THE EFFECTS ON SENTENCE RECALL OF VARYING AGE, MEAN DEPTH, AND SENTENCE TYPE

by

DIANA LEE MITCHELL

B.Ed., University of British Columbia, 1966

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department of

EDUCATIONAL PSYCHOLOGY

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

February, 1968

Department of Education Psychology

The University of British Columbia
Vancouver 8, Canada

Date February 30, 1968

ABSTRACT

The purpose of this study was to investigate the effects on recall of varying sentence complexity, sentence type, and age of the subjects. The measure of sentence complexity used was Martin and Roberts' (1966) adaptation of Yngve's (1960) depth. The six sentence types involved were kernel (K), negative (N), passive (P), negative-passive (NP), passive truncated (P_T), and negative-passive truncated (NP_T).

One hundred twenty children, and one hundred twenty adults, were exposed, ten at a time, to an orally-presented example of each sentence type. After hearing the six sentences, the subjects were instructed to write as many as they could recall. Six such trials were effected with each subject. Twelve sets of six sentences each were used such that six sets were of the lesser mean depth (1.29) and six sets were of the greater mean depth (1.71). Sixty subjects of each age were exposed to one of either the six sets of sentences having mean depth 1.29, or to one of the six sets of sentences having mean depth 1.71.

The data supported two of three stated hypotheses, that is:

i) The likelihood of recall of a sentence is inversely related to the mean depth of that sentence when both children and adults are subjects.

ii) More sentences of all types and of either depth are recalled by adults than by children.

A third hypothesis that::

iii) Kernel sentences are recalled better than non-kernel sentences by children and adults, was not supported by the data. It was found, rather, that both adults and children correctly recalled sentences involving the negative better than those which did not. This finding, although supported by neither psycholinguistic theory nor by the experimental literature, was interesting. It was suggested that a study be performed in an attempt to reproduce these results, and that an investigation be made to determine if a tendency to respond correctly more often to negative stimuli is a culturally-determined factor. Further, it was suggested that study be made of the significance of the mean depth factor, of transformations, and of their interaction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In recent years, considerable interest has arisen over changes in descriptive linguistics designed to provide psychologists with more adequate descriptions of language than were previously available. The present study is an attempt to provide findings which may contribute to an understanding of language behavior. In order that the problem to be discussed may be understood by those who are unfamiliar with the psycholinguistic domain, certain important principles will first be introduced.

Noam Chomsky is generally acknowledged as having pioneered the psycholinguistic movement in the 1950's by introducing his principles for a generative grammar. Essentially, Chomsky proposed that a grammar "can be viewed as a device of some sort for producing the sentences of the language under analysis" (Chomsky, 1957, p. 11). The grammar establishes a small set of rules from which are generated an infinite number of grammatical sentences of the language and no non-sentences. Chomsky proposed that there are three broad classes of rules that can be established. The present study is chiefly concerned with the following two: (a) formation rules--from which the terminal strings underlying the kernel of basic sentences are derived, and (b) transformation rules-- which convert terminal strings to various sentence types which are derivatives of the underlying kernel.

Miller (1962, p. 749) described how Chomsky's formation rules can be used to derive the terminal string underlying the sentence, "Bill hit the ball". The constituent units used in the analysis are symbolized as follows: T = article; N = noun; NP = noun phrase; V = verb; VP = verb phrase; S = sentence.

> Figure 1 illustrates how a small fragment of English grammar might be expressed in this manner. The basic axiom is S. The rewriting rules F1-7 permit us to form the sentence "Bill hit the ball" in a sequence of steps. First S is rewritten as NP + VP, according to rule F1. Then we can rewrite NP as "Bill" according to rule F4. Since there is not any rule available for rewriting "Bill", we are forced to stop at this point. We can, however, rewrite VP according to rule F3, thus getting "Bill" + V + NP. In this way we can proceed as indicated by the tree graph on the right until the desired sentence is derived...
>
> The set of rewriting rules on the left of Figure 1 can be conveniently referred to as the grammar, and the set of sentences that the grammar generates defines the language. It is an important feature of this kind of grammar that there are terminal symbols that cannot be rewritten, and these comprise what we ordinarily recognize as the vocabulary of the language. According to this way of representing it, the vocabulary is included in the grammar.

Such rules as are used by Miller in the preceding quotation are called "phrase structure" rules by Chomsky (1957), and may only be used to rewrite single symbols, never strings of symbols. Transformation rules, however, are used to write a string of symbols so that it may be transformed into another kind of string. Chomsky (1957) proposes a variety of transformation rules, each applying to strings of certain specified structural descriptions. Clifton and Odom (1966, p. 2) give an example of how one kind of transformation rule may be applied:

> The passive transformation specifies that a string of symbols (morphemes or names of classes of morphemes) having the structural description NP_1 - Aux - V - NP_2 can be rewritten as NP_2 - Aux - V - NP_2 en - V - By + NP_1. The transformation can be used to rewrite the string of symbols that underlies, for instance, the sentence "John hit the ball" into the string that underlies "The ball was hit by John".

S

F1. S ⟶ NP + VP

F2. NP ⟶ T + N

F3. VP ⟶ V + NP

F4. NP ⟶ Bill, John

F5. T ⟶ the, a

F6. N ⟶ boy, girl, ball

F7. V ⟶ hit

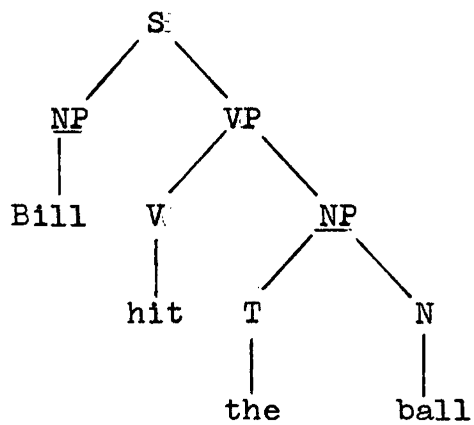

FIG. 1. A fragment of English grammar, phrased in terms of rewriting rules, illustrating a generative grammar. (Miller, 1962, p. 749)

Although there are different forms of transformation rules, a detailed description falls outside the scope of this paper. What is important is to explain the role played by what Chomsky has described as the "terminal strings" underlying "kernel" sentences (Chomsky, 1957). Phrase-structure rules plus the appropriate transformation rules together will generate terminal strings which, Chomsky states, are the fundamental basis of the simplest form of sentence, the kernel. A kernel sentence is commonly referred to in English usage as a simple active, declarative sentence such as "Bill hit the ball". According to Chomsky, application of various transformation rules to the terminal string underlying any kernel sentence will generate strings underlying more complex sentences. These complex sentences--the passive (P), the negative (N), the negative-passive (NP)*, the question (Q), the negative question (NQ), the negative-passive question (NPQ), the passive question (PQ), along with the kernel (K)-- go to make up sentence families. Miller (1962, p. 760) graphically represents the relationships between the members of a sentence family in a way which parallels Chomsky's description. Miller's interpretation of these relationships is presented in Figure 2.

The transformation theory proposes that, as the string underlying the K also underlies the other sentence types, varying numbers of transformations must be made upon the underlying string in order to generate the other sentence types. The closeness of the relationship between any two sentence types is indicated in Figure 2 by the number of lines in the shortest path connecting them. Thus, to move from a K to a P requires apparently one transformation, whereas to move from a K to a NP requires two. The step-wise property of transformational grammar is clearly indicated in this diagram.

* Throughout this paper, <u>NP</u> refers to noun phrase whereas NP refers to negative passive sentence type.

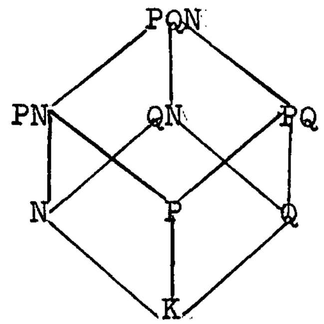

FIG. 2. Graph indicating relations among eight types of sentences formed by negative (N), passive (P), and interrogative (Q) transformations. (Miller, 1962, p. 760)

Miller hypothesizes that:

> ...what people remember is the kernel sentence, but that when you ask them to recite the original sentence exactly, they supplement their memory of the kernel with a footnote about the syntactic structure (Miller, 1962, p. 760).

The results of a number of current studies seem to corroborate this hypothesis (Mehler, 1963; Menyuk, 1963; Gough, 1965; Slobin, 1966). The apparent simplicity of the K, along with the implication that it is easier to recall than any other sentence type, will be shown later in this paper to be questionable.

Certainly there is more than one theory regarding the derivation of sentences. Lees (1960), as well as Katz and Postal (1964) suggest that the kernel, the passive, and the negative have separate derivations. Katz and Postal argue that there are certain morphemes which are "universal". To be "universal", they say, a morpheme must be specified by a general theory of linguistic descriptions, rather than being a morpheme specific only to certain languages. The negative morpheme, the passive morpheme, and the question morpheme are examples of "universal morphemes", as they are most likely included in the underlying structure of many languages. Thus, Katz and Postal see the relationships between sentences as depending upon the number of universal morphemes separating them. The active and the passive, differing in only one universal morpheme, are supposedly very closely related. The hypothesis that the kernel, the negative, and the passive have separate derivations has been exposed to very little experimentation (Lees, 1960; Katz and Postal, 1964) in comparison with the transformation hypothesis put forth by Chomsky (1957). Nevertheless, as this question of derivation has as yet not been settled, both hypotheses have heuristic value.

A brief description of another interpretation of modern linguistic theory should provide sufficient background from which to explore the problem to be investigated in this paper. In 1960, Victor Yngve described "A model and an hypothesis for language structure" which dealt with generative grammar in a different manner, and with different emphasis than did Chomsky's model. The differences, being of a philosophical nature, will not be dealt with here. What is important, is to describe what Yngve calls the "depth hypothesis" (Yngve, 1960; 1964) and the manner in which depth of sentences is assigned. Yngve states (1964):

> It seems, that as we speak, we incur commitments to finish our sentences in certain ways in order to make them grammatical. As an example of what these commitments are, take the following sentence: "When the president spoke, the people listened". As we start this sentence with the word "when" we have two commitments, one to finish a dependent clause "when the president spoke", and the other to follow this with an independent clause "the people listened". Then, as we start to fulfill the first commitment with the word "the", we have two new

> commitments making a total of three: the original one, that is to finish up with an independent clause, another to finish the subject of the dependent clause we started with the word "the", and the third to follow this with a predicate like "spoke". Apparently we cannot cope with more than about seven such commitments at any one time without forgetting what it was we were going to say... The maximum number of commitments at any one time in a sentence is called the depth of the sentence.

To complete Yngve's depth analysis of the sentence "When the president spoke, the people listened", observe that when the speaker says "president" he commits himself to finish up with an independent clause, as well as to supply a predicate for the dependent clause. Thus, "president" can be said to be embedded in this sentence to the depth of 2. Upon saying "spoke", however, the speaker is committed only to completing the sentence with an independent clause, so that "spoke" is embedded to the depth of 1. But, upon saying "the", the speaker commits himself further. He must furnish a subject of the independent clause, as well as follow this with a predicate, thus involving "the" to the depth of 2. "People" requires only that a predicate follow, and so is embedded to the depth of 1, and "listened", being that predicate, finishes the independent clause, so fulfilling the final commitment of the speaker. Voice intonation indicates that there is nothing to follow the independent clause, and the speaker's commitments are met. Therefore, "listened" is embedded to the depth of 0.

The words of the sentence "When the president spoke, the people listened" can thus be said to possess the Yngve numbers

2,3,2,1,2,1,0, respectively. (By definition (Yngve, 1960; 1964), the depth of this sentence is 3.) Martin and Roberts (1966) proposed that the mean of the Yngve numbers be taken as being a measure of the "structural complexity" of the sentence as a whole. Thus, the "mean depth", or structural complexity of the sentence "When the president spoke, the people listened" is 1.57.

Always, when determining depth of words in a sentence, phrase-structure rules must be applied to discover the constituents involved. For example, Figure 3 represents the formal analysis of the sentence "When the president spoke, the people listened". The figure assumes the form of a phrase-structure tree which was first presented by Lukasiewicz and Tarski (1930) and later adopted by Chomsky (1957), Yngve (1960), Miller (1962) and other contemporary psycholinguists to diagram the constituents of English sentences. The principal units are written on the horizontal "nodes" of the tree, and appropriate examples of the sorts of words represented by these units are written at the bottom of the vertical "branches". Yngve states (1960, p. 455) that the English language works in such a way as to limit depth.

> The mechanism that English uses to limit depth is a restricted relabeling scheme involving sentence, clause, noun phrase, primary attribute (adjectival), secondary attribute (adverbial), and tertiary attribute (adverbial).

Secondary units in the system are verb phrases, verbs, adjectives, adverbs, articles (labeled "T" on the tree), and coordinators (Co). Coordinators serve the function of joining what

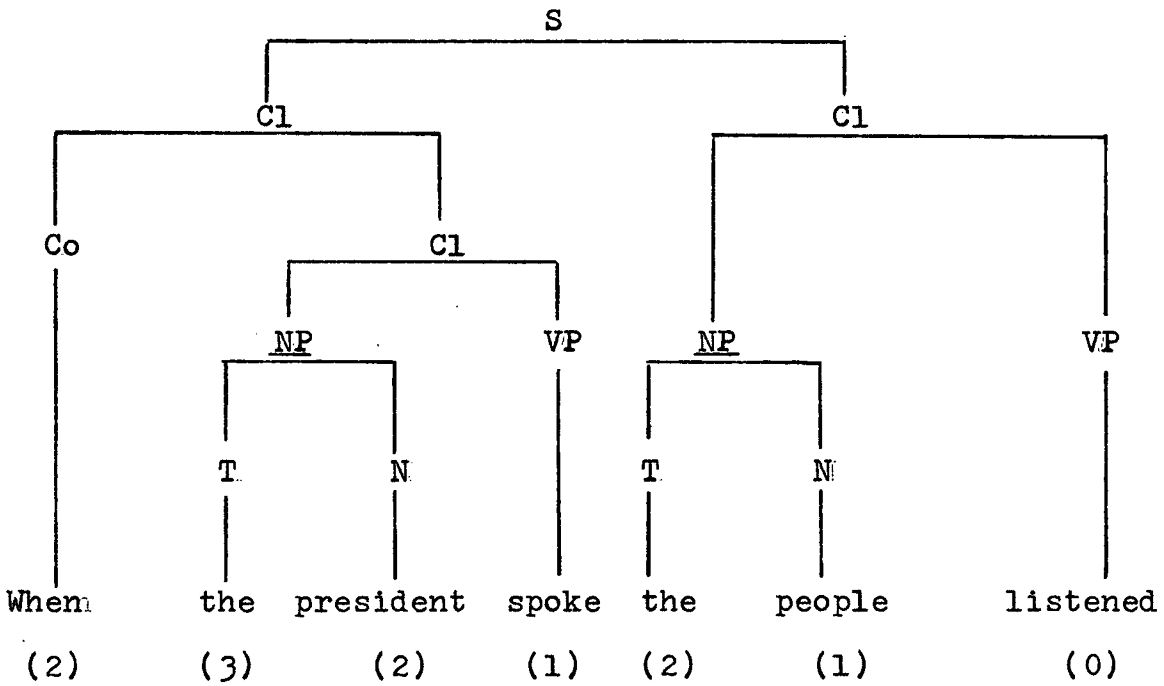

FIG. 3. Phrase-structure analysis and Yngve's measure of sentence depth.

are analysable as two separate sentences. For example, in the sentence "She hit him when he laughed", "when" is the coordinator (symbolized as Co on the tree). Tertiary attributes, secondary attributes, and primary attributes are involved in sentences such as "When very neatly dressed children rose, she smiled". The tertiary attribute is "very", which modifies the adverb "neatly". The secondary attribute comprises "neatly", together with the tertiary attribute. The primary attribute comprises the secondary attribute plus the adjective "dressed". This pattern of subsuming constituents under broader constituents is what Yngve means by a "restricted relabeling scheme". It is possible to determine the depth of a sentence by counting off the number of principal constituents involved in its makeup.

As can be seen in Figure 3, the principal constituents involved in the example sentence are marked off by nodes labeled S (Sentences), CL (Clause), NP (Noun phrase). Thus, the depth of this sentence is 3. The mean depth is computed by summing the depths for each word and dividing this sum by the number of words in the sentence (11/7 = 1.57).

Yngve, after introducing his concept of depth, hypothesized "that there would be an easily observable effect of the depth limitation in the grammar of English." Explicitly, he expands this hypothesis thus (Yngve, 1960, p. 452):

a) Although all languages have a grammar based on constituent structure,

b) the sentences actually used in the spoken language have a depth that does not exceed a certain number

c) equal or nearly equal to the span of immediate memory (presently assumed to be 7 ± 2).

d) The grammars of all languages will include methods for restricting regressive constructions so that most sentences will not exceed this depth,

e) and they will include alternative constructions of lesser depth that would maintain the power of expression of the language.

f) For all languages, much of the grammatical complexity over and above the minimum needed for the signaling function can be accounted for on this basis.

g) When languages change, depth phenomena will frequently be involved, and will often play an important role.

Yngve goes on to state that constructions of lesser depth are expected to be preferred over equivalent constructions of greater depth. This is a very exciting hypothesis, for it predicts what transformation theories do not. Specifically, Yngve's hypothesis would predict that, given two sentences of the same type, but of different depths, the sentence of lesser depth would be more likely, say, to be spoken by a speaker or recalled by a listener. Transformation theories (ie., Chomsky's transformational grammar) predict ease of one sentence type over another, but make no predictions concerning the ease of recall of a sentence of a given type over another sentence of the same type.

In order to investigate the possibility that depth rather than sentence type might be the significant factor in a sentence recall situation, Martin and Roberts (1966) devised an interesting study. 120 college students were randomly assigned to one of two treatment conditions. The first condition involved the recall of sets of sentences of mean depth 1.29, whereas the second involved the recall of sets of sentences of mean depth 1.71. Each set of sentences included one each of six sentence types: Kernel (K), Negative (N), Passive (P), Negative-Passive (NP), Passive Truncated (P_T), and Negative-Passive Truncated (NP_T). (A P_T sentence is one in which the agent is not specified, as in "Lights are turned off in the evening"). Sentence length was held constant. A set of sentences of either mean depth was read to a group of ten subjects, who were then instructed to write down as many of the sentences as they could remember. This procedure was repeated with the same set of subjects and sentences until six free-learning trials were effected.

Martin and Roberts found that sentences of lesser mean depth were recalled significantly more often than were sentences of greater mean depth, regardless of sentence type. Further, kernel sentences were less easily remembered than were non-kernel sentences. The conclusion that the likelihood of recall of a sentence was inversely related to mean depth of that sentence strongly corroborates Yngve's depth hypothesis. However, the conclusion that kernel sentences were harder to recall than non-kernel sentences is directly opposed to the results discussed earlier (Mehler, 1963; Menyuk, 1963; Gough, 1965; Slobin, 1966) which seemed to corroborate Chomsky's transformation hypothesis. Further, Martin and Roberts' finding of no differences between the various non-kernel sentence types is curious, in light of the results mentioned in the previous sentence.

Considering the conflicting results to be found in the data with respect to both the depth hypothesis and the transformation hypothesis, as well as the paucity of data regarding their interaction, further investigation seemed merited. It is possible that investigation along these lines could contribute to a better understanding of language development. If this goal is to be achieved, data must be gathered regarding both children's and adult's language behavior. Some of the questions which might be answered by such research are:

1. Is there any one sentence type which is easier than others for adults to recall?

2. Is there any one sentence type which is easier than others for children to recall?

3. Is sentence mean depth a significant factor in sentence recall when adults are subjects?

4. Is sentence mean depth a significant factor in sentence recall when children are subjects?

5. What is the effect of the interaction of mean depth and sentence type on recall of sentences by both adults and children?

6. Will adults correctly recall more sentences than will children?

With respect to questions 1 and 2, one would expect, after considering the wealth of supportive data, that Kernels would be the simplest sentences to recall for both adults and children. This prediction would have to be made, at least until Martin and Roberts' (1966) results were replicated.

Considering Yngve's depth hypothesis and the results of Martin and Roberts' study, one would expect that sentences of a lesser mean depth would be recalled more often than those of a greater mean depth regardless of the subject's age.

As answers to any or all of the above questions would be illuminating, the present study was proposed to investigate the effects on recall of varying sentence mean depth and sentence kind when sentence length is controlled. Two age groups were considered in order to determine if, in the controlled situation described, there were age differences in recall of sentences at two mean depths and of varying kind.

It was hypothesized that:

1. The likelihood of recall of a sentence is inversely related to the mean depth of that sentence when both children and adults are subjects.

2. Kernel sentences are recalled better than non-kernel sentences by children and by adults.

3. More sentences, regardless of depth, are recalled by adults than by children.

CHAPTER II

METHOD

On each of six free-learning trials, the same six sentences were read to the subjects (*S*s), who then proceeded immediately to recall them as best they could. The experimental design was a 2 x 2 x 6 factorial design comprising six sentence types and two age groups within *S*s, and two levels of sentence depth between *S*s.

Materials

Twelve sets of six sentences each were constructed such that any one set contained the following six sentence kinds: kernel (K), passive (P), passive truncated (P_T), negative (N), negative-passive (NP), and negative-passive truncated (NP_T). The sentences were all seven words long and within each set were unrelated in meaning in the judgment of the investigator. The twelve sets of sentences appear in Appendix A.

The Thorndike-Lorge (1944) counts for all sets of sentences was uniformly high so as to maximize the probability of the words' being equally familiar to all *S*s. The words of all sentences appear in the Copp-Clark readers, Grades 1 through 3, which are used in the British Columbia schools.

Of the twelve sets of sentences, six were made up of sentences each of which had Yngve numbers 1,3,2,1,1,1,0 for the seven words, hence mean depth 1.29. The other six sets had Yngve numbers 1,4,3,2,1,1,0, and hence mean depth 1.71. Depth of each word was computed according to the phrase-structure rules described by Yngve (1960). Mean depth of each sentence was computed according to the method described by Martin and Roberts (1966).

Subjects

The *S*s were 120 children who were attending either Morley-Brantford Elementary School or Kitchener Elementary School in Burnaby, B.C., during the school year 1966-67 and 120 adults who were attending summer school at the University of British Columbia during July and August, 1967. All the *S*s were at least second-generation Canadians who spoke English as a first language. The children's parents were all either engaged in professional work or skilled labour. No child was classified by his school as being in need of remedial work in any subject. The mean age of the children was 9 years 5 months, and the mean I.Q. was 115. All the adults were teachers in either the British Columbia elementary or secondary school systems. The mean age of the adults was 30 years 3 months. I.Q. data were available only for younger *S*s.

*S*s were run in groups of ten. Ten *S*s from each age group were randomly assigned to each of the 12 sets of sentences. Altogether, there were four main groups of 60 *S*s each. The groups were named according to the age of the *S*s as well as the mean depth of the sentences being presented to that particular group. Thus, the groups were named: 9 (1.29); 9 (1.71); 30 (1.29); 30 (1.71).

Procedure

Ten Ss at a time were seated in standard classroom desk chairs in an otherwise empty room (except for the experimenter's table). Each S was given one sheet of blank paper and the following instructions were read aloud by the experimenter (E):

> I am going to read six ordinary English sentences to you. When I am finished reading, I will say "start" at which time I want you to write or print as many of these sentences as you can remember. Don't write before I say "start". I am not interested in how you spell the words or in your handwriting. If you don't know how to spell a word, write it down the way it sounds, and that will be fine. If I don't understand it, I will ask you later to explain. Are there any questions? Are you ready?

Following the instructions, the E read six sentences aloud to the Ss, after which they wrote them on their papers, taking as much time as they wished. As soon as an S finished writing, the E picked up the sheet on which the recalled sentences were written. After all ten Ss had finished, fresh sheets of paper were distributed, and the following instructions were read to the Ss by the E:

> I am going to read the same six sentences again. This time they will be in a different order, but that doesn't matter. I want you to listen to them and when I say "start", write down as many of them as you can remember. We will do this 4 more times after this. Are there any questions? Are you ready?

In all, six trials were effected with each group of ten _S_s. A Latin square was used to regulate ordering of sentences. As a result, every sentence occured in each ordinal position over six trials, and no sentence followed another in the same order twice.

CHAPTER III

RESULTS

Correct recalls

For each combination of sentence mean depth, sentence type, and age, the mean number of correct recalls over six trials per group of Ss was computed (see scoring procedures, Appendix B). These results are presented in Table 1. Each mean correct recalls score represents 60 Ss.

A 2 x 2 x 6 analysis of variance was performed (age X mean depth X sentence type) with repeated measures on sentence type (see Winer, 1962, p. 337). It was found that older Ss made significantly more correct recalls, regardless of the depth or the kind of sentence, than did younger Ss ($F = 35.92$, $df = 1/236$, $p < .0001$). It was also found that all Ss recalled significantly more sentences, regardless of kind, at mean depth 1.29 than they did at mean depth 1.71 ($F = 47.904$, $df = 1/236$, $p < .0001$). Further, sentence type was significant ($F = 12.102$, $df = 5/1180$, $p < .0001$), as was the interaction between age and sentence type ($F = 4.749$, $df = 5/1180$, $p < .001$). The interaction between mean depth and sentence type, however, was not significant ($F = .918$, $df = 5/1180$, $p > .25$).

Figure 4 represents the age X type interaction and indicates that sentences vary in their availability for immediate recall, depending upon the S's age. Simple Negative sentences were re-

TABLE 1

MEANS AND STANDARD DEVIATIONS

OF CORRECT RECALL SCORES N=60

	Sentence Type											
	K		N		P_T		NP_T		P		NP	
GROUP	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
9(1.29)	2.97	1.87	3.83	1.78	2.62	1.89	3.32	1.85	2.83	1.85	3.40	2.08
9(1.71)	1.97	1.81	2.93	1.62	2.17	1.91	2.22	1.94	1.80	1.57	2.50	1.91
30(1.29)	3.97	1.70	4.28	1.44	3.43	1.60	4.27	1.39	3.93	1.45	4.72	1.50
30(1.71)	3.12	1.81	2.75	1.81	2.52	1.70	3.38	1.67	2.98	1.73	3.45	1.63

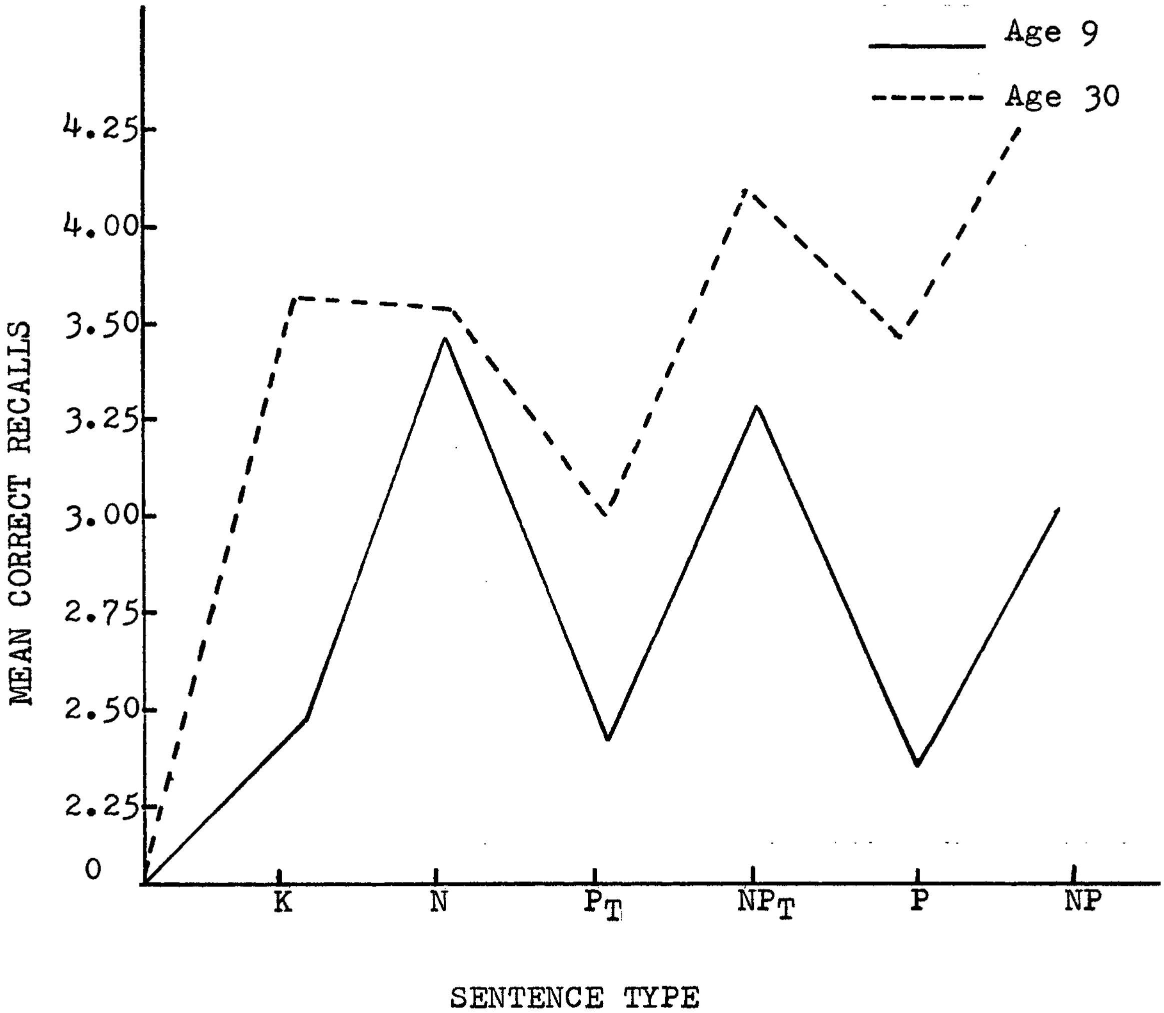

FIG. 4. Mean correct recalls of 9 year old and 30 year old <u>S</u>s for 6 sentence types n = 120.

called correctly most often by 9 years old Ss (mean correct recalls = 3.40), whereas Negative-Passive sentences were recalled correctly most often by 30 years old Ss (mean correct recalls = 4.09).

In view of the significance of sentence types a Tukey (b) test on ordered pairs of correct recall scores for each sentence type was computed (see Winer, 1962, p. 87). The results of the Tukey analysis are presented in Table 2. Negative-Passive and simple Negative sentences were recalled significantly more often than Passive Truncated, Passive, and Kernel sentences ($p < .01$). Negative-Passive Truncated sentences were recalled significantly more often than were Passive Truncated sentences ($p < .01$) or Passive sentences ($p < .05$). There were no significant differences in ease of recall of K vs. NP_T; K vs. P; K vs. P_T; N vs. NP_T; N vs. P; or NP_T vs. NP ($p > .05$).

Errors

Table 3 summarizes the number and kinds of errors arising from each combination of mean depth and sentence type for Groups 9(1.29) and 9(1.71). For example, in the first column, out of the (60 Ss) x (6 trials) = 360 response opportunities for recalling a K of mean depth 1.29, there occurred 182 errors. Of these errors, 71 occurred in Ks (errors of the same type as presented sentence), 1 occurred as a N transformation, 2 were P_Ts, 65 were Incomplete sentences, 7 were Ungrammatical sentences, and 36 were Omissions of sentences. There were no NP_T, P, or NP errors where a K of mean depth 1.29 was presented.

TABLE 2

Tukey (b) Test on All Ordered Pairs of Total Correct Recalls of Each Sentence Type By All Subjects N=240

Sentence Type	Total Number Correct Recalls					
	644	693	721	791	831	844
	Passive Truncated	Passive	Kernel	Negative-Passive Truncated	Negative	Negative Passive
Passive Truncated	-	49	77	147**	187**	200**
Passive		-	28	98*	138**	151**
Kernel			-	70	110**	123**
Neg. Pass. Truncated				-	40	53
Negative					-	13
Negative Passive						-

** $p < .01$
* $p < .05$

The most frequent errors were omissions, followed by errors of the same sentence type as the presented sentence, and then by incomplete sentences. Exceptions to this pattern occurred when Ks of mean depth 1.71, NP_Ts of mean depth 1.29, and Ps of mean depth 1.71 were presented. Where Ks of mean depth 1.71 and NP_Ts of mean depth 1.71 were presented, the most frequent errors were errors of the same sentence type, followed by omissions, and then by incomplete sentences. While omissions were the most frequent errors where Ps of mean depth 1.71 were presented, incomplete sentences were the next most frequent errors, followed by errors of the same sentence type. Where Ks of mean depth 1.29 were presented, errors of the same type were most frequent, followed by incomplete sentences, and then by omissions. Also, when Ks of mean depth 1.71 and NP_Ts of mean depth 1.29 were presented, the most frequent errors were again those of the same sentence type as the presented sentence.

Recall errors of the same syntactic type as the presented sentence are entered on the main diagonal and are attributable chiefly to substitutions of pronouns, nouns, and verb tenses. All grammatical, complete-sentence errors off the main diagonal are transformations of the presented sentence. To determine at which mean depth level the greatest number of transformations occurred, a proportion analysis was made (Ferguson, p. 177). The proportion of grammatical, complete-sentence errors that are transformations is .24 for presented sentences of mean depth 1.29

TABLE 3

ERRORS IN RECALL MADE BY GROUPS 9(1.29) AND 9(1.71) CLASSIFIED BY SENTENCE TYPE OF THE PRESENTED SENTENCE N=60

	Presented Sentence											
	K		N		P_T		NP_T		P		NP	
Type of Error / Group	9 (1.29)	9 (1.71)	9 (1.29)	9 (1.71)	9 (1.29)	9 (1.71)	9 (1.29)	9 (1.71)	9 (1.29)	9 (1.71)	9 (1.29)	9 (1.71)
K	71	110	1	8	32	13	5	2	13	30	1	11
N	1	2	48	66	1	1	7	19	0	0	11	17
P_T	2	10	0	0	58	55	3	1	2	1	1	0
NP_T	0	0	0	2	2	3	64	67	1	0	3	2
P	0	0	0	0	0	0	0	1	46	49	8	3
NP	0	0	0	2	0	0	0	3	5	0	59	65
Incomplete	65	30	23	16	32	37	15	23	29	65	14	24
Ungrammatical	7	5	5	6	12	7	9	4	9	6	3	9
Omissions	36	85	50	84	66	114	58	107	85	101	58	79
Total	182	242	127	184	203	230	161	227	190	252	156	210

and .22 for presented sentences of mean depth 1.71. This difference is not statistically significant ($z = .74$, $p > .05$). When the proportion of off-diagonal errors (transformation errors) of any one type for presented sentences of mean depth 1.29 is compared with the proportion of off-diagonal errors of the same type for presented sentences of mean depth 1.71, no statistically significant differences are found ($p > .05$).

Table 4 summarizes the number and kinds of errors arising from each combination of mean depth and sentence type for Groups 30 (1.29) and 30 (1.71). When Ks and P_Ts of both mean depths, Ns of mean depth 1.71, and NPs of mean depth 1.29 were presented, the most frequent errors were of the same type as the presented sentences followed by omission errors. In all other cases, the most frequent errors were omissions, followed by errors of the same type as the presented sentences.

The proportion of grammatical, complete-sentence errors that are transformations is .34 for presented sentences of mean depth 1.29, and .39 for presented sentences of mean depth 1.71. This difference is not statistically significant ($z = 1.62$, $p > .05$). Of the total number of off-diagonal errors (transformation errors), the proportion that are NPs is .02 for presented sentences of mean depth 1.29, and .19 for presented sentences of mean depth 1.71. This difference is statistically significant ($z = -4.94$, $p < .001$). Of the total off-diagonal errors, the proportion that are NP_Ts is .22 for presented sentences of mean

TABLE 4

ERRORS IN RECALL MADE BY GROUPS 30(1.29) AND 30(1.71)

CLASSIFIED BY SENTENCE TYPE OF THE PRESENTED SENTENCE N=60

	Presented Sentence											
	K		N		P_T		NP_T		P		NP	
Type of Group / Error	30 (1.29)	30 (1.71)	30 (1.29)	30 (1.71)	30 (1.29)	30 (1.71)	30 (1.29)	30 (1.71)	30 (1.29)	30 (1.71)	30 (1.29)	30 (1.71)
K	69	67	13	25	21	26	5	6	9	23	2	1
N	4	12	29	87	1	2	6	2	2	0	3	9
P_T	7	6	2	1	57	82	9	15	2	6	3	2
NP_T	2	3	0	11	27	4	30	41	0	7	0	0
P	1	2	0	0	2	16	2	3	46	48	9	12
NP	0	6	1	7	0	8	0	14	2	12	36	45
Incomplete	12	22	1	2	4	6	2	3	3	14	1	8
Ungrammatical	1	3	0	5	0	4	3	2	2	0	0	7
Omission	26	52	57	57	42	61	47	71	58	71	23	69
Total	122	173	103	195	154	209	104	157	124	181	77	153

depth 1.29, and .10 for presented sentences of mean depth 1.71. This difference, also, is statistically significant ($z = 3.00$, $p < .001$). No other comparisons of this type reached statistical significance ($p > .05$).

When comparing the total number of written responses made by children with those made by adults (that is, total correct responses + errors - omissions), it was found that adults made proportionately more responses out of a total possible of 8640 responses than children ($z = 5.00$, $p < .001$). The proportion of all written responses made by children that are transformation errors in .07, whereas for adults this proportion is .10. This difference is statistically significant ($z = 4.28$, $p < .001$).

CHAPTER IV

DISCUSSION

The present experiment corroborates the finding of Martin and Roberts (1966) that sentence mean depth is a definitive factor in sentence retention. Both age Ss consistently recalled sentences of mean depth 1.29 correctly more often than they did sentences of mean depth 1.71. Furthermore, as was hypothesized, older Ss consistently scored more correct recalls over all sentence types, regardless of sentence mean depth, than did younger Ss. In fact, adults made more responses overall than did children. The adult pattern of responding, however, differed somewhat from that of the children. The most frequent responses, other than correct ones, made by adults were errors of the same type as the presented sentence (syntactically-correct responses) followed by transformation errors. The most frequent responses made by children, after correct recalls and errors of the same type as the presented sentence, were incomplete sentences. In other words, adults attempted more responses, made quite a few transformational errors, but wrote very few incomplete responses. Children attempted fewer responses, made fewer transformational errors, but wrote many more incomplete sentences than did the adults. All of this seems to indicate either that children are less likely to write down something they aren't sure of than adults, or perhaps that children's memory spans are considerably shorter than adults' in situations such as is described in the present study.

The possibility that children are more conservative in their responding behavior in sentence recall situations such as was described by this experiment should be investigated further. This is true also for the possibility that children's memory spans for sentence recall tasks are shorter than adults'. Piaget's developmental theory implies that increased memory span is required for tasks performed regularly by older children than is required by tasks regularly performed by younger children (Flavell, 1964). More experimental evidence is required before conclusions can be firmly drawn from such results as are presented here regarding either the "conservative" hypothesis or the memory-span hypothesis.

The data also indicated that the semantic variable of affirmation-negation was extremely important. All sentence types involving the negative were recalled best by younger Ss, while NPs and NP_Ts were recalled best by older Ss. It is interesting that, when all the data for both age groups were analyzed (excluding P_T and NP_T data for the sake of comparison with other studies), the sentence types in order from that recalled correctly most often to that recalled correctly least often were: $N = NP > K = P$.

Although Martin and Roberts found that kernels were recalled correctly less frequently than non-kernels, they did not find a similar, strong negative factor as did this investigator. Nor is this result one which would be predicted by either the transformational grammar model or the phrase structure model. There are several possible explanations:

1. Even though the instructions emphasized the non-test aspects of the experimental situation, the Ss, all of whom came to the experimental sittings directly from classroom situations, may have been unconvinced that they were not expected to meet some undefined standard. They thus may have been personally motivated to "pass" on what they indeed felt was a test. If this were so, the negative morpheme, so clearly lacking in ambiquity, may have been an obvious aid to recall of those sentences in which it was heard. Although this is conjecture on the E's part, it is both a logical and testable hypothesis.

2. As there were three sentences out of every group of six which included some form of negation, an induced response bias may have existed toward the Negative, Negative-Passive, and Negative-Passive Truncated sentences which was solely due to the presence of negation. If the induced-response-bias hypothesis is true, then it is curious that Martin and Roberts' data did not show a similar trend.

3. Katz and Postal's separate-derivations hypothesis with regard to Ks, Ns, and Ps may be closer to reality than Chomsky's

transformation hypothesis. The results of the present study indicate that the relationship between affirmative and negative sentences is very close, but that, contrary to what Katz and Postal (1964) propose, the negative sentence is simpler than the affirmative sentence.

4. Perhaps, if the results regarding ease of recall of negative sentences were replicated, it might be shown that this is a culturally-determined tendency. It may be that there are cultures in which the negative morpheme plays a very small role. Ss living in such cultures may respond quite differently in a recall situation such as was described in this study. A comparison might be made, then, of the responses to sentences in a sentence family of Ss drawn from a community in which the negative morpheme seldom occurs, and the responses of Ss drawn from such a community as was described in the present study. One might hypothesize that Ss from that culture in which the negative plays a minor role would correctly recall affirmative sentences much more frequently than negatives. This could be an interesting line of investigation to pursue.

Although each of these explanations is possible and testable, it is the opinion of this investigator that much can be gained here from pursuing yet another argument. Martin and Roberts (1966) describe a study in which "structural complexity", or structural embeddedness, rather than sentence type, influenced behavior:

Further evidence on these matters comes from a study by the present writers in which Ss made judgments as to which of two seven-word sentences they thought would be easier to remember if they had to relay one of them to someone else after a delay of several minutes. Since the sentence pairs were in writing in front of them, and since their judgments were not paced, Ss need not have stored any information in memory; only their intuitions regarding ordinary English sentences presumably were in effect. Twelve Ks were paired with 12 Ns and with 12 Ps, thus giving 144 K-N and 144 K-P sentence pairs. Of the 12 sentences of each kind, there were two each of mean depths 1.00, 1.14, 1.29, 1.43, 1.57, and 1.71. Thus the difference in mean depths for pairs of sentences ranged from -.71 to .71; a difference of -.71 obtained when, for example, the K member of a K-N pair had mean depth 1.00 and the N member had mean depth 1.71. Which kind of sentence occurred first in a pair was balanced over the magnitudes of difference in mean depth. The judgments were made immediately following the free-learning experiment, and by the same Ss. The sentences of the two experiments were completely unrelated semantically.

With respect to the resulting judgments of retainability as a function of sentence kind, Ns were chosen over Ks 54% of the time and Ps were chosen over Ks 50% of the time. As for judgments as a function of difference in sentence mean depths, Ss consistently chose the sentence of lesser mean depth as easier to remember: for K-N pairs, Kendall's tau between proportion of choices of K over N and the -.71 to .71 range of mean depth differences is .54 ($p = .01$); for K-P pairs, tau is .85 ($p < .001$). In a similar study comparing Ps and P_Ts, Ps were judged easier to remember than P_Ts 51% of the time; the corresponding tau is .93 ($p < .001$).

Thus, in a situation where Ss are asked to make judgments regarding storage and subsequent reproduction, that is, to exercise their knowledge of the language without an actual test of ability, one again finds that it is structural complexity and not sentence kind per se that accounts for behavior.

This is very convincing evidence, seemingly indicating the strength of the depth factor. Further evidence in support of the depth hypothesis is given by Martin and Roberts (1966). When the K, N, P, and NP sentences used by Mehler (1963) were analyzed for mean depth, it was found that the order in which sentences had been recalled (from easiest to hardest), that is, K < P < N < NP, was a direct reflection of the mean depths of the sentences: 1.17, 1.38, 1.43, and 1.67, respectively. It does not seem unlikely, then, that the mean depth factor played a part, perhaps even a very significant part, in determining the results of other studies which have helped to build the theories that predict K < N < P < NP or that K < P < N < NP.

It is advisable to recommend that not only should a replication of the present study be attempted, but also, an analysis of the stimulus sentences used by some of the investigators mentioned in this paper should be made using the mean depth measure. Perhaps, then, the roles of both sentence kind and sentence depth and their interaction will appear clearer.

CHAPTER V

SUMMARY

A study was performed which utilized Martin and Roberts' adaptation of Yngve's depth measure. Subjects aged 9 and 30 were exposed to sentences of either one of two mean depths and of six different types. Four groups, comprising 60 Ss each were named according to the age of the Ss in the group as well as for the depth of the sentences to which the members of the group were exposed. Ten Ss at a time from the groups 9(1.29), 9(1.71), 30 (1.29), and 30(1.71) heard the E read six sentences out loud. Each set of six sentences comprised a Kernel (K), a Negative (N), a Passive (P), a Negative-Passive (NP), a Passive Truncated (P_T), and a Negative-Passive Truncated (NP_T) sentence. There were 12 sets of sentences -- six sets at mean depth 1.29 and six sets at mean depth 1.71.

After hearing the set of sentences once, the Ss were asked to write as many of the sentences as they could remember. Six such trials were effected per 10 Ss with a different set of sentences being used for every 10 Ss.

It was found that all Ss recalled sentences of mean depth 1.29 correctly more often than sentences of mean depth 1.71. Further, Ss aged 30 scored more correct recalls at both mean depths and on each sentence type than did Ss aged 9. It was also found that, when the four main sentence types (K, N, P, NP) were considered, negatives of both types were recalled better than were non-negatives by both aged Ss.

The first-mentioned finding corroborates results of the Martin and Roberts (1966) study and indicates that sentence depth is a highly significant factor in a sentence recall situation. The second finding appears to give some indication that children aged 9 have either a shorter memory span in sentence recall situations than adults, or else are less likely to put down "answers" unless they are sure of being "correct" than adults are. It was suggested that both the memory span hypothesis and the "conservatism" hypothesis be investigated further. In the case of the last finding -- that regarding the affirmation-negation dichotomy found to exist between sentence types -- it was suggested that a mean depth analysis of the sentences used in the Gough, 1965, and Slobin, 1966 studies be made. It was felt that such a depth analysis might reveal that the Gough and Slobin results (that K sentences were easier to recall than Ns, which were easier than Ps, which in turn were easier than NPs) may have occurred as a result of the Ks having lesser mean depths than the other sentence types.

The Katz and Postal hypothesis that Ks, Ns, and Ps had separate derivations was considered as a possible explanation, but it was suggested that if this were so, the relationship between affirmative and negative sentences was more significant than relationships between other sentences separated by only one "universal morpheme".

It was further suggested that a replication of the present study be attempted and that another study be devised to test the hypothesis that the strong tendency to respond correctly to negative sentences is culturally determined.

REFERENCES

Chomsky, N. Syntactic structures. The Hague: Mouton, 1957.

__________ On the "notion rule of grammar". In R. Jakobson (Ed.), Structure of language and its mathematical aspects. (Proceedings 12th symposium in applied mathematics) Providence, R.I.: American Mathematical Society, 1961, 6-24.

Clifton, C., and Odom, P. Similarity relations among certain English sentence constructions. Psych. Monog., 1966, 80, 1-35.

Ferguson, G.A. Statistical analysis in Psychology and Education. New York, McGraw-Hill, 1959.

Flavell, J.H. The developmental psychology of Jean Piaget. Toronto, Van Nostrand, 1964.

Gough, P.B. Grammatical transformations and speed of understanding. J. verb. Learn. verb. Behav., 1965, 4, 107-111.

Harman, G.H. Generative grammars without transformation rules. Language, 1963, 39, 597-616.

Katz, J.J., and Postal, P.M. An integrated theory of linguistic descriptions. Cambridge: M.I.T. Press, 1964.

Lees, R.B. The grammar of English nominalizations, International J. Amer. Linguistics, 1960, 26, 1-25.

Lukasiewicz, J., and Tarski. Untershungen uber den Aussagenkalkul, Comptes Rendus des Seances de la Societe des Sciences et des Lettres de Varsovie, 23, CI III: 31-32, 1930.

Martin, E., and Roberts, K.H. Grammatical factors in sentence retention. J. verb. Learn. verb. Behav., 1966, 5, 211-218.

Mehler, J. Some effects of grammatical transformations on the recall of English sentences. J. verb. Learn. verb. Behav., 1963, 2, 346-351.

Menyuk, P. Syntactic structures in the language of children. Child Development, 1963, 34, 407-422.

Miller, G.A. Some psychological studies of grammar. Amer. Psychol., 1962, 17, 748-762.

Slobin, D.I. Grammatical transformations and sentence comprehension in childhood and adulthood. J. verb. Learn. verb. Behav., 1966, 5, 219-227.

Thorndike, E.L., and Lorge, I. The teacher's word book of 30,000 words. New York, Columbia University, 1944.

Winer, B.J. Statistical principles in experimental design. New York, McGraw-Hill, 1962.

Yngve, V.H. A model and an hypothesis for language structure. Proc. Amer. Phil. Soc., 1960, 104, 444-466.

__________ Implications of mechanical translation research. Proc. Amer. Phil. Soc., 1964, 108, 275-281.

APPENDIX A

STIMULUS SENTENCES

SETS I - VI

Mean Depth 1.29

I
1. Boys kick stones hard with their boots.
2. Cats can never escape from my dog.
3. Doors are always locked in our house.
4. Eggs were not laid in the haystack.
5. Ball is often played in the street.
6. Fires are not set by a fireman.

II
1. Men pile bricks fast on the wall.
2. She was not dressed for the cold.
3. Eggs were sometimes laid in the haystack.
4. Doors were not locked in our house.
5. Beets are often grown by a farmer.
6. Ball is never played in the street.

III
1. Pigs wash themselves clean in the tub.
2. They were not prepared for rainy weather.
3. Boats are often left at the river.
4. Meals were never eaten in the kitchen.
5. Prizes are sometimes given by the school.
6. Frogs are not liked by my sister.

IV
1. He sings songs loudly in the night.
2. He is not swimming in our pool.
3. Wood was burned quickly in the stove.
4. Boats are not left on the bank.

5. We were often met by our children.

6. Orders were not given by the boss.

V 1. Baby drinks milk slowly from a cup.

2. John was never happy in the rain.

3. Meals were always eaten in the kitchen.

4. Bells were not rung on 9 Monday.

5. Orders were quickly given by the boss.

6. We were not met by our children.

VI 1. He bought milk yesterday at the store.

2. I could never jump over the ditch.

3. Music was sometimes played on the radio.

4. Bags were not made for carrying stones.

5. Songs were loudly sung by the class.

6. Candy is not made by a baker.

SETS VII - XII

Mean Depth 1.71

VII 1. Children will sometimes go out after dark.

2. We can not make diamonds from coal.

3. She was gladly given help by Bob.

4. Webs were not spun carefully by spiders.

5. Dinner is served late here for guests.

6. Help is never hired here in summer.

VIII 1. Men were busily making fires from trash.

2. I can not win money at cards.

3. Jobs were cheerfully finished today by Mary.

4. Lunch was not eaten quickly by Bill.

5. Children are sometimes allowed out after dark.

6. Books were never stacked up in piles.

IX 1. We can sometimes win money at cards.

2. Hens will not lay two eggs daily.

3. Lunch was eaten quickly today by Bill.

4. Lights are not put out by mother.

5. Help is often hired here in summer.

6. Prizes were not given out at school.

X 1. Hens will sometimes lay two eggs daily.

2. He will not throw stones at you.

3. Webs are spun carefully here by spiders.

4. She was not given help by Bob.

5. Books were carefully stacked up in piles.

6. Dinner is not served late for guests.

XI 1. Animals will often creep about at night.

2. Children do not go out after dark.

3. Letters often come here now for Jane.

4. Jobs were not finished today by Mary.

5. Prizes were given out fairly at school.

6. Boys should not set fires in barns.

XII 1. He was certainly throwing stones at you.

2. Birds do not fly around at night.

3. Lights are put out here by mother.

4. Letters never come here now for Jane.

5. Skies everywhere are lit up by rockets.

6. Children are not allowed out after dark.

APPENDIX B

SCORING PROCEDURES

A. <u>Correct Recalls</u>

1. Any sentence which, word for word, matches the stimulus sentence, is a correct recall.

2. Sentences in which the following changes occur are correct recalls:

   a) where "a" was presented and "the" was written by the <u>S</u>, and vice versa.

   b) where an auxiliary (ie., "is") was presented and another form of the auxiliary (ie., "was") was written by the <u>S</u>.

   c) where "never" was presented and "not" was written by the <u>S</u>, and vice versa.

   d) where two words were presented but were written as grammatically correct contractions by the <u>S</u> (ie., "can't").

   e) where personal pronoun (ie., "she") was presented and a different personal pronoun was written by the <u>S</u> (ie., "he").

   f) where a possessive pronoun was presented (ie., "her") and a different possessive pronoun was written by the <u>S</u> (ie., "their).

B. <u>Errors</u>

1. A sentence is scored as a <u>transformation error</u> if it is, for example, presented as a K and reproduced as a N by the <u>S</u>.

2. A sentence is scored as an <u>error of the same type</u> as the presented sentence if, for example, a K is presented and a different K is reproduced by the <u>S</u>. ("Different" in that a change has been made which was not included in the changes listed under "Correct recalls" 2.)

3. A sentence is marked as being <u>incomplete</u> if there are words deleted which were included in the presented sentence.

4. A sentence is marked as being <u>ungrammatical</u> if it does not follow the established rules of traditional English grammar (see Chomsky, 1957) supposedly exemplified in the presented sentences.

5. <u>Omissions</u> occur when a presented sentence is not included at all in the sentences reproduced by the <u>S</u>.