highly significant, where $p < .05$.

Despite its proficiency at estimating treatment response times, several criticisms must be directed at the proposed model. Failure to completely counterbalance Late and On Time in Stage 1 has previously been cited as a design deficiency resulting in inability to interpret slower RT's in the presence of Late. Furthermore, the change in stimulus (i.e., Late in Sentence 2) is theoretically responsible for more complex, time consuming processing. Whether this is due to variation in inner string propositions (Inner String 1 $\neq$ Inner String 2) or is simply the result of the presence of Late cannot be determined from the present experiment, again due to the failure to counterbalance.

Examination of within-subject RT's variability in each condition also raised some doubts as to subject processing. Specifically, there were indications that individual subjects may have been differentially responding to a given condition. That is, response to a particular treatment was not necessarily constant across eight presentations.

Finally, the proposed model yields an incomplete factorial design. Examination of the model reveals that there is no condition characterized by $(t_o + b + k)$, $(t_o + b + 2k)$, $(t_o + b + c)$, or any combination of $(t_o + n + c + \ldots)$, of $(t_o + n + b + \ldots)$, for example. Furthermore, if the model is an accurate representation of subject processing, conditions such as these could not be constructed manipulating On Time and Late. It is possible, therefore, that the model produced is the outcome of a host of interacting artifactual variables.

While these deficiencies have made it more difficult to analyze and interpret the findings of the pilot study, the ability of the model to characterize trends in the data, particularly in the case of the false double negative Assertion treatment, leads one to believe that if such deficiencies were minimized or eliminated, the model would be representative of the pattern of RT's yielded, for complex processing (e.g., the abstraction of Inferences) necessitates the use of more complex strategies for successful task completion. To test this hypothesis, the following experiments have been conducted.

Experimental Extensions

A major criticism of the Clark and Chase (1972) approach to information processing is that very simple tasks have been used to illustrate principles of analysis theoretically applicable to any task. Yet, little research has been carried out to validate this generalization. The results of the pilot study indicate that in more complex situations, only peripheral aspects of the pure true or conversions models can be found. It would seem that processors adapt the complexities of their strategies to meet the demands of the task at hand. It is important to determine whether subjects in fact do use highly complex strategies in difficult problems, or whether all verification strategies can be reduced to the true or conversion strategies. Experiment 1 examines this problem.

An auditory sentence-sentence verification of Assertions and Inferences task is used in an attempt to replicate the experimental findings of the pilot study; specifically, that a highly complex form of conversion (i.e., subject to two conditions of occurrence) operates in the verification of Assertions and Inferences; and that the false double negative Assertion, involving two conversion operations, is the most difficult verification to make. As in the pilot study, negative entries have been input to Stage 1, Stage 2, and Stages 1 and 2 encodings.

Experiment 1 differs from the pilot study in three ways: first, there is complete counterbalancing of On Time and Late in Sentences 1 and 2. This will allow one to determine whether lexical marking exists along the dimension of time. Secondly, all the subjects are tested in three individual sessions. It is hypothesized that the problem of within-subject variability will be remediated by multiple session testing of subjects. It is thought that an individual's performance pattern will stabilize across three sessions of testing such that, by session 3, a subject will be processing a given condition consistently on all presentations.

It is predicted that Experiment 1 will reproduce the findings of the pilot study using eight parameters: $\underline{t}_o$, base processing time; $\underline{a}$, Stage 2 Late encoding time; $\underline{b}$, Stage 2 negation encoding time (Assertions); $\underline{n}$, negation encoding time in Stage 2 (Inferences); $\underline{i}$, Stage 2 Inference encoding time; $\underline{k}$, conversion operation in Stage 2 substage; $\underline{c}$, change in value of truth index in Stage 3 due to the presence of an inner string mismatch; and $\underline{d}$, change in the value of the truth index in Stage 3 due to the occurrence of an outer string mismatch. Since in the pilot study, no lexical marking effects were predicted for Late, in Experiment 1 prediction of parameter $\underline{a}$ is expected to decrease the pilot study values of both $\underline{c}$ and $\underline{k}$. Conversion strategies are expected to be used in conditions in which inner string mismatches occur with a negative in Representation 1 or an Inference in Representation 2.

It is quite possible, however, that complex operations are used in any higher order language comprehension task. Experiments 2 and 3 are designed to test this hypothesis in the auditory and visual modalities, respectively. Both experiments investigate subjects' verification of simple Assertions (21), that is, information directly relayed in a sentence.

(21) He was <u>on</u> <u>time</u> today.
He was <u>on</u> <u>time</u> today.

It is predicted that in both experiments the effects of encoding a negative and a Late in Sentence 2 will be seen. Furthermore, it is predicted that inner string mismatches will be more time consuming than inner string identities.

If in simple language comprehension tasks the same rules are used to verify information as in the simple sentence-picture verification tasks, one would expect the results of Experiments 2 and 3 to meet the predictions of either the true model of negation or the conversion model. It may be, however, given the complexities of language, that different rules (which may be characterized by another serial stage processing model, e.g., the pilot model, or which may be executed in parallel operations) are used in processing. Finding an answer to this question is one of the purposes of Experiments 2 and 3.

The final aim is to determine whether differences exist in how the same information is processed in different modalities. It is possible that conversion strategies would not be used in a visual modality sentence-sentence verification task, as conversion does not seem to be a necessary routine (since the stimuli are always available); that is, the routine would not seem to maximize processing time. Conversion with subsequent mismatch checks may be less economical than simple immediate inner and outer string checks when stimuli are always available. On the other hand, if conversion operations were executed in an auditory task, they could serve to minimize the number of elements being held in the short term store at any one time, thus easing the load of the processor. Hence, conversion in auditory tasks may be an optimal strategy.

Summary

The experiments are investigations of negative conversion operations in a sentence-sentence verification of Assertions and/or Inferences task. Their basic premise is the determination of whether conversion operations or complex rule systems are automatically used in higher order processing as a function of the task being a language comprehension problem.

Experiment 1 addresses itself to the problems of lexical marking and conversion operations in a complex task (the verification of Assertions, and the construction and verification of Inferences). Experiment 2 examines whether complex operations are simply a function of higher order processing (in situations in which Inferences must be made, as in Experiment 1) or are rather a function of the complexity of the task used. Finally, Experiment 3 attempts to determine whether complex comprehension operations are equally employed in both auditory comparison tasks and visually presented sentence-sentence verifications.

CHAPTER 2

Experiment 1

Experiment 1 is a strict test of the model generated from the pilot study's results. Essentially, the experiment purports to test the hypothesis that complex conversion operations are employed in more demanding problem solving settings. Investigation of negative conversion operations will be made using an auditory sentence-sentence verification task. Verification of both directly conveyed information (Assertions) and abstracted information (Inferences) will be examined.

Method

Subjects. The subjects were 12 male and female undergraduates participating for credit in an introductory psychology class at Michigan State University. All the subjects were also paid $2.00 for their participation.

Stimuli. Two hundred fifty-six sentences recorded by a female speaker were presented in two-sentence test sequences. Sentence 1 constructed a context concerning an individual's activity, closing with a phrase indicating what a particular actor did today. One-half of the phrases contained the term On Time, the remaining half contained the term Late. In addition, one-half of the sentences were affirmaitve and one-half negative. Sentence 2 was a simple verification statement; one-half of the sentences contained the term On Time; one-half, Late; one half were affirmative statements, with the remaining half being negative. Sentence 2 also varied as to the nature of the information tested: one-half of the sentences referred to information directly conveyed in Sentence 1 (Assertions), while the remaining half tested Inferential judgments. Table 2.1 contains the 32 possible stimulus combinations.

Apparatus. Sentence stimuli, recorded on a TEAC 2300S stereo recorder, were presented to the subjects through stereo headphones. A click inaudible to the subject was placed on the tape approximately 200 msec. after the beginning of clause two of Sentence 1 (...surprisingly...) and approximately 200 msec. after the onset of speech in Sentence 2. These clicks automatically activated Hunter timers which recorded the subjects' RT's. A three button panel was located directly in front of the subjects: a button marked "C" (Comprehension) was centrally positioned, with a "T" (True) or "F" (False) button to either side. Depression of any button automatically stopped the Hunter timers and triggered the presentation of the next stimulus.

Procedure. All the subjects were individually tested in three performance sessions. The order of stimulus presentation was randomly varied across the three sessions, so that no subject received the same input order twice. The participants were instructed via the tape recorder to listen carefully to the contents of each sentence, for they would be asked to make true-false judgments of each sentence-sentence pair. Following Sentence 1, the subjects were required to press the comprehension ("C") button as quickly as possible, indicating that they had understood the sentence. Sentence 2 immediately followed. The subjects were then required to depress either the "T" (True) or "F" (False) button, depending upon their evaluation of the comparison. The position of

TABLE 2.1

Conditions of Experiment 1

Sentence 1	Sentence 2
. . . surprisingly, he was . . .	Mr. Smith . . .
. . . *on time* today.	. . . was *on time* today.
	. . . was *late* today.
	. . . was *not on time* today.
	. . . was *not late* today.
	. . . is *usually on time*.
	. . . is *usually late*.
	. . . is *not usually on time*.
	. . . is *not usually late*.
. . . *late* today.	. . . was *on time* today.
	. . . was *late* today.
	. . . was *not on time* today.
	. . . was *not late* today.
	. . . is *usually on time*.
	. . . is *usually late*.
	. . . is *not usually on time*.
	. . . is *not usually late*.
. . . *not on time* today.	. . . was *on time* today.
	. . . was *late* today.
	. . . was *not on time* today.
	. . . was *not late* today.
	. . . is *usually on time*.
	. . . is *usually late*.
	. . . is *not usually on time*.
	. . . is *not usually late*.
. . . *not late* today.	. . . was *on time* today.
	. . . was *late* today.
	. . . was *not on time* today.
	. . . was *not late* today.
	. . . is *usually on time*.
	. . . is *usually late*.
	. . . is *not usually on time*.
	. . . is *not usually late*.

the "T" and "F" buttons was randomly varied between subjects. All the subjects were instructed to answer each comparison as quickly and as accurately as possible. A five second interval occurred between each test comparison. The experimenter recorded all response times.

All the subjects received eight practice trials prior to actual testing, and were advised of their performance. In addition, after each day of testing, the subjects were advised of their overall performance for the day. Upon completion of the third day of testing, the subjects were asked to specify strategies used during processing, if any. The subjects were also asked to indicate whether they had changed their strategy across the three testing sessions, to describe this change, and to explain it.

Results and Discussion

In this section, performance over all three sessions, as well as performance in Session 3 alone, will be discussed. In addition, an attempt will be made to interpret the Day 3 results in terms of an additive factor (independent, serial stage) model of processing. The rationale for focusing on the Day 3 results was as follows. It was felt that if the subjects did choose a common strategy which directed their information processing, it wuold be most obvious and stable by the Day 3 performance. Thus, it was hoped that analysis of the Day 3 data would yield the most clear-cut insights into the processes underlying performance in the conditions of Experiment 1. It was also hoped that the within-subject variability in responding seen in the pilot study would decrease over the three testing sessions.

Since the major results to be reported occurred in both the Day 1 and Day 2 performances, these data will not be discussed. Mean RT per condition for Days 1 and 2 are contained in Appendices A and B, respectively. $\underline{F}$ values of the significant results of within-subject analyses of variance performed on these data can be found in Appendix C.

Overall. Only correct responses were included in the means; individual data were eliminated if the were $\pm$ 3 standard deviations from the mean. The error rate for the overall performance in Experiment 1 was 7.2%.

A 2 x 2 x 2 x 2 x 2 x 3 within-subjects analysis of variance was performed on the data. The factors were: Condition (Assertion vs. Inference), Term 1 (On Time_1 vs. Late_1), Negation 1 (Affirmative_1 vs. Negative_1), Term 2 (On Time_2 vs. Late_2), Negation 2 (Affirmative_2 vs. Negative_2), and Day (1, 2, or 3). Significant results of this analysis, where $\underline{p} < .05$, in comparison with the significant results of Days 1, 2, and 3, are indicated in Table 2.2.

In general, Inferences were more difficult than Assertions, Condition, $\underline{F}(1,11) = 231.94$, $MS_e = .410$; the presence of a negative in either sentence significantly increased processing time, Negation 1, $\underline{F}(1,11) = 9.415$, $MS_e = .143$, and Negation 2, $\underline{F}(1,11) = 158.987$, $MS_e = .217$; and the subjects were able to perform the verifications more rapidly over time, Day, $\underline{F}(2,22) = 126.54$, $MS_e = .201$. Contrary to predictions, comparisons containing an On Time_2 were more difficult than Late_2 comparisons, Term 2, $\underline{F}(1,11) = 77.981$, $MS_e = .140$.

The presence of an Inference served in several cases to inflate the effect of other factors paired with it. As indicated in Tables 2.3 and 2.4, the effect of a negative in either Sentence 1 or 2 was greater for

TABLE 2.2

Significant Main Effects and Interactions of the Analyses of Experiment 1

Condition	Overall	Day 1	Day 2	Day 3
Negation 1	*	*	*	M
Term 2	*	*	*	*
Negation 2	*	*	*	*
Condition	*	*	*	*
Day	*			
Condition x Negation 1	*	*	*	
Condition x Term 2	*		*	M
Condition x Negation 2	*			
Negation 1 x Negation 2	*	*	M	
Term 1 x Negation 1	M			
Term 1 x Term 2	*	*	*	*
Term 1 x Term 2 x Negation 2	*	*		M
Condition x Term 1 x Negation 1 x Negation 2	*			
Condition x Term 1 x Term 2 x Negation 2			*	
Condition x Term 1 x Negation 1 x Term 2 x Negation 2	*	*		*
Condition x Term 1 x Negation 1 x Negation 2 x Day	*			
Condition x Term 1 x Negation 1 x Term 2 x Day	*			

Where * denotes a significant result ($p < .05$); and M denotes a marginally significant result.

TABLE 2.3

Mean RT in Msec. for the Condition x Negation 1 Interaction

Condition	Negation 1		
	$Affirmative_1$	$Negative_1$	Mean
Assertion	1627	1647	1637
Inference	2154	2270	2213
Mean	1891	1958	

TABLE 2.4

Mean RT in Msec. for the Condition x Negation 2 Interaction

Condition	Negation 2		
	$Affirmative_2$	$Negative_2$	Mean
Assertion	1489	1786	1638
Inference	2015	2410	2213
Mean	1752	2098	

TABLE 2.5

Mean RT in Msec. for the Negation 1 x Negation 2 Interaction

Negation 2	Negation 1		
	$Affirmative_1$	$Negative_1$	Mean
$Affirmative_2$	1702	1800	1751
$Negative_2$	2078	2116	2097
Mean	1890	1958	

Inferences than for Assertions, Condition x Negation 1, $F(1,11) = 12.965$, $MS_e = .050$, and Condition x Negation 2, $F(1,11) = 7.71$, $MS_e = .090$. Also, the effect of an On Time in Sentence 2 was greater among Inferences (mean RT of On $Time_2$-$Late_2$ = 245 msec.) than among Assertions (mean RT difference between On $Time_2$-$Late_2$ = 144 msec.), Condition x Term 2, $F(1,11) = 27.547$, $MS_e = .027$.

As can be seen in Table 2.5, the effects of two negatives were not additive, Negation 1 x Negation 2, $F(1,11) = 8.031$, $MS_e = .032$. That is, the effect of a $Negative_2$ was greater in the presence of an $Affirmative_1$ than a $Negative_1$. It could be the case that negation in Sentence 1 primed the subject for another negative entry, thus facilitating a response. An affirmative entry, on the other hand, would not prepare the subject for a negative, thus resulting in a greater slowdown of processing.

Consistent with the predictions, inner string mismatches were more time consuming than inner string identities, Term 1 x Term 2, $F(1,11) = 12.459$, $MS_e = .171$. That is, $Late_1$-On $Time_2$ comparisons had slower RT's than On $Time_1$-On $Time_2$ comparisons (2069 msec. vs. 1974 msec., respectively) and On $Time_1$-$Late_2$ comparisons were longer than $Late_1$-$Late_2$ comparisons (1866 msec. vs. 1790 msec., respectively). This effect is illustrated in Figure 2.1.

As seen in Table 2.6, the effect of a $Negative_2$ was greater in the presence of an inner string mismatch than in the presence of a match, Term 1 x Term 2 x Negation 2, $F(1,11) = 7.535$, $MS_e = .036$. When interpreted in light of the Term 1 x Term 2 interaction reported above, one might conclude that as the processing demands of a comparison are increased (for example, due to the occurrence of an inner string

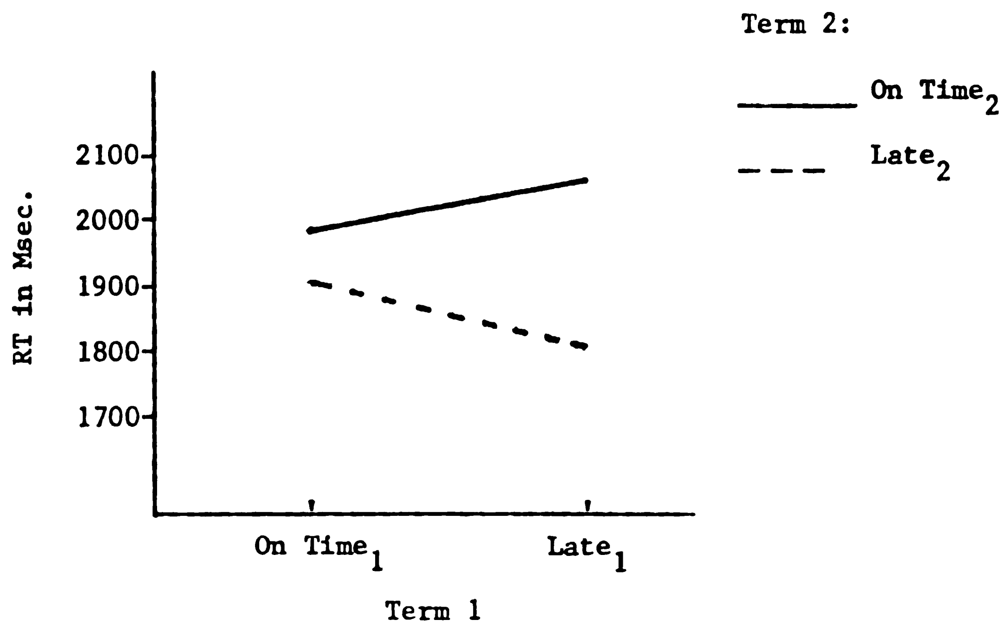

Figure 2.1 Mean RT in Msec. to Term 2 as a function of Term 1

TABLE 2.6

Mean RT in Msec. for the Term 1 x Term 2 x Negation 2 Interaction

Negation 2	Match		Mismatch		Mean
	On $Time_1$- On $Time_2$	$Late_1$- $Late_2$	$Late_1$- On $Time_2$	On $Time_1$- $Late_2$	
$Affirmative_2$	1815	1633	1888	1671	1751
$Negative_2$	2134	1945	2251	2061	2098
Mean	1975	1789	2070	1866	

mismatch), the effects of the Negative_2 will be exaggerated, resulting in slower RT's.

The remaining significant interactions of the overall analysis were the following five-way ones: Condition x Term 1 x Negation 1 x Term 2 x Negation 2, $F(1,11) = 18.608$, $MS_e = .072$; Condition x Term 1 x Negation 1 x Negation 2 x Day, $F(2,22) = 5.393$, $MS_e = .028$; and Condition x Term 1 x Negation 1 x Term 2 x Day, $F(2,22) = 3.632$, $MS_e = .042$. Basically, these indicated that comparisons requiring an Inference judgment, and which contained a Negative_2, an On Time_2, and in which the inner strings did not match were most time consuming. Furthermore, it appeared that while the subjects' performance did improve over days, Assertion judgments showed greater improvement over time than did Inference conditions. All Assertion conditions were executed more rapidly on Day 3 than on Day 1; however, only certain of the Inference conditions improved over time.

Day 3. Mean RT per condition for the Day 3 performance is contained in Table 2.9 (see p. 47). The average error rate for session 3 was 5.5%.

A 2 x 2 x 2 x 2 x 2 within-subject analysis of variance was performed on the data. The factors investigated were: Condition (Assertion vs. Inference), Term 1 (On Time_1 vs. Late_1), Negation 1 (Affirmative_1 vs. Negative_1), Term 2 (On Time_2 vs. Late_2), and Negation 2 (Affirmative_2 vs. Negative_2). The significant results of this analysis, where $p < .05$, are indicated in Table 2.2.

As in the overall analysis, main effects of Term 2, $F(1,11) = 60.006$, $MS_e = .058$; Negation 2, $F(1,11) = 270.205$, $MS_e = .042$; and Condition, $F(1,11) = 133.133$, $MS_e = .281$ were significant. However, the main effect of Negation 1 was only marginally significant, $\underline{F}(1,11) = 4.681$, $MS_e = .061$. Why this was so is not readily understood. It is possible that subjects can more efficiently encode and operate on a Negative_1 (i.e., the Negative_1 makes fewer processing demands) after three days of testing.

Only two interactions were significant in the Day 3 data: Term 1 x Term 2, $\underline{F}(1,11) = 5.513$, $MS_e = .158$, and Condition x Term 1 x Negation 1 x Term 2 x Negation 2, $\underline{F}(1,11) = 14.313$, $MS_e = .201$. As in the overall analysis, mismatches were more difficult than matches, and negative Inferences containing inner string mismatches and an On Time_2 were the most difficult comparisons to make.

We will now turn to an evaluation of how well the data of Experiment 1 fit the model derived in the pilot study.

Evaluation of the Pilot Study Model. The predictions of the pilot study model are not supported by the results of the present experiment. It was predicted that Late would be marked with respect to On Time; however, the opposite effect was obtained. In each Experiment 1 analysis, On Time_2 comparisons took significantly longer than Late_2 comparisons. Whether this effect was due to the lexical marking of On Time or was an artifact of the stimulus itself (On Time_2 stimuli were approximately 297 msec. longer than Late_2 stimuli) cannot be determined at present.

Furthermore, recall that in the pilot study, Affirmative_2 Assertion conditions ordered themselves in the following manner: TA < FN < FA < TN. Inclusion of Negative_1 Assertions yielded the following ordering: TA < FA ≃ TN < FN. In the present experiment for Day 3 data, when only Affirmative_1 Assertion comparisons were considered, it was found that TA (1286 msec.) < FA (1447 msec.) < FN (1632 msec.) < TN (1771 msec.). When Negative_1 Assertion conditions were included in the analysis, the following ease of processing ordering was produced: TA (1286 msec.) < FA (1447 msec.) < TN (1603 msec.) ≈ FN (1609 msec.). So on either count, the results of the pilot study do not match those of Experiment 1.

The pilot model encounters more trouble when one compares the predicted and actual orderings of conditions on Days 1 and 3. The most noticeable violation of the pilot model's predictions was the finding that the false double negative Assetion was not the most time consuming condition on either day. Examination of Table 2.9 reveals that by Day 3, 13 of the 16 Inference conditions were as or more time consuming than the false double negative Assertion.

In order to more fully assess the inadequacies of the pilot model, ordinal predictions of the model were compared with the obtained orders of the conditions. A tabulation of the comparisons between the predicted and obtained data is shown in Table 2.7. In each section of this table are listed two or more conditions which should significantly increase in difficulty as we move from left to right according to the pilot model because of an increase in the number of parameters. For instance, according to the pilot model, Late_1-On Time_2 should take significantly longer to process than On Time_1-On Time_2 because one more parameter is required ($\underline{t_o} + \underline{a} + \underline{c}$ vs. $\underline{t_o} + \underline{a}$).

It should be noted that in virtually every ordering, the model makes an error in prediction for one of the conditions. It can also be seen that the misrepresentation of the data usually occurs when parameter $\underline{d}$ (outer string mismatch) is predicted to be the only factor

TABLE 2.7

Pilot Model Predictions and Obtained Values in Msec. for Selected Conditions of Days 1 and 3

	On Time$_1$- On Time$_2$	Not On Time$_1$ On Time$_2$	On Time$_1$- Not On Time$_2$	Late$_1$- Not On Time$_2$
	t_o+a	t_o+a+d	t_o+a+b+d	t_o+a+b+c+d
Day 1	1654	1661*	1986	1993*
Day 3	1383	1480	1683	1759

	On Time$_1$- On Time$_2$	Late$_1$- On Time$_2$
	t_o+a	t_o+a+c
Day 1	1654	1703
Day 3	1383	1519

	On Time$_1$- On Time$_2$	Not On Time$_1$- Not On Time$_2$	On Time$_1$- Not On Time$_2$	Late$_1$- Not On Time$_2$
	t_o+a	t_o+a+b	t_o+a+b+d	t_o+a+b+c+d
Day 1	1654	1862	1986	1993*
Day 3	1383	1681	1683*	1759

	On Time$_1$- On Time$_2$	Not Late$_1$- On Time$_2$
	t_o+a	t_o+a+k
Day 1	1654	1787
Day 3	1383	1517

	Not On Time$_1$- Usually On Time$_2$	On Time$_1$- Usually On Time$_2$
	t_o+i+a	t_o+i+a+d
Day 1	2234	2056*
Day 3	2116	2064*

TABLE 2.7(cont'd.)

	Not On Time$_1$- Usually On Time$_2$	On Time$_1$- Not Usually On Time$_2$	Not On Time$_1$- Not Usually On Time$_2$	Late$_1$- Not Usually On Time$_2$
	t_o+i+a	t_o+i+a+n	t_o+i+a+n+d	t_o+i+a+n+d+k
Day 1	2234	2397	2599	2721
Day 3	2116	2397	2518	2462*

	Not On Time$_1$- Usually On Time$_2$	Late$_1$- Usually On Time$_2$
	t_o+i+a	t_o+i+a+k
Day 1	2234	2043*
Day 3	2116	2070*

	Not Late$_1$- Usually On Time$_2$	Late$_1$- Not Usually On Time$_2$
	t_o+i+a+k+d	t_o+i+a+k+d+n
Day 1	2307	2721
Day 3	2195	2462

	Not Late$_1$- Not Usually On Time$_2$	Late$_1$ Not Usually On Time$_2$
	t_o+i+a+n+k	t_o+i+a+n+k+d
Day 1	2747	2721*
Day 3	2471	2462*

Where * indicates an inaccurate prediction differentiating two conditions. This was true of both Day 1 and Day 3 performances. In conclusion, the pilot model is _not_ a good predictor of either Day 1 or Day 3 performance.

It is clear that practice is not responsible for the dramatic change of events between the pilot study and Experiment 1. It is possible that the change in the point from which one measured RT in the pilot study and in Experiment 1 is in part responsible. Since RT in the pilot study was measured from speech offset, a great deal of the subjects' encoding and processing could have been lost. RT measurement from Sentence 2 onset in Experiment 1 is undoubtedly a more sensitive measure of processing.

Evaluation of the True Model of Negation. Recall from Chapter 1 that the true model of negation (Clark & Chase, 1972) predicts that TA < FA < FN < TN. Thus, the true model of negation was able to account for the true-false, affirmative-negative orderings cited above (p. 35). for the Affirmative_1-Assertion conditions of Experiment 1. However, as can be seen in Table 2.8 (which is similar to Table 2.7), this model has just as many problems as the pilot model in predicting overall performance on Days 1 and 3. For this model, both parameters _c_ (inner string mismatch) and _d_ (outer string mismatch) appear to have questionable roles. Thus, the true model must also be discarded as an accurate representation of the subjects' processing.

Development of a New Model of Information Processing. In developing a model of information processing in the present paradigm, it would be most efficient to assume that the verification of Assertions and Inferences are subject to the same set of comparison rules, that is, entailing no extra processing time beyond encoding due to the nature of

TABLE 2.8

Predictions of the Clark and Chase Model and Obtained Values in Msec. for Selected Conditions of Days 1 and 3

	Late$_1$-Late$_2$	On Time$_1$-Late$_2$	Not On Time$_1$-Late$_2$	On Time$_1$ Not Late$_2$
	t_o	t_o+c	t_o+c+d	$t_o+b+c+d$
Day 1	1381	1569	1550*	1839
Day 3	1189	1374	1323*	1785
	Late$_1$-Late$_2$	On Time$_1$-Late$_2$	Not On Time$_1$-Not Late$_2$	On Time$_1$-Not Late$_2$
	t_o	t_o+c	t_o+b+c	$t_o+b+c+d$
Day 1	1381	1569	1970	1839*
Day 3	1189	1374	1635	1785
	Late$_1$-Late$_2$	Not On Time$_1$-Not On Time$_2$	On Time$_1$-Not On Time$_2$	On Time$_1$-Not Late$_2$
	t_o	t_o+b	t_o+b+d	$t_o+b+c+d$
Day 1	1381	1862	1986	1839*
Day 3	1189	1681	1683*	1785
	Late$_1$-Late$_2$	Not On Time$_1$-On Time$_2$	Not On Time$_1$-Late$_2$	
	t_o	t_o+d	t_o+c+d	
Day 1	1381	1661	1550*	
Day 3	1189	1480	1323*	
	Not Late$_1$-Usually Late$_2$	Late$_1$-Usually Late$_2$	On Time$_1$-Usually Late$_2$	Not On Time$_1$-Not Usually Late$_2$
	t_o+i	t_o+i+d	$t_o+i+c+d$	$t_o+i+b+c+d$
Day 1	2057	1958*	1821*	2359
Day 3	1817	1822*	1802*	2353

TABLE 2.8(cont'd.)

	Not Late$_1$- Usually Late$_2$ t_o+i	Late$_1$- Not Usually Late$_2$ t_o+i+b	On Time$_1$- Not Usually Late$_2$ $t_o+i+b+c$	Not On Time$_1$- Not Usually Late$_2$ $t_o+i+b+c+d$
Day 1	2057	2102	2481	2359*
Day 3	1817	2102	2308	2353

	Not Late$_1$- Usually Late$_2$ t_o+i	Not On Time$_1$- Usually Late$_2$ t_o+i+c	On Time$_1$- Not Usually Late$_2$ $t_o+i+b+c$
Day 1	2057	1964*	2481
Day 3	1817	1980	2308

	Not Late$_1$- Usually Late$_2$ t_o+i	Late$_1$- Not Usually Late$_2$ t_o+i+b	Not Late$_1$- Not Usually Late$_2$ $t_o+i+b+d$	Not On Time$_1$- Not Usually Late$_2$ $t_o+i+b+d$

Where * denotes an inaccurate prediction their being an Assertion or an Inference. This, of course, would make processing most economical.

In order to determine whether the differences between Inferences and Assertions, and negatives and affirmatives, are confined to the encoding stage of processing or whether time consuming differences in processing occur in both encoding and comparison stages, six subjects participated in a language perception experiment. It was reasoned that time to recognize a stimulus is a fundamental part of the encoding process. Therefore, if one parcels out the time it takes to identify a stimulus from the overall processing time, one should be able to identify processes affecting the comparison stage. It was also hoped that this experiment would shed some light on the Late_2-On Time_2 difference which was not predicted.

Six college students participated in a language perception experiment in fulfillment of an introductory psychology course requirement. The stimuli were Assertion and Inference test sentences appearing in Experiment 1.

Subjects were instructed to depress a button as soon as they recognized the target phrase for which they were primed: On Time, Late, Not On Time, Not Late, Usually On Time, Usually Late, Not Usually On Time, or Not Usually Late. All the subjects were tested in blocks under all priming conditions. In each condition, one-half of the stimuli contained the term On Time, one-half, Late; one-half were affirmative, the remaining half, negative; finally, one-half of the stimuli were Assertions, and the remaining half tested Inference constructions. Under Assertion instructions, the subjects made 80 judgments; 96 judgments were made under Inference instructions.

Mean RT per condition for correct identifications were tabulated: On Time = 1094 msec.; Late + 959 msec.; Not On Time = 1227 msec.; Not Late = 1145 msec.; Usually On Time = 1670 msec.; Usually Late = 1407 msec.; Not Usually On Time = 1846 msec.; and Not Usually Late = 1679 msec. The overall error rate was 2.0%. Mean RT's were then subtracted from the overall condition mean RT's of Experiment 1 and the Day 3 data were reanalyzed according to the analysis of variance described earlier.

The effect of Term 2 was no longer significant ($F(1,11) = 1.373$, $p > .05$), nor was the Condition main effect ($F(1,11) = 2.198$, $p > .05$). It must be concluded then, that the effects of an Inference and an On Time_2 are effects which must be restricted to the encoding stage of processing. The effects of negation (Negation 2, $F(1,11) = 52.824$, $p < .005$) as well as the effects of inner string mismatch (Term 1 x Term 2, $F(1,11) = 5.513$, $p < .039$) were found to exceed the bounds of the encoding stage.

In conclusion, any model which purports to represent the data must posit encoding parameters for On Time_2 and the Inference, and both encoding and comparison parameters for negation. The following model is proposed as an adequate characterization of the Day 3 data. Of the models tested, it was found to be the most representative of actual performance; furthermore, it was able to restrict the effect of the Inference to the encoding stage alone, as mandated by the last reported analysis.

As in the previously discussed models, processing will be broken down into four successive stages. In Stage 1, subjects will encode Sentence 1. The following representation (22) is suggested but not

(22) . . . surprisingly, he was $\underbrace{\underline{\text{not}}\ \underline{\text{on}}\ \underline{\text{time}}}_{\text{false (on time)}}$ today.

necessarily the only possible one. As mentioned earlier, however, time to encode this sentence is not part of the measured and analyzed RT.

In Stage 2, the subject will encode the second entry in a representation similar to that of Stage 1 (23). The encoding of this sentence

(23) Sentence: Mr. Smith is <u>not</u> <u>usually</u> <u>on</u> <u>time</u>.
Representation: false (usually (on time))

is part of the overall RT recorded, therefore certain encoding parameters must be posited.

Given that On Time_2 sentences took consistently more time to verify than Late_2 sentences, parameter *a* will be the time taken to encode an On Time_2, parameter *b* will reflect the time taken to encode a Negative_2, and *i* will be the time needed to encode an Inference. Finally, since in both the Assertion and Inference conditions two negative entries appeared to be most difficult, it is proposed that parameter *n* will reflect the extra processing time due to the occurrence of a second negative.

It is only during Stage 2 that parameters *a*, *b*, *i*, and *n* function. If example (23) were encoded in the presence of (22), one would predict that parameters *a*, *b*, *i*, and *n* would all be operating. Once the second entry has been entirely encoded, processing enters the comparison phase, or Stage 3.

It is during this stage that the *c*, *d*, and *x* parameters operate. It is proposed that in both Assertion and Inference comparisons, subjects will verify first the contents of the inner strings, and then proceed to verify outer strings. Finding mismatches in either of these two comparisons will result in an increase in processing time by +*c* and +*d*, respectively.

Following inner and outer string comparisons, the Inference or Assertion condition will determine the final decision rule. It is believed that subjects will not automatically make changes in the value of the truth index upon finding a mismatch. Rather subjects will temporarily store the fact of the mismatch in a hypothetical memory buffer. Where Inference and Assertion operations differ will be in their interpretation of the contents of the memory buffer.

In the Assertion condition, if the number of mismatches recorded in the buffer is an even number (0 or 2), the subject will respond True. If an odd number of mismatches has been found, however, the subjects will change the value of the truth index to its opposite value (False), requiring additional processing time, $+\underline{x}$. As in other models, the value of the truth index is assumed to be set at True at the outset.

In the Inference comparisons, decision making is governed by a converse set of rules. In this instance, if the value stored in the buffer is an even number, the subject will respond false. It is only when the buffer is set at an odd value that the subject will respond True. As in the Assertion comparison, the changing of the truth index to its opposite polarity will increase processing time by $+\underline{x}$. Upon completion of comparison operations, the output of Stage 3 is executed in the response phase, Stage 4.

The additive factor model assumes that each stage is independent of all others, and that the sum of all operations in a given condition will reflect the total processing time. In the example cited below (24), the response time of that comparison is predicted to be the sum

(24) false (on time)
false (usually (on time)) $t_o+i+a+b+n+x$

of t_o, a base processing time, and the necessary parameters appearing in

Stages 2 and 3. Figure 2.2 displays a flowchart of the possible alternatives one has in such a paradigm.

A least squares analysis was performed using subjects' mean RT per condition in order to determine parameter estimates. The eight parameters were estimated to be: $\underline{t}_o$, 1169 msec.; $\underline{i}$, 625 msec.; $\underline{a}$, 191 msec.; $\underline{b}$, 285 msec.; $\underline{c}$, 95 msec.; $\underline{d}$, 72 msec.; $\underline{n}$ 118 msec.; and $\underline{x}$, 52 msec. Observed scores, predicted values, and their deviations from the actual scores per condition are contained in Table 2.9.

In order to test the goodness of fit of the model, the RMSD was obtained, RMSD = 53 msec. Since parameters $\underline{t}_o$, $\underline{a}$, $\underline{b}$, $\underline{i}$, $\underline{c}$, $\underline{d}$, and $\underline{n}$ were all greater than the RMSD, these parameter estimates were interpreted as being reliable features of information processing (see Sternberg, 1969a,b). Parameter $\underline{x}$ is questionable, however, since it is equivalent to the RMSD.

Parameter $\underline{a}$, as represented by the Term 2 main effect was highly significant. Parameter $\underline{b}$, represented by the Negation 2 main effect was also significant. In addition, parameter $\underline{i}$ was shown to be a reliable effect through the main effect for Condition, and parameter $\underline{c}$ significant by the Term 1 x Term 2 interaction.

Unfortunately, significant effects corresponding to the remaining three parameters, $\underline{d}$, $\underline{n}$, and $\underline{x}$ could not be isolated. No two conditions could be found which differed only in terms of <u>one</u> of these parameters. However, it is the case that unlike other models, the proposed model is able to predict the T-F, A-N ordering found in the data, that is, that $TA < FA < TN \simeq FN$.

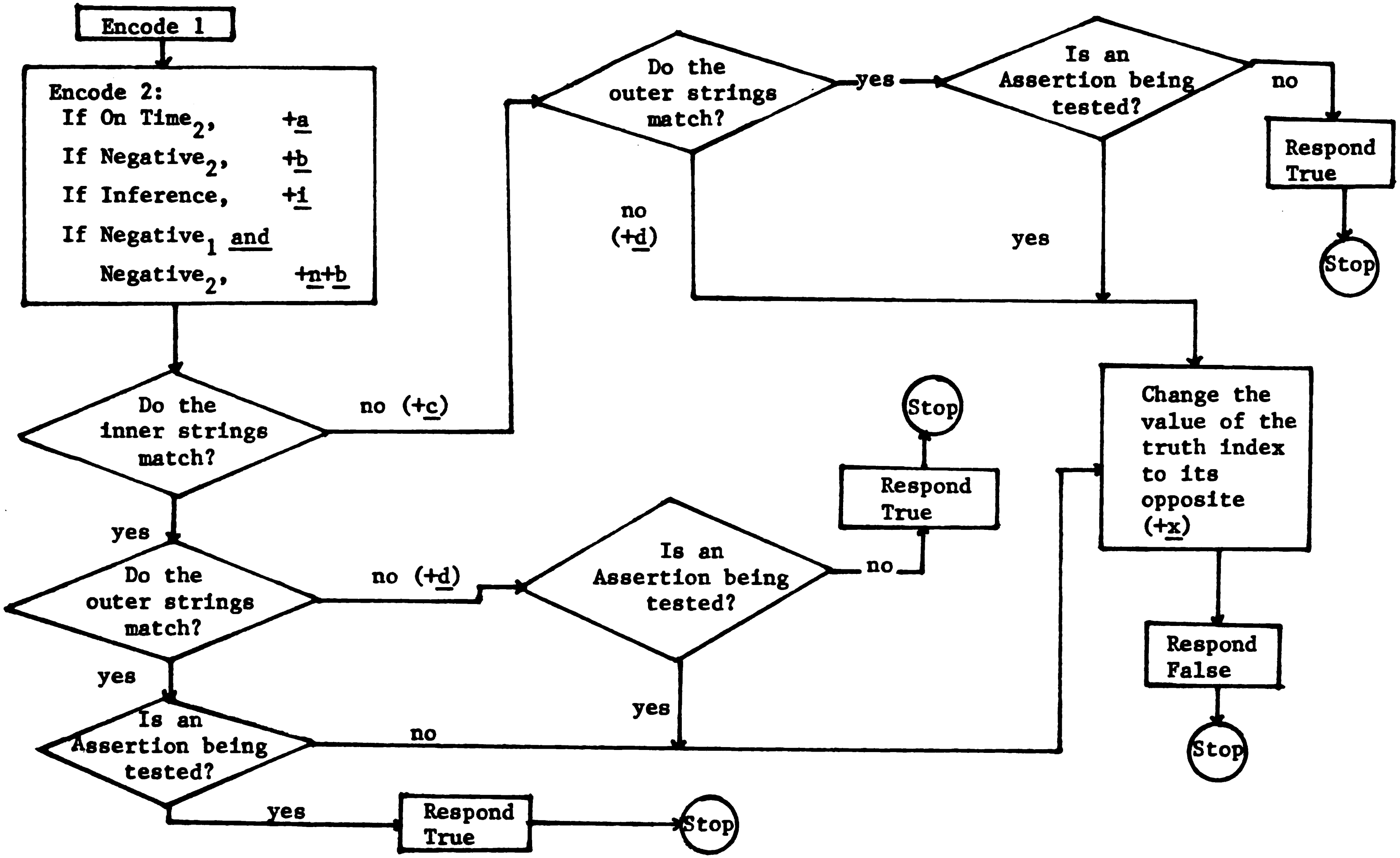

Figure 2.2. Processing Alternatives of the Model of Experiment 1

TABLE 2.9

Mean RT in Msec. Per Condition, Estimated, Observed and Deviations for Experiment 1-Day 3 Data

Conditions	Projected Parameters	Estimated	Observed	Deviation
On Time-On Time	t_o+a	1360	1383	- 23
On Time-Late	t_o+c+x	1316	1374	- 58
On Time-Not On Time	t_o+a+b+d+x	1769	1683	+ 86
On Time-Not Late	t_o+b+c+d	1621	1785	-164
Late-On Time	t_o+a+c+x	1507	1519	- 12
Late-Late	t_o	1169	1189	- 20
Late-Not On Time	t_o+a+b+c+d	1812	1759	+ 53
Late-Not Late	t_o+b+d+x	1578	1580	- 2
Not On Time-On Time	t_o+a+d+x	1484	1480	+ 4
Not On Time-Late	t_o+c+d	1336	1323	+ 13
Not On Time-Not On Time	t_o+a+b+n	1763	1681	+ 82
Not On Time-Not Late	t_o+b+c+n+x	1719	1635	+ 84
Not Late-On Time	t_o+a+c+d	1527	1517	+ 10
Not Late-Late	t_o+d+x	1293	1334	- 41
Not Late-Not On Time	t_o+a+b+c+n+x	1910	1945	- 35
Not Late-Not Late	t_o+b+n	1572	1552	+ 20

TABLE 2.9(Cont'd)

Conditions	Projected Parameters	Estimated	Observed	Deviation
On Time- Usually On Time	t_o+i+a+x	2037	2064	- 27
On Time- Usually Late	t_o+i+c	1889	1802	+ 87
On Time- Not Usually On Time	t_o+i+a+b+d	2342	2397	- 55
On Time- Not Usually Late	t_o+i+c+b+d+x	2298	2308	- 10
Late- Usually On Time	t_o+i+a+c	2080	2070	+ 10
Late- Usually Late	t_o+i+x	1846	1822	+ 24
Late- Not Usually On Time	t_o+i+a+b+c+d+x	2489	2462	+ 27
Late- Not Usually Late	t_o+i+b+d	2151	2102	+ 49
Not On Time- Usually On Time	t_o+i+a+d	2057	2116	- 59
Not On Time- Usually Late	t_o+i+c+d+x	2013	1980	+ 33
Not On Time- Not Usually On Time	t_o+i+a+b+n+x	2440	2518	- 78
Not On Time- Not Usually Late	t_o+i+b+c+n	2292	2353	- 61
Not Late- Usually On Time	t_o+i+a+c+d+x	2204	2195	+ 9
Not Late- Usually Late	t_o+i+d	1866	1817	+ 49
Not Late- Not Usually On Time	t_o+i+a+b+c+n	2483	2471	+ 12
Not Late- Not Usually Late	t_o+i+b+n+x	2249	2254	- 5

Summary

It is clear from Experiment 1 that Inference verifications take considerably more time than Assertion judgments. Furthermore, the increase in processing time in this case is restricted to the encoding stage. Consistent with earlier studies, our findings confirmed that negatives are more difficult than affirmatives, their effects being seen in both the encoding (+*b*) and the comparison (+*d*) stages; and that inner string mismatches are psychologically more difficult than identities. Furthermore, despite the subjects' variation in processing across the three sessions, these features are characteristic of each day's performance. Thus, we have been able to isolate some critical features of processing in a sentence-sentence verification task of Inferences and Assertions.

CHAPTER 3

Experiment 2

Experiment 2 investigates whether subjects use complex strategies in a simple sentence-sentence verification task. In the present study, all verifications of Inferences have been eliminated; only simple declarative sentences counterbalanced for Term (On Time vs. Late) and sentence polarity (affirmative vs. negative) will be examined. If, the data cannot be adequately represented by simple models, it would appear that the comprehension processes involved in a strictly linguistic task, either in higher order situations (formulating Inferences) or simple Assertion comprehension, are more complex than those involved in previous sentence-picture matching experiments.

Method

Subjects. The subjects were 12 male and female undergraduates participating for credit in an introductory psychology at Michigan State University.

Stimuli. Two hundred fifty-six declarative sentences recorded by a female speaker were presented in two-sentence test sequences. Sentence 1 was a simple statement of fact. One-half of these sentences contained the term On Time, the remaining half, Late. One-half of the sentences were affirmative, one-half, negative. Sentence 2 was a simple verification sentence, one-half being affirmaitve, and one-half, negative. One-half of the sentences contained the term On Time, the remaining ones, Late. The design yielded 16 possible combinations, as seen in Table 3.1.

Apparatus. Identical to that used in Experiment 1.

Procedure. All the subjects were subject to the same experimental procedures used in Experiment 1.

Results and Discussion

The significant results of an overall analysis of the Experiment 2 data, and of a separate analysis of the Day 3 data will be presented and discussed in the following section. As in Experiment 1, an attempt will be made to interpret the Day 3 data in terms of an additive factor model.

TABLE 3.1

Conditions of Experiments 2 and 3

Sentence 1	Sentence 2
Mr. Smith was on time today.	. . . on time today. . . . late today. . . . not on time today. . . . not late today.
. . . . late today.	. . . on time today. . . . late today. . . . not on time today. . . . not late today.
. . . . not on time today.	. . . on time today. . . . late today. . . . not on time today. . . . not late today.
. . . . not late today.	. . . on time today. . . . late today. . . . not on time today. . . . not late today.

No specific discussion will be made of the results of Day 1 and Day 2 performance analyses, as all the significant results of these two performances were also significant on Day 3. A complete listing of the mean RT's per condition on Days 1 and 2 can be found in Appendices D and E, respectively. The significant main effects and interactions in the analyses of variance for these two days are indicated in Appendix F.

Overall. Only correct responses were included in the analysis; individual data $\pm$ 3 standard deviations from the mean were interpreted as errors. The overall error rate was 4.1%.

A 2 x 2 x 2 x 2 x 3 within-subjects analysis of variance was performed on the data. The factors were: Term 1 (On Time_1 vs. Late_1); Negation 1 (Affirmative_1 vs. Negative_1); Term 2 (On Time_2 vs. Late_2); Negation 2 (Affirmative_2 vs. Negative_2); and Day (1, 2, or 3). The significant results of this analysis, where $p < .05$, in relation to the significant results of Days 1, 2, and 3, are indicated in Table 3.2.

In general, On Time_2 conditions were more difficult than Late_2 judgments, Term 2, $F(1,11) = 43.230$, $MS_e = .036$; and, conditions containing negative second sentences were more time consuming than those containing affirmative ones, Negation 2, $F(1,11) = 214.667$, $MS_e = .060$. The subjects' execution of responses also became more rapid over time, Day, $F(2,22) = 38.195$, $MS_e = .062$.

Consistent with the findings of Experiment 1, inner string mismatches were more difficult to process than inner string matches, Term 1 x Term 2, $F(1,11) = 14.958$, $MS_e = .185$. This interaction can be seen in Table 3.3. While this interaction was significant in all three days analyses (as indicated in Table 3.2), the differences between match and mismatch conditions on Day 1 (189 msec.) was greater than on Day 3

TABLE 3.2

Significant Results of the Analyses of Variance in Experiment 2

Condition	Overall	Day 1	Day 2	Day 3
Term 1	M			*
Negation 1		M		
Term 2	*	*	*	*
Negation 2	*	*	*	*
Day	*			
Term 1 x Term 2	*	*	*	*
Negation 1 x Day	*			
Term 2 x Negation 2	*		*	M
Term 1 x Negation 1 x Term 2				*
Term 1 x Term 2 x Negation 2	*	*	*	*
Term 1 x Term 2 x Day	*			
Term 1 x Negation 1 x Term 2 x Negation 2	*	*	*	*
Term 1 x Negation 1 x Term 2 x Day	*			

Where * denotes a significant result, $\underline{p} < .05$; and M denotes a marginally significant result

TABLE 3.3

Mean RT in Msec. to Verify
Inner String Matches and Mismatches

Term 2	Term 1		
	On $Time_1$	$Late_1$	Mean
On $Time_2$	1550	1713	1632
$Late_2$	1585	1470	1527
Mean	1568	1470	

TABLE 3.4

Mean RT in Msec. to Verify
Inner String Matches and Mismatches Over Days

	Day			
	Day 1	Day 2	Day 3	Mean
Match	1606	1490	1435	1510
Mismatch	1795	1626	1527	1649
Mean	1705	1558	1481	

(92 msec.), Term 1 x Term 2 x Day, $F(2,22) = 10.862$, $MS_e = .010$, as seen in Table 3.4.

While the main effect of Negation was not significant in the overall analysis, $F(1,11) = 1.015$, $MS_e = .035$, there was a significant Negation 1 x Day interaction, $F(2,22) = 4.056$, $MS_e = .008$. The effect of Negation 1 was marginally significant on Day 1, $F(1,11) = 4.145$, $MS_e = .02$, but decreased over days.

The two-way Negation 2 x Term 2 interaction was also significant, $F(1,11) = 14.132$, $MS_e = .010$. As can be seen in Table 3.5, the presence of a Negative_2 decreased the difference between On Time_2 and Late_2 which existed in the affirmative conditions (On Time_2-Late_2 = 136 msec.; vs. Not On Time_2-Not Late_2 = 73 msec.).

Negation 2 was also found to interact with inner string congruity, Term 1 x Term 2 x Negation 2, $F(1,11) = 21.047$, $MS_e = .024$. As indicated in Table 3.6, the difference between Affirmative_2 and Negative_2 conditions was greater when inner strings did not match (358 msec.) than when they were identical (241 msec.).

The remaining significant interactions were the following four-way ones: Term 1 x Negation 1 x Term 2 x Negation 2, $F(1,11) = 39.599$, $MS_e = .883$; and Term 1 x Negation 1 x Term 2 x Day, $F(2,22) = 9.611$, $MS_e = .009$. As can be seen in Table 3.2, the latter interaction represents the fact that the three-way interaction, Term 1 x Negation 1 x Term 2, was significant in the daily analyses only on Day 3.

In summary, comparisons which contained an On Time_2, a Negative_2, an inner string mismatch, or some combination of these factors were more time consuming for the subjects to process than their counterpart conditions (Late_2, Affirmative_2, inner string match, etc.).

TABLE 3.5

The Effect of Negation on RT in Msec.
in On $Time_2$ and $Late_2$ Conditions

Negation 2	Term 2		
	On $Time_2$	$Late_2$	Mean
$Affirmative_2$	1498	1362	1430
$Negative_2$	1766	1693	1730
Mean	1632	1528	

TABLE 3.6

The Effect of Negation of RT in Msec. to Conditions
Containing Inner String Matches and Mismatches

Negation 2	Terms 1 and 2		
	Match	Mismatch	Mean
$Affirmaitve_2$	1390	1470	1430
$Negative_2$	1631	1828	1730
Mean	1511	1649	

Day 3. Mean RT's per condition for Day 3 performance is contained in Table 3.8. Since the subjects's protocols did not provide information as to what kinds of strategies they may have adopted throughout processing, all the subjects' data were pooled. The overall error rate for Day 3 performance was 2.4%.

A 2 x 2 x 2 x 2 within-subjects analysis of variance was performed on the data. The variables were: Term 1 (On Time_1 vs. Late_1), Negation 1 (Affirmative_1 vs. Negative_1), Term 2 (On Time_2 vs. Late_2), and Negation 2 (Affirmative_2 vs. Negative_2). Significant results of this analysis are contained in Table 3.2.

As in the overall analysis, conditions containing On Time_2 and a Negative_2 were more time consuming than Late_2 and Affirmative_2 judgments, Term 2, $F(1,11) = 22.378$, $MS_e = .032$, and Negation 2, $F(1,11) = 156.468$, $MS_e = .032$, respectively. The main effect of Term 1 was also significant, $F(1,11) = 14.231$, $MS_e = .008$, indicating that On Time_1 conditions were more difficult to process than Late_1 conditions.

Inner string mismatches required significantly more processing time than inner string identities (Mean RT difference for Mismatch-Match comparisons = 92 msec.), Term 1 x Term 2, $F(1,11) = 9.09$, $MS_e = .044$. The effect of a Negative_2 was also found to be greater among inner string mismatch conditions (Mean RT difference between Negative_2-Affirmative_2 = 396 msec.) than among inner string identities (Mean RT difference between Negatvie_2-Affirmative_2 = 246 msec.), Term 1 x Term 2 x Negation 2, $F(1,11) = 21.409$, $MS_e = .013$. The four-way Term 1 x Negation 1 x Term 2 x Negation 2 interaction, $F(1,11) = 15.832$, $MS_e = .018$, was also significant.

The Term 1 x Negation 1 x Term 2 interaction was the remaining significant effect, $\underline{F}(1,11) = 14.784$, $MS_e = .013$. In those conditions in which the inner strings matched, Negative_1 conditions were longer than Affirmative_1 conditions by 107 msec. However, in conditions in which the inner strings did not match, Affirmative_1 conditions were 148 msec. longer than Negative_1 judgments.

This is a most interesting finding. To begin with, it is an aspect of the subjects' processing that is seen only in Day 3 performance. Furthermore, it suggests that when a negative appears in Sentence 1, and an inner string mismatch occurs, a subject may opt for a strategy that would ease his processing load, perhaps a conversion strategy.

Consider the following encoded comparison (25). The Term 1 x

(25) false (on time)
(late)

Negation 1 x Term 2 interaction suggests that "false(on time)" would be converted to "(late)"; the inner strings would then be compared and finding no mismatch, a true response would be executed (thus, no additional time for a truth index change would accrue). This operation is to be contrasted with the true model of negation's comparison of both inner and outer strings, resulting in two changes in the truth index and an increase in processing time.

It is important to note, however, that the Term 1 x Negation 1 x Term 2 interaction does not specify when the conversion operation takes place. It is possible that after both sentences are encoded, the subject will convert Sentence 1. It would be equally plausible to assume that the subject encodes a Negative_1 sentence as an affirmative and then continues processing. It may also be the case that a subject has both affirmative and negative representations equally available, the one chosen being dependent upon the second representation.

We shall now examine the results of the Day 3 analysis in light of the pilot study's findings.

Evaluation of the Pilot Model. Recall that in the pilot study data TA < FN < FA < TN in Affirmative_1-Assertion conditions. In the present experiment when Negative_1 entries were excluded from analysis, TA (1255 msec.) < FA (1412 msec.) < FN (1562 msec.) < TN (1715 msec.). The inclusion of Negative_1 conditions yielded the following ordering: TA (1255 msec.) < FA (1412 msec.) < TN (1519 msec.) < FN (1567 msec.). This finding was also not consistent with the results of the pilot study (TA < FA ≈ TN < FN). Thus a discrepancy exists between the pattern of RT's to equivalent conditions of the pilot study and Experiment 1.

Table 3.7 contains predictions of the pilot model regarding the ordering of certain conditions. Within each row of comparison, as one moves from left to right, RT is predicted to increase by virtue of adding one more parameter. As can be seen in the table, the pilot model encounters some difficulty in predicting conditions containing either a $\underline{d}$ or a $\underline{k}$ parameter. The model errs in predicting an increase in processing due to $\underline{d}$ when the preceeding condition contains a double negative. As for parameter $\underline{k}$, when $\underline{k}$ is associated with an On Time_2, errors in prediction occur.

Furthermore, the pilot model predicts that false double negative Assertions will be the most difficult condition. However, in Day 3 performance, the Late_1-Not On Time_2 condition (1811 msec.) was more time consuming than either false double negative Assertion (Not On Time_1-Not Late_2, 1683 msec.; Not Late_1-Not On Time_2, 1784 msec.).

TABLE 3.7

Pilot Model Predictions and Obtained Values in Msec.
for Selected Conditions of Day 3

Late$_1$-Late$_2$ t_o 1189	Late$_1$-On Time$_2$ t_o+c 1499		
Late$_1$-Late$_2$ t_o 1169	Not Late$_1$-On Time$_2$ t_o+k 1354		
Late$_1$-Late$_2$ t_o 1189	Not Late$_1$-Not Late$_2$ t_o+b 1528	Late$_1$-Not Late$_2$ t_o+b+d 1564*	
Late$_1$-Late$_2$ t_o 1189	Not Late$_1$-Late$_2$ t_o+d 1315	Late$_1$-Not Late$_2$ t_o+b+d 1564	
On Time$_1$- On Time$_2$ t_o+a 1320	Not On Time$_1$- On Time$_2$ t_o+a+d 1424	On Time$_1$- Not On Time$_2$ t_o+a+b+d 1559	Late$_1$- Not On Time$_2$ t_o+a+b+c+d 1811
On Time$_1$- On Time$_2$ t_o+a 1320	Late$_1$- On Time$_2$ t_o+a+c 1499		
On Time$_1$- On Time$_2$ t_o+a 1320	Not On Time$_1$- Not On Time$_2$ t_o+a+b 1581	On Time$_1$- Not On Time$_2$ t_o+a+b+d 1559*	Late$_1$- Not On Time$_2$ t_o+a+b+c+d 1811
On Time$_1$- On Time$_2$ t_o+a 1320	Not Late$_1$- On Time$_2$ t_o+a+k 1354*		

Where * denotes an inaccurate prediction

Given these considerations, the model constructed from the results of the pilot study was rejected as a candidate for Experiment 2 data representation. An attempt was made to characterize the data using the Young and Chase (1971) conversion model. However, this also proved fruitless. The model stipulates that all negatives would be converted to their equivalent affirmative form. As such, one would predict that, for example, Late_1-Not On Time_2 < Late_1-Not Late_2, or that On Time_1-Not Late_2 < On Time_1-Not On Time_2. Consulting Table 3.8 one will see that this is not the case (1811 vs. 1564 msec.; and 1619 vs. 1581 msec., respectively). Therefore, this approach was also abandoned.

Development of a New Model of Information Processing. It was decided to test a modified version of the Experiment 1 model in which *i* is eliminated (no encoding time for Inferences) and in which *x* is eliminated (since there is no *i*, it would not be necessary to have a s parate truth index changer for the Assertions alone). This version is pictured in Figure 3.1. Six parameters are proposed: t_o, a base processing time common to all comparisons; in accordance with the main effect of Term 2, parameter *a* will represent the difference between Late_2 and On Time_2 encoding times; parameter *b* will represent the amount of time needed to encode a Negative_2; parameter *n* will represent the additional time taken to encode a Negative_2 in the presence of a Negative_1 (that is, $+b+n$); parameter *c* will represent the amount of time taken to detect an inner string mismatch and change the value of the truth index; and parameter *d* will represent the amount of time needed to change the value of the truth index upon finding an outer string mismatch.

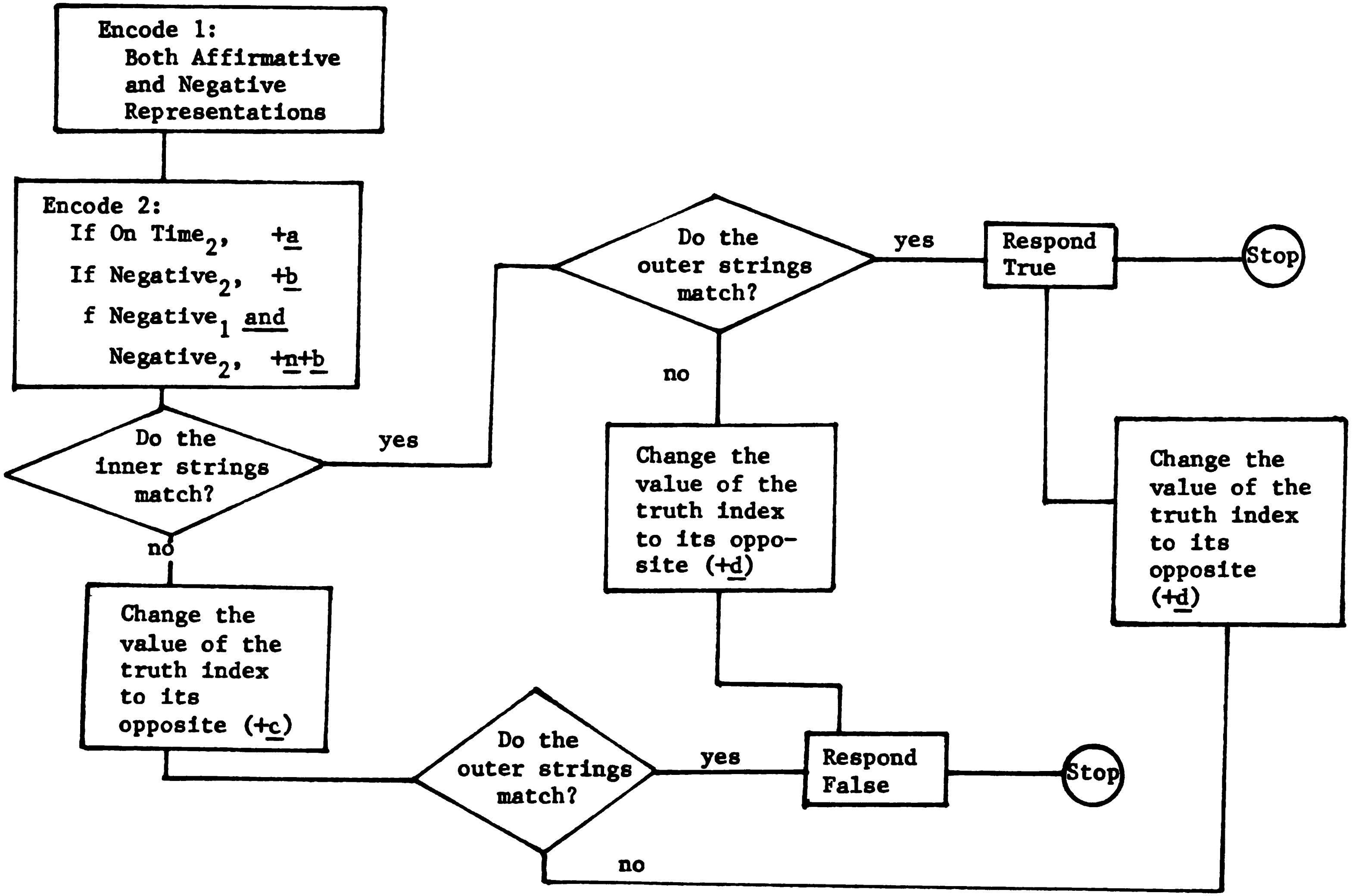

Figure 3.1. Processing Alternatives of the Model of Experiment 2

Processing time will be divided into four stages: Encode Sentence 1; Encode Sentence 2; Compare the encoded representations; Execute a response. According to an additive factor model, each stage is presumed to be independent of each other.

In Stage 1, subjects will encode Sentence 1. This particular stage is not usually the subject of detailed discussion; rather, the proposed internal representation is generally the only feature of this stage which is mentioned, for example (26).

(26) . . . not on time today.
false (on time)

Recall that the Term 1 x Negation 1 x Term 2 interaction suggested that subjects were using some sort of conversion strategy to ease processing demands. However, the rejection of a complete conversion model (that of Young and Chase) indicates that subjects do not convert negatives in every case. The following interpretation of these data was drawn. When encoding a Negative_1 sentence, it is proposed that two representations are constructed: the negative and its affirmative equivalent. Both representations are proposed to be equally available. Thus, one need not propose a conversion parameter to account for processing time. Which representation will be used in the comparison will be entirely determined by the nature of Sentence 2. Post-hoc examination of the Day 3 data indicated that subjects opted for the affirmative representation in the following two conditions: Not On Time_1-Late_2 and Not Late_1-On Time_2. It thus appeared that whenever an affirmative term was encoded in Sentence 2 which matched the affirmative representation in Store 1, subjects opted for the affirmative representation.

Sentence 2 will be encoded immediately following the end of Stage 1. It is only during this stage that the effects of parameters $\underline{a}$, $\underline{b}$, and $\underline{n}$ are seen. It is also important to note that the polarity of Sentence 1 differentially affects the encoding of Sentence 2. If Sentence 1 is affirmative, encoding a Negative_2 will take $+\underline{b}$ msec.; if a Negative_1 is stored in representation 1, encoding a Negative_2 will require $+\underline{b}+\underline{n}$ msec. more.

With the completion of the encoding phases, comparison operations to determine truth validity are commenced. It is proposed that the subject will first compare the inner strings of the representations. Finding an inner string mismatch and changing the value of the truth index will increase processing by time $+\underline{c}$. The outside strings will then be examined, with a mismatch and index change increasing processing time by time $+\underline{d}$.

The value of the truth index at the end of Stage 3 is then input to Stage 4 at which point a response (pushing a true or false button) is executed.

There are three major differences between the models of Experiments 1 and 2. First of all, with the elimination of the Inference condition, there was no longer any need to include a $+\underline{i}$ encoding parameter; similarly, there was no longer any need to include $+\underline{x}$ which operated differentially on Assertions and Inferences. The major difference between the models, and one which cannot readily be interpreted as the result of the exclusion of Inference conditions, is the generation of two Sentence 1 representations in Stage 1. It could be the case, however, that a lighter processing load (i.e., no Inferences) increased the likelihood that the subjects would generate more than one representation.

A least squares analysis using mean RT per condition was employed to estimate each of the proposed parameters. Parameter estimates to the nearest msec. were: $\underline{t_o}$, 1210; $\underline{a}$, 111; $\underline{b}$, 191; $\underline{c}$, 156; $\underline{d}$, 104; and $\underline{n}$, 110. Observed scores, condition estimates and their deviation can be found in Table 3.8. The RMSD was calculated in order to estimate the goodness of ift of the model. Since the RMSD was equal to 32 msec., and all parameter estimates were found to be well above this value, the model was interpreted as being a reliable characterization of subjects' processing (see Sternberg, 1969a,b).

As a further test of the parameters' reliability, an attempt was made to relate analysis of variance effects to the proposed parameters. Related measures $\underline{t}$-tests were also performed where necessary. Parameter $\underline{a}$, the time to encode an On Time_2, was significant according to the Term 2 main effect, as well as parameter $\underline{b}$ (Negation 2 main effect), parameter $\underline{c}$ (Term 1 x Term 2 interaction), and paramter $\underline{d}$ ($\underline{t}(11) = 4.200$, $\underline{p} < .01$). Unfortunately, the effect of parameter $\underline{n}$ could not be isolated in an interaction or by a related measures $\underline{t}$-test.

Upon finding that parameter $\underline{n}$ could not be isolated and tested for its significance, a least squares analysis using only parameters $\underline{t_o}$, $\underline{a}$, $\underline{b}$, $\underline{c}$, and $\underline{d}$ was performed on the data. This is, of course, the Clark and Chase (1972) model of negation. However, elimination of $\underline{n}$ resulted in a negative $\underline{d}$ value, a totally unacceptable result. Thus, the Clark and Chase model was abandoned in favor of maintaining $\underline{n}$.

While the model is able to represent the data of Experiment 2, there is a flaw in this approach. The model itself violates a basic assumption of the additive factor framework in at least two instances. The additive factor model has as its root the notion that each stage of

TABLE 3.8

Mean RT in Msec. Per Condition, Predicted Scores, and Their Deviations for Day 3 Performance

Conditions	Predicted Parameters	Predicted	Observed	Deviation
On Time-On Time	t_o+a	1321	1320	+ 1
On Time-Late	t_o+c	1366	1325	+ 41
On Time-Not On Time	t_o+a+b+d	1616	1559	+ 57
On Time-Not Late	t_o+b+c+d	1661	1619	+ 42
Late-On Time	t_o+a+c	1478	1499	- 21
Late-Late	t_o	1210	1189	+ 21
Late-Not On Time	t_o+a+b+c+d	1772	1811	- 39
Late-Not Late	t_o+b+d	1505	1564	- 59
Not On Time-On Time	t_o+a+d	1425	1424	+ 1
Not On Time-Late	t_o	1210	1220	- 10
Not On Time-Not On Time	t_o+a+b+n	1621	1581	+ 40
Not On Time-Not Late	t_o+b+c+n	1667	1683	- 16
Not Late-On Time	t_o+a	1321	1354	- 33
Not Late-Late	t_o+d	1314	1315	- 1
Not Late-Not On Time	t_o+b+c+d+n	1778	1784	- 6
Not Late-Not Late	t_o+b+n	1510	1528	- 18

processing is independent of every other stage. However, in order to represent the data, the model posits that the final representation of Stage 1 will not be chosen until Stage 2 is completed. Thus, Stage 1 is dependent upon Stage 2.

Furthermore, the encoding of a Negative_2 is dependent upon the contents of Stage 1. If an affirmative is in Store 1, encoding Negative_2 requires additional time $+b$ msec. If, however, a Negative_1 appears in Store 1, encoding is slowed down by $+n$ msec. (that is, Negative_2 requires $+b+n$ msec.).

In this experiment, we have been able to represent the data by a model. However, given that we have violated a basic tenet of the serial process position, one must be cautious in concluding that the additive factor approach is an adequate framework in which to interpret the results of all sentence-sentence verification studies.

Summary

Using an auditory sentence-sentence verification task to test people's comprehension of Assertions, it was found that On Time_2 comparisons were more difficult than Late_2 comparisons, and that comparisons containing a Negative_2, or an inner string mismatch were more time consuming than Affirmative_2 or inner string identity comparisons, respectively. These findings, in conjunction with the findings of Experiment 1, provide support for the psychological reality of these factors.

It was also the case that aspects of Sentence 1 played a larger role in subjects' processing in Experiment 2. It was suggested that as processing capacity was freed during the task, subjects' RT's may have become more sensitive to new stimulus parameters.

An attempt was made to model the data of Experiment 2 according to an additive factor model. While a model was constructed that was able to capture the overall condition orderings, problems inherent in the model force one to interpret with caution the modeling attempts of Experiment 2.

CHAPTER 4

Experiment 3

The results of Experiment 2 indicated that even simple auditory language comprehension tasks involve more complex operations than the sentence-picture tasks. A feature of the preceding experiments which may be responsible for those findings is the nature of the task used. In both Experiments 1 and 2, auditory presentations were employed. Recall, however, that earlier sentence-picture models (Clark & Chase, 1972; Young & Chase, 1971) were based on RT to visual stimuli. Therefore, it is possible that the differences in processing that were found in Experiments 1 and 2 are the result of modality differences rather than the result of the use of complex language materials.

Experiment 3 was designed to investigate sentence-sentence verification processing using a visual RT task. With the exception of the change in stimulus modality, Experiment 3 is essentially a replication of Experiment 2.

Method

Subjects. The subjects were 12 male and female undergraduates participating for credit in an introductory psychology class at Michigan State University.

Stimuli. Two hundred fifty-six declarative sentences in pica-size type print (black print, white background) appearing on slides served as stimuli, and were presented in two sentence test sequences. Sentence 1 was a simple declarative statement of fact. One-half of the sentences contained the term On Time, the remaining sentences included Late. Sentence 1 also varied as to the presence or absence of negation: one-half of the sentences were affirmative, one-half, negative. Sentence 2 was a simple verification sentence varying with respect to the same factors manipulated in Sentence 1. The design yields 16 possible combinations (see Table 3.1, page 51).

Apparatus. Two slide projectors arranged vertically were used to project slides onto a screen located four feet in front of the subjects. Projector 1 projected Sentence 1 onto the top of the screen; Projector 2 projected the second sentence onto the lower portion. Sentences were separated by 18 inches. Automatically operated Hunter timers were used to record the subjects' RT's, recordings beginning concurrently with each stimulus onset. A three button panel was located directly in front of the subject: a button marked "C" (Comprehension) was centrally located, with a "T" (True) or "F" (False) button positioned to either side. Depression of the "C" button activated Projector 2. Depression of any button stopped the Hunter timers.

Procedure. All the subjects were individually tested in three performance sessions. Order of the stimulus presentations was randomly varied over the three sessions so that no subject received the same input order twice. All instructions were administered verbally by the Experimenter. Participants were instructed to carefully read each sentence presented, in order to make accurate true-false judgments of each sentence-sentence pair. The Experimenter initiated Trial 1. The subjects were required to depress the "C" button as quickly as possible following the presentation of Sentence 1, indicating that they understood the sentence. Sentence 2 immediately followed. The subjects were instructed to depress either the T or F button, depending upon their evaluation of the comparison. The position of the T and F buttons was randomly varied between subjects. Both the speed and accuracy of the response were emphasized. A five second interval occurred between each test comparison. The Experimenter recorded all responses.

Eight practice trials were administered prior to actual testing. The subjects were advised of their accuracy before proceeding to the actual test. In addition, prior to each session, the subjects were advised of their previous performance. Upon completion of the third session of 128 sentence comparisons, the subjects were asked to specify strategies used during processing, if any. The subjects were also asked to indicate whether they were aware of any change in the strategy selected across the three sessions of testing, to state the nature of this change, and the proposed reason for it.

Results and Discussion

The significant results of an analysis of the Experiment 3 data across all three days, and of a separate analysis of the Day 3 data will be presented and discussed in the following section. As in the previous experiments, an attempt will be made to interpret the Day 3 data in terms of an additive factor model. Day 3 data was specifically examined because it was felt that if the subjects were processing information according to a common strategy, this pattern would most likely be seen, and be most stable, in the most practiced Day 3 data.

No specific discussion will be made of the results of the Day 1 and Day 2 performance analyses, as all the significant results of these two performances were significant on Day 3. However, the mean RT's per condition for Day 1 and Day 2 can be found in Appendices G and H, respectively. The significant $\underline{F}$ values of the within-subject analyses of variance are reported in Appendix I.

Overall. Only correct responses were included in the analysis; individual data $\pm$ 3 standard deviations from the mean were interpreted as errors. The overall error rate was 6.2%.

A 2 x 2 x 2 x 2 x 3 within-subject analysis of variance was performed on the data. The factors were: Term 1 (On Time_1 vs. Late_1); Negation 1 (Affirmative_1 vs. Negative_1); Term 2 (On Time_2 vs. Late_2); Negation 2 (Affirmative_2 vs. Negative_2); and Day (1, 2, or 3). The significant results of this analysis, where $\underline{p} < .05$, are indicated in Table 4.1. The significant results of the analyses on the Day 1, Day 2, and Day 3 data are also included in this table.

In general, conditions which contained a negative in either Sentence 1 or 2 were more time consuming than their affirmative counterpart conditions, Negation 1, $\underline{F}(1,11) = 24.476$, $MS_e = .163$; and Negation 2, $\underline{F}(1,11) = 47.816$, $MS_e = .186$. Although the Negation 2 main effect was significant in each separate day analysis, the difficulty of a Negative_2 was greater on Days 1 and 2 than on Day 3, Negation 2 x Day, $\underline{F}(2, 22) = 4.958$, $MS_e = .070$ (Mean RT difference between Negative_2-Affirmative_2 for Day 1 = 328 msec.; for Day 2 = 260 msec.; and for Day 3 = 158 msec.).

As found in Experiments 1 and 2, On Time_2 conditions were more time consuming than Late_2 conditions, Term 2, $\underline{F}(1,11) = 47.552$, $MS_e = .037$. The subjects' performance was also found to become significantly faster over days, Day, $\underline{F}(2,22) = 36.180$, $MS_e = .249$.

TABLE 4.1

Significant Results of Analyses of Experiment 3

Condition	Overall	Day 1	Day 2	Day 3
Term 1				*
Negation 1	*	*	*	*
Term 2	*	*	*	*
Negation 2	*	*	*	*
Day	*			
Term 1 x Term 2	*	*	*	*
Negation 1 x Negation 2				M
Term 2 x Day	M			
Negation 2 x Day	*			
Term 1 x Negation 1 x Term 2	M			
Term 1 x Negation 1 x Negation 2				*
Term 1 x Term 2 x Negation 2	*	*	*	*
Term 1 x Negation 1 x Term 2 x Negation 2	*	*	*	*
Term 1 x Term 2 x Negation 2 x Day	*			

Where * denotes a significant result ($p < .05$); and M denotes a marginally significant effect

As indicated in Table 4.2, inner string mismatches were more difficult than inner string identities, Term 1 x Term 2, $\underline{F}(1,11) = 62.373$, $MS_e = .628$. Furthermore, the effect of a $Negative_2$ was greater among the mismatch conditions than in identity conditions, Term 1 x Term 2 x Negation 2, $\underline{F}(1,11) = 21.501$, $MS_e = .200$. This effect is illustrated in Table 4.3.

While the Term 1 x Term 2 x Negation 2 interaction was significant in all performances (see Table 4.1), the effect of a $Negative_2$ on the difference between Match and Mismatch conditions was greater on Day 1 (Mean RT difference of Mismatch-Match = 499 msec.) than on either Day 2 or Day 3 (280 msec., and 256 msec., respectively), Term 1 x Term 2 x Negation 2 x Day, $\underline{F}(2,22) = 4.696$, $MS_e = .045$.

The remaining significant result of the overall analysis was the four-way Term 1 x Negation 1 x Term 2 x Negation 2 interaction, $\underline{F}(1,11) = 31.822$, $MS_e = .285$. This effect was also observed in the analyses of the Experiment 2 data.

Day 3. Examination of the subjects protocols did not yield a suitable criterion by which to divide the subjects' performances according to strategies employed. Therefore, all the subjects' data were pooled for the Day 3 analysis. Mean RT's for each condition are contained in Table 4.6. The overall error rate for the Day 3 performance was 5.1%.

A 2 x 2 x 2 x 2 within-subjects analysis of variance was performed on the data. The factors were: Term 1 (On $Time_1$ vs. $Late_1$); Negation 1 ($Affirmaitve_1$ vs. $Negative_1$); Term 2 (On $Time_2$ vs. $Late_2$); and Negation 2 ($Affirmative_2$ vs. $Negative_2$). The significant results of this analysis, where $\underline{p} < .05$, are indicated in Table 4.1.

TABLE 4.2

Mean RT in Msec. to Inner String Identities and Mismatches

Term 2	Term 1		
	On $Time_1$	$Late_1$	Mean
On $Time_2$	1555	2042	1799
$Late_2$	1966	1410	1688
Mean	1761	1726	

TABLE 4.3

Effect of a $Negative_2$ on RT in Msec. to
Inner String Match and Mismatch Conditions

Negation 2	Inner Strings		
	Match	Mismatch	Mean
$Affirmative_2$	1445	1794	1620
$Negative_2$	1521	2215	1868
Mean	1483	2005	

As in the overall analysis, conditions containing either a Negative_1 or a Negative_2 were more time consuming than Affirmaitve_1 and Affirmative_2 conditions, respectively, Negation 1, $F(1,11) = 14.930$, $MS_e = .096$; and Negation 2, $F(1,11) = 23.438$, $MS_e = .051$. On Time_2 conditions were also found to be more difficult than Late_2 conditions, Term 2, $F(1,11) = 7.660$, $MS_e = .033$. In the Day 3 analysis, however, conditions containing an On Time in Sentence 1 were also more time consuming than those containing Late_1, Term 1, $F(1,11) = 9.581$, $MS_e = .042$.

Inner string mismatches were more time consuming than inner string identities, Term 1 x Term 2, $F(1,11) = 47.339$, $MS_e = .271$. Once again, the effect of a Negative_2 was greater upon the mismatch conditions (Mean RT difference between Negative_2-Affirmative_2 conditions = 287 msec.) than the inner string identities (30 msec.), Term 1 x Term 2 x Negation 2, $F(1,11) = 11.461$, $MS_e = .069$.

The Term 1 x Negation 1 x Negation 2 interaction was also significant, $F(1,11) = 5.572$, $MS_e = .023$. As indicated in Table 4.4, the difference in mean RT between outer string mismatches and identities was greater when an On Time appeared in Sentence 1 (Mean RT difference between identities and mismatches = 131 msec.) than when a Late_1 appeared there (Mean RT difference = 27 msec.).

The four-way Term 1 x Negation 1 x Term 2 x Negation 2 interaction was the remaining significant result, $F(1,11) = 10.806$, $MS_e = .211$. The two-way Negation 1 x Negation 2 interaction was also marginally significant, $F(1,11) = 4.112$, $MS_e = .072$.

<u>Comparison of Experiment 3 with Experiments 1 and 2</u>. In all three experiments which we have discussed, certain aspects of processing were constant across a change in task demands and a change in stimulus

TABLE 4.4

Effect of the Level of Term 1 on the Verification of Outer String Matches and Mismatches in Msec.

Term 1	Outer Strings		
	Match	Mismatch	Mean
On $Time_1$	1531	1662	1597
$Late_1$	1492	1519	1506
Mean	1512	1591	

modality: On $Time_2$ conditions were more difficult than $Late_2$ conditions; conditions containing a $Negative_2$ were more difficult than those containing an $Affirmative_2$; and inner string mismatches were more difficult than inner string identities.

We saw, however, that as the processing load of the subject was reduced, in terms of the task demands of Experiment 1 and Experiment 2 (i.e., the removal of Inference conditions), the polarity of the first sentence came to play an important role in processing.

With the change from auditory to visual stimuli, it is now the case that both the term and the polarity of Sentence 1 have important effects on subsequent processing. In Experiment 3, Sentence 1 was present during the presentation of Sentence 2. Therefore, it is possible that the effect of Term 1 was really the result of its continuous presence, rather than the result of it being a visual stimulus. Contrasting the results of the present experiment with a visual RT experiment in which the first sentence is removed from view following comprehension would provide some insight to this question. In terms of this study, however, the key feature to note is that aspects of Sentence 1 play a greater role in determining the ultimate processing time.

Another interesting feature of the change from an auditory to a visual modality task is the increased importance of pattern matching. In Experiment 3, by Day 3 <u>all</u> conditions in which Sentence 1 and Sentence 2 matched were faster than any other condition. This was not the case in the previous experiments. In fact, in the earlier studies On Time_1-Late_2 mismatches were generally as fast or faster than On Time_1-On Time_2 identities.

<u>Applicability of Past Models: The Pilot Study</u>. Recall that in the pilot study the following ordering of data among Affirmative_1-Assertions was found: TA < FN < FA < TN. In the present study on Day 3, TA (1083 msec.) < FN (1409 msec.) < FA (1611 msec.) < TN (1758 msec.). This same ordering was obtained with the Day 1 data.

Inclusion of Negative_1 conditions produced the following ordering of the pilot study conditions: TA < FA ≈ TN < FN. Among the Experiment 3-Day 3 data, TA (1083 msec.) < TN (1562 msec.) < FA (1611 msec.) < FN (1677 msec.). The ordering was also found among the Day 1 data. While the latter comparison did not match the results of the pilot study, it must be noted that for at least the Affirmative_1-Assertion conditions, subjects in the pilot study and subjects in Experiment 3 seemed to be responding in the same manner. It should also be noted that the false double negative Assertion which was the most difficult condition of the pilot study was also the most difficult condition of the Experiment 3, Day 1 and Day 3 data.

The Young and Chase Conversion Model. Recall that this conversion model predicted that a negative entry would automatically be encoded as its affirmative equivalent. As such, if a negative were in Store 1, it should automatically be changed to an Affirmative_1; for example, Not On Time_1 would be converted to Late_1. Therefore, one would predict that there would be no difference in RT between Not On Time_1-Late_2 comparisons and Late_1-Late_2 comparisons. Table 4.5 contains some of the predictions that this model would make. As can be seen from the table, in none of the cases cited do the results of Experiment 3 match the predictions. Therefore, the Young and Chase conversion model was abandoned as a candidate for characterizing the results of Experiment 3.

TABLE 4.5

Predictions of the Young and Chase (1971) Conversion Model and the Actual Results of Experiment 3

Prediction			Obtained Times (Msec.)
Not On Time_1- Late_2	=	Late_1 Late_2	1754 ≠ 1011
Not On Time_1 On Time_2	=	Late_1 On Time_2	1566 ≠ 1680
Not Late_1- On Time_2	=	On Time_1 On Time_2	1690 ≠ 1154
Not Late_1 Late_2	=	On Time_1 Late_2	1379 ≠ 1541

The True Model. The true model of negation predicts that TA < FA < FN < TN. As indicated in the evaluation of the pilot model, this ordering was not obtained. Furthermore, the true model predicts that conditions which contain both an inner string and an outer string mismatch will be more time consuming than false double negative Assertions. This was not the case in Day 1, 2, or 3. Thus, the true model of negation was also rejected.

The Models of Experiments 1 and 2. Recall from Experiment 1 that Affirmative_1-Assertion verifications were ordered in the following manner: TA < FA < FN < TN, while the inclusion of Negative_1-Assertions yielded: TA < FA < TN ≈ FN. In Experiment 2, TA < FA < FN < TN among Affirmative_1-Assertions and TA < FA < TN < FN when Negative_1-Assertions were included in the comparison. Since none of these orderings match the obtained condition orderings of Experiment 3, it must be concluded that neither model (constructed to characterize each experiment's pattern of responses) will be able to accurately account for the data of the present experiment. Thus, both models were rejected as possible candidates.

The Development of a New Model. Although the pilot study model proved to be a better characterization of Experiment 3 performance than any of the other models considered, its failure to adequately characterize condition orderings including Negative_1-Assertions indicated that certain aspects of processing were being overlooked. Hence, an attempt was made to construct a model which would more accurately reflect features of the subjects' processing.

Past models of information processing have all concluded that it takes one significantly longer to encode a negative than an affirmative;

it should, therefore, take longer to encode "Not On Time" than "On Time", and encoding "Not Late" should take more time than "Late". In the present experiment since RT was recorded from the onset of Sentence 2, it should be the case that: On $Time_1$-Not On $Time_2$ RT > Not On $Time_1$-On $Time_2$ RT; that $Late_1$-Not $Late_2$ RT > Not $Late_1$-$Late_2$ RT; that $Late_1$-Not On $Time_2$ RT > Not On $Time_1$-$Late_2$ RT; and that On $Time_1$-Not $Late_2$ RT > Not $Late_1$-On $Time_2$ RT by at least a factor of encoding a negative, or, as referred to in most models, $+\underline{b}$ (that is, the second comparison of each pair should be longer than the first due to the presence of a negative in Sentence 2).

Inspection of Table 4.6 will reveal that this ordering was only true of the On $Time_1$-Not $Late_2$, Not $Late_1$-On $Time_2$ judgments. Thus, any model which attempts to characterize the data of Experiment 3 will have to deal with this result.

The finding that three pairs of the conditions were the reverse of what was predicted or that members of the pairs were equivalent is counterintuitive. The only explanation which seems feasible is that since Sentence 1 remained in view during the entire course of processing, the subjects encoded Sentence 2 and then returned to Sentence 1 to carry out further operations. Thus, if a negative appeared in Sentence 1, additional time would be reflected in the overall RT for the re-encoding of $Negative_1$.

If this assumption is correct, then if one were to contrast a condition in which Sentence 1 is affirmative and Sentence 2 is negative, with the reverse condition where Sentence 1 is negative and Sentence 2 is affirmative, there would be little difference in processing time, as the $Negative_2$ of the former condition would cancel out the $Negative_1$

of the latter. The major problem with this explanation is that it does not explain why On $Time_1$-Not $Late_2$ > Not $Late_1$-On $Time_2$. That is, there is no a priori rule by which to explain subjects' processing.

Keeping these findings in mind, as well as the failure of other models to predict the results of Experiment 3, an attempt was made to construct a model which would characterize the Day 3 results. A four-stage, five parameter model was devised. The processing alternatives of the model are featured in Figure 4.1.

Information processing is divided among two encoding stages (a separate stage for each sentence), a comparison stage, and a response execution stage. In Stage 1, the subjects encode Sentence 1. It is assumed that the subjects do not perform any conversion operations during this stage, given that the term and polarity values of Sentence 1 may be rechecked at any time during the verification process. Furthermore, the three-way Term 1 x Negation 1 x Term 2 interaction, which suggested alternative representations in Experiment 2, was not significant in the Experiment 3-Day 3 results. Upon completion of Sentence 1 encoding, the subject will begin Stage 2 processing. Since Sentence 1 encoding is not part of the RT of interest in the present study, no parameters accounting for its time course are proposed.

In Stage 2, Sentence 2 is encoded. It is proposed that sentences containing On $Time_2$ will take $+\underline{a}$ msec. longer to encode than $Late_2$ sentences, as suggested by the Term 2 main effect.

In trying to fit the data to particular models, it was noticed that the Not On $Time_1$-$Late_2$ and Not On $Time_1$-Not $Late_2$ conditions were usually poorly predicted. In each case, the error in the predictions was enough to suggest that another parameter was involved. In order to

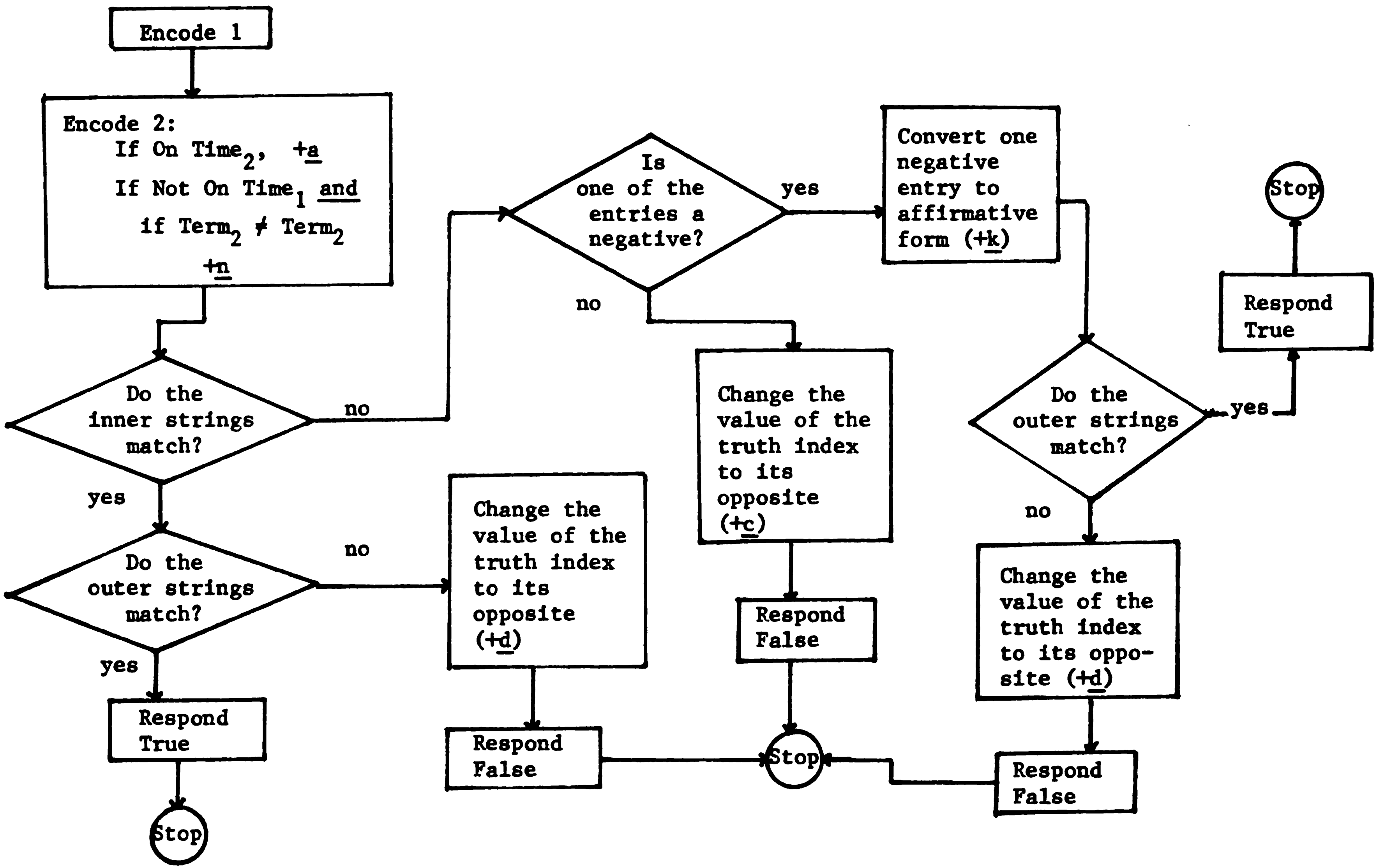

Figure 4.1. Processing Alternatives of the Model of Experiment 3

accommodate this problem, paramter *n* is proposed. If a subject has stored both a negative and an On Time in Store 1, and if the visual properties of the Term 2 factor are different from Term 1, second stage encoding is believed to be prolonged by +*n* msec.

The current model does not propose a negation encoding parameter. As suggested earlier, the effects of negation encoding tend to wash out over comparable conditions, for example, Not On Time_1-On Time_2 and On Time_1-Not On Time_2. However, an attempt was made to model the data using a negative encoding parameter which would capture Negative_2 encoding and the re-encoding of Negative_1. The effort was unsuccessful, resulting in negative estimates of both t_o, the base processing time, and *d*, the outer string mismatch parameter. A model showing no differential effects of negation was therefore decided upon.

Once the encoding of the second sentence is completed, the subject will begin Stage 3 or the comparison process. The inner strings of the two representations will first be compared. If a match is found, the values of the outer strings will be compared. If a match is again found, the subject will respond True. This comparison, for example (27), is

(27)	(late) (late)	t_o

proposed to take t_o time, a base processing time. If, however, a mismatch of the outer strings is detected, as in (28), it is proposed that

(28)	(late) false (late)	t_o+d

the subject will change the value of the truth index to False, a process requiring +*d* time.

At the time of the inner string comparison, if the subject had found a mismatch, it is proposed that he would have then checked to see if one of the entries was a negative. A no response (29) would result

(29) (on time) t_o+c
(late)

in the subject changing the value of the truth index to False, an operation requiring +c msec. If a negative had been found, it would have been converted to its affirmative equivalent (+k time) and the outer strings would then be checked. A match in this case (30) would

(30) (on time) t_o+k
false (late)

result in a true response. A mismatch (31), however, would result in a

(31) false (on time) t_o+n+k+d
false (late)

change of the truth index value, adding +d msec.

Finally, in Stage 4, the value of the truth index is indicated by the depression of a True or False button.

A least squares analysis was performed on the actual data in order to arrive at parameter estimates. The analysis yielded the following estimates to the nearest msec.: t_o, 1076; a, 103; c, 483; d, 329; k, 598; and n, 121. Observed scores, predicted values for each condition and their deviation from the actual case can be seen in Table 4.6. The RMSD was obtained to test the goodness of fit of the model, RMSD = 66 msec. Since the RMSD was lower than any of the predicted parameters, the parameters and their condition estimates were believed to be accurate reflections of the subjects' processing.

An attempt was made to test the significance of the individual parameters. Parameter a corresponded to the significant Term 2 main effect, and parameter c was reflected in the significant Term 1 x Term 2 interaction. While parameters d and k were not reflected directly in the analysis of variance effects, related measures t-tests did show

TABLE 4.6

Mean RT in Msec. Per Condition, Predicted Scores and Their Deviation for Day 3 Performance

Condition	Predicted Parameters	Observed	Predicted	Deviation
On Time-On Time	t_o+a	1154	1179	+ 25
On Time-Late	t_o+c	1541	1559	+ 18
On Time-Not On Time	t_o+a+d	1513	1508	- 5
On Time-Not Late	t_o+k	1814	1674	-140
Late-On Time	t_o+a+c	1680	1662	- 18
Late-Late	t_o	1011	1076	+ 65
Late-Not On Time	t_o+a+k	1701	1777	+ 76
Late-Not Late	t_o+d	1305	1405	+100
Not On Time-On Time	t_o+a+d	1566	1508	- 58
Not On Time-Late	t_o+k+n	1754	1795	+ 41
Not On Time-Not On Time	t_o+a	1265	1179	- 86
Not On Time-Not Late	$t_o+d+k+n$	2166	2125	- 41
Not Late-On Time	t_o+a+k	1690	1777	+ 87
Not Late-Late	t_o+d	1379	1405	+ 26
Not Late-Not On Time	$t_o+a+d+k$	2131	2107	- 24
Not Late-Not Late	t_o	1146	1076	- 70

each to be significant effects (d: $t(11) = 4.205$, $p < .01$; k: $t(11) = 7.953$, $p < .001$).

A problem arose in the test of parameter n. N was not directly reflected in the significant interactions of the analysis of variance, and while four test conditions were predicted to contain parameter n, its significance could only be tested using a related measures t-test contrasting t_o+k+n (Not On Time$_1$-Late$_2$) and t_o+k (On Time$_1$-Not Late$_2$). Parameter n was not significant, $t(11) = .58$, $p > .05$. Examination of Table 4.6 reveals that of all of the conditions, the On Time$_1$-Not Late$_2$ condition was the one most poorly predicted by the model (being underestimated by 140 msec.). It would appear that some other parameter is operating in this condition, thus making the Not On Time$_1$-Late$_2$:On Time$_1$-Not Late$_2$ contrast inappropriate for the testing of n.

However, elimination of paramter n is not an acceptable solution. This procedure resulted in parameters a and d becoming unreliable predictors of behavior; furthermore, estimations of conditions become in error by as much as 290 msec. in some cases. The existence of parameter n is indeed questionable; however, since its elimination resulted in a poorer prediction of performance, the decision was made to allow parameter n to remain as an encoding parameter of the model.

Summary

Important aspects of visual sentence processing have been demonstrated in Experiment 3: encoding a negative does consume a significant portion of a subjects' processing time and mismatches of either term or polarity are more difficult than identities. This last point was especially interesting in this experiment for all identities (regardless of whether they contained a negative) were processed more rapidly than any mismatch. This was not always the case in the two previous experiments. Thus, the subjects may be able to make better use of pattern matching strategies in the visual than the auditory modality.

Failure to find strong support for the true or conversion models of negation in the results of Experiment 3 suggests that sentence-sentence verification regardless of modality is a more complex task than the sentence-picture verification task. As an alternative, a four-stage, five-parameter model was proposed to account for the data of Experiment 3. The model employed conversion operations, and was able to give a reasonably good representation of the results of the present study.

CHAPTER 5

General Discussion

An attempt will now be made to integrate the findings of Experiments 1, 2, and 3, so as to better understand sentence processing in general, and to appreciate the issues which underlie its study.

The Psychological Reality of Processing Variables

It is clear that the results of Experiments 1, 2, and 3 lend credence to the psychological reality of certain psychological phenomena, given the variety of experimental manipulations in which they are found.

In all three experiments, the paramters of Sentence 2 had dramatic effects on processing times. On Time_2 comparisons were more time consuming than Late_2 comparisons, and Negative_2 comparisons required more processing time than Affirmative_2 judgments. While it cannot be directly shown in the present findings, the former result is probably largely the result of the presentation time parameters of the Term 2 stimuli (On Time_2 being approximately 300 msec. longer than Late_2).

That an effect of negation was seen in Sentence 2 was a most encouraging result. Clark and Chase (1972) had posited separate encoding and comparison stage effects of negation, but the paradigm employed in that study did not permit their isolation. In the present series of experiments, this problem was avoided; yet, the effects of negation in encoding and/or comparison stages remained. Thus, negation regardless of which stage of the verification process it affects, is psychologically more complex to comprehend that its affirmative counterpart.

In addition, Experiment 1 of the series provided evidence for the psychological reality of Inferences. It took time to verify information which was not directly presented in a story, even when that information could be readily and easily deduced from the facts presented.

The results of the pilot study indicated that the majority of subjects generated Inferences appropriate to a particular context, and then executed True-False judgments on the basis of that deduction. Experiment 1's results, on the other hand, argued for a decision rule in which generation of the actual Inference was not necessary. Rather, decision making depended upon the total number of mismatches encountered in a comparison and upon whether an Inference or an Assertion judgment was being called for.

While there were methodological differences between the two studies, both strategies are viable options for successful completion of the task. Thus, it would be interesting to determine under what conditions a particular strategy would be chosen, and what advantages (in terms of time gained, memorability of items, etc.) exist in opting for one strategy as opposed to the other. At the very least, the existence of two very different strategies as a means of solving the task at hand emphasizes the complex nature of the language comprehension task.

Another phenomenon reported by Clark and Chase (1972) and Young and Chase (1971) consistently found in the present experiments was the Term 1 x Term 2 interaction, or the inner string match-mismatch effect. Simply stated, recognizing an inner string mismatch (that On $\text{Time}_1 \neq \text{Late}_2$) is a more time consuming process than recognizing an inner string match (that On $\text{Time}_1 = \text{On Time}_2$). The results of the present series of experiments were interesting in that this effect was found to be significant in both the auditory and visual modalities. One can conclude, therefore, that in general, difference relationships are cognitively more complex to the subject (or the information processor) than sameness relationships.

The results of the experiments yielded other interesting findings. Despite the similarities in responding across the two modalities tested (e.g., Term 2 and Negation 2 main effects, the Term 1 x Term 2 interaction), there were some strong differences. In the visual modality, the effects of pattern recognition were clearly seen: all true comparisons in which Sentences 1 and 2 were identical were responded to more rapidly than any other comparisons. This was not true of the auditory modality, in which some FA comparisons (On Time_1-Late_2) were responded to more rapidly than True identical comparisons (On Time_1-On Time_2).

It is possible that this difference was not the result of a change in modality, but rather was due to a change in stimulus presentation co-occurring with the modality switch. Visual presentation lends itself to either a simultaneous or successive framework, and in Experiment 3, simultaneous presentation of the stimuli was used. Auditory presentation in sentence-sentence verification tasks, on the other hand, demands that the stimuli occur successively, thus, the first stimulus could never co-occur with the second. Research is currently underway to determine whether the pattern recognition effect observed in Experiment 3 was the result of the manner of stimulus presentation or the result of the modality employed.

If we examine the task demands of Experiments 1, 2, and 3, it appears that with each successive experiment, the processing demands made upon the subject decreased. Experiment 1 required that the subjects make verification judgments of both explicit and implicit information, while Experiment 2 eliminated the verification of all implicit information. Experiment 3, while in spirit a replication of Experiment 2, was constructed so that subjects were not required to store _any_ information in their heads, that is, the stimuli were always available to them.

An interesting corollary of this trend was the growing effect of Sentence 1 parameters on processing time from Experiment 1 to Experiment 3. In the first experiment, no parameters of Sentence 1 significantly affected processing on Day 3, whereas in Experiment 2, Term 1 was significant. In Experiment 3-Day 3 performance, both Term 1 and Negation 1 significantly affected processing time. Furthermore, in Experiment 1, only two higher order interactions were significant on Day 3, while in the other experiments, twice that number were significant.

It is possible that as the processing demands of an experimental paradigm are lightened, freed capacity is then directed to intricate aspects of the stimuli. Experiments which by their very nature consume all of the subjects' available capacity (e.g., Experiment 1) would not be expected to be as sensitive to all stimulus parameters. Further experimentation manipulating task demands would shed light upon this hypothesis.

Usefulness of the Additive Factor Model Approach

It was the case in the experiments described herein that while the true and conversion models of negation were able to capture certain aspects of the data in each instance, they were not as accurate in their predictions as they were in the sentence-picture verification experiments. In each experiment, new models of information processing were generated which yielded better representations of the condition orderings.

There was a problem with the violation of the independent stage principle in the model of Experiment 2. Recall that it was proposed that before one could choose between an affirmative and negative representation in Stage 1, Sentence 2 encoding had to be completed. Thus, Representation 1 was found to be dependent upon Representation 2. This could be a serious problem depending upon how intricate an additive factor one wishes to construct.

If one wishes to account for every aspect of a subject's processing, then this finding would raise some doubt concerning the applicability of additive factor models to sentence-sentence verification tasks. However, if one is interested in capturing basic features of processing, then this finding does not affect additive factor model applicability.

The four models proposed herein have one thing in common: they all can be represented in terms of an encoding phase, a comparison phase, and a response execution phase. The intricate differences between the models arose from an attempt to capture the details of processing corresponding to different task demands. If one examines the nature of the independent stage principle violation, one sees that it occurs <u>within</u> a stage, thus preserving the overall serial structure of processing. It would appear, therefore, that between stage processing occurs in a serial stage fashion, whereas the possibility exists for either serial or parallel processing within a stage.

In conclusion, additive factor models are useful tools by which to interpret the results of sentence-sentence verification tasks in so far as they are able to capture the most fundamental features of processing: encoding and comparing.

Other Issues

The issue of the nature of the internal representation of stimuli is always a concern in verification studies, but the problems associated with this issue were particularly relevant in the consideration of the results of Experiment 2. It is clear that considerably more research must be devoted to the problem of internal representation. At present, we have no rule system available which defines a priori how a subject will encode a sentence. Indeed, we do not even know the variety of ways a sentence may be encoded, and whether experimental paramters, if any, would favor one format over another. How one encodes information, however, is a critical factor in the subsequent processing of that information. Thus, in order to gain insight to the operations involved in sentence comprehension, knowledge of encoding formats must be obtained.

Another issue which the series of experiments raises is what is meant by language comprehension. Clark and Chase (1972) assumed that the operations reported in the sentence-picture studies were basic to any language comprehension situation. The results of the present study support this assumption to a certain degree: negatives and inner string mismatches were more difficult than affirmatives and inner string identities, respectively; and processing could be divided into encoding and comparison stages. If the modeling interpretations of the pilot study and Experiments 1, 2, and 3 are correct, however, it must be the case that sentence-sentence processing is a more complex task than sentence-picture verification, as evidenced by the variety of operations and combinations of operations employed. At the very least, in terms of the alternative operations available to the processor, sentence-sentence processing appears to be quite sensitive to particular task demands.

What is needed is a delineation of those processes which are necessary for comprehension to occur. The issue now becomes one of whether or not the complexities of the individual sentence-sentence task schemes (i.e., models) were essential for comprehension. In other words, one must determine whether comprehension in the sentence-picture task can be defined in the same manner (or equated with) comprehension in the sentence-sentence task.

APPENDICES

APPENDIX A

APPENDIX A

Mean RT in Msec. Per Condition on Day 1
Experiment 1

		Assertion				Inference			
		On $Time_1$		$Late_1$		On $Time_1$		$Late_1$	
		Aff_1	Neg_1	Aff_1	Neg_1	Aff_1	Neg_1	Aff_1	Neg_1
On $Time_2$	Aff_2	1654	1661	1703	1787	2056	2234	2043	2307
	Neg_2	1986	1862	1993	1983	2397	1862	2721	2747
$Late_2$	Aff_2	1569	1550	1381	1672	1821	1964	1958	2057
	Neg_2	1839	1970	1844	1584	2481	1970	2102	2611

APPENDIX B

APPENDIX B

Mean RT in Msec. Per Condition on Day 2
Experiment 1

		Assertion				Inference			
		On $Time_1$		$Late_1$		On $Time_1$		$Late_1$	
		Aff_1	Neg_1	Aff_1	Neg_1	Aff_1	Neg_1	Aff_2	Neg_1
On $Time_2$	Aff_2	1419	1526	1531	1593	2050	2140	2104	2288
	Neg_2	1788	1726	1871	1970	2460	2507	2540	2547
$Late_2$	Aff_2	1416	1411	1324	1414	1820	2018	1789	1839
	Neg_2	1763	1754	1690	1612	2216	2271	2115	2294

APPENDIX C

APPENDIX C

Significant (p < .05) Results of the 2x2x2x2x2 Within-Subjects ANOVA on Days 1 and 2 in Experiment 1

Condition	Day 1	Day 2
Negation 1	$F(1,11)=10.981$	$F(1,11)=6.824$
Term 2	$F(1,11)=29.624$	$F(1,11)=67.488$
Negation 2	$F(1,11)=54.893$	$F(1,11)=104.366$
Condition	$F(1,11)=363.120$	$F(1,11)=147.442$
Condition x Term 2		$F(1,11)=14.526$
Condition x Negation 1	$F(1,11)=10.532$	$F(1,11)=6.698$
Negation 1 x Negation 2	$F(1,11)5.410$	$F(1,11)=4.768$*
Term 1 x Term 2	$F(1,11)=10.191$	$F(1,11)=14.041$
Term 1 x Term 2 x Negation 2	$F(1,11)=5.178$	
Condition x Term 1 x Negation 1 x Negation 2	$F(1,11)=6.169$	
Condition x Term 1 x Term 2 x Negation 2		$F(1,11)=5.136$
Condition x Term 1 x Negation 1 x Term 2 Negation 2	$F(1,11)=15.409$	$F(1,11)=4.181$*

*denotes a marginally significant result

APPENDIX D

APPENDIX D

Mean RT in Msec. Per Condition on Day 1
Experiment 2

		On $Time_1$		$Late_1$	
		$Affirmative_1$	$Negative_1$	$Affirmative_1$	$Negative_1$
On $Time_2$	$Affirmative_2$	1493	1594	1670	1715
	$Negative_2$	1804	1684	1958	2045
$Late_2$	$Affirmative_2$	1620	1540	1342	1537
	$Negative_2$	1803	2004	1745	1646

APPENDIX E

APPENDIX E

Mean RT in Msec. Per Condition on Day 2
Experiment 2

		On $Time_1$		$Late_1$	
		$Affirmative_1$	$Negative_1$	$Affirmative_1$	$Negative_1$
On $Time_2$	$Affirmative_2$	1372	1503	1526	1502
	$Negative_2$	1658	1611	1827	1865
$Late_2$	$Affirmative_2$	1392	1358	1258	1329
	$Negative_2$	1749	1789	1615	1573

APPENDIX F

APPENDIX F

Significant ($p < .05$) Results of the 2x2x2x2 Within-Subjects ANOVA on Days 1 and 2 in Experiment 2

Condition	Day 1	Day 2
Term 2	$F(1,11)=30.686$	$F(1,11)=23.932$
Negation 1	$F(1,11)=4.145$*	
Negation 2	$F(1,11)=312.941$	$F(1,11)=108.022$
Term 1 x Term 2	$F(1,11)=23.679$	$F(1,11)=9.924$
Term 2 x Negation 2		$F(1,11)=7.612$
Term 1 x Term 2 x Negation 2	$F(1,11)=9.343$	$F(1,11)=8.059$
Term 1 x Negation 1 x Term 2 x Negation 2	$F(1,11)=30.649$	$F(1,11)=7.308$

*denotes a marginally significant result

APPENDIX G

APPENDIX G

Mean RT in Msec. Per Condition on Day 1
Experiment 3

		On $Time_1$		$Late_1$	
		$Affirmative_1$	$Negative_1$	$Affirmative_1$	$Negative_1$
On $Time_2$	$Affirmative_2$	1528	1929	2112	1950
	$Negative_2$	1849	1786	2323	2965
$Late_2$	$Affirmative_2$	1901	1860	1431	1804
	$Negative_2$	2203	2639	1761	1611

APPENDIX H

APPENDIX H

Mean RT in Msec. Per Condition on Day 2
Experiment 3

		On $Time_1$		$Late_1$	
		$Affirmative_1$	$Negative_1$	$Affirmative_1$	$Negative_1$
On $Time_2$	$Affirmative_2$	1255	1661	1769	1823
	$Negative_2$	1672	1483	1874	2488
$Late_2$	$Affirmative_2$	1774	1669	1162	1452
	$Negative_2$	1996	2278	1517	1337

APPENDIX I

APPENDIX I

Significant (p < .05) Results of the 2x2x2x2 Within-Subjects ANOVA on Days 1 and 2 in Experiment 3

Condition	Day 1	Day 2
Negation 1	$F(1,11)=35.293$	$F(1,11)=8.067$
Term 2	$F(1,11)=33.613$	$F(1,11)=22.182$
Negation 2	$F(1,11)=45.041$	$F(1,11)=20.152$
Term 1 x Term 2	$F(1,11)=50.493$	$F(1,11)=59.654$
Term 1 x Term 2 x Negation 2	$F(1,11)=31.706$	$F(1,11)=7.487$
Term 1 x Negation 1 x Term 2 x Negation 2	$F(1,11)=25.739$	$F(1,11)=15.177$

BIBLIOGRAPHY

BIBLIOGRAPHY

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 1975, 82, 45-73.

Chase, W. G., & Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), Cognition in learning and memory. New York: Wiley, 1972.

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

Clark, H. H. How we understand negation. Paper presented at Workshop on Cognitive Organization and Psychological Processes, University of California, Irvine, 1970.

Clark, H. H. Semantics and comprehension. In T. A. Sebeok (Ed.), Current trends in linguistics, Vol. 12: Linguistics and adjacent arts and sciences. The Hague: Mouton, 1974.

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

Cornish, E. R. Pragmatic aspects of negation in sentence evaluation and completion tasks. British Journal of Psychology, 1971, 62, 505-511.

Gough, P. B. Grammatical transformations and speed of understanding. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 107-111.

Gough, P. B. The verification of sentences: The effects of delay of evidence and sentence length. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 492-496.

Greene, J. M. Syntactic form and semantic function. Quarterly Journal of Experimental Psychology, 1970, 22, 14-27. (a)

Greene, J. M. The semantic function of negatives and passives. British Journal of Psychology, 1970, 61, 17-22. (b)

Just, M. A., & Clark, H. H. Drawing inferences from presuppositions and implications of affirmative and negative sentences. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 21-31.

Klima, E. S. Negation in English. In J. A. Fodor & J. J. Katz (Eds.), _The structure of language: Readings in the philosophy of language._ Englewood Cliffs, N. J.: Prentice Hall, 1964.

Slobin, D. I. Grammatical transformations in childhood and adulthood. _Journal of Verbal Learning and Verbal Behavior_, 1966, _5_, 219-227.

Sternberg, S. Memory scanning: Mental processes revealed by reaction time experiments. _American Scientist_, 1969, _57_, 421-457. (a)

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), _Attention and performance II_. Amsterdam: North Holland, 1969. (b)

Trabasso, T., Rollins, H., & Shaughnessy, E. Storage and verification stages in processing concepts. _Cognitive Psychology_, 1971, _2_, 239-289.

Wason, P. C. The processing of positive and negative information. _Quarterly Journal of Experimental Psychology_, 1959, _11_, 92-107.

Wason, P. C. Response to affirmative and negative binary statements. _British Journal of Psychology_, 1961, _52_, 133-142.

Wason, P. C. The contexts of plausible denial. _Journal of Verbal Learning and Verbal Behavior_, 1965, _4_, 7-11.

Wason, P. C., & Johnson-Laird, P. N. _Psychology of reasoning: Structure and content_. Cambridge, Massachusetts: Harvard University Press, 1972.

Wason, P. C., & Jones, S. Negatives: Denotation and connotation. _British Journal of Psychology_, 1963, _54_, 299-307.

Young, R., & Chase, W. G. Additive stages in the comparison of sentences and pictures. Paper presented at Midwestern Psychological Association meetings, Chicago, 1971.