LISTENING COMPREHENSION STRATEGIES
IN NATIVE AND SECOND LANGUAGE

By

Linda Conrad

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of English

1981

ABSTRACT

LISTENING COMPREHENSION STRATEGIES
IN NATIVE AND SECOND LANGUAGE

By

Linda Conrad

Studies of reading have shown that, when reading in a
non-native language, the reader typically cannot make full
use of the semantic-level cues in the text. Non-native
readers have been shown to direct proportionately more
attention to syntactic and grapho-phonic information relative
to the meaning they extract from a passage than native
readers. The present study attempts to explore whether this
processing difference can also be shown to characterize con-
trasts between first- and second-language listening compre-
hension.

Three groups of subjects were used in the study, a
native group, a German group proficient in English, and a
mixed-language-background group of intermediate proficiency
in English. Based on the general assumption that much of
the difficulty in comprehending an aural message in a second
language lies in the unavailability to the second-language
listener of semantic information in the text, the present
research made use of three experimental designs to test
three hypotheses regarding listening comprehension.

Experiment one tested the hypothesis that with increased proficiency in a language, a listener's processing of an incoming message shows a greater attention to the semantic cues than to syntactic or phonological cues. Subjects were given a 55-item post-listening cloze test, and responses were scored using a diagnostic scoring procedure developed by M. A. Clarke and L. Burdell (On TESOL 1977. Washington, D. C.: TESOL, 1977, 131-143); these were then submitted to a multiple analysis of variance. Results confirmed our hypothesis. The number of semantically acceptable responses showed a significant increase for the mixed, German, and native groups respectively while responses based on syntactic structure decreased.

Experiment two replicated the well-known idea-acquisition experiment of J. D. Bransford and J. J. Franks (Cognitive Psychology, 1971, 2, 331-350) to test our second hypothesis that more proficient listeners in a language are progressively less apt to remember the syntactic details of sentences relative to the amount of the content they remember. Although for native listeners our results supported our hypothesis, for non-native listeners our hypothesis was not confirmed. The German group performed similar to the native group on the test. The pattern for the mixed group was to reject sentences more complex than those they had actually heard during acquisition and to retain less of the semantic content. Results for this group, however,

were insufficient to conclude that retention for syntactic features was greater than that of the other two groups.

Experiment three tested the hypothesis that the recall of connected discourse is facilitated if the listener thinks the discourse fits together into a story, and that this facilitation increases with increasing proficiency in the language. In an experiment similar to one by P. A. de Villiers (Journal of Experimental Psychology, 1974, 103, 263-268), one group of subjects from each language group heard nineteen sentences with definite articles and one group heard the same sentences with indefinite articles. All subjects were asked to retell the passage and to indicate whether they thought the sentences formed a story or were unrelated. Mann-Whitney U tests confirmed as predicted that story- versus non-story-processing had a highly significant effect on the amount recalled by the native group, a slightly lesser effect on the amount recalled by the German group, and no significant effect for the mixed group.

The results of the present study lend empirical evidence to our general assumption that difficulty in listening to a second language is related to difficulty in extracting semantic information from the text. Our results cannot be used, however, to pinpoint the exact source of this difficulty or to define the precise processes involved.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Setting and Aims of the Present Study

One way of characterizing the process of listening
comprehension is to picture the listener as busy trying to
"unlock" the message or meaning in a flow of speech sounds.
However, the more closely psychologists have looked at the
process of listening, the more their studies have shown that,
at best, this analogy only represents a small part of what
goes on when a listener comprehends a spoken message. It
is an oversimplification to see the sounds of a message
linearly affecting meaning changes in the perception of the
listener. More accurately, the role of the listener should
be viewed as a very constructive one, which might be defined
as follows: On the basis of his knowledge of the language
of the message and his past experience and knowledge of the
world, the listener actively selects information from a mes-
sage in an attempt to recreate and modify his previous
knowledge.

On the basis of this definition, communication between
a speaker and a listener becomes a very complex process.
To be effective it would seem to depend on a shared knowl-
edge of what the world is like and a shared language which

expresses this knowledge. It is not enough to view comprehension as the passive effect of a message on a listener. One has to see it as part of a dynamic and interrelated communication process in which the knowledge and expectations of both the speaker and the listener have to be taken into account as well as the changing interrelations between speaker and message and listener and message (see figure 1).

```
┌──────────┐   ┌──────────┐   ┌──────────┐
│ Receiver │◄─►│ Message  │◄─►│  Sender  │
└──────────┘   └──────────┘   └──────────┘
```

Figure 1. A simple model of communication. (The simplest model of communication must take into account at least all the above components--receiver, message, and sender, as well as the relations among them. The boxes signify a common linguistic and cultural context).

Given this understanding of listening comprehension as part of a complex and dynamic interaction of the speaker, the message, and the listener, it is not surprising that psychologists are just beginning to understand how listeners comprehend messages. Indeed, presently, there is probably more disagreement than agreement on what kind of role each speaker, message, and listener has to play in the process and how important that role is in the three-way exchange. If complicating the analysis of the listening comprehension process has complicated research and the interpretation of research findings, it has also led to the important but simple realization that it is misleading to study the role

of the speaker, the message, or the listener as if that role were static or could be isolated from its part in the communication process. The study of comprehension seems to lead to interesting results only when one proceeds from complex to simple or from whole to part.

If this methodology of approach is correct, the present study need not apologize for trying to test the system when one further complication is added. What happens to the process of listening comprehension when the language of the message is changed from one common to both the speaker and listener to one that is native to the speaker but foreign to the listener? Because it need not be the case that a model of second- or foreign-language comprehension be a simple variation of native comprehension, it would seem desirable, using what is known about comprehension, to attempt to approach second language comprehension directly. Indeed, this seems to be the only viable means of approach; for, returning to our earlier model of communication, it becomes evident that altering one of the three participants, in this case the message, is not a simple change. The dynamic relationships among the three variables in communication will also be altered. As illustrated in figure 2, the introduction of a foreign language may well alter the process of communication as a whole.

Within this larger framework of communication the present study hopes to be able to contribute to a better
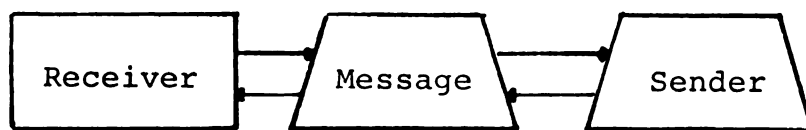
Figure 2. A simple model of communication in a foreign language. (The differently shaped boxes represent divergent linguistic and cultural contexts. It has not yet been shown that the arrows signify relationships similar to those in figure 1.

understanding of the process of second-language listening comprehension and to an assessment of the difficulties faced by a listener encountering a foreign language message. A series of listening tests in the English language were ad- ministered to groups of native and non-native language sub- jects. Through a comparison of the results of the two groups on each of the tests, the present research hopes to focus on ways in which the role of the foreign language listener in comprehension is different from that of the native language listener.

Three experimental designs which have successfully been used in past research to characterize aspects of native language listening or native language comprehension were chosen for the present study. The first experiment makes use of the cloze procedure and adapts it to a post- listening exercise. Each cloze test response was analyzed according to whether it conformed to the syntactic and semantic restrictions of the individual test sentence as well as to those restrictions imposed by the passage as a

whole. In addition, results from the experimental group were compared to results obtained from a control group that had taken the same cloze test without the advantage of having heard the passage. A second experiment was based on the well-known study conducted by Bransford and Franks (1971) to measure how well subjects are able to identify familiar sentences on the basis of their syntactic form. A third experiment, based on a study by de Villiers (1974), attempted to measure how a feature of discourse (in this case, the use of the definite or indefinite article) influenced the recall of nineteen sentences. Discussion of the results for all three experiments compares native-language-listener performance to that reported in the earlier research and discusses how this is similar to, or different from, the performance of the non-native listeners on the tests.

The three groups of subjects who participated in the present study included a group of native language subjects, a group of foreign language subjects with German as their common native language, and a group of foreign language subjects from mixed language backgrounds. This third group of subjects had a slightly lower level of proficiency in English than the German language group. To allow closer comparison to the native language listening task, all of the foreign language subjects who participated had reached a high level of proficiency in the language. All had scores

of between 70% and 96% on the Michigan State University English Language Examination, which also includes a test of listening comprehension.

It is hoped that a comparison of the groups' comprehension results on the administered tests will lead to an understanding of how a foreign language input alters the strategies a listener uses in comprehending a message. Recalling the complexity of our model, we can only hope to deal with one small area of second-language listening comprehension. Sociolinguistic, cultural, and educational factors, all of which influence the respective relationships among listener, speaker, and message lie beyond the scope of the present work. Based on the three experiments we have chosen, our main concern will be on how a listener processes a foreign language message. To construct a firm base for the present findings, we must first examine the work already carried out by many other researchers.

In Search of a Model for Foreign
Listening Comprehension

The first extensive work in the methodology of listening comprehension was carried out at the instructional level by foreign language teachers educated in the audiolingual school of foreign language teaching and at the theoretical level by followers of the structuralist school of linguistics. The concern of psychology at this time was for observable empirical relations governing the acquisition, transfer, and extinction of skills. Given our original

analogy of a listener decoding a message from a flow of speech sounds, it is not surprising that language psychologists and linguists first directed their attention to an analysis of the observable features in the message, which the listener was thought to decode. When one altered the listening model to include a foreign language instead of the listener's native language, it seemed that the most productive results could be achieved from contrasting the sound systems of the two languages, which made up the differing stimuli presented to the listener. This is not to say that one excluded the possibility of rules at other levels which might affect the phonological form of an utterance. It was a question of approach, however, that caused syntactic and even semantic rules to be discovered at the level of the phonological input to the listener, because this level formed the most immediate and measurable component of comprehension.

It is not surprising then that early studies of foreign language listening comprehension became almost preoccupied with contrasting the phonological systems of two languages. Likewise, the teaching methodology which accompanied these studies concerned itself with materials and techniques for teaching the foreign language learner the skill of sound discrimination in his new foreign language. The gains made in the field of phonology at this time were considerable and certainly invaluable to any study of either

native or second-language listening comprehension. The re-
alization that this approach, at best, could never lead to
a complete understanding of the process came when psycholo-
gists and linguists began questioning the foundations of the
model and also the nature of language and its role in com-
prehension.

One of the most important contributions of the new
cognitive school of research was to show the inadequacies of
the model of comprehension that underlay past experimenta-
tion and observations. It had been a fallacy to think that
human perception could be understood by ignoring the role
the human mind played in the process and looking only at
the external stimulus and an external response. After all,
cognitive psychologists insisted, it was not an event in
the world that a perceiver remembered, but rather his en-
coding or interpretation of that event. One could only hope
to understand listening comprehension when one focused on
the role played by the listener. This does not undermine
the findings of earlier studies of contrastive phonology and
listening discrimination. Rather these must be seen in
terms of an expanded and much more complex model of what
listening comprehension involves.

What Can the Foreign Language Teacher Contribute?

Unfortunately for the foreign language teacher, most
of the materials presently available for teaching listening

comprehension have been based on earlier contrastive studies, and as a result they concentrate primarily on teaching listening comprehension to the beginning foreign language student. To date there are still few listening materials available for the intermediate or advanced student. Materials by Crymes, James, Smith, and Taylor (1974), Morley (1972, 1976), Plaister (1976), and So (1979), however, are evidence that research in psycholinguistics and the psychology of perception are providing the framework for trying to better understand what listening comprehension in the foreign language classroom means at all levels of instruction.

Recently there has been growing sentiment that foreign language teaching cannot wait for its theoretical basis to be handed down from the fields of linguistics and psychology, but that any full understanding of second-language learning would have to be a collaborative effort by psychologists, linguists, and foreign language teachers. In the wake of psycholinguistic research with native language subjects, foreign language researchers have already begun to explore second-language acquisition and the process of reading comprehension in a second language. Although the field of listening has as yet received little attention in research, the increasing number of native-language-listening studies leads one to predict that the present situation is only a temporary one. The knowledge that has been gained from native-language-listening research combined with the

findings of foreign language research in the area of second language reading could serve as a guide to approaching the problem of second-language listening comprehension. The succeeding sections discuss recent findings in the three research areas that most directly bear on the present investigation: psycholinguistic research, research in discourse analysis, and second-language reading research.

Psycholinguistic Research

Early studies in psycholinguistics found evidence for a more abstract level of representation in memory than could be accounted for by the surface structure or deep structure of transformational-generative grammar. The seminal experimental work of Sachs (1967) showed that subjects in a recognition task could only very poorly detect various syntactic transformations made in sentences, but that they had no difficulties in spotting changes that had been made in the meanings of the test sentences. In research published in 1972, Bransford, Barclay, and Franks further discounted the hypothesis that what is stored in memory is a representation of the deep structure of a sentence. When their subjects identified two syntactically different sentences such as sentences one and two as one and the same sentence, the researchers concluded that, rather than remembering the deep structure of the test sentences, subjects were basing their responses on their memory of the overall semantic situation.

1. Three turtles rested on a floating log and
   a fish swam beneath them.

2. Three turtles rested on a floating log and
   a fish swam beneath them.

Further studies by Garrett, Bever, and Fodor (1966) at the level of phonology; Mehler, Segui, Pittet, and Barrière (1978) at the level of syntax; and Bransford and Franks (1971) and Kintsch and Monk (1972) at the level of semantics and text have confirmed the hypothesis that the listener does not process language at one level but rather uses his total knowledge at all levels to predict and confirm a message. Just how important each of these levels of knowledge is for an act of perception is a question currently being investigated.

According to the presently accepted view of perception, the listener stores in memory information from any or all of the levels of language as a result of the act of comprehension. Although a listener may process the language input at various levels, comprehension is usually equated with processing at the deepest semantic level (Mistler-Lachman, 1974). This seems to be a result of the storage and temporal capacity limitations of short-term memory.

It is now well-established that the phonemic and syntactic information, which might be seen as forming the medium for sentence meaning, almost completely disappears in long-term memory, whereas semantic information preferentially remains. The capacity limitations of short-term memory to four or five elements at a time (Miller, 1956)

can only be overcome by the listener's "chunking" the language input into elements larger than single sounds or syntactic features. The listener uses the syntactic and semantic knowledge he already has in his long-term memory to predict these larger units, which results in an increased memory span and a more efficient use of an otherwise limited storage capacity. This ability to "chunk" an incoming speech signal, on the other hand, is largely based on the perceiver's expectations at the phonological, syntactic, and semantic levels of language.

Kintsch (1974, pp. 256-259) has suggested that multiple traces are formed in memory during the comprehension of a message. Although the exact nature of these traces has not yet been established, some seem to be the result of a rather superficial encounter with the language of the input. These consist of samples from the phonetic, syntactic, semantic, and contextual elements processed in short-term memory. Kintsch predicts that as the listener accumulates more information about an aural message, he performs a deeper semantic analysis of the material, bringing to bear his past knowledge and whatever information the context has to offer. On the basis of this expanded knowledge the listener relates and organizes semantic information from the input, making inferences where necessary, and drawing on his full knowledge in comprehending the message. According to this theory, only this latter semantic information remains in long-term memory. The many clues about the surface features used in

short-term memory to arrive at this semantic framework are
generally lost from short-term memory as the listener
processes new cues from the incoming message. As a result,
these lower-level cues generally do not become fully rep-
resented in long-term memory and are thus subsequently no
longer retrievable, at least not in their complete form.

One can conclude that, given these memory-processing
constraints, the effective listener would attend only as
much as necessary to surface feature cues. He would use
instead the knowledge he has in long-term memory to enable
him to "chunk" large segments of the incoming aural message
in terms of its semantic components. His knowledge of the
language would allow him to predict the form and content
of the message and optimally to concentrate on extracting
important or new information from the input.

Discourse Analysis: The Role of the Text

The process of comprehension can best be viewed as an
exchange between the perceiver's knowledge and the informa-
tion from an incoming message. Even though there is as yet
no complete model of psychological decoding and retrieval
processes, psychologists have also begun to turn their
attention to the role the text plays in comprehension.

The full meaning of any sentence in a discourse can
be determined only in the context of the information pro-
vided by the other sentences, and the psychologist asks how a

perceiver is able to make use of this textual coherence in comprehension. By testing the effects different inputs have on memory, psychologists have been able to show that certain linguistic rules and conventions for organizing discourse are an important part of the expectations of the perceiver. Where these expectations are met, comprehension is enhanced; likewise, where these expectations are not met, comprehension is impaired.

Two levels of text organization are currently under study. Very high-level conventions often referred to as schemata (following Bartlett [1932]) or macrostructures determine how the content of large units of prose such as entire stories is organized. This underlying organization is what identifies a sequence of sentences as a story rather than as an assortment of unrelated events and details. It is assumed when a reader or listener has processed a story, that he or she has extracted the macrostructure from the text and not just the structure of the individual sentences. In addition, this structure is thought to be hierarchical in nature. At the top levels are the more abstract components of stories in general, for example, the setting, the main characters, the plot. At the lower levels of the hierarchy are the more elaborative details of a particular story. A reader or listener may miss or forget some of the structure of the text; but, in essence, a psychologist's description of the structure of a story is also meant to be

a description of the knowledge of a story which a perceiver
stores in memory.

Research seems to bear out the hypothesis that the
macrostructure of stories is a psychologically valid con-
struct. Studies of recall (Kintsch, Kozminsky, Streby,
McKoon, & Keenan, 1978; Rumelhart, 1977) have shown that
subjects do in fact retain the hierarchical structure of
the original text in their retellings of stories. The
propositions that stand at the higher levels in the text
structure are recalled much better than those at the lower
levels. Subjects are twice as likely to recall sentences
high in the macrostructure than those low in the macro-
structure.

In general, recall seems to be guided by the general
structure of stories as well as by the specific information
stored in memory about a particular story. A study by
Black, Turner, and Bower (1979) shows that recall for
stories that fail to develop from a consistent point of
view is much poorer than for narratives that meet this sty-
listic criterion. A consistent point of view may well be
a universal requirement for stories. Some of the features
of narrative structure, on the other hand, seem to vary
from culture to culture. In a study conducted by Kintsch
and Greene (1977), American university students showed
highly accurate recall for European fairy tales but gross
distortions and omissions in their attempted recall of

Indian myths.

At least two studies (Cirilo & Foss, 1980; Kintsch &
Kintsch, 1979) have shown that the formation of macro-
propositions or expectations about the structure of a story
is an essential part of the comprehension process and not
just an organizational technique applied when a reader or
listener is asked to give a summary of a particular story.
Cirilo and Foss measured reading times for individual sen-
tences in a story and found that high-level sentences took
longer to read than low-level sentences. Kintsch and
Kintsch compared summaries of scrambled and unscrambled ver-
sions of the same story and found that scrambling did not
affect their subjects' ability to summarize effectively but
it significantly slowed down their reading speeds.

All of this research emphasizes the "height" of a sen-
tence in the hierarchical structure of a story and how this
influences the way the text of the story is processed.
Analysis proceeds from high-level expectations down to the
sentences in the text. Another level of discourse currently
under investigation (Carpenter & Just, 1977; Cirilo & Foss,
1980; Haviland & Clark, 1974; Kintsch & Van Dijk, 1978;
Lesgold, Roth, & Curtis, 1979) examines the serial position
of a sentence in the text and attempts to determine the ease
of processing that sentence by the connections it has to
earlier sentences. This research focuses primarily on the
way syntactic information in a discourse directs the per-
ceiver's integration of textual information. Presumably,

where a listener cannot immediately identify the macro-
structure of a text or where he has few expectations about
the organization of its content he would have to rely on
lower-level syntactic cues.

Haviland and Clark (1974), using the linguistic termi-
nology of Halliday (1967), have proposed a set of general
conventions, the given-new contract, which states that each
sentence in a discourse should provide some new information
about an already known given object. In order to compre-
hend each sentence in a text, the listener searches memory
for the representation of the designated given object and
then at this point integrates the new information into his
already existing memory. Comprehension runs smoothly as
long as the information marked given by the speaker actually
has an antecedent in the working memory of the listener.
Where this is not the case the listener is forced to search
memory for a possible antecedent he can reinstate, or else
he uses inference to create a new antecedent. Both of these
conditions slowed comprehension in the research.

A study by Lesgold (1972) gives further proof that
the linguistic organization of individual sentences in a
discourse has an important influence on comprehension. This
study showed that when two sentences were syntactically re-
lated by pronominal reference, they tended to be compre-
hended and remembered together. After listening to compound
sentences such as sentences three and four, whose clauses

were linked by the conjunction <u>and</u> or by pronominal refer-
ence, subjects were asked to recall the gist of the sen-
tences given a prompt word. Findings showed that subjects'
recall was better in two ways for sentences like sentence
four. First, subjects recalled more words from the sen-
tences with pronominal references. Secondly, they also
recalled more words from the clause of the sentence which
did not contain the sentence prompt word. Lesgold concludes
that information from two clauses seems more likely to be
integrated in memory when that information is linked to a
common pronominal referent than when merely joined by the
conjunction <u>and</u>.

      3.   The blacksmith was skilled and the anvil was
          dented and the blacksmith pounded the anvil.

      4.   The blacksmith was skilled and he pounded the
          anvil which was dented.

In a later series of studies Lesgold, Roth, and Curtis
(1979) expand this finding to show that entire concepts
which were related to a target concept earlier in a dis-
course became reinstated in memory when the target concept
was mentioned again in a passage. The researchers explain
their findings in terms of <u>foregrounding</u>, a contextual fea-
ture Chafe introduced into discourse analysis. Chafe (1973)
used the term <u>foregrounding</u> to refer to the possible status
of a concept in a discourse after that concept has been in-
troduced. A concept is <u>foregrounded</u> at a particular point
in a discourse if common language conventions can assume
that the listener has the concept actively in mind at that

point. Chafe notes that the contextual feature <u>foregrounded</u>
tends to dissolve as more and more sentences in a discourse
are uttered in which the <u>foregrounded</u> item does not appear.
The earlier a concept has been mentioned without being an
active part of the discourse, the more likely it will be
treated by the listener as <u>backgrounded</u> (Chafe, pp. 50-53).
The Lesgold, et al. study shows that comprehension time in-
creases with the amount of memory one has to search in order
to find an antecedent that has become <u>backgrounded</u>. On the
other hand, mentioning a concept that is related to a target
subject that has become <u>backgrounded</u> in a passage facili-
tates comprehension as if that subject itself had been di-
rectly mentioned. These findings underline the importance
of the limits of the human processing capacity and also in-
dicate how the organization of discourse either contributes
to or detracts from the amount of information a perceiver
has in working memory at each point in processing a message.

A study by Kintsch and Van Dijk (1978) suggests that
the readability of prose passages can be successfully pre-
dicted on the basis of how well consecutive propositions
in the passage can be referenced or matched in memory by
<u>foregrounded</u> and/or high-level propositions in the discourse.

Although they do not directly use the term <u>foreground-</u>
<u>ing</u>, Carpenter and Just (1977) introduce a similar con-
struct, the <u>discourse pointer</u>, which they define as "a sym-
bol in the comprehender's mind that indicates the current
topic of the discourse or the perceptual context" (p. 217).

For example, if the discourse pointer is set by the following opening sentence: The ballerina captivated a musician in the orchestra during her performance, a reader or listener would easily integrate the additional subsequent information: The one who the ballerina captivated was the trombonist. The discourse pointer would have trouble, however, integrating either of the following inappropriate second sentences: The one who arranged the choreography was the stagehand or The one who captivated the trombonist was the ballerina. Neither of these sentences would fill the comprehender's expectations at this point in the text (p. 223).

In a series of experiments, Carpenter and Just demonstrate how traditional rules of writing and paragraph structure can be explained in terms of how they ease the reader's integrating task. Their studies show that linguistic devices direct the reader's, and presumably also the listener's, search processes. Prose which does not make use of linguistic devices to guide comprehension from given to new seems less readily comprehended than prose that delivers the information predicted by the discourse pointer.

All of these studies point to the fact that, given the limitations of the human processing capacity, linguistic elements in a discourse are an important aid to comprehension. These linguistic elements are part of the listener's knowledge of the language and in a discourse they act as cues to help the listener integrate the information from

the incoming message.

How Does Foreign Language Comprehension Differ from Native Language Comprehension?

If listening comprehension is understood in terms of the role it plays in the larger process of communication, common sense would lead us to predict differences between native and second-language listening comprehension. When a native language listener remarks about a message in his native language, "I didn't understand a word the speaker was saying," the difficulty does not usually lie in the fact that the listener actually had trouble understanding the words of the message. Rather, understanding failed at the more general level of understanding the gist or content of the speaker's message. These difficulties arose from not understanding how the words and clauses in the input sentences were related to other sources of information such as the general context of the message or previous sentences in the discourse. One may conclude that the discourse was poorly organized; perhaps it did not incorporate an accepted macrostructure or it did not follow the development from given to new information.

There are also times, however, when a native language listener actually does have trouble understanding the words a speaker is using. This could happen, for example, listening to a lecture in an unfamiliar field of science. Because many of the concepts are new, the listener does not have the prerequisite knowledge to follow the otherwise

familiar structure of the discourse. The lecture may have had excellent organization and the listener may recall some of the unfamiliar words at the end of the lecture when asked what the lecture was about: but because the listener cannot provide the background knowledge that the speaker is assuming, he fails to comprehend the message in the form it was presented.

What does it mean for a second-language listener to make the claim, "I didn't understand"? Even if we consider this response as an answer to a question about our science lecture, which we will suppose our second-language listener had the necessary academic background to have understood in his native language, it becomes apparent that the second-language listener is faced with even more possible reasons than the native language listener for failing to understand the content of a message. Although the concepts are familiar, the non-native listener may have had trouble with the words used for the concepts in the second language; and, in fact, this is probably only part of the problem. The syntax that joins the words into a message and the overall structure of the discourse might also be unfamiliar. Perhaps a combination of too many new pieces of information prevented the second language listener from keeping up with the speed of the delivery of the message. All of these potential comprehension barriers are independent of considerations of the form and organization of the message, which psychologists have shown to be important to native language

comprehension. Clearly nonverbal aspects of the message might also be unfamiliar and might also contribute to a second-language listener's misunderstanding a message.

Ideally, one would like to find a limited number of factors that would be able to account for the differences between native and non-native listening comprehension, but research is far from being able to state what these might be. For the present, one can only hope to explore the differences through many careful observations and examinations of both processes.

The only research data available on how foreign language learners process discourse comes from the area of second-language reading research. If we consider that both reading and listening are types of comprehension performance based on the same overall language proficiency or competence. one would expect second-language listening and second-language reading to have much in common. In fact, given that reading is a skill that is traditionally taught whereas listening is acquired, Bernard Spolsky (Jones & Spolsky, 1975, p. 58) has claimed that after a certain amount of exposure time, a learner's listening comprehension would most closely reflect his overall language proficiency. Of course any attempt to generalize from second-language reading to second-language listening on the basis of an overall language proficiency would have to first ensure that the foreign language learner had had the necessary exposure to the spoken language. Another important difference between the

listening process and the reading process, which would have to be kept in mind, is that in listening, the language learner has little control over the intake of language since this is controlled by the speaker. Otherwise, one would expect that many of the findings of second-language reading research would also be applicable to second-language listening.

The very diverse experimental studies carried out in second-language reading research all seem to point to similar conclusions about the nature of the second-language reading process as it compares to native language reading. Both the native and second-language reader are seen as sampling, predicting, and confirming from the three levels of cueing systems in the text: The grapho-phonic, the syntactic, and the semantic. The efficient reader predictably uses the least amount of information from the text. Since the semantic level is less redundant, in other words contains fewer overlapping sources of information than either of the other two levels, the efficient reader relies more heavily on these higher level cues. The redundancy in the lower levels of the language serves to confirm the reader's predictions of meaning or to notify him when they are in error.

The more proficient the reader, the more familiar he is with information at all levels of the text. Therefore, it would appear that with increasing proficiency, both native and second-language readers should be able to process

units of language of increasing length. (The origin and complete discussion of the theory underlying present reading research can be found in Goodman, [1967]).

An important distinction between native and second-language reading centers around the use of semantic-level information in the text. The second-language reader seems to differ from the native language reader in his use of semantic cues. A study by Macnamara (1970) showed that bilinguals given similar problem-solving tasks in their weaker and stronger languages were able to perform the operations described in the native language texts but not in the foreign language texts. Macnamara concluded that in their weaker language, subjects had to concentrate on visual or syntactic information in the message at the expense of full semantic comprehension.

In a study of eye-movement comparing native and non-native readers of English, John Oller (1972a) similarly attributed the more frequent and longer fixations of his non-native subjects to their need to concentrate on lower level cuing systems in the text.

Devine (1979), analyzing miscues of Spanish readers reading English as a foreign language, found that higher reading proficiency was characterized primarily by a reader's processing larger units of language and by an increase in the semantic acceptability of his responses. In a study of reading French, Cziko (1978), using random, anomalous, and meaningful passages modeled after those of

Miller and Isard (1963), also showed that below a certain proficiency level, the semantic contextual cues in a text were unavailable to the second-language reader. This study, comparing native readers and second-language readers at different levels of proficiency, found that although all of the groups seemed able to make use of the syntactic constraints present in the anomalous texts, only the native language readers and the most proficient non-native readers were able to take advantage of the semantic constraints that were present in the meaningful texts.

These findings suggest that while even readers with only beginning knowledge of a foreign language are able to use the syntactic constraints of a passage as a source of information in reading, only readers with a relatively high level of competence in a language are able to use the semantic constraints. Studies in psycholinguistics and discourse analysis have shown that access to the meaning of a passage is determined by the reader's familiarity with the language of the text at all levels, familiarity with the semantic content and macrostructures as well as familiarity with the linguistic forms (syntactic structures, words, graphemes) encoding the meaning of the message. A non-native reader may be handicapped in comprehension at several or all of these levels. It would seem reasonable to conclude, however, from the studies conducted so far, that the second-language reader first tries to cope with the more immediate and obvious lower-level cues in the

message. Comprehension becomes more difficult for him the more attention he has to devote to these linguistic forms. Without the advantage of an overall semantic idea to help "chunk" or integrate this large information load, information at these levels usually becomes lost in short-term memory and the reader has difficulty retaining meaningful information from the passage.

Hypotheses of the Present Study

In listening, a second-language learner has even less control over the input message than in reading. The reader can choose his reading speed or go back and re-read whereas the listener usually has little influence over the speed of an aural message. Listening is more easily distracted by outside noise. Furthermore, the phonological surface form of an aural message contains a far greater range of variation due to dialect and idiolect than the graphic surface form of a printed text. Without the advantage of an overall semantic idea to help integrate this increased information load, the learner, in listening to a foreign language even more so than in reading that language, would be expected to have to attend more carefully to lower-level cues in the message.

Because little research has been conducted in the area of foreign language listening comprehension, the present study must be viewed as an attempt to gather data in this field rather than as an attempt to confirm a

specified theory of listening comprehension. In the experiments that follow, a number of hypotheses are tested which elaborate on the differences between native and second-language listening comprehension as these can be predicted by our review of recent research findings.

Three experiments were chosen for the present research because they had all been successfully used in the past to characterize aspects of native language listening comprehension. Based on this available methodology and our general assumption that an important difference between native and non-native listening lies in the amount of attention each group spends on the various levels (phonological, syntactic, semantic) in the incoming message, this study hopes to test the following three hypotheses.

Hypothesis one tests the most obvious prediction of our basic general assumption, that with increased proficiency in a language, a listener's expectations of an incoming message as measured by a Reading Miscue Inventory technique (Goodman & Burke, 1972) should show greater attention to the semantic cues in the text than to the syntactic or phonological cues.

Hypothesis two states that more proficient listeners in a language will progressively be less apt to remember details of syntax in sentences they have heard. Past studies of native language listening have shown that listeners recall the semantic content but not the syntactic form of

sentences. Our assumptions about non-native listeners predict that with increasing attention to the grammatical form of a message, they would be more likely to recall specific linguistic patterns. The present measurement is not sensitive enough to allow us to distinguish between a subject who did not remember syntactic form because he processed the sentence semantically and the one who did not remember syntactic form because he could not decode the syntax. We assumed, however, that all the non-native students in the study were advanced enough to be able to decode the sentences syntactically, and that if they were attending, as hypothesized, to the syntax of the sentence they would be able to retrieve relatively larger amounts of it from long-term memory than the native language subjects in the study.

Hypothesis three specifies in more detail the advantage the native listener has over the non-native listener in processing discourse at the semantic level. This hypothesis predicts that the recall of connected discourse is facilitated if the listener thinks the discourse fits together into a story, and that this facilitation increases with increasing proficiency in the language. If we can assume that second-language listeners must spend more time processing lower-level cues in the message, they will be able to direct less attention to extracting semantic information from the input. For this reason, we can expect that at the lower levels of proficiency in our population, where subjects are dedicating large portions of their attention

to the linguistic patterns in the text, there will be lower

recall of semantic information. We expect this to be true

even for those in the lower-proficiency group who claim to

be able to detect a story in the text.

The three hypotheses are listed below in table one.

The experiments that follow will each be described in terms

of one of these hypotheses. In summary we will review the

evidence gathered in the data and discuss how our findings

support the predictions we have made.

Table 1

Hypotheses of the Present Study

---

1. With increased proficiency in a language, a
    listener's expectations of an incoming message
    as measured by a modified RMI technique should
    show greater attention to the semantic cues in
    the aural input than to the syntactic or phono-
    logical cues.

2. More proficient listeners in a language will
    progressively be less apt to remember details
    of syntax in sentences they have heard.

3. The recall of connected discourse is facili-
    tated if the listener thinks the discourse
    fits into a story; this facilitation increases
    with increasing proficiency in the language.

Limitations of the Present Study

Because there has been little past research in the area

of listening comprehension, the experiments that follow must

be viewed as purely exploratory in nature. At present a

theory of listening comprehension does not exist, much less

a theory of second-language listening comprehension. Exploring the process of second-language listening in light of recent research in native language comprehension might, however, help clarify just what a theory of second-language listening comprehension should include. Since work has just begun in the area of native language listening comprehension, clearly this study can touch only the surface of the problems involved in comprehending a second-language message.

All of the present testing was carried out in university foreign language classes or in university English classes so that the results can only be generalized to this student population. In addition, the classroom test situation affected test results, and the exact nature of this influence is obviously impossible to pinpoint.

Two other difficulties in interpreting the test results arise from the choice of subject groups who participated in the test. Ideally the two non-native groups of subjects should have had the same level of English proficiency. However, the Michigan State University English Language Examination, which both groups had taken, showed that the German language group had a higher level of proficiency in English than the mixed language group. The desired homogeneous-heterogeneous language background differences is therefore compounded by a difference in English proficiency. Thus caution in interpreting the differences between these two groups must be exercised.

The size of the test groups (18-34 subjects) is a
further limit to the strength of the conclusions drawn from
our results. Although each student recorded sex, age, na-
tive language, knowledge of foreign languages, length of
stay in an English-speaking country, it was impossible sta-
tistically to estimate the effects of these variables on
the subsequent test results because test groups were too
small. Rather than attempt to draw conclusions from too
small a data base, I have chosen merely to report these per-
sonal data with the description of the subjects partici-
pating in each of the following experiments.

CHAPTER II

EXPERIMENT ONE

Purpose

Evidence accumulated in recent testing research supports our earlier model of communication as a dynamic process consisting of more than the sum of its linguistic components. Because listening comprehension is now seen as strongly related to overall language proficiency, test researchers have tried to replace former listening tests, which attempted to measure listening comprehension by evaluating isolated skills attributed to the listener, with tests which attempt to measure comprehension in the larger context of communication. One test that seems to be a valid and reliable test of overall proficiency in a language is the cloze test. Research has shown cloze tests to be valid in measuring reading ability (Oller and Conrad, 1971) and to be sensitive to discourse constraints that range across sentence boundaries (Chihara, Oller, Weaver, & Chavez-Oller, 1977).

In a cloze test, the subject faces a meaningful passage in which blanks have been substituted for some of the words. Because a reader only uses some of the cues in a text to determine its meaning, his response to missing words should lend insight into the strategies he is using to

33

interpret other cues available to him. Examining these responses allows the researcher to observe the strategies a reader uses in gathering information from the material he is reading. Clarke and Burdell (1977) have suggested evaluating criteria for judging the acceptability and quality of cloze-test responses in reading.

So far there has been no test comparable to the cloze test which would provide insight into the ongoing process of listening comprehension as a subject encounters an aural message. A study by Page (1975) found that subjects' scores on a conventional cloze test, where subjects had had no prior contact with the passage, and on a post-oral reading cloze test of matched difficulty correlated significantly. After a prior oral reading, his subjects showed an average gain of 10.45 percentage points over their scores on the conventional test (Page, 1975, p. 385). Page concludes that a post-oral reading cloze test can be used as an indicator of the comprehension that took place during the initial oral reading.

The present study seeks results similar to Page's by using a cloze test of a text as an indicator of the comprehension that took place during listening comprehension. Although the test does not yield access to a listener's sampling, predicting, and confirming strategies during actual listening, one would expect that a subject's strategies in a second encounter with the same passage could be

similar to those he used in his first encounter with the text. Assuming that the cloze procedure can be used as an indirect method of measuring listening comprehension, experiment one seeks to confirm our first hypothesis that with increased proficiency in a language, a listener's expectations of an incoming message should show greater attention to the semantic cues. In addition, based on Page's findings, we will test the prediction that listening to a text before reading it significantly improves comprehension of the passage.

For this experiment, conventional cloze tests were administered to control groups from the three language groups participating in the study: a native language group, a foreign language group with German as the common native language, and a foreign language group of mixed language background. Post-listening cloze tests were administered to experimental groups from the same three language groups. The error analysis device developed by Clarke and Burdell was used to evaluate all responses from both the control and experimental groups.

The Clarke and Burdell scoring procedure provides a scale for rating the semantic and syntactic acceptability of each cloze test response both on the immediate sentence level and in the larger context of the text as a whole. We have assumed that whereas native language listeners process a message using primarily semantic units, second-language listeners must direct more of their attention to lower-level language units. Furthermore, we expect that cloze

test scores will be higher for subjects more proficient in English than for subjects with lower proficiency.

After all responses have been evaluated on the basis of semantic and syntactic acceptability, a comparison of the native and non-native listener scores based on the Clarke and Burdell rating scale should show significant group differences on the mean number of semantically acceptable responses and the mean number of responses acceptable only on the syntactic level. Results are expected to show that native and more proficient second-language subjects are better able to make use of the semantic and discourse constraints in the text than are subjects less proficient in English as a foreign language.

As an additional test of our use of the cloze procedure as a post-listening exercise, it is predicted that, as in Page's study, the subjects in the experimental groups, having already encountered the text in its complete form, will have higher comprehension as demonstrated by their cloze test scores than subjects in the control groups, who did not have this advantage.

Method

Subjects. Three groups of university students participated in the test; each group of subjects was divided into an experimental group and a control group, making six groups of subjects in all. Classroom teachers asked for students who would volunteer to participate in an

experimental study of listening comprehension. The native speaker subjects were volunteers from three freshman English courses at Michigan State University; 34 students participated as the experimental group and 28 participated in the role of native control group.

The two groups of German subjects consisted of 50 freshman English majors enrolled at the University of Salzburg, Austria. The 22 students who volunteered to participate as the experimental group were enrolled in an elective course designed to improve pronunciation. The students from this group, like the 18 control group participants, were attending their first semester of university English classes. Because the Austrian high school English program is state-supervised, it can be assumed that both groups of students had similar preparation for university. Most of the students had at least visited either Great Britain or the United States. The experimental group of subjects was given the Michigan State University English Language Examination one week before they participated in the present test, and their total average scores on tests of listening comprehension, grammar-vocabulary, and writing ranged between 83.7 percent and 96 percent. Their test results would place these students in the E-level of the Michigan State University Language Center based on their mastery of English.

D- and E-level students at the Michigan State University English Language Center formed the two groups of

foreign language subjects of mixed language background. The twenty-five D-level students who comprised the experimental group were tested instead of attending their regularly held five-day-a-week language laboratory course. Thirty E-level students wrote the control test as an optional exercise in their normally scheduled English class, which meets five times a week. D-level students are the most advanced students at the English Language Center who receive full instruction in English, four hours a day, five days a week. Their scores on the Michigan State University English Language Examination fall between 70 percent and 80 percent. Because the number of D-level students was too small to yield two groups of subjects, E-level students were asked to participate as a control group for the test. E-level students rank higher than D-level students in overall proficiency but they have one or two skills that fall below this level.

The breakdown of the experimental mixed language group by native language is as follows: five speakers of Arabic, one of Greek, six of Japanese, three of Korean, six of Persian, seven of Spanish, one of Thai, and one of Turkish. Not all students in the control group turned in a personal data sheet along with their test results. For those who included this personal information, the languages which were represented in this group are as follows: two speakers of Arabic, one of German, one of Greek, four of Japanese,

two of Persian, two of Portuguese, and two of Spanish.

All of the students who participated in the experiment were asked to fill out a page of personal information, which was thought to be valuable in analyzing the results of the test. All students were asked to give their age, sex, father's occupation, and other foreign languages they knew along with a rating of their fluency in each. Non-native speakers were further asked to give their native language and how much time they had spent in an English-speaking country.

Because my original test groups were too small to allow a meaningful statistical analysis of these personal effects, I have presented this data in tables 2 through 4. The reader is asked to refer to these tables for a characterization of the subject population in the subsequent tests.

Design and materials. A cloze test of fifty-five items was constructed from a passage taken from a listening comprehension text which forms part of the Michigan State University English Language Center D-level language laboratory course (see appendix A). The first sentence of the text was left intact, and subsequently every fifth word of the passage was removed and replaced by a blank of uniform length. A numerical rather than a selective deletion procedure was chosen so that the Clarke and Burdell means of analyzing students' responses could be applied. Instead of

Table 2

Breakdown of Test Groups by Age and Sex

Test group	Mean age	Age range	No. of females	No. of males
Native experiment Experiment 1 n=34	21.0	18 - 31	28	6
Native experiment Experiments 2 & 3 n = 29	21.6	19 - 29	25	4
Native control n=28	21.5	17 - 33	20	8
German experiment* n=22	19.3	18 - 23	13	9
Mixed experiment n=25	25.6	19 - 36	7	18
Mixed control** n=30	26.0	20 - 37	3	11

* The only personal information available for the German control group with 18 S's is that 13 of these were females and 5 were males.

** There are 16 missing values for this group.

Table 3

Breakdown of Test Groups by Length of Stay
in an English-Speaking Country

Test group	Length of stay	No. of subjects
Native experiment Experiment 1 n=34	---	---
Native experiment Experiment 2 & 3 n=29	---	---
Native control n=28	---	---
German experiment* n=22	1 2	7 15
Mixed experiment n=25	0 2 3 4 5	2 9 8 5 1
Mixed control** n=30	0 2 3 4 5	15 2 4 5 3

Note. Numerical values in the table for length of stay refer
to the following categories:
0 = missing value   4 = 6 months to 2 years
1 = none   5 = more than 2 years
2 = visit less than 6 weeks
3 = live less than 3 months

   * This information is not available for the 18
     members of the German control group.
  ** There are 16 missing values for this group.

Table 4

Breakdown of Test Groups by Knowledge of One
or More Foreign Languages

Test group	Foreign language	No. of subjects
Native	0	1
experiment	1	10
Experiment 1	2	16
n=34	3	6
	5	1
Native	1	5
experiment	2	17
Experiment 2 & 3	3	7
n=29		
Native	0	2
control	1	12
n=28	2	13
	3	1
German	4	2
experiment*	5	19
n=22	6	1
Mixed	0	4
experiment	2	15
n=25	3	4
	5	2
Mixed	0	15
control**	2	10
n=30	3	2
	5	2

Note. Numerical values in the table for foreign language
refer to the following categories:

0 = missing value	4 = mastery of 1
1 = none	5 = mastery of 1 and intermed.
2 = intermed. knowl. of 1	knowl. of 2
3 = intermed. knowl. of 2	6 = mastery of 2 and more
and more	

* This information is not available for the 18
members of the German control group.
** There are 16 missing values for this group.

selecting the words to be deleted, which could unevenly weight the test in favor of either the syntactic or semantic cues available to the student, this test followed the procedure of deleting every fifth word.

Procedure. Those students who took part in the experiment as members of a control group were tested as a group by their classroom instructor. The students were given the instructions in English to fill in each blank with what they considered to be the best single word. They were told to guess if they were not sure of the right word. The German control group was told that they would only be allowed ten minutes to complete the exercises. The two other control groups of native and mixed language background subjects were allowed as much time as they needed to complete their versions of the text.

This difference in test times among the control groups parallels similar test-time differences among the three experimental groups. The German experimental group participated in all three tests included in the present study in one of their 90-minute laboratory periods. Because time was just sufficient to complete all three of the tests, only ten minutes were allotted for the written exercise in experiment one.

The native experimental group which participated in the first experiment was not the same as the one which participated in experiments two and three. Both native groups

of subjects were allowed as much time as they needed to complete the exercise. The mixed language experimental group took the three tests included in the present study on two consecutive days in a 50-minute laboratory period. The cloze test for experiment one was administered the first day, and students were allowed as much time as they needed to complete the test.

The three experimental groups all took the test in a language laboratory. Taped instructions in English were played from the main console over student headphones. Each student had a test booklet containing a personal data sheet and answer sheets for experiments one, two, and three. Students were asked to keep their answer sheets closed while they listened to a short lecture on nonverbal behavior and communication. After the lecture, students were asked to turn to experiment one in their worksheets, where they would find a printed text of the lecture in which some of the words had been left out. They were asked to take as much time as they needed to fill in as many blanks as they could with the best single word. Again, the German experimental group was told they would only be allowed ten minutes writing time to complete the test. I noted during testing that all students had finished writing within this time limit.

Scoring. The cloze test responses were scored following the procedure suggested by Clarke and Burdell for evaluating the quality of errors produced in a cloze test

administered for reading diagnosis. Based on the Goodmans'
reading miscue analysis (Goodman, 1967; Goodman & Burke,
1972), this procedure allows the researcher to evaluate
each subject response on the basis of its syntactic accept-
ability (SYNAC), semantic acceptability (SEMAC), and the
semantic change from the original text (SEMCH).

For the present research, one minor alteration was
made in the Clarke and Burdell response codes so as also to
include a rating for the sound similarity of a response to
a word in the original text: The SEMAC 9 rating, which
their research had used for responses with indeterminate
contextual motivation, was adapted in this study to score
responses which seemed to be based on phonological cues a
subject had encountered in the aural text. With this excep-
tion, the codes taken from Clarke and Burdell are listed
below with example responses. Following each example re-
sponse is the score it was given on the present test.

Syntactic Acceptability (SYNAC)

4. totally acceptable:

    Try to picture in your mine how the policeman
    walked as he approached. (SYNAC 4 SEMAC 6-2)
    (looked)

3. acceptable in the sentence; the response satis-
   fies sentence level syntactic constraints, but
   violates discourse constraints:

    Is his posture erect and defiant?
    (Was)
    (The passage requires the past tense). (SYNAC 3
    SEMAC 3)

2. acceptable only with the following portion of
   the sentence; from the response on, the sen-
   tence is syntactically acceptable:

   The picture showed him begging, almost
   ___chat___ with the driver. (SYNAC 2 SEMAC 2)
   (pleading)

1. acceptable only with the preceding portion of
   the sentence; the sentence is syntactically
   acceptable up to and including the response:

   You __are__ seldom see an American policeman
        (will)
   approach someone this way. (SYNAC 1 SEMAC 1)

0. totally unacceptable:

   Perhaps the Italian policeman would speak more
   firmly than his ___wear___ posture suggests.
                   (pleading)
   (SYNAC 0 SEMAC 0)

Semantic Acceptability (SEMAC)

6. totally acceptable:

   Perhaps you have been pulled over yourself
                         (stopped)
   because you didn't see a stop sign or drove too
   fast. (SYNAC 4 SEMAC 6-3)

5. totally acceptable if syntactic constraints are
   ignored; the sentence and/or response requires
   minor syntactic changes:

   Think back and try to picture in your mind
   how the policeman ___walk___ as he approached.
                     (looked)
   (SYNAC 4 SEMAC 5)

4. acceptable in the sentence; the response violates
   some passage-level meaning constraints:

   How did he walk as __you__ came to the car and
                      (he)
   how did he stand? (SYNAC 4 SEMAC 4)

3. acceptable in the sentence if syntactic con-
   straints are ignored; the sentence and/or re-
   sponse requires minor syntactic changes to be-
   come acceptable at the sentence level:

   Was his posture erect and defiant or did he
   look __the car__ if he was afraid to go near
         (as)
   the driver? (SYNAC 4 SEMAC 3)

2. acceptable only with the following portion of
   the sentence; from the response on, the sen-
   tence is semantically acceptable:

   The picture showed him begging, almost
   quarreling with the driver. (SYNAC 4 SEMAC 2)
   (pleading)

1. acceptable only with the preceding portion of
   the sentence; the sentence is semantically
   acceptable up to and including the response:

   The picture showed him begging, almost
   __crying__ with the driver. (SYNAC 4 SEMAC 1)
   (pleading)

0. totally unacceptable:

   The picture showed him begging, almost
   __killing__ with the driver. (SYNAC 4 SEMAC 0)
   (pleading)

9. sound similarity: it is difficult to determine
   a contextual motivation for the response; it
   seems to be based on phonological information in
   the text:

   Perhaps the Italian policeman would speak more
   firmly than his __bleeding__ posture suggests.
             (pleading)
   (SYNAC 4 SEMAC 9)

All responses receive a SYNAC rating and a SEMAC rating.
If a response is coded SEMAC 6, indicating that it is to-
tally acceptable semantically, it is further evaluated for
semantic change.

Semantic Change (SEMCH)

text. the response is identical to the original
     text.

3. no change: the exact response and the observed response are synonymous, given the context:

Perhaps you have been <u>ticketed</u> yourself
(stopped)
because you didn't see a stop sign or drove too fast.

2. minor change: there is a change in connotation:

Was his posture erect and ___<u>tall</u>___?
(defiant)

or the response is logical, but changes the meaning of the sentence slightly:

<u>Probably</u> the Italian policeman would speak
(Perhaps)
more firmly than his pleading posture suggests. . .
(SYNAC 4 SEMAC 6-2)

1. major change: the response conveys the author's meaning, but is "unusual" or "non-native":

From our experience in other similar
<u>examples</u>
(situations)
we can be sure that someone who approaches us
with the gesture of pleading is going to be
easier to deal with than someone with his hands
at his side ready to shoot or fight. (SYNAC 4
SEMAC 6-1) (Clarke & Burdell, pp. 6-7)

Scoring followed conventions used in the Clarke and Burdell research. Misspellings were scored correct unless the response could be construed to be another English word, in which case the resulting word was rated in the given context. Multiple word responses and blanks left uncompleted were evaluated using the same procedure as the one-word responses.

It is important to point out that the scoring procedure used here and by Clarke and Burdell does impose certain restrictions on the results. Both studies have

followed Goodman's principle (Goodman & Burke, 1976, p. 173) that the semantic acceptability of a response never be allowed to exceed its syntactic acceptability. In this way partial objectivity is aimed for by always assessing meaning within a syntactic framework. Two codes, SEMAC 3 and SEMAC 5, designate responses which, except for minor syntactic error,would be coded semantically acceptable at the sentence level (SYNAC 4) or semantically acceptable at the text level (SEMAC 6) respectively. These allow non-native subjects to make such errors as non-agreement of subject and verb, an incorrect verb tense, incorrect use of a function word, or incorrect choice of preposition, etc., without detracting from the semantic acceptability of their response. However, responses that violate the acceptable clause structure rules in English, inasmuch as these syntactic features also carry a heavy semantic load, cannot receive a higher semantic rating than that given to the syntax of the response in question.

A response was given the rating SEMAC 9, that it was based on phonological cues in the text, only when it was a deviation from the original text, or in Goodman's terms, a miscue. Responses coded SYNAC 4 SEMAC 4-6-text although they may contain phonological information from the aural message, presumably are based on syntactic and semantic cues as well.

Analysis of data. After a student's 55 responses were scored and categorized according to the scoring procedure outlined above, responses of the various categories were totaled for the entire text. These total scores yield for each subject a profile which shows how many of his responses fall into each of the nine following code categories:

1. exact word: number of responses to restore the original word from the text (Responses coded SYNAC 4 SEMAC 6-text).

2. exact content word: number of responses to restore the original content word from the text (Responses coded SYNAC 4 SEMAC 6-text. This total is a sum over only the content words, the nouns, verbs, adjectives in the text).

3. exact final word: number of responses to restore the original word from the final paragraph in the text (Responses for answers 50-55 coded SYNAC 4 SEMAC 6-text).

4. no meaning loss: number of responses that indicate no meaning loss (Responses coded SYNAC 4 SEMAC 6-text, SYNAC 4 SEMAC 6-3, SYNAC 4 SEMAC 6-2, SYNAC 4 SEMAC 6-1, SYNAC 4 SEMAC 5).

5. minor syntax error: number of responses that indicate no meaning loss but show minor syntactic error (Responses coded SYNAC 4 SEMAC 6-1 and SYNAC 4 SEMAC 5).

6. strong context: number of responses that indicate an effective use of contextual cues from the text (Responses coded SYNAC 4 SEMAC 6-text, SYNAC 4 SEMAC 6-3, SYNAC 4 SEMAC 6-2, SYNAC 4 SEMAC 6-1, SYNAC 4 SEMAC 5, SYNAC 4 SEMAC 4, SYNAC 4 SEMAC 4, SYNAC 3 SEMAC 3).

7. weak context: number of responses that indicate a weak use of contextual cues from the text (Responses coded SYNAC 0 SEMAC 0, SYNAC 1 SEMAC 0, SYNAC 1 SEMAC 1, SYNAC 2 SEMAC 0, SYNAC 2 SEMAC 1, SYNAC 2 SEMAC 2).

8. syntax over semantic: number of responses that
    indicate a reliance on syntactic cues over seman-
    tic cues in the text (Responses coded SYNAC 4
    SEMAC 2, SYNAC 4 SEMAC 1, SYNAC 4 SEMAC 0, SYNAC
    3, SEMAC 2, SYNAC 3 SEMAC 1, SYNAC 3 SEMAC 0).

9. sound similarity: number of responses that indi-
    cate a reliance on phonological cues from the
    text (Responses coded SYNAC 4 SEMAC 9, SYNAC 3
    SEMAC 9, SYNAC 2 SEMAC 9, SYNAC 1 SEMAC 9, SYNAC
    0 SEMAC 9).

A generalized multivariate analysis of variance was then
performed based on these subject profiles to determine
whether there were significant differences among the six
test groups.

Results and Discussion

For each of the six groups participating in the test,
the mean number of responses falling in each of the nine
above categories, along with their standard deviations, are
given in table 5. An initial generalized multivariate
analysis of variance to measure the effects of the nine
variables demonstrated that the ninth variable sound simi-
larity was completely linearly dependent on the other eight
variables and was not reporting anything not already re-
ported by the other variables. Sound similarity was
therefore not included in the subsequent analysis.

The remaining eight variables, as might be expected
from the scoring method we used for the cloze responses,
were found to show a high degree of correlation among them-
selves. Table 6 gives the correlation matrix of the eight
scores given each individual on the 55-item cloze test

Table 5

Mean Number of Responses (and Standard Deviations) by Group for Nine Response Types on a 55-Item Cloze Test

Response Type / Group	Exact word	Exact content word*	Exact final word**	No meaning loss	Minor syntax error	Strong context	Weak context	Syntax over semantic	Sound similarity
Native exptl. n=34	34.68 (4.62)	12.88 (2.18)	1.26 (.99)	50.47	2.00 (1.41)	51.88 (2.20)	2.18 (1.75)	.91 (1.06)	.03 (.17)
Native control n=28	29.25 (6.19)	11.61 (2.81)	1.39 (1.20)	49.61 (4.50)	2.07 (1.44)	51.96 (3.27)	2.18 (2.51)	.89 (1.20)	0 (0)
German exptl. n=22	29.59 (6.20)	11.04 (3.52)	0.86 (1.12)	46.00 (5.45)	3.68 (1.86)	48.36 (4.61)	4.04 (3.40)	1.54 (1.56)	0.13 (.35)
German control n=18	22.89 (3.56)	7.22 (1.40)	.44 (.98)	43.28 (3.75)	3.33 (1.33)	46.61 (3.31)	5.33 (2.91)	3.06 (1.59)	0 (0)
Mixed exptl. n=25	16.20 (4.67)	4.72 (1.93)	.64 (.91)	33.6 (5.39)	6.04 (2.01)	36.76 (4.46)	13.44 (3.23)	4.60 (2.10)	.20 (.41)
Mixed control n=30	14.13 (4.36)	3.56 (1.98)	.03 (.18)	34.63 (5.80)	6.83 (1.93)	38.83 (4.94)	11.57 (4.07)	4.57 (1.98)	.03 (.18)

* There are 23 blanks for content words in the text.
** This refers to the last 5 blanks in the text (the final paragraph).

(variables 1-8).

Table 6

Correlation Matrix for the
Eight Individual Cloze Test Scores

		2	3	4	5	6	7	8
1.	exact word	.94	.60	.89	-.68	.87	-.81	-.74
2.	exact content word		.56	.84	-.67	.83	-.76	-.73
3.	exact final word			.53	-.39	.51	-.46	-.48
4.	no meaning loss				-.52	.98	-.92	-.77
5.	minor syntax error					-.52	.47	.50
6.	strong context						-.97	-.71
7.	weak context							.52
8.	syntax over semantic							

The unusually high correlation coefficients among
almost all of the eight original variables demonstrated
that, in fact, these variables all tended to measure the
same thing. In other words, it would be impossible to
assess the independent contributions of each of the eight
variables. A principal component analysis was carried out
to test whether and how these eight variables could be con-
densed to a smaller number that would result in the same
information but provide a more economical scoring measure
for the test. The results of this analysis are presented
in table 7. As evident from the table, the first component

Table 7

Correlations of Principal Components with Eight Cloze Test Scores

Component	Variance (Eigen-value)	% of variation accounted for	Component Weights							
			exact word	exact content word	exact final word	no meaning loss	minor syntax error	strong context	weak context	syntax over context over seman-tic
1	5.90	73.7	.96	.93	.64	.96	-.68	.95	-.88	-.80
2	.73	9.1	-.07	-.10	-.38	.24	.51	.29	-.39	.12
3	.63	7.8								
4	.47	5.9								
5	.20	2.5								
6	.05									
7	.02									
8	.00									

accounts for 73.7 percent of the between-individual varia-
tion on the cloze test scores in the present test. This
component reflects a contrast between the measures minor
syntax error, syntax over semantic, and weak context and
the remaining measures. The factor loadings for exact word,
exact content word, no meaning loss, strong context show
that these measures are all common to the component.
Exact final word shows a slightly lower positive correlation
with the other four measures, but all of them are clearly
measuring the same thing, something that is highly nega-
tively related to scores on the measures minor syntax error,
syntax over semantic, and weak context.

A graphic representation of how the original variables
contribute to make up this new variable (table 8) furnishes
evidence for calling the new variable comprehension. Test
scores of exact word, exact content word, no meaning loss,
strong context are all measuring the effect of this princi-
pal component with correlation values greater than .93.
Test scores of syntax over semantic and weak context are
negatively measuring the new variable with correlation
values of -.80 and -.88 respectively. Perhaps the rela-
tively high negative correlation of minor syntax error and
the new comprehension variable at first seems surprising,
especially in view of the fact that in our scoring procedure
a minor error in syntax still resulted in a score of no
meaning loss. Possibly this test result is an artifactual
effect of our subject populations. If we predict that both

Table 8

Relative Weights of the
Original Eight Variables in the Principal Component Comprehension

+ exact word	+ exact content word	+ no meaning loss	+ strong context	+ exact final word	− minor syntax error	− syntax over semantic	− weak context
====	=======	=======	=======	=======	-------	-------	=======

Table 9

Relative Weights of the
Original Eight Variables in the Second Principal Component

+ minor syntax error	+ strong context	+ no meaning loss	+ syntax over semantic	− exact word	− exact content word	− weak context	− exact final word
======	------	=======	=======	=======	=======	=======	=====

groups of native language subjects will score higher on
comprehension and all non-native subjects will have lower
comprehension scores, one would also expect to find a higher
number of minor syntax errors in the non-native group scores
than in the native group scores.

It is interesting that the second principal component
(table 7), which accounts for only 9.1 percent of the
between-individual variance, seems to be a contrast of just
these measures: minor syntax error and strong context on
the one hand and weak context and exact final word on the
other. The relative weights of the original eight variables
of this second component are depicted in table 9. As shown
by the values of the factor loadings in table 7, the con-
tribution of some of the original eight variables to this
component, in contrast to their contributions to the first
principal component, is very small. Because of this, and
because the component accounts for less than 10 percent of
the between-individual variation in the test, the second
principal component will not be included in the succeeding
analysis of variance of our six test groups. Nevertheless,
the contrast which the component expresses may show an in-
teresting if minor trend to the data. The grouping minor
syntax error and strong context is most probably due to the
definition of both of these measures. The grouping exact
final word and weak context may be some evidence for our
prediction that a failure to be able to make use of the
semantic cues in a text will force the listener to rely on

lower-level syntactic or phonological cues. Ability to re-call the final paragraph (in the case of our text, the most difficult paragraph) might be showing a recency effect and reliance on phonological cues.

Based on our first principal component, which compo-nential analysis has shown to be measuring 73.9 percent of the between-individual scores of the cloze test, we can now express our original test scores in a much simpler form. Instead of giving each student eight different scores for each of the eight variables, we can measure the same entity using the best combination of these variables as these are expressed by the first principal component. Because of the composition of this component (tables 7 and 8), we have chosen to call this single combined score comprehension.

It is worth noting that our factor loadings show the measure exact word, with a loading of .96 to have along with no meaning loss the highest loading on the first principal component. If one were looking for the easiest scoring procedure which would yield the same test results, our test shows that one would be justified in scoring only those responses as correct which supplied the exact word from the original text. This is the traditional scoring method used for cloze tests with native language subjects (Taylor, 1953). Although most studies in second-language research have used a scoring method that gives the non-native subject credit for contextually acceptable responses, research has

indicated that such a scoring method does not lead to different results (Oller, 1972b). The componential analysis of the eight original variables in the present study bears out this conclusion from earlier research. However, also in line with earlier research, there seems to be a psychological reason to choose a scoring method which attempts to be more sensitive to the comprehension proficiency of the second-language learner. Furthermore, the present study demonstrates that such a scoring method is more exact. Even though in the present study all of the eight variables, with the possible exception of exact final word, contribute highly to our single whole factor, which like exact word scoring also yields only one score, it should be noted that these other variables are not perfectly correlated with the original variable exact word. It would seem then that there still remain differences in our test groups which would not be measured using an exact word scoring procedure. For this reason we chose to use the new variable comprehension as the single sampling characteristic in the subsequent analysis of variance to test for group differences among the language groups and between the experimental and control conditions.

The arrangement of our three language groups by treatment is diagrammed in table 10, with the respective test group numbers. Table 11 gives the observed means (and standard deviations) for the new test measure comprehension for each of the six groups.

Table 10

Sampling Diagram for Cloze Test Study

| | | Treatment | |
		Experimental	Control
	Native	34	28
Language	German	22	18
	Mixed	25	30

Table 11

Observed Means (and Standard Deviations) for <u>Comprehension</u>

| | | Treatment | |
		Experimental	Control
	Native	.91 (.41)	.77 (.53)
Language	German	.45 (.65)	-.07 (.38)
	Mixed	-1.01 (.46)	-1.08 (.47)

A test to determine whether there were interactions between groups showed that there were none. We then used the model to explore the following contrasts: Are there differences which are statistically different from zero among the population means, first for the main effect of treatment (experimental versus control), and secondly for the main effect of language, for the native group versus the two non-native groups (German + mixed) and for the German group versus the mixed language group?

The calculation of the estimated effects for the first contrast, experimental versus control, shows a relatively small effect of treatment, however, one which proves to be significant [F (1,151) = 9.25, p< .0028]. This supports our original hypothesis that listening to a test prior to taking a cloze test on the same text significantly improves comprehension.

The second contrast we were interested in was whether there is a significant effect for language between the native and non-native test groups. The language difference proves to be highly significant [F (1,151) = 245.76, p<.0001] supporting our prediction that increased proficiency in a language leads to better comprehension, as measured by a cloze test.

Although we would not normally predict differences between the German and mixed language groups, we have had to do so in this study because the groups were shown to

differ in language proficiency. (Michigan State English Language Examination scores for the mixed experimental group were all under 80 percent. In contrast the German experimental group had scores ranging from 83.7 to 96 percent). Based on this difference, we would expect the German group to score higher in comprehension on the cloze test. This difference is also shown to be significant [F (1,151) = 146.19, p<.0001]. Again, our hypothesis that increased proficiency in a language leads to better listening comprehension is supported.

Our analysis of variance of the cloze test results shows that comprehension as measured by this test increases with proficiency in the language. We had also hoped that the cloze test would be able to tell us something about how native language listening comprehension is better than second-language listening comprehension. Do we have any support for our hypothesis that native language listeners rely mainly on semantic and discourse constraints in the text whereas non-native listeners still have to direct a good deal of attention to syntactic constraints?

If we look back to table 7 to the factor loadings of our new variable comprehension, it becomes evident that our analysis of variance supports this definition of group differences. Our original variable strong context, or relying on higher-level constraints in responding to the missing cues in the text, has a factor loading of .95, whereas

weak context and syntax over semantic have loadings of -.88 and -.80 respectively. These measures are included by definition in our first principal component, which as we have seen, has pointed out significant differences among the test groups. We can more directly examine group differences for the ability to make use of semantic and discourse contraints by referring to the group means of the original eight variables taken from Clarke and Burdell's diagnostic evaluation of cloze responses (table 5).

We have predicted that the groups native, German, and mixed would respectively show increasingly lower scores on the measure strong context. In contrast, their scores should respectively be increasingly higher on the measures syntax over semantic and weak context. Table 5 shows that the data lie in the direction of these predictions. The mean group scores for strong context for the native, German, and mixed groups are as follows in this order: 51.88, 47.17, and 36.76. While scores for the ability to make use of context are lowest for the mixed experimental group, scores increase for the measures syntax over semantic and weak context over the groups, native, German, and mixed respectively. Mean group scores for syntax over semantic are: .91, 1.48, and 4.60. Mean scores for weak context are: 2.18, 6.22, 13.44.

We have too little evidence to interpret the listeners' reliance on phonological cues in the text although the

mean scores (table 5) for non-native listeners for <u>sound similarity</u> are higher than the mean score for the native group (native, .03; German, .13; mixed, .20).

The question might now be raised whether we have been justified in using the cloze test results to make generalizations about listening strategies and whether our prediction was correct that listening to a text before reading it, significantly improves comprehension of the passage. Our analysis of variance showed a significant effect of treatment: Subjects in the experimental groups, who listened to the text prior to writing the cloze test, scored significantly higher than subjects in the control groups, who had not listened to the text. (The reader is asked to refer to table 11 for a listing of the mean group <u>comprehension</u> scores). The smallest advantage of listening is demonstrated by the mixed group. This is to be expected as the control group (E-level students at the Michigan State University English Center) as a group had higher proficiency in English than the experimental group (D-level students). It seems that in the mixed group the experimental group almost loses the predicted listening advantage it would be expected to have over a group of like proficiency. The differences between the experimental and control groups of the other two language groups seem to bear out Page's findings that a previous encounter with a text improves cloze test scores by about 10 percent above what they would be on a

conventional cloze test (Page, 1975, p. 385).

The question of how well a cloze test measures listening comprehension is beyond the scope of the present study. We have merely made use of the test to investigate whether it shows the same differences in listening comprehension between native and non-native groups of subjects that reading research comparing these groups has shown. Our results support our hypothesis that, as in reading, native language subjects process a message using primarily semantic units. Second-language subjects direct more attention to lower-level syntactic information. With increased proficiency in the language, the listener seems progressively better able to attend to the semantic cues in the text.

Experiment one dealt with a text; our second test in experiment two is an attempt to look more closely at the differences between native and second-language subjects in processing single sentences.

CHAPTER III

EXPERIMENT TWO

Purpose

This experiment is an extension of the well-known
study conducted by Bransford and Franks (1971) which demon-
strated that in a memory task, subjects recalled the com-
plete ideas of sentences they had heard rather than the syn-
tactic form of these ideas. Subjects in the Bransford and
Franks study were native speakers of English. To test our
second hypothesis that more proficient listeners in a lan-
guage will progressively be less apt to remember details of
syntax in sentences they have heard, a test similar to the
one in the original research was given to groups of both
native and non-native listeners.

The Bransford and Franks study was designed to present
four semantic ideas to each subject participating in their
study. Each semantic idea in its most complex form could
be represented by a single complex sentence (e.g., The young
boy walking alongside the road could not see the dust storm
approaching from the southwest). During an acquisition
phase of the experiment, subjects were never presented with
sentences expressing the complete semantic idea, but only
with less complex sentences (The young boy walking alongside

the road could not see the dust storm; The dust storm was
approaching from the southwest). Taken together, these sen-
tences presented all of the information in the four idea
sets, but in syntactically simpler form.

In a recognition phase of the experiment, which fol-
lowed the acquisition phase, subjects heard a second list
of sentences presenting the same four semantic ideas. This
group of sentences contained some sentences from the acqui-
sition list (OLD sentences), some sentences which expressed
the same ideas as the acquisition sentences but which were
syntactically different (NEW sentences), and some sentences
which were both syntactically and semantically different
from the acquisition sentences (NONCASE sentences). Sub-
jects were asked to indicate which of the sentences in the
recognition list they had heard during the acquisition exer-
cise and how confident they were of their answer. Results
of the study showed that subjects were unable to differen-
tiate syntactically different NEW sentences from the OLD
sentences heard in the acquisition exercise. Surprisingly,
although no sentences appeared in their most complex form
in the acquisition phase of the experiment, subjects were
more confident that they had actually heard these sentences
than they were about sentences containing a lesser amount
of the information in any idea set. Subjects appeared to
be retaining the information in an idea set as a semantic
unit rather than the syntactic form in which this idea was

presented.

The present experiment attempts to test these results on non-native as well as native speakers of English. Based on our second hypothesis, we expect that the native listeners, like Bransford and Franks' subjects, will integrate the information expressed in the individual acquisition sentences to construct wholistic semantic ideas. If our assumption holds that second-language listeners will process and store lower-level language units, then, given a similar acquisition and recall task, they should show a higher degree of retention for the syntactic form of the sentences presented during acquisition. Unlike the native listeners, they should be able to discriminate on the basis of syntax between sentences which they had actually heard during acquisition and sentences which they had not heard. To test whether native and non-native listeners differ in the amount of semantic and syntactic information they retain after a sentence acquisition task, the present experiment adapts the research methodology of Bransford and Franks' semantic integration study.

Method

Subjects. The seventy-six subjects who participated in this experiment were from the same pool of subjects from which the test groups from experiment one were taken. The twenty-two German subjects were the same students who made up the German experimental group in the first experiment.

The twenty-two students of mixed foreign language back-
ground were likewise the same students who made up the mixed
language background experimental group in experiment one.
Three students in this group who wrote the cloze test were
not in class for the second part of the testing, experi-
ments two and three. The mixed group, therefore, has
twenty-two instead of twenty-five subjects for these two
tests. Because of scheduling considerations, the twenty-
nine native speakers who took part in this second experi-
ment were not the same students who participated as the na-
tive experimental group in experiment one. But they were
volunteers from the same Michigan State University freshman
English class. As a group, these native speakers partici-
pated in experiments two and three of the present study.
The reader is asked to refer to tables two through four for
a breakdown of this fourth experimental group by sex, age,
and knowledge of a foreign language. As with all of the
experiments in this study, the groups of subjects were too
small to do any statistical analysis of these personal vari-
ables. Tables two through four present a numerical record
of group similarities and differences.

Design and materials. As in the Bransford and Franks
study, materials consisted of a set of English sentences
constructed from semantic information from four idea sets.
Each of the idea sets included four units of information,
which were represented by four simple declarative sentences

(ONES). These units of information were combined by sentence embedding to produce sentences of varying complexity. In its most complete and complex form, an idea set was represented by combining the four units of information from the simple declarative sentences (ONES) into one complex sentence, called a FOUR. Likewise, less complex sentences, which only partially expressed the complete information in the idea set were constructed by combining or embedding two or three of the simple sentences from the same idea set (TWOS AND THREES respectively).

The four complex FOUR sentences in the test completely exhaust the semantic information included in the set of test sentences. These are:

Idea Set A: The young boy walking alongside the road could not see the dust storm approaching from the southwest.

Idea Set B: An old pick-up truck hauling a trailer was having trouble getting up the ice-covered hill.

Idea Set C: The tax collector found himself talking on the doorstep of a shabby house to a woman with two children tugging at her skirts.

Idea Set D: Along an interstate highway a motorist with a Citizens Band radio reported an accident involving a fuel truck and a small postal vehicle.

A set of twenty-eight sentences was constructed for each of these idea sets. These twenty-eight sentences can be called CLEARCASE sentences. For each idea set they include: (a) a complex sentence which exhaustively represents

the four units of semantic information in an idea set (FOUR);
(b) four simple declarative sentences (ONES); (c) four sen-
tences combining two simple sentences belonging to the idea
set (TWOS); (d) three sentences combining three simple sen-
tences from the idea set (THREE).

An example of a complete set of sentences defining
one particular idea is given in table 12. The sets of sen-
tences for the other three idea sets were constructed in an
analogous manner. These can be found in appendix B.

The four complex sentences used in the present study
were not constructed for the present test, but taken from
English as a foreign language reading texts. It should be
noted that they are more complex than the sentences used
in the original study by Bransford and Franks. While the
four simple declarative sentences making up each idea set
in the earlier research contained no further embedded sen-
tences, ONE sentences in the present study in some cases
still contain embedded attributes.

In addition to the ONES, TWOS, THREES, and FOURS
expressing the information of the four idea sets in experi-
ment two (CLEARCASE sentences), six NONCASE sentences were
also constructed. These combined information across idea
sets rather than within each idea set as did the CLEARCASE
sentences. Each NONCASE sentence represented four simple
declarative sentences, precisely the same number that were
represented in the four complex FOUR sentences. However,

Table 12

Sentences for Idea Set A

FOUR: The young boy walking alongside the road could not
          see the dust storm approaching from the southwest.
                (On Recognition Only)

THREES: The young boy could not see the dust storm ap-
          proaching from the southwest.
                (On Acquisition Only)
          The young boy walking alongside the road could
          not see the dust storm.
                (On Both Acquisition and Recognition)
          The boy walking alongside the road could not see
          the dust storm approaching from the southwest,
                (On Recognition Only)

TWOS: The young boy could not see the dust storm.
                (On Acquisition Only)
          The young boy was walking alongside the road.
                (On Acquisition Only)
          The boy could not see the dust storm approaching
          from the southwest.
                (On Recognition Only)
          The boy walking alongside the road could not see
          the dust storm.
                (On Recognition Only)

ONES: The boy was young
                (On Acquisition Only)
          The boy was walking alongside the road.
                (On Acquisition Only)
          The boy could not see the dust storm.
                (On Recognition Only)
          The dust storm was approaching from the southwest.
                (On Recognition Only)

---

in the NONCASE sentences the four simple declarative sentences came from different idea sets, thus violating semantic relationships represented in the acquisition sentences. The six NONCASE sentences used in experiment two are as follows: The tax collector walking alongside the road reported an accident involving a fuel truck and a small postal vehicle; Along an interstate highway a young boy was having trouble getting up the ice-covered hill; A motorist with a Citizens Band radio found himself talking on the doorstep of a shabby house to a woman with two children; The young tax collector reported an accident along an interstate highway; A boy walking alongside the road reported an accident involving a fuel truck and a pick-up truck hauling a trailer; An old fuel truck and a small postal vehicle were having trouble along an interstate highway getting up the ice-covered hill.

From these original sentences two lists of sentences were made, one for the acquisition part of the test and one for the recognition exercise. The acquisition list consisted of twenty-four sentences in all. Only CLEARCASE sentences were used in this list. From each idea set these represented two ONES, two TWOS, and two THREES. No FOUR sentences were used. (The acquisition sentences for idea set A can be found in table 12; the list for the three remaining idea sets can be found in appendix B). The acquisition sentences were chosen to exhaust all the information

in each of the idea sets. The sentences were presented so that one sentence from each idea set would appear in each successive sequence of four sentences. Within these blocks of four sentences the order was mixed. However, no two sentences from the same idea set occurred consecutively on the list. The ONES, TWOS, and THREES were distributed randomly across the entire list of twenty-four sentences.

The recognition list included five sentences from the acquisition list plus twenty-nine NEW sentences, making a total of thirty-four sentences in all. The five OLD sentences from the first list include: Idea set A--<u>The young boy walking alongside the road could not see the dust storm</u> (THREE); idea set B--<u>An old pick-up truck was hauling a trailer</u> (TWO); idea set C--<u>The two children were tugging at their mother's skirts</u> (ONE); idea set D--<u>Along an interstate highway a motorist with a Citizens Band radio reported an accident</u> (THREE); <u>The motorist had a Citizens Band radio</u> (ONE).

Of the twenty-nine NEW sentences in the recognition list, twenty-three of the sentences were from the original four idea sets. These included the following number of sentences from each idea set: two ONES, two TWOS, one THREE, and one FOUR (the only FOUR). The sentences used in the recognition exercise from idea set A can be found in table 12. The sentences used in the remaining idea sets can be found in appendix B. (Only five instead of six NEW

sentences were used from idea set B). In addition to the twenty-three NEW CLEARCASE sentences from the original four idea sets, six NEW sentences on the recognition list were NONCASE sentences, which combined semantic units from two or more idea sets.

The order of presentation of the thirty-four sentences was similar to that of acquisition. One sentence from each idea set and one NONCASE sentence appeared in each successive sequence of five sentences. Within these blocks of five sentences the order was mixed. The five OLD sentences were randomly assigned positions in this list. No two sentences from the same idea set were allowed to be consecutive. ONES, TWOS, THREES, and FOURS were distributed over the entire list.

Procedure. The subjects in each of the three language groups were divided into two groups to counterbalance the effects of presentation order for the acquisition sentences. Subjects in group one (17, 11, and 12 subjects from the native German and mixed groups respectively) heard the acquisition sentences in the order 1-24. Subjects in group 2 (2, 11 and 10 subjects from the native, German, and mixed groups respectively) heard the sentences presented in the reverse order 24-1.

The test consisted of an acquisition and a recognition exercise. Both were tape recorded. All groups began with the acquisition sentences. Over laboratory headsets the

students were given instructions to answer questions about some sentences they would hear on tape. For each sentence the acquisition exercise proceeded as follows:

1. A sentence was read on tape.

2. The student was asked to perform a distracting task which consisted of being asked to write a certain number (1-5) of letter a's on the worksheet. A four-second pause followed.

3. The student heard an elliptical question asking him to recall a sentence element which had been replaced by a wh-question word. Examples of possible questions for the sentence The young boy was walking alongside the road are: Who was walking? Walking where? The young boy was doing what? Only one question was asked for each sentence, and sentence elements were questioned equally often.

4. The student wrote the answer to the question. A four-second pause followed.

Before beginning the acquisition exercise, students were forewarned that they might find the exercise frustrating because of the distracting task. Between each sentence and question they would be asked to answer, there would be an interruption. After each sentence they would be asked to write a given number of a's on their worksheet in the appropriate space. This would impose quite a strain on their memory, but students were asked to try to answer the questions as accurately as possible. Students were not told that there would be a second, recognition, phase to the exercise; indirectly they were encouraged to process the sentences semantically.

Following acquisition, students were given a three-minute break in which they were asked to recognize the

melodies of three popular songs: "Where Have All the
Flowers Gone?", "Let It Be," and "Jingle Bells." Then the
recognition part of the exercise began with the tape-
recorded instructions that students would be asked to listen
to a new set of some of the same sentences they had heard
in the first part of experiment two. After each sentence,
they would be asked to underline yes if they had actually
heard the sentence before or no if they were hearing it for
the first time. Students were also asked to circle a con-
fidence rating from 1-5 (from very unsure to very sure) in-
dicating how confident they were of their answer. The ex-
ercise began with an example sentence to insure that the
students understood the marking procedure for the exercise;
then the recognition sentences were presented to all groups
in the order 1-34. The students heard each sentence in
the recognition list only once.

Analysis of data. Subjects' ratings, like in the
Branford and Franks study, were converted into numerical
values. A yes response received a plus rating, and a no
response received a minus rating. These prefixed the con-
fidence score a subject had given each of his responses.
Thus a yes response with a very high confidence rating be-
came a +5, a yes response with a high confidence rating
became a +4, and so on down to a very low yes rating of +1.
The scale continued in the same fashion for the no or nega-
tive responses, building a 10-point rating scale: +5, +4,

+3, +2, +1, -1, -2, -3, -4, -5. All data are reported in terms of the rating which was computed for each of the three groups.

Results and Discussion

Figure 3 summarizes the mean scores for the five sentence types (FOURS, THREES, TWOS, ONES, and NONCASE) for each of the three groups tested. Before we look at how results for the three groups differ, we will compare our results for the native group of subjects with the results Bransford and Franks found for their native language subjects.

Bransford and Franks report two trials for their test, one with and one without the NONCASE sentences. Mean recognition ratings which they reported as characteristic for individual sentence types are: FOUR, +4.26; THREE, +3.63; TWO, -0.73, +2.93, +3.59; ONE, -2.66, -.50. NONCASE sentences had a mean rating of -4.0 (Bransford & Franks, 1971, pp. 339-341). In discussing their findings, Bransford and Franks report that OLD sentences, on the whole, received lower ratings than the NEW sentences. Among NEW sentences FOURS, THREES, AND TWOS received positive recognition ratings. In fact their results showed recognition ratings clearly ordering the FOURS > THREES > TWOS > ONES down to the NONCASES, which received the lowest rating of all.

By comparison, (see figure 3) native language subjects in the present test showed the same tendency to rate more
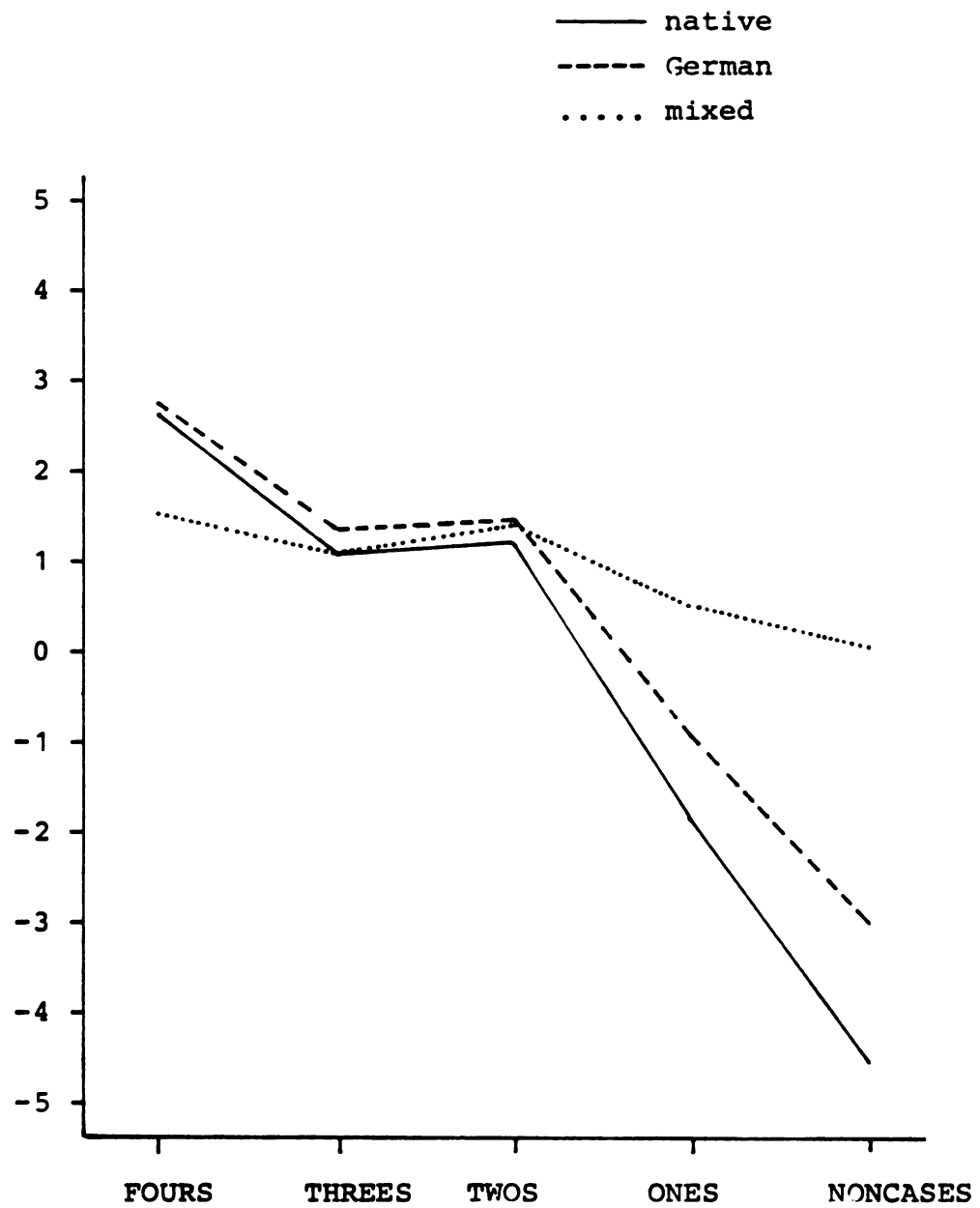
Figure 3. Mean group ratings for NEW sentences.

complex sentences, sentences containing more of the semantic units of an idea set, higher than the less complex sentences. NONCASE sentences, which combined information across idea sets thus confusing original semantic relationships, received considerably lower ratings than any of the CLEAR-CASE sentences. Our data differ from the Bransford and Franks results in the height of the mean rating scores. The Bransford and Franks data show only a slight mean difference between TWO and THREE sentences although their difference is in the direction of the prediction THREE > TWO. Our findings also show only a slight difference between the mean ratings of TWO and THREE sentences; however, this difference goes in the other direction TWO > THREE. FOURS in the present study have a lower mean rating than the Bransford and Franks study, as do both THREES and TWOS; but they were nevertheless clearly given higher ratings than any of the other sentence types, even though, as in the earlier study, no FOUR sentences were included in the acquisition list.

Like the Bransford and Franks study, subjects in the present study seemed to show a slight recognition advantage for ONE sentences. NEW ONES received a mean rating in the earlier study of slightly less than 0. In our study the mean rating was even lower at -1.79. Because our test sentences were markedly longer than those in the Bransford and Franks study, subjects were conceivably more certain of not having heard syntactically simple ONE sentences in the

acquisition list.

OLD sentences displayed the same characteristics in our data as in the earlier study. Bransford and Franks report that 7 (of 24) NEW sentences received higher ratings than their lowest ranking OLD sentence and 15 NEW sentences were higher than their lowest ranking OLD sentence (Bransford & Franks, 1971, p. 340). The mean recognition ratings for the OLD sentences in our study for each of the three language groups are given in figure 4. It must be kept in mind that these are means for only one or two sentences. Nevertheless, it is clear that native language subjects did not distinguish between syntactically OLD and NEW sentences in recall. In fact, the OLD sentences are rated much like the NEW sentences of the same type (THREE, TWO, ONE). The highest ranking OLD sentence in our study was a ONE, which received the high ranking of +3.11. This high ranking may have been due to the slight recall advantage which subjects in both studies showed for ONE sentences. Only 2 (out of 23) NEW sentences received higher ratings in our study than +3.11, but 14 NEW sentences received ratings higher than the lowest OLD.

This analysis of the data supports the general hypothesis that, during the acquisition phase of the experiment, subjects were integrating the information rather than the form of the sentences they heard. They were more certain of recognizing sentences that expressed a complete semantic idea than they were of recognizing sentences they had
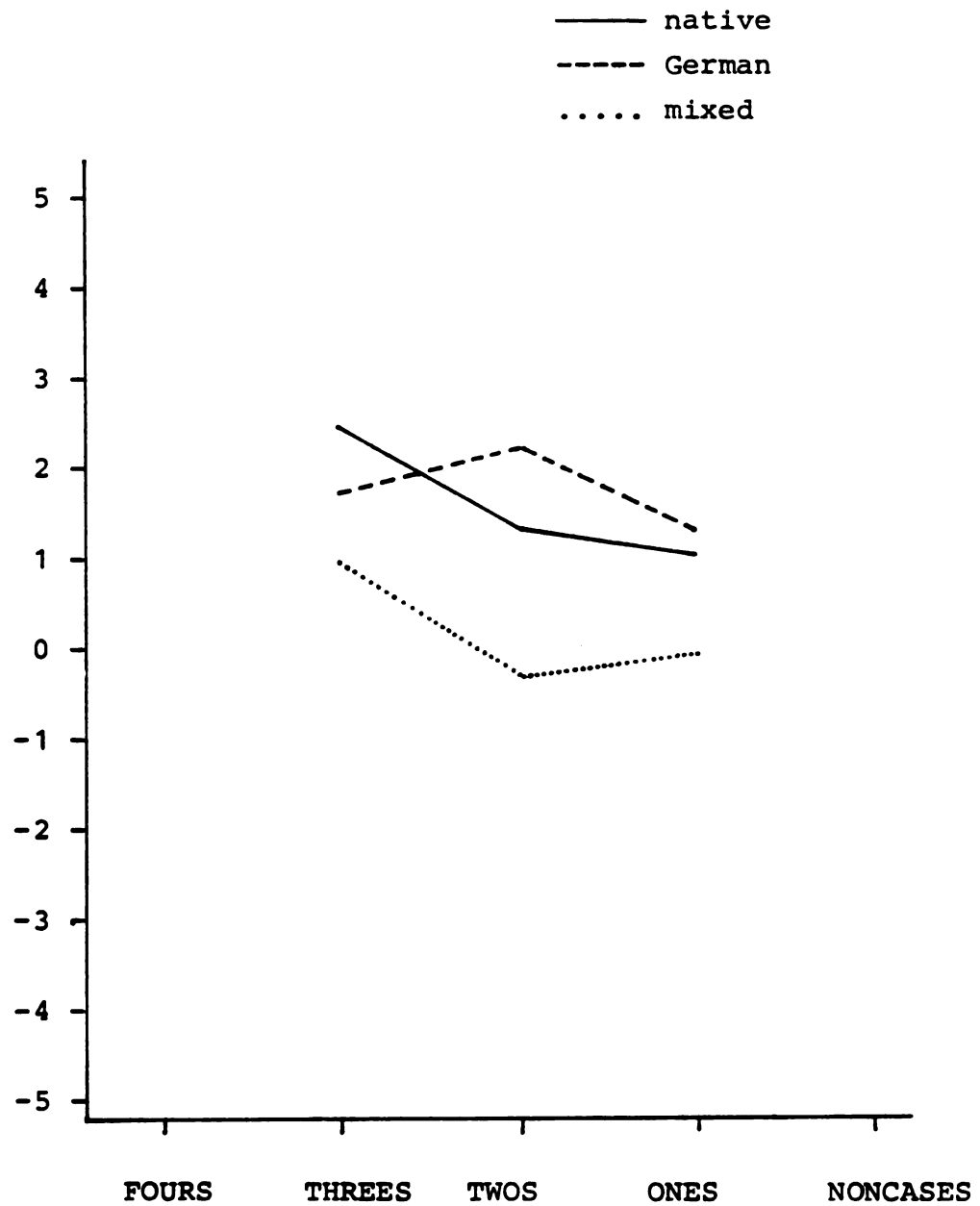
Figure 4. Mean group ratings for OLD sentences.

actually heard. In fact, FOUR sentences, which were not represented at all in the acquisition list, received as a group the highest mean ratings. NONCASE sentences, which combined information across two or more of the original acquisition sentences, received very low negative ratings ranging from -4.14 to -4.89. Even though subjects could no longer recall the syntactic form of the original acquisition sentences, they were still very clear about the semantic content of these sentences.

Although there was only a slight mean difference between their subjects' ratings of THREE and TWO sentences, the Bransford and Franks study showed relative ratings of the recognition sentences to be FOURS > THREES > TWOS > ONES > NONCASES. Nevertheless, the emphasis of that study was on abstracting semantic ideas and not showing that subjects would give higher recognition ratings to sentences as they increased in syntactic complexity. Therefore, it need not be the case that our recognition ordering FOURS > TWOS > THREES > ONES > NONCASES be construed as contradicting the earlier findings. What seems to be most important in both studies is that FOURS were given higher mean ratings than any of the other sentence types. Because the information in the FOUR sentences was never presented during acquisition, subjects could only have arrived at the complete idea by integrating information which had been presented in a different syntactic form.

In their research, Bransford and Franks suggest that the instructions given to the students before the acquisition exercise might influence recognition results. Both the present experiment and the Bransford and Franks study encouraged semantic processing; students were not told before the acquisition exercise that they were going to be later asked to recognize the sentences they had heard. The researchers also suggest that memory for the syntactic form of the acquisition sentences might be influenced by the general style of these sentences, i.e., recognition may have been different, for example, had all of the acquisition sentences been ONES. The more accurate ratings (with respect to the acquisition list) for the ONES, especially in the present study, seem to indicate that subjects do have some recognition for syntactic form, even if this is only a general feeling for the style of the acquisition sentences.

Both studies seem to show, however, that memory for syntactic form is secondary to the integration of new information to form complete semantic ideas. Subjects retain information in the form of wholistic ideas, or as Bransford and Franks describe it "abstracted" from the original syntactic form of the sentences from which the information was acquired. To more accurately test the phenomenon of semantic integration, Bransford and Franks subjected their data to a second analysis, which examined each recognition sentence, not by the number of clauses it contained, but by its information load within its respective idea set.

Figure 5 presents such a breakdown of how the sentences in idea set B of the present test contribute to the semantic idea expressed completely by the FOUR sentence. The researchers predict that where two sentences deal with the same semantic content, the sentence which contains all of the information of the other sentence plus one or more additional propositions will receive the higher recognition rating. To take an example from the present test, the THREE sentence, <u>Along an interstate highway a motorist with a Citizens Band radio reported an accident</u>, should receive a higher rating than the ONE sentence, <u>The motorist had a Citizens Band radio</u>, not because of its more complex syntactic structure, but because it more completely expresses the information in the idea set. The same THREE sentence need not, however, receive a higher rating than the TWO sentence, <u>A motorist reported an accident involving a fuel truck and a small postal vehicle</u>, because the TWO sentence contains information not included in the THREE sentence (<u>accident involving a fuel truck and a small postal vehicle</u>).

Using this type of analysis, outside interference is excluded from the semantic relations between the sentences; and one arrives at a set of predictions for each idea set about which sentences will be rated higher on the recognition rating. In this type of comparison, information in two sentences is always qualitatively the same. Two sentences differ only by virtue of the fact that one of the

+3.9, +1.8, +2.0
1. Along an interstate
   highway a motorist
   with a Citizens Band
   radio reported an acci-
   dent involving a fuel
   truck and a small
   postal vehicle. (FOUR)

+1.9, +0.4, +1.2
OLD 2. Along an interstate
   highway a motorist
   with a Citizens Band
   radio reported an
   accident. (THREE)

+0.5, -0.4, +2.0
3. Along an interstate
   highway was an
   accident involving
   a fuel truck and a
   small postal vehicle.
   (TWO)

+1.1, +1.4, +4.5
4. A motorist reported
   an accident involving
   a fuel truck and a
   small postal vehicle.
   (TWO)

-2.3, -1.7, +2.0
5. There was an acci-
   dent involving a
   fuel truck and a
   small postal
   vehicle.
   (ONE)

+3.1, +1.6, +0.5
OLD 6. The motorist
   had a Citizens
   Band radio.
   (ONE)

-0.7, +0.2, +2.6
7. A motorist
   reported
   an accident.
   (ONE)

Predictions

1 > 3 > 5   1 > 2 > 6   2 > 7   1 > 4 > 7

Figure 5. Recognition sentences for idea set B. [The numeri-
cal values above each sentence represent the mean
recognition ratings for (a) the native group,
(b) the German group, (c) the mixed group.
Sentences 2 and 6 in this idea set are OLD; the
remaining sentences are NEW.]

two sentences has more information. A higher rating for this second sentence more clearly indicates that the result is due to integration and a preference for retaining a more complete semantic unit.

The valid predictions for idea set B are presented at the bottom of figure 5. It should be pointed out that these predictions are transitive. Thus the prediction for idea set B 1 > 3 > 5 not only compares 1 with 3 and 3 with 5, but also states that 1 > 5. Predictions were formed for all four of the idea sets in experiment two. The predictions for idea sets A, C, and D can be found in appendix C. For the four idea sets there were 45 ordinal predictions in all. Table 13 presents the number of these predictions that were confirmed for each of the three language groups participating in the present test.

Table 13

Number of Predictions Confirmed for Each Group

Native	German	Mixed
37 of 45	34 of 45	32 of 45

For the native language group, 37 of the 45 predictions proved to be in the predicted direction. These results are similar to those reported by Bransford and Franks, who reported the following proportion of ordering predictions confirmed for two trials on three different tests: 47/49,

43/49; 35/41, 39/41; 135/140, 134/140. (Bransford & Franks, 1971, p. 342).

The results of our test with native language listeners seem to lend further support to Bransford and Franks' hypothesis that, for the most part, memory is a function of the ideas acquired during an acquisition task, not a function of the syntactic form in which these ideas are presented.

The results for the German group closely resemble the results obtained for the native listener group. German subjects also gave the highest mean rating +2.75 to FOUR sentences, although in fact, they had heard no FOUR sentences in the acquisition list. Like the native group, they made no distinction in their ratings for THREE and two sentences, and they gave ONES the lowest rating among the CLEARCASE sentences. It is worth noting that this rating is also the most accurate with regard to the original acquisition list. Among all sentence types both the native and German groups seem best able to distinguish which ONE sentences they had heard. Both groups gave NEW ONES slightly negative confidence ratings, while they rated OLD ONES with slightly positive ratings. Bransford and Franks have also noted this advantage for ONE sentences in their study. It is possible that the longer sentences used in the present study serve to increase the slightly improved recall for the syntax of the simple sentences.

German subjects, like native subjects, seem more confident of having heard sentences that express a complete semantic idea than they are of recognizing the syntactic form of the sentences they had actually heard. Like the native language listeners, they gave NONCASE sentences much lower ratings than any of the other sentence types. These had a mean rating of -3.07 and a range of -1.32 to -4.36. These values are not quite as low as those given to the NONCASE sentences by the native language subject group, but they seem to show that the German subjects were sure they had not heard sentences that violated the semantic relationships of the original acquisition sentences.

In a comparison of the number of predictions confirmed by the German subject data (table 13) with those by the native subject group, the German subjects differ only slightly from the native language listeners: 34 of 45 predictions were confirmed, only three less than by the native group.

It would appear, from both the ratings for the recognition of sentences by sentence type (FOURS, THREES, TWOS, ONES, AND NONCASES) (figure 3) and from the analysis of the ordering predictions for semantic completeness (table 13), that the subjects in the German group were not retaining syntactic information about the acquisition sentences. Rather, they seem to have integrated the information from the sentences into semantic units. Like the native subjects, they too were most confident of having heard FOUR

sentences, which never occurred in the acquisition list but which contained the most complete information from each idea set. They can only have acquired this information through integrating the individual sentences into a semantic unit. Individual sentences seem to have lost their status in memory. Contrary to our predictions about second-language listeners at the outset of this experiment, this group was not more accurate in retaining syntactic information from the acquisition sentences than the native subjects in the present study or in the study by Bransford and Franks.

Before we discuss the mixed language group, we must point out that for part of the group, those who were presented with the acquisition sentences in the order 24-1, the data are not complete. These students only completed rating 22 of the 34 recognition sentences. This was due to a five-minute delay in starting the test. To insure that these students would still have the same amount of time as the other group for writing the test for experiment three, the final sentences in the recognition list for experiment two were cut from the test. For part of the group, this omission affects the following NEW sentences: 2 FOURS, 3 TWOS, 5 ONES, and 1 NONCASE. One OLD ONE sentence was also in this incompleted sequence. For these twelve sentences we have mean ratings based on ratings from only half of the subjects in the mixed language group. Because all of the ratings are expressed in means, however, and sentences types were distributed evenly throughout the entire

recognition list, no one sentence type was affected un-
evenly, and this omission should not have altered the rat-
ings for the remaining sentences.

Figures 3 and 4 clearly show differences between the
mixed language group results and the results of both the
German and the native language groups. Figure 3 shows how
their responses differed for the NEW sentences. In con-
trast to both the previous groups, the mixed group does
not show higher ratings for the FOUR sentences in the recog-
nition list. Although they were rated slightly higher than
THREES, they were rated slightly lower than TWOS. When we
compare this rating for the FOURS with the only slightly
negative mean rating for the NONCASE sentences (Native and
German groups had considerably higher negative ratings for
these sentences), one can conclude that this group was for
some reason not able to integrate the information from the
acquisition sentences into wholistic semantic ideas. They
apparently did not recognize the NONCASE sentences violated
the semantic content of the acquisition sentences. They
gave these sentences the mean rating of -.80, the same
rating they gave to OLD ONE sentences. It was not as clear
to them as to the other groups that NONCASE sentences, due
to their semantic content, could not have been on the
acquisition list. If we look at the low rating for FOUR
sentences as well, it seems that this group showed less
tendency than the other groups to process the acquisition
sentences semantically.

If one considers that the majority of the sentences on the recognition list were NEW, the low ratings of the mixed group are actually more accurate than higher ratings given by the other two groups. One might argue that the mixed group was demonstrating a better retention for the syntactic information in the acquisition list. However, it seems more likely that this group of subjects had more trouble comprehending and retaining the sentences in general and, therefore, was less confident about whether a sentence was NEW or OLD.

The ratings from the mixed language group for the OLD sentences (figure 4) seem to support this second explanation. These ratings show no evidence that the mixed language group was able to recognize these sentences as OLD on the basis of their syntactic form. In fact, they rated OLD sentences slightly lower than the NEW sentences they had not heard before.

An examination of this group's results on the ordering predictions (table 13) seems to indicate some semantic processing, however. Of the forty-five predictions in the test, thirty-two were in the predicted direction. This result is not considerably lower than that of either the German or native group.

If the mixed group processed the acquisition sentences using syntactic cues rather than semantic cues, they were not able to retain this information. Our hypothesis that

native and non-native listeners differ in the relative amounts of semantic and syntactic information they retain is thus only partly supported by our findings. Of the two non-native groups, only the mixed group, which is also the group with the lowest English proficiency, differed markedly from the native group of listeners. Their results showed that, while they were unable to retain the semantic information in the acquisition sentences, they were also unable to discriminate syntactically between sentences they had actually heard during acquisition and those they had not heard.

In light of the present findings, it seems that a more plausible explanation than our original hypothesis for the behavior of both non-native language groups might be as follows: Second-language listeners attempt to integrate the information in an acquisition task semantically. Due to the need to concentrate on syntactic cues as well, attempting to integrate all this information has the result that very little of either the semantic or syntactic information in the sentences is retained.

A second explanation for the failure of the non-native listeners to have retained the syntactic information in the acquisition sentences might be that they had more trouble understanding the syntactic form in the first place. It should be pointed out again that our scoring of recall in the test is not sensitive enough to distinguish between

subjects who did not remember the syntactic form because they processed the sentences semantically and subjects who did not recall the syntactic form because they had been unable to decode the syntax of the original sentence. Although none of the sentences used in the present test should have been beyond the proficiency of any of the students participating, we cannot exclude the possibility that this factor may have influenced our results.

CHAPTER IV

EXPERIMENT THREE

Purpose

This experiment addresses our third hypothesis that
the recall of connected discourse is facilitated if the
listener thinks the discourse fits together into a story,
and that this facilitation increases with increasing pro-
ficiency in the language. In one of the few existing tests
of the effects of discourse structure on native language
comprehension, de Villiers (1974) showed how the use of
definite articles instead of indefinite articles in an
aural text enabled his native language subjects to perceive
the passage as a story. His study also showed that proces-
sing the passage as a story significantly improved recall
of the text. Our third hypothesis predicts that merely
knowing that the text is a story should not significantly
improve the recall of non-native listeners. Experiment
three follows the methodology of de Villiers' earlier study
and hopes to test whether the perception of a story struc-
ture in a passage does in fact facilitate recall and, if
so, if it differentially affects the recall of non-native,
as opposed to native, listeners.

In the de Villiers' test, two groups of subjects lis-
tened to different versions of the same text, one with

definite articles and one with indefinite articles. As opposed to the indefinite article group, the subjects who listened to the text with the definite articles tended to see the passage as a unified story and showed significantly better recall. Furthermore, the sentences these subjects recalled correlated highly with sentences that an independent control group had rated as most central to the theme of the passage.

The subjects who heard the text with indefinite articles, on the other hand, tended to view the sentences as unrelated. These subjects recalled significantly less of the passage than the other group. In addition, their recall seemed to depend on factors shown in earlier studies to be important to the recall of isolated sentences, namely, that a sentence occurred at the beginning or end of the presented list or that it had a high image-evoking quality. In fact, as expected, the sentences recalled by the indefinite article group showed a high negative correlation with the order of presentation of the set of sentences. They also showed a high positive correlation with a visual imagery rating that had been given to each sentence by an independent control group.

De Villiers argues that the use of the definite article in the sentences directs the listener to connect a hierarchy of propositions around what become, by virtue of the direct reference, shared referents in a narrative. The listener is thus guided by the surface structure of the

text to organize the arguments of the message into hierar-
chical schemata, which, in turn, mediate and enhance his
recall of the sentences.

Using a different text, but a testing procedure based
on the one used by de Villiers, the present experiment in-
vestigates how these results vary over our three popula-
tions. Our third hypothesis predicts that the perception
of a story structure in the test passage will facilitate
recall and will do so progressively with increased profi-
ciency in the language. In addition, we expect that our
native language listener group will show the same recall
patterns that characterized de Villiers' findings:

1. The contextual feature <u>definite</u> marked by
   the definite article should aid the listener
   in organizing the passage into a story.

2. The listener's recall of details should be
   directed by how these relate to the central
   theme in the story. On the other hand, lis-
   teners who hear the same passage with indefi-
   nite articles are expected to treat the same
   group of sentences as if they were unrelated
   to each other. The recall of this group is
   not expected to be dictated by centrality to
   theme, but rather by visual imagery and
   recency in presentation.

It is predicted that neither of the two factors, pres-
ence of the definite article nor the listener's organiza-
tion of the passage into a story, will have a strong effect
on the recall of the non-native listener groups. In the
case of the definite article, one would expect the non-
native listener to be less familiar with the discourse
structure of English and less able to take advantage of the

contextual information marked by the definite article. Of course, this will depend in part on the similarities between articles in the listener's native language and in English. Likewise, if we have correctly assumed that the non-native listener is forced to concentrate his attention on lower-level cues in the aural input, the effect of story organization on the quality and quantity of his recall should also be less.

The non-native listener's expectations on the levels of phonology and syntax should prove to be more important than his expectation that he is processing a story. As a result, in contrast to the native listener, one would expect the non-native listener to show similar recall patterns for both the definite and indefinite article passages. For both passages it is expected that recency and visual imagery will be the prime determinants for whether a sentence is recalled or not. This stands in contrast to native listener results, where these factors operate only for the recall of unrelated sentences.

The mixed language background group of second-language listeners is expected to demonstrate these characteristics most strongly. The German students in the present study should show results that are closer to those of the native language group because of their higher average overall proficiency in English and because rules for story structure and rules for the use of the definite and indefinite

article in discourse are very nearly the same in English and German. All of these factors should aid the German listener in organizing the definite article sentences as connected discourse.

Method

Subjects. Because experiments two and three were administered in all three language groups in the same testing session, subjects in the experimental groups for this experiment are the same as those who participated in experiment two. The present experiment also includes three control groups, one from each of the three language backgrounds. These subjects came from the same groups who participated as control groups in experiment one. The resulting six groups of subjects for experiment three include the following number of participants from each language group for the experimental and control groups respectively: native language group, 28 and 29; German group, 22 and 21; mixed group, 20 and 25.

Design and materials. Two sets of the following nineteen sentences were used in the experiment. With the definite articles in them, the sentences are a very free adaptation of a short story by William Saroyan "Locomotive 38: The Ojibway." Admittedly much of Saroyan's narrative art has been lost in adapting the sentences to the test exercise; but, nevertheless, in this form they do present

a somewhat stilted story. When all of the definite articles
are replaced by indefinite articles, however, the sentences
appear to be largely unrelated to each other.

> The stranger came to town on a donkey. The tall
> young Indian was an Ojibway. The animal was
> struck by a trolley in Tulare Street. The boy in
> the drugstore was eating a banana split. The in-
> jured animal fell and passed away of internal in-
> juries. Everyone thought the man behaved strangely.
> The man wanted to buy a car. Crazy Horse was a
> millionaire. The bank called Pacific Southwest to
> check the account. The man asked the boy at the
> counter to be the driver. School was closed for
> summer vacation. The Indian bought a set of fish-
> ing gear. The boy was delighted over the gift.
> The twosome drove to Mendota right up to the river
> bank. The Indian couldn't drive the car. The
> brother came in November from Oklahoma. The man
> in the hotel was a witness. The brothers left
> the valley in the Packard sedan. The Indian from
> the boarding house drove.

Procedure. Control groups: Students in the three
control groups were given both versions of the above pas-
sage as a short reading exercise. Fourteen students in the
native language group, eleven students in the German group,
and ten students in the mixed language background group re-
ceived the story version of the text with the definite ar-
ticles. These students were asked to read the story and
then indicate the five sentences they considered most cen-
tral to the theme of the story and five sentences least im-
portant to the theme of the story. In addition, a blank
was left after each sentence of the passage where the stu-
dent was asked to go back through what he had read and rate
each sentence for the strength of the visual image it
evoked. Sentences that evoked no visual image were rated

zero, and with increasing visual imagery a sentence could be rated 1-7.

Different students from the native, German, and mixed control groups were asked to rate the visual imagery in the indefinite article version of the same set of sentences. These students; 15, 10, and 15, from the native, German, and mixed groups respectively; were asked to rate each of the nineteen indefinite article sentences using the same 0-7 visual imagery rating scale.

Experimental groups: Each group of subjects in the three experimental groups was divided into two sub-groups to hear the two versions of the test passage. Students who heard the definite article version (D) were 12, 9, and 9 students from the native, German, and mixed groups respectively. Students who heard the indefinite article version (ID) were 16, 11, and 10 students from the same respective groups. (Three of the twenty-two original students in the German group had been in the class from which the control group was taken. Although they wrote test three with the rest of the group, the results of their exercises have not been recorded).

Each D-group heard a tape recording of the definite sentences and each ID-group heard a tape recording of the indefinite sentences. Both versions of the text were recorded with three-second pauses between sentences. Instructions for doing the exercise were included on the tape.

Students were told that after they had listened once to a set of nineteen sentences, they were going to be asked to do a written recall exercise over what they had heard.

When the final sentence had been played on the tape, students were told to take as much time as they needed to write down as much as they could remember of the nineteen sentences. When all of the students had finished writing, they received the additional instruction to indicate on the bottom of their worksheet with the words story or unrelated whether they felt that the sentences they had heard formed a connected story or whether they were unrelated to each other.

Analysis of data. Control groups: A relative ranking of the centrality of each sentence to the theme of the story was obtained by giving each sentence the value of the number of subjects who had placed it among the five most central sentences minus the number of subjects who had placed it among the five most unimportant or peripheral sentences. For example, if ten students in a group of fifteen rated a sentence as central and only one rated the same sentence as peripheral, then that sentence received the rating of nine. On the basis of this value, the nineteen sentences received a rank order for their centrality to the theme of the story.

Imagery values for each of the nineteen sentences were established for both the definite and indefinite

sentences rated by each group. A mean imagery value was calculated for each group for each of the nineteen sentences, both with a definite article and the same sentence with an indefinite article.

Experimental groups: Recall data from each subject in both the D- and ID-groups were scored on six different measures for each of the nineteen sentences:

1. Was the sentence attempted at all in the recall exercise? (Were any of the content words from the original sentence recalled?)

2. Where in a student's retelling did the sentence appear?

3. How many content words were correctly recalled from the original sentence?

4. How many synonyms of words semantically equivalent to words in the original sentence were recalled?

5. What percentage of content words or their semantic equivalents was recalled from the original sentence?

6. Was the meaningful content or gist of the sentence recalled? (This value is important in cases such as the following: Original sentence--The man asked the boy at the counter to be the driver; student retelling--The boy asked the man at the counter to drive. Here the gist of the original sentence is not rendered in the retelling even though the sentence otherwise has an accuracy criterion of 100 percent).

Results and Discussion

Rank orders for the nineteen sentences for centrality and imagery were calculated for each language group using the mean ratings given each sentence by the respective control groups. Table 14 presents the three rank orders for

Table 14

Control Group Rank Orders for Centrality to Theme
for 19 Definite Sentences

Sentence	Rank Orders Native	German	Mixed
1. The stranger came to town on a donkey.	2.5	2	1
2. The tall young Indian was an Ojibway.	9	15	10
3. The animal was struck by a trolley in Tulare Street.	4	1	7
4. The boy in the drugstore was eating a banana split.	19	19	18.5
5. The injured animal fell and passed away of internal injuries.	7	4.5	2
6. Everyone thought the man behaved strangely.	9	9.5	7
7. The man wanted to buy a car.	1	18	10
8. Crazy Horse was a millionaire.	16.5	9.5	13.5
9. The bank called Pacific Southwest to check the account.	11	13	15.5
10. The man asked the boy at the counter to be the driver.	2.5	3	3.5
11. School was closed for summer vacation.	18	16.5	18.5
12. The Indian bought a set of fishing gear.	14	13	10
13. The boy was delighted over the gift.	12	11	12
14. The twosome drove to Mendota right up to the river bank.	6	7.5	3.5
15. The Indian couldn't drive the car.	5	4.5	5
16. The brother came in November from Oklahoma.	16.5	13	17
17. The man in the hotel was a witness.	14	16.5	7
18. The brothers left the valley in the Packard sedan.	9	6	15.5
19. The Indian from the boarding house drove.	14	7.5	13.5

Note. Ranks 1-19 indicate the range "most central" – "most
peripheral."

centrality to theme. Table 15 presents the three rank orders from the indefinite article control groups for imagery.

Spearman rank-order correlations were calculated for the two image-ratings (definite versus indefinite sentences) and among the image-rating, the rating for centrality to theme, and the order of presentation of the nineteen sentences. The rank-order correlations between the two image orderings for the native, German, and mixed groups respectively were .88, .90, and .34. There seemed to be consensus for the native and German subjects between definite and indefinite article groups as to which of the nineteen sentences present strong visual images. The positive correlation for the mixed group failed to reach significance.

For the set of definite sentences, the Spearman rank-order correlations among the original order of presentation of the nineteen sentences, the image-ratings and the ratings for centrality to theme ranged from .06 to .48 for all three groups. Table 16 shows the correlations among these values for each of the three control groups who rated the definite article sentences. We shall use these values as evidence that, in their roles in the text, presentation order, centrality to theme and imagery all act independently of one another. If the two experimental groups from each language (definite versus indefinite) show differences in what they recall of the passage; if they show a preference for either

Table 15

Control Group Rank Orders for Imagery
for 19 Indefinite Sentences

Sentence		Rank Orders	
	Native	German	Mixed
1. A stranger came to town on a donkey.	1	2	3
2. A tall young Indian was an Ojibway.	6	8	14
3. An animal was struck by a trolley in Tulare Street.	2	3	1.5
4. A boy in a drugstore was eating a banana split.	3	1	1.5
5. An injured animal fell and passed away of internal injuries.	9	5	8
6. Everyone thought a man behaved strangely.	13	19	9.5
7. A man wanted to buy a car.	15	16	15
8. Crazy Horse was a millionaire.	16	14.5	16
9. A bank called Pacific Southwest to check an account.	19	11.5	18
10. A man asked a boy at a counter to be driver.	12	4	9.5
11. School was closed for summer vacation.	10.5	13	19
12. An Indian bought a set of fishing gear.	8	7	12
13. A boy was delighted over a gift.	4	6	6
14. A twosome drove to Mendota right up to a river bank.	6	9	7
15. An Indian couldn't drive a car.	17	11.5	11
16. A brother came in November from Oklahoma.	18	17	13
17. A man in a hotel was a witness.	14	18	17
18. Brothers left a valley in a Packard sedan.	6	14.5	5
19. An Indian from a boarding house drove.	10.5	10	4

Note. Ranks 1-19 indicate the range "strongest imagery" -
"weakest imagery."

Table 16

Spearman Rank-Order Correlations among Centrality,
Image, and Order of Presentation for the Three Control Groups

			Centrality (C)	Image (I)	Order of Presentation
	Native	C		.15	.30
		I			.38
Language	German	C		.36	.06
		I			.48*
	Mixed	C		.24	.27
		I			.15

* Significant at $p < .05$; none of the rho values are
significant at $p < .01$.

the highly rated image-evoking sentences or for the highly
rated central sentences, or for some particular ordering
of the nineteen sentences; we will then want to conclude
that this difference is related to one or the other of the
between-group differences. Results will have been influ-
enced either by (a) whether a subject belonged to a group
that heard the sentences with the definite articles or
whether he belonged to a group that heard the sentences with
the indefinite articles, or (b) whether a subject thought
the sentences formed a story or whether he thought the nine-
teen sentences were unrelated.

Before we make use of the control group data to ana-
lyze how the variables definite/indefinite or story/non-
story influenced the quality of recall among the experimen-
tal groups, we will attempt to determine whether there were
group differences in the amount of the story that was re-
called. Table 17 shows how much of the story was recalled
(mean number of words, synonyms, and sentences) by each
group of subjects: those who heard the definite article
version of the text, those who heard the indefinite article
version, those who thought the sentences formed a story,
and those who thought they did not. A Mann-Whitney U test
was used to determine whether group membership had a sig-
nificant effect on recall. We will first compare the mean
amount recalled by our four groups of native language sub-
jects with the results of de Villiers' study. Slightly

Table 17

Mean Amount Recalled by <u>Story</u> and <u>Non-Story</u> Subjects
in Each Language Group for 19 Definite or
Indefinite Sentences

Group	n	M no. of words recalled	M no. of synonyms recalled	M no. of sentences recalled
Native - Definite				
Story	6	24.8	6.8	9.7
Non-Story	5	11.6	1.8	3.8
Native - Indefinite				
Story	9	22.2	4.1	7.3
Non-Story	8	9.6	1.8	3.3
German - Definite				
Story	3	22.0	2.0	6.0
Non-Story	5	9.0	0.5	2.5
German - Indefinite				
Story	5	13.2	3.0	4.0
Non-Story	6	10.5	1.2	3.5
Mixed - Definite				
Story	3	14.3	0.7	2.7
Non-Story	6	10.3	1.2	2.8
Mixed - Indefinite				
Story	2	9.0	3.5	4.0
Non-Story	9	9.0	1.5	2.2

less than half of de Villiers' subjects in the definite article group recalled the sentences as a story. None of the subjects in the indefinite article group thought they had heard a story. De Villiers found that his subjects who recalled the definite sentences as a story showed significantly better recall than those who recalled the passage as unrelated sentences. For non-story groups there was no significant difference between subjects who belonged to the definite or the indefinite groups. While the definite article seemed to facilitate processing the passage as a story, de Villiers concludes that definite reference alone does not improve recall. The use of the definite article only seems to have this effect when subjects process the passage as a story (de Villiers, 1974, p. 266).

Definite reference in the present study also failed to have an effect on recall when viewed separately from how a subject had processed the nineteen sentences. Mann-Whitney U tests showed that differences in recall between the definite and indefinite article groups were not significant on any of the three measures of recall (mean number of words, mean number of synonyms, or mean number of sentences recalled). This is true for all three of the language groups tested. On the other hand, subjects who processed the sentences as a story seemed to have a significant advantage in two of the language groups over subjects who processed the sentences as unrelated.

It should be pointed out that, in contrast to the de
Villiers study, we are unable to predict from the article
used in the present text, how subjects will process the
passage. De Villiers found among his subjects that there
was a strong correlation between having heard the definite
article passage and processing the passage as a story. Al-
though our test was unable to pinpoint just which feature
or features of the discourse facilitated processing the pas-
sage as a story, de Villiers' finding that story processing
depends on the presence of the definite article does not
generalize to the present results. In the present test,
whether a subject viewed the input as a story or not seemed
to be independent of whether he had heard the definite or
the indefinite sentences. This was true for all three of
the language groups in the study (table 17). Among the na-
tive language subjects, in the definite article group, six
subjects processed the sentences as a story and five as a
non-story. In the indefinite article group there was a
similar division: nine subjects viewed the sentences as a
story and seven as a non-story. The other two language
groups were similar. Definite article groups for both
German and mixed subjects contained three members who proc-
essed the sentences as a story and six as a non-story.
Indefinite article groups were comprised of five and two
story members and six and nine non-story members for the
German and mixed language groups respectively.

Because the difference definite versus indefinite failed to be significant either in affecting how subjects processed the nineteen sentences or in influencing how much they recalled of the sentences, this distinction was not retained for further analysis of the test. Instead groups were consolidated under the larger divisions story versus non-story. Mann-Whitney U tests show that there significant differences among the language groups based on this measure. The results of these tests appear in table 18.

For the native group of subjects, differences between story and non-story groups were significant on all measures of recall. Subjects who processed the sentences as a story recalled significantly more words and sentences and used more synonyms than those subjects who did not process the sentences as a story. This was true to a lesser extent for the German language group. Only the measure mean number of sentences failed to reach significance between story and non-story subjects in this group. German subjects who processed the sentences as a story, like native subjects, recalled significantly more words and used more synonyms than those subjects who processed the sentences as unrelated. For the mixed language group story processing did not seem to affect the amount of the text recalled. Differences in recall between story and non-story groups were not significant for this group of subjects on any of the measures of recall. Thus the advantage of story processing seems

Table 18

Results of Mann-Whitney U Tests
(with Probability Levels)
Comparing the Mean Amounts Recalled
by Story and Non-Story Subjects

Group	n	U score for words recalled	U score for syno-nyms recalled	U score for sen-tences recalled
Native				
Story vs.	15	25.5	26.5	20.5
Non-Story	13	(.0009)	(.001)	(.0004)
German				
Story vs.	8	21.0	21.0	27.0
Non-Story	11	(.057)	(.049)	(.152)
Mixed				
Story vs.	5	27.0	34.0	29.0
Non-Story	15	(.358)	(.751)	(.442)

greatest for the native language subjects, less for the German group, and non-existent for the mixed language group.

If there is a difference in the amount recalled between story and non-story groups, is there also a difference in what kind of information each group recalls from the nineteen sentences? How does this vary over the three language groups tested? For his native language subjects, de Villiers used Spearman rank-order correlations between the mean group accuracy of recall for each sentence and each of the ratings given by the control groups for centrality and imagery to test for differences in the quality of recall of both story and non-story groups. De Villiers found sentence accuracy to show significant positive correlations with centrality within the story group. In contrast, sentence accuracy showed a significant positive correlation with imagery in the non-story group. Sentence accuracy for the non-story group also correlated negatively with the order of presentation of the original fifteen sentences. These results suggest that non-story subjects relied on strong imagery and how recently they had heard a sentence in writing their recalls. In contrast, subjects who recalled the sentences as a story based their recall on how central a sentence was to the theme of the story (de Villiers, 1974, pp. 266-267).

Spearman rank-order correlations between sentence accuracy and each of the respective control group ratings

for centrality and imagery in the present study failed to
show a satisfactory number of correlations which could sig-
nificantly characterize the story and non-story processing
differences among the three language groups. Table 19 pre-
sents the correlation coefficients and p-values between
the two recall measures, sentence accuracy and recalled
order, and each of the ratings, centrality and imagery, from
the respective control groups. Rank-order correlations were
also calculated between the two measures of experimental
group recall and original order of presentation of the nine-
teen sentences.

The only correlations which reached significance were:
(a) between the original order of presentation and the re-
called order of presentation for the native story group,
.82, p = .001; and (b) between sentence accuracy and imagery
for both story and non-story mixed group subjects, .32,
p < .10. The remaining Spearman rank-order correlations were
not able to determine significant differences in the quality
of recall between story and non-story groups among any of
the three language groups tested.

Because of the lack of significant correlations with
the control group ratings, we attempted to analyze experi-
mental group recall independently of the centrality and
image ratings the control groups had given the nineteen
sentences. Instead we compared story and non-story subjects
for each of the three language groups on the basis of how
many subjects in each group recalled each of the nineteen

Table 19

Spearman Rank-Order Correlations
between Centrality, Image, and Order of Presentation
of Original Sentences and Experimental Groups' Recalled
Sentence Accuracy and Recalled Order of Presentation

Group		Order of presentation	Centrality	Image
Native				
Story	sentence accuracy	-.11	-.01	.04
	recalled order	.82***		
Non-Story	sentence accuracy	.13	.17	-.19
	recalled order	.29		
German				
Story	sentence accuracy	.02	.17	-.10
	recalled order	.30		
Non-Story	sentence accuracy	.13	.001	.03
	recalled order	.32		
Mixed				
Story	sentence accuracy	-.22	.06	.32*
	recalled order	.21		
Non-Story	sentence accuracy	.10	.41	.32*
	recalled order	.39**		

* Significant at $p < .10$.
** Significant at $p < .05$.
*** Significant at $p < .01$.

test sentences. Tables 20 through 22 present for each language group the proportion of story and non-story subjects that attempted to recall each of the sentences. Values for z are presented at the right of the table for each sentence and critical values for significance are noted.

The greatest number of significant differences between story and non-story subjects are found within the native language group (table 20). Significant differences are shown between the two processing groups for sentences 1, 7, 11, 15; $p < .01$ and sentences 2, 3, 5, 16, 17; $p < .05$. The mean number of subjects in the story group to recall each of these sentences was always significantly higher than the mean number of subjects in the non-story group for the same sentence. If we examine these sentences more closely, we find that they include all but one of the five sentences the native control group rated most central to the theme of the story. These are: 1. The stranger came to town on a donkey (rank, 2.5). 3. The animal was struck by a trolley in Tulare Street (rank, 4), 7. The man wanted to buy a car (rank, 1). 10. The man asked the boy at the counter to be the driver (rank, 2.5), 15. The Indian couldn't drive the car (rank, 5). Sentence 19 just fails to show a significant difference between story and non-story groups at $p < .05$. All of the other sentences predict significant differences between the two groups.

We can conclude for the native language subjects that there are significant differences between those who

Table 20

Proportions of Native <u>Story</u> and <u>Non-Story</u> Subjects
to Recall Each of the 19 Test Sentences

Sentence	Group <u>Story</u>	<u>Non-Story</u>	Value for <u>z</u>
1	13 / 15	3 / 13	2.65**
2	8 / 15	2 / 13	2.09*
3	6 / 15	1 / 13	1.97*
4	11 / 15	7 / 13	1.07
5	9 / 15	2 / 13	2.41*
6	2 / 15	1 / 13	0.48
7	9 / 15	1 / 13	2.88**
8	8 / 15	8 / 13	-0.44
9	6 / 15	1 / 13	1.97*
10	10 / 15	4 / 13	1.89
11	8 / 15	1 / 13	2.58**
12	7 / 15	2 / 13	1.77
13	5 / 15	2 / 13	1.09
14	5 / 15	2 / 13	1.09
15	10 / 15	2 / 13	2.73**
16	8 / 15	2 / 13	2.09*
17	9 / 15	2 / 13	2.41*
18	8 / 15	5 / 13	0.79
19	6 / 15	4 / 13	0.51

* Significant at $p < .05$.
** Significant at $p < .01$.

Table 21

Proportions of German Story and Non-Story Subjects
to Recall Each of the 19 Test Sentences

| Sentence | Group | | Value |
	Story	Non-Story	for z
1	5 / 8	3 / 11	1.54
2	0 / 8	0 / 11	1.00
3	0 / 8	0 / 11	1.00
4	6 / 8	5 / 11	1.29
5	0 / 8	1 / 11	-0.88
6	1 / 8	2 / 11	-0.34
7	4 / 8	1 / 11	2.00*
8	2 / 8	6 / 11	-1.29
9	2 / 8	1 / 11	0.94
10	5 / 8	3 / 11	1.54
11	5 / 8	0 / 11	3.05**
12	3 / 8	1 / 11	1.50
13	1 / 8	2 / 11	-0.34
14	2 / 8	0 / 11	1.75
15	7 / 8	7 / 11	1.17
16	2 / 8	1 / 11	0.94
17	1 / 8	4 / 11	-1.17
18	1 / 8	1 / 11	0.24
19	2 / 8	3 / 11	-0.11

* Significant at $p < .05$.
** Significant at $p < .01$.

Table 22

Proportions of Mixed Story and Non-Story Subjects
to Recall Each of the 19 Test Sentences

Sentence	Group Story	Non-Story	Value for z
1	1 / 5	2 / 15	0.36
2	1 / 5	1 / 15	0.86
3	2 / 5	1 / 15	1.81
4	3 / 5	8 / 15	0.26
5	0 / 5	1 / 15	-0.60
6	1 / 5	2 / 15	0.36
7	4 / 5	4 / 15	2.11*
8	2 / 5	3 / 15	0.89
9	1 / 5	4 / 15	0.30
10	2 / 5	2 / 15	1.29
11	1 / 5	7 / 15	-1.05
12	1 / 5	3 / 15	0.00
13	1 / 5	4 / 15	-0.30
14	1 / 5	0 / 15	1.78
15	4 / 5	7 / 15	1.30
16	1 / 5	7 / 15	-1.05
17	1 / 5	3 / 15	0.00
18	0 / 5	0 / 15	1.00
19	0 / 5	1 / 15	-0.59

* Significant at $p < .05$.

processed the passage as a story and those who did not, not
only in the amount each group recalled, but also differences
in the particular sentences each group recalled. Story sub-
jects were significantly better than non-story subjects at
recalling the sentences the control group had rated as most
central to the theme of the story. As the Spearman rank-
order correlations showed, this group also recalled the
story sentences in the original order in which they were
presented. This was not true for the non-story group sub-
jects.

We have already noted that group differences in the
amount of the passage recalled are smaller for story and
non-story subjects in the German and mixed language groups
than in the native group. (In fact, differences were not
significant for the mixed group). Likewise, fewer signifi-
cant differences in the quality of recall emerge between
story and non-story subjects in the two non-native groups.
German story and non-story groups (table 21) showed sig-
nificant differences on only two of the sentences in the
passage. Their story recall was only significantly better
for sentences 7. The man wanted to buy a car and 11.
School was closed for summer vacation. Mixed group story
recall (table 22) was only significantly better for sen-
tence 7. Neither of these sentences was included in the
respective control group ratings for the five sentences
most central to the theme of the story. We can conclude
that non-native story and non-story subjects in the present

test did not differ in their ability to recall the sentences thematically most central to the passage.

We will want to examine more closely the sentences recalled by the non-native story and non-story groups to see if any differences emerge in the quality of what each group recalled. Spearman rank-order correlations have already indicated that both story and non-story mixed subjects show a slightly significant correlation between sentence accuracy and strong imagery. Table 23 compares the six most accurately recalled sentences for all groups of story and non-story subjects. This comparison shows that, except perhaps for a slight preference among the non-story groups for the final sentences in the story, there seems to be little difference among subjects as to which sentences they recalled accurately. For the most part, the sentences most accurately recalled were the same for both story and non-story subjects.

Our test confirms our third hypothesis that the recall of connected discourse is facilitated if the listener thinks the discourse fits together into a story, and that this facilitation increases with increasing proficiency in the language. Native subjects who processed the text as a story recalled significantly more than native subjects who did not process the text as a story. Retention in the German group was also higher among those who claimed to find a story in the text than for those who did not find a story,

Table 23

Sentences Most Accurately Recalled
by Story and Non-Story Groups

		Story		Non-Story	
		M % Accuracy	Sentence	M % Accuracy	Sentence
		71	1	32	1
		60	4	42	4
		51	7	50	8
	Native	53	8	26	10
		49	10	21	18
		66	15	25	19
		50	1	20	1
		65	4	40	4
		50	7	54	8
Group	German	57	10	21	10
		49	11	63	15
		87	15	33	17
		20	1	53	4
		40	4	20	7
		65	7	46	11
	Mixed	40	8	19	13
		20	13	46	15
		73	15	38	16

although the effect of this processing difference was less than for the native group. The mixed group showed no significant processing effects whether they claimed to have seen a story in the text or not. In fact, as table 19 points out, there is an inverse relationship in correctly ordering the sentences in recall between those who claim to have perceived the text as a story and those who said they saw the sentences as unrelated.

Based on de Villiers' findings, we made two further predictions, (a) that the absence of definite articles in a passage significantly changes the perception of story structure and (b) that the perception of a story aids in the recall of sentences centrally related to the theme of the story, while the recall of unrelated sentences is determined by recency or strong imagery. Neither of these predictions was confirmed. In the present experiment, definite reference did not facilitate processing the passage as a story. Many subjects who heard the indefinite-article sentences processed the passage as a story, and many subjects who heard the definite-article sentences thought the sentences were unrelated. This was true for native as well as non-native groups.

It is also difficult to assess the role imagery played in the present results. Although recall of the mixed language group for both story and non-story subjects correlated positively with the control group rating for imagery,

non-story subjects in the other two language groups did not show this expected effect. Perhaps the sentences in the text were partly responsible for the inconsistent results obtained. The list of the most accurately recalled sentences (table 23) includes sentences 4. The boy in the drugstore was eating a banana split and 7. The man wanted to buy a car for both story and non-story subjects. Sentence 7 was given high rankings for centrality and image by all of the control groups. Sentence 4 was given extremely low control group ratings for centrality and high ratings for imagery. Nevertheless, there is little difference in how these sentences were recalled by the experimental groups; they were both among the most accurately recalled sentences in all groups for both story and non-story subjects.

Although our experiment was unable to specify which features of sentences facilitate story processing, as predicted by our hypothesis the test has shown that native language subjects who process a passage as a story show significantly better recall than subjects who process a passage as unrelated. Among our non-native groups, German story subjects also showed an advantage over non-story subjects in the number of words they recalled and the number of synonyms they used. Mixed language subjects showed no significant effects of story or non-story processing. Both story and non-story subjects in this group relied on the image-evoking quality of the sentences in their retellings more

than on their perception of a story structure in the pas-

sage.

CHAPTER V

CONCLUSIONS AND GENERAL DISCUSSION

We now turn to an evaluation of our results from ex-
periments one, two, and three with respect to our three
original major hypotheses. The results of all three experi-
ments show that semantic and discourse constraints in an
aural message serve as important sources of information for
the native language listener. The two non-native listener
groups in the present study seem to have less access to
semantic information in a message, and the amount of seman-
tic information they use seems to decrease with decreased
proficiency in English. In all three experiments, the
second-language group with the higher level of proficiency
in English as measured by the Michigan State University
English Language Examination, that is, the German group,
showed results closer to the results of the first-language
listeners than did those of the other non-native experi-
mental group.

Our three experiments were an exploratory attempt to
characterize the differences between strategies used by
native and non-native language listeners. Based on pre-
vious studies in reading comprehension, we made the general
assumption that an important difference between first- and
second-language listening comprehension lies in the native

language listener's ability to make use of the semantic cues
in the aural message and, contrastively, the second-language
listener's difficulty in attending to this information be-
cause of his excessive preoccupation with the more surface
features of the text. Guided by this general assumption,
we made use of three tests which had successfully been used
in native language research to test the following three
hypotheses.

Experiment one was addressed to our first hypothesis
that, with increased proficiency in a language, a listener's
expectations of an incoming message should show greater
attention to the semantic cues in the aural input than to
the syntactic or phonological cues. The experiment diag-
nosed post-listening cloze scores and found that native,
German, and mixed language groups scored increasingly higher
on the diagnostic scoring categories weak context and
syntax over semantic. Our test results also showed a high
negative correlation between scores for weak context and
syntax over semantic and comprehension scores on the cloze
test. This would indicate that along with poorer compre-
hension, attention to syntactic cues and a failure to make
use of the overall semantic context were characteristic of
the non-native listener scores. Again, this proved to be
more characteristic of the scores for the mixed group,
which was less proficient in English, than for the German
group.

The contrasting results between first and second-
language subjects on our post-listening cloze test show par-
allels to results on traditional cloze tests administered
to these two groups and to results obtained on tests of
first- and second-language reading. Traditional cloze tests
given to native and non-native subjects (Chihara et al.,
1977; Cziko, 1978; Lapkin & Swain, 1977) have repeatedly
shown that non-native subjects are characteristically less
sensitive to the discourse constraints in a text than they
are to the syntactic constraints of each immediate phrase.
Chihara et al. have shown that discourse constraints in a
cloze test have a relatively smaller facilitating effect
on the responses of non-native subjects than of native sub-
jects. Similarly, Devine's (1979) reading research using
the Reading Miscue Inventory (Goodman & Burke, 1972) showed
that while non-native readers of English with differing
levels of proficiency made equal numbers of miscues preserv-
ing syntactic structure, miscues made by the most proficient
readers tested were characterized by their preservation of
the meaning of the text.

From our findings for experiment one, we assume that
processing differences that exist in reading between native
and second-language learners can be extended to character-
ize the two groups' skills in listening comprehension. In
listening to connected discourse, we can conclude that the
non-native listener directs more attention to the syntactic

cues in the text than the native listener does.

A second experiment tested our second hypothesis that more proficient listeners in a language are progressively less apt to remember details of syntax in sentences they have heard. This experiment replicated the well-known idea-acquisition experiment of Bransford and Franks (1971), which used a sentence acquisition and recognition task to show that native language listeners retain little of the syntactic form of individual sentences but instead seem to retain whole semantic units. Our experiment two, modeled on the earlier study, showed similar results for the native group of test subjects. Like the subjects in the Bransford and Franks study, when asked to identify syntactically different sentences with the same semantic content, this group of subjects was unable to differentiate between sentences they had actually heard before and sentences they were hearing for the first time.

Our second hypothesis predicts that non-native listeners, in contrast to native listeners, should be better able to discriminate between sentences on the basis of their syntactic form. Because we expected these subjects to direct more attention to the syntactic features of the individual sentences, we predicted that they would not demonstrate the native-listener tendency to retain semantic information at the expense of a nearly total memory loss for the syntactic form.

Our results for experiment two only partially support this prediction. Contrary to our expectations, the German group of non-native listeners showed the same tendency exhibited by the native listeners of integrating the semantic information from several acquisition sentences and then no longer being able to recognize the syntactic form in which this information had initially been presented. Like the native listeners, this group was unable to discriminate between novel sentences expressing the same semantic content (NEW) and the original sentences they had heard during the acquisition exercise (OLD).

The results of the mixed group of non-native subjects are more supportive of our original hypothesis. In rating whether they had heard a sentence before or not, this group was more accurate in picking out the OLD and NEW sentences. This was most notable in their ratings of THREE and FOUR sentences, sentences which were designed to most completely integrate the semantic ideas in the original acquisition sentences. Whereas native and German subjects were strongly convinced that they had heard these semantically complex sentences before, the mixed group of subjects gave these sentences lower confidence ratings.

There was not enough evidence, however, to conclude that this group was retaining the syntactic form of the original acquisition sentences. On the contrary, they, like the other groups, were still quite unable to accurately

recognize OLD sentences among the sentences presented in the recognition list. In addition, analysis of their test results using a second technique suggested by Bransford and Franks, showed that this group, although to a lesser extent than the other two test groups, was also demonstrating a tendency to integrate semantic information from the acquisition sentences. On the recognition task, between two sentences with equivalent qualitative semantic content, they tended to give a higher confidence rating to the sentence with more semantic content than to the semantically less complete sentence. This would indicate that semantic integration was also part of the sentence processing demonstrated by this non-native group.

What is most interesting perhaps is that semantic processing for the mixed non-native group of listeners seems to be less effective than for subjects from the other two test groups. The mixed group of listeners, on the whole, was less accurate about the semantic information they retained. They were less confident than native or German subjects of having heard sentences that semantically contradicted information from the acquisition sentences (NONCASE).

We would conclude from our results for experiment two that, although non-native listeners show less of a tendency than native listeners to process individual sentences semantically, contrary to our predictions, second-language subjects did not show a significant advantage over the

native subjects in retaining the syntactic form of the test sentences. The most reasonable explanation for the failure of our prediction that non-native listeners should be better at retaining syntactic features of a message would seem to be processing capacity limitations. Based on Miller's (1956) theory of "chunking," reading researchers have proposed that familiarity with the language of the input at the grapho-phonic and syntactic levels allows the reader to construct hypotheses about the incoming information. In this way, the vast number of diverse and short-lived bits of information in short-term memory are "chunked" into larger units which do not exceed the processing capacity limitations of long-term memory. The information in short-term memory can thus lead to a higher-level processing, freeing the lower level of memory for further input. With no adequate semantic framework, however, the syntactic information in short-term memory is quickly lost. As a result, little information from the text ends up getting stored in long-term memory, and the ineffective syntactic processing characteristically leads to much poorer retention of the passage (Smith, 1971, pp. 185-211). Processing limitations seem a plausible explanation for the failure of our mixed non-native group to have the predicted advantage in retaining the syntactic information of the acquisition sentences in experiment two and also this group's failure to have retained the semantic information in the sentences.

Our test measure has not been able to exclude the possibility that non-native listeners in the mixed group had more trouble understanding the syntactic form of the acquisition sentences. It should be pointed out that our scoring of recall in the test is not sensitive enough to distinguish between subjects who did not remember the syntactic form because they processed the sentences semantically and those who were unable to recall the syntax because they were unable to decode it in the first place. Although none of the sentences used in the present study should have been beyond the proficiency of any of the advanced-level students participating, we cannot exclude the possibility that this factor may have influenced our results.

Experiment three tested our third hypothesis that the recall of connected discourse is facilitated if the listener thinks the discourse fits together into a story, and that this facilitation increases with increased proficiency in the language. Based on an earlier study by de Villiers (1974), this experiment attempted to test how retention for nineteen sentences differed depending upon whether these sentences contained definite or indefinite articles and whether subjects processed the sentences as a story or whether they processed them as if they were unrelated to one another. The experiment was designed so that the definite articles served to link the sentences into a story and the use of indefinite articles gave the impression that

the sentences were unrelated. After writing a retelling
of the nineteen sentences, subjects were asked to indicate
whether they had heard nineteen isolated sentences or a
short story.

Like the results obtained for a similar task by de
Villiers, the fact that subjects viewed the nineteen sen-
tences as a story had a significant effect on their recall
of the passage. An additional finding of the present ex-
periment was to show that the advantage of semantic proces-
sing seemed to be a function of proficiency in English.
Native subjects who processed the story semantically re-
called significantly more of the text than native subjects
who did not process the sentences as a story. Story proces-
sing improved retention for the German group of listeners,
but to a lesser extent than for the native group. The
mixed group showed no significant effects of processing the
text as a story or of processing it as isolated sentences.
Retention for the <u>story</u> group of mixed language subjects
did not differ from retention for the <u>non-story</u> group.

De Villiers' original experiment was also designed
to test two further hypotheses: (a) that the presence of
definite articles in a passage significantly affects the
perception of story structure and (b) that the perception
of a story aids in the recall of sentences centrally re-
lated to the theme of the story, while the recall of unre-
lated sentences is determined by recency or strong imagery.

De Villiers' study confirmed both of these hypotheses. The discourse feature of definiteness carried by the article had a significant effect on the recall of his native group of subjects.

Our experiment predicted results similar to those of de Villiers for the native group of subjects participating in our test. For our non-native groups, we predicted that especially the mixed group would show little effect of whether they had heard the passage with the definite articles or the passage with the indefinite articles; we predicted that recall for this group would not be significantly different for either passage.

Contrary to our expectations based on de Villiers' first prediction, the present study failed to show a significant effect of the use of the definite article for any of the groups tested. It should be noted that de Villiers' test also never found an effect for definite reference when this was not combined with subjects' processing the sentences as a story. In the present test it was story-processing alone which significantly affected recall. We were unable, however, to link story-processing to the presence of the definite article.

As we expected, based on de Villiers' second prediction, our native language subjects, and to a lesser extent our German subjects, who processed the sentences as a story were shown to be better at retaining the thematically important sentences in the passage. We were unable to

characterize, however, the sentences recalled by the non-story native or German subjects. Spearman rank-order correlations failed to show significant effects for recency or imagery on the sentences recalled by either of these groups. Recency also failed to affect the recall of either group of mixed language subjects. However, recall for all of the mixed language subjects seemed to be affected by sentence imagery, which can probably best be considered to be a lower-level type of semantic processing. Significant positive correlations were obtained between how many times a sentence was recalled by this group and an image rating that had been given to the sentences by an independent control group. There were no differences among the mixed subjects between those who thought they had heard a story and those who thought they had heard unrelated sentences in the listening exercise.

The results of experiment three further support our more general prediction that much of the difficulty in listening to a second language may be due to the inability of the non-native listener to extract semantic-level cues from the message, and that this ability, at least in part, develops with proficiency in the language. Our first experiment demonstrated that the non-native listener attends more to the syntactic-level cues in the text than to semantic information. Experiment three yields some evidence that sentence imagery, a very low level of semantic processing,

also plays more of a role in non-native listening comprehension than does the thematic, textual organization of semantic content. Because both of these tendencies were stronger for the less proficient of our two non-native listener groups, and because the group with the higher proficiency in English showed results more like those of the native language group, if one is allowed to speculate, this might suggest that there is a developmental order in the non-native language listeners' ability to make use of the semantic constraints in an aural message. Such a claim goes beyond the realm of the present study, but it is interesting that several researchers have come to a similar conclusion about the development of skill in reading a second language, namely, that a sensitivity to syntactic constraints develops before a sensitivity to semantic constraints in a text (Cziko, 1978; Hanzeli, 1979; Lapkin & Swain, 1977).

If further research should support the hypothesis that the second-language learner develops a reliance on syntactic information in the new language before developing a reliance on semantic information, this may be a result of language instruction or this may be a result of the nature of comprehension itself. Carroll (1972) has proposed that comprehension consists of two stages: first, the apprehension of the linguistic information in an aural or written message; and second, only when this first stage has been achieved, the relating of this information to a larger semantic

context (Carroll, 1972, p. 13). From the test results of the present study, it would seem that much of the difficulty in second-language listening lies with getting past the first of Carroll's stages, the apprehension of the linguistic information in the message. For native language listeners much of this information is redundant; and native listeners can readily move back and forth between the two levels or stages, going back to the linguistic information only when their expectations at the semantic level are not met and require further confirmation. In contrast, Carroll's stage one for the non-native listener seems often to act as an unfilled prerequisite for stage two. Faced with so much new information at the first stage of comprehension, the second-language reader or listener is still unable to retain and integrate this over a longer text.

Research into the nature of the comprehension process has clearly only just begun, and up to the present, development in the research of listening comprehension is much behind progress that has been made in reading-comprehension research. Not unexpectedly, the present research leaves us with as many unanswered questions as answered ones. We have concluded that in listening to a message in a foreign language, the listener's comprehension breaks down because he is unable to form the semantic framework necessary for efficient language processing. Our data have not been able to conclusively show just where the second-language

listener's difficulties lie. They could be at all or any of the levels in the aural text: at the semantic or contextual level, at the level of text organization, at the vocabulary level, or at the lower level of processing syntax and phonology.

Our test results have suggested that at least part of the problem can be found at this last level, the amount of attention the non-native listener has to direct to lower-level syntactic cues in the text. What we have missed most in the present research is a testing device that would be capable of directly measuring this attention. All of the present experiments have had to make use of recall tests, which test language production, when what we were trying to test was language reception. Although our second experiment attempted to measure the relative amounts of attention native and non-native listeners direct to syntactic and semantic information in sentences, we were unable to pinpoint where the comprehension breakdown occurred for the non-native listeners because we attempted to measure this in a recall task. We were able to show that comprehension for this group was less than for the native group. However, we were unable conclusively to show that the reason for comprehension failure was this group's processing too much syntactic detail, as we had no direct way of measuring the processing time spent.

Our third experiment presented a similar difficulty in drawing conclusions about how non-native listeners

process a story. Our test showed that recall of a passage for the less-proficient non-native group of listeners was not affected by whether they perceived a story structure in the passage or not. However, again we were unable to conclusively determine from our data whether this group was less able to make use of the story structure in comprehending the passage or only in reproducing it.

These are just some of the questions our research leaves unanswered. Clearly, there is still much to be learned about both native- and second-language listening comprehension. In an exploratory study the present experiments have presented some empirical evidence for the hypothesis that much of the difficulty in second-language listening lies in the unavailability to the second-language listener of semantic information in the text. The preliminary nature of the present inquiry is only too obvious. Both the many questions our study has raised and the tentative conclusions we have drawn will need the support of further research.

APPENDICES

APPENDIX A

CLOZE TEST FOR EXPERIMENT ONE

APPENDIX A

CLOZE TEST FOR EXPERIMENT ONE

Nonverbal Behavior and Communication

Have you ever seen a driver stopped by a traffic
policeman in the United States? Perhaps you have been
__(stopped)__ yourself because you didn't __(see)__
a stop sign or __(drove)__ too fast. Think back
__(and)__ try to picture in __(your)__ mind how
the policeman __(looked)__ as he approached. What
__(kind)__ of facial expression did __(he)__
have? Was he smiling __(and)__ friendly or did he
__(look)__ firm or even angry? __(How)__ did he
walk as __(he)__ came to the car,__(and)__ how
did he stand? __(Was)__ his posture erect and
__(defiant)__ or did he look __(as)__ if he was
afraid __(to)__ go near the driver? __(What)__
about his hands and __(arms)__ --were they swinging
freely __(at)__ his side or were they straight and
tense, ready __(to)__ draw his gun?

Now __(let's)__ compare this image of __(an)__
American policeman with one __(I)__ saw of an
Italian __(policeman)__ some time ago. The __(picture)__
showed him begging, almost __(pleading)__ with the driver--
his __(posture)__, his hands, his face __(all)__
suggested a very different, __(more)__ polite relation-
ship to the __(driver)__ of the car. You __(will)__
seldom see an American policeman approach someone this way.

__(Perhaps)__ what each policeman would __(have)__
to say to the __(driver)__ would change the judgment
__(or)__ inference we have made __(about)__ each

142

of these officers. _____(Perhaps)_____ the Italian policeman would speak more firmly than his __(pleading)____ posture suggests, or the American might be more friendly;

_____(we)_____ don't know. But from ____(our)_____ experience in other similar ___(situations)___ we can be fairly __(confident)____ that someone who approaches _____(us)_____ with the gesture of ___(pleading)_____ or begging is going ____(to)_____ be much easier to ____(deal)_____ with than someone with ___(his)_____ hands at his side ____(ready)_____ to shoot or fight.

_____(Not)_____ all the gestures we ___(use)_____ are easy to __(interpret)_____, though, nor are all ____(gestures)____ to which we may ___(give)_____ meaning intended to communicate ___(anything)____.

Note.   Permission to use the tape recording and script of this passage was kindly granted by Mrs. Wu Yi So of the Michigan State University English Language Center. The text is an excerpt from a listening comprehension instructional lecture given by Elwood Friesen. Words in parentheses were deleted for the written part of the test.

APPENDIX B

SENTENCES FOR EXPERIMENT TWO

APPENDIX B

SENTENCES FOR EXPERIMENT TWO

Idea Set B

FOUR: Along an interstate highway a motorist with a
           Citizens Band radio reported an accident involving
           a fuel truck and a small postal vehicle.
                (On Recognition Only)

THREES: Along an interstate highway a motorist reported an
           accident involving a fuel truck and a small postal
           vehicle.
                (On Acquisition Only)
           A motorist with a Citizens Band radio reported an
           accident involving a fuel truck and a small postal
           vehicle.
                (On Acquisition Only)
           Along an interstate highway a motorist with a
           Citizens Band radio reported an accident.
                (On Both Acquisition and Recognition)

TWOS: A motorist with a Citizens Band radio reported an
           accident.
                (On Acquisition Only)
           Along an interstate highway was an accident invol-
           ving a fuel truck and a small postal vehicle.
                (On Recognition Only)
           A motorist reported an accident involving a fuel
           truck and a small postal vehicle.
                (On Recognition Only)

ONES: The accident was along an interstate highway.
                (On Acquisition Only)
           The motorist had a Citizens Band radio.
                (On Both Acquisition and Recognition)
           There was an accident involving a fuel truck and a
           small postal vehicle.
                (On Recognition Only)
           A motorist reported an accident.
                (On Recognition Only)

144

Idea Set C

FOUR: An old pick-up truck hauling a trailer was having trouble getting up the ice-covered hill.
          (On Recognition Only)

THREES: An old pick-up truck hauling a trailer was having trouble getting up the hill.
          (On Acquisition Only)
An old pick-up truck has having trouble getting up the ice-covered hill.
          (On Acquisition Only)
A pick-up truck hauling a trailer was having trouble getting up the ice-covered hill.
          (Recognition Only)

TWOS: An old pick-up truck was having trouble getting up the hill.
          (On Acquisition only)
An old pick-up truck was hauling a trailer.
          (On Both Acquisition and Recognition)
A pick-up truck was having trouble getting up the ice-covered hill.
          (On Recognition Only)
A pick-up truck hauling a trailer was having trouble getting up the hill.
          (On Recognition Only)

ONES: A pick-up truck was having trouble getting up the hill.
          (On Acquisition Only)
A pick-up truck was hauling a trailer.
          (On Acquisition Only)
The hill was ice-covered.
          (On Recognition Only)
The pick-up truck was old.
          (On Recognition Only)

Idea Set D

FOUR: The tax collector found himself talking on the
            doorstep of a shabby house to a woman with two
            children tugging at her skirts.
                (On Recognition Only)

THREES: The tax collector found himself talking on the
            doorstep of a shabby house to a woman with two
            children.
                (On Acquisition Only)
            The tax collector found himself talking to a woman
            with two children tugging at her skirts.
                (On Acquisition Only)
            On the doorstep of a shabby house was a woman with
            two children tugging at her skirts.
                (On Recognition Only)

TWOS: The tax collector found himself talking to a
            woman on the doorstep of a shabby house.
                (On Acquisition Only)
            The woman on the doorstep of the shabby house had
            two children.
                (On Acquisition Only)
            The two children on the doorstep of a shabby house
            were tugging at their mother's skirts.
                (On Recognition Only)
            The tax collector found himself talking to a woman
            with two children.
                (On Recognition Only)

ONES: The tax collector found himself talking to a woman.
                (On Acquisition Only)
            The two children were tugging at their mother's
            skirts.
                (On Both Acquisition and Recognition)
            The woman had two children.
                (On Recognition Only)
            The tax collector found himself on the doorstep of
            a shabby house.
                (On Recognition Only)

APPENDIX C

RECOGNITION SENTENCE RATINGS FOR EXPERIMENT TWO

APPENDIX C

RECOGNITION SENTENCE RATINGS FOR EXPERIMENT TWO

Idea Set A

+2.7, +2.8, +1.2
1. The young boy walking
alongside the road
could not see the dust
storm approaching from
the southwest. (FOUR)

+3.0, +2.2, +1.5
2. The boy walking along-
side the road could
not see the dust storm
approaching from the
southwest. (THREE)

OLD

+2.7, +0.3, +1.2
3. The young boy walking
alongside the road
could not see the
dust storm. (THREE)

+1.0, -5.0, 0.0
4. The boy could not
see the dust storm
approaching from the
southwest. (TWO)

+3.0, +2.6, +1.3
5. The boy walking
alongside the road
could not see the
dust storm. (TWO)

-3.2, -1.4, -0.3
6. The dust storm was
approaching from
the southwest. (ONE)

+1.3, +1.0, +1.3
7. The boy could not
see the dust storm.
(ONE)

Predictions

1 ⟩ 2 ⟩ 4 ⟩ 6   1 ⟩ 3 ⟩ 5 ⟩ 7   2 ⟩ 5   4 ⟩ 7

Note. The numerical values above each sentence represent
the mean recognition ratings for (a) the native group,
(b) the German group, (c) the mixed group. Sentence
3 in this idea set is OLD; the remaining sentences
are NEW.

147

Idea Set C

+1.8, +3.7, +3.5
1. An old pick-up truck hauling a trailer was having trouble getting up the ice-covered hill. (FOUR)

+1.1, +2.7, +1.9
2. A pick-up truck hauling a trailer was having trouble getting up the ice-covered hill. (THREE)

+2.7, +3.0, +1.8
3. A pick-up truck was having trouble getting up the ice-covered hill. (TWO)

+3.8, +4.1, +2.5
4. A pick-up truck hauling a trailer was having trouble getting up the hill. (TWO)

+1.3, +2.2, -0.3
OLD 5. An old pick-up truck was hauling a trailer. (TWO)

-3.6, -3.3, -1.1
6. The hill was ice-covered. (ONE)
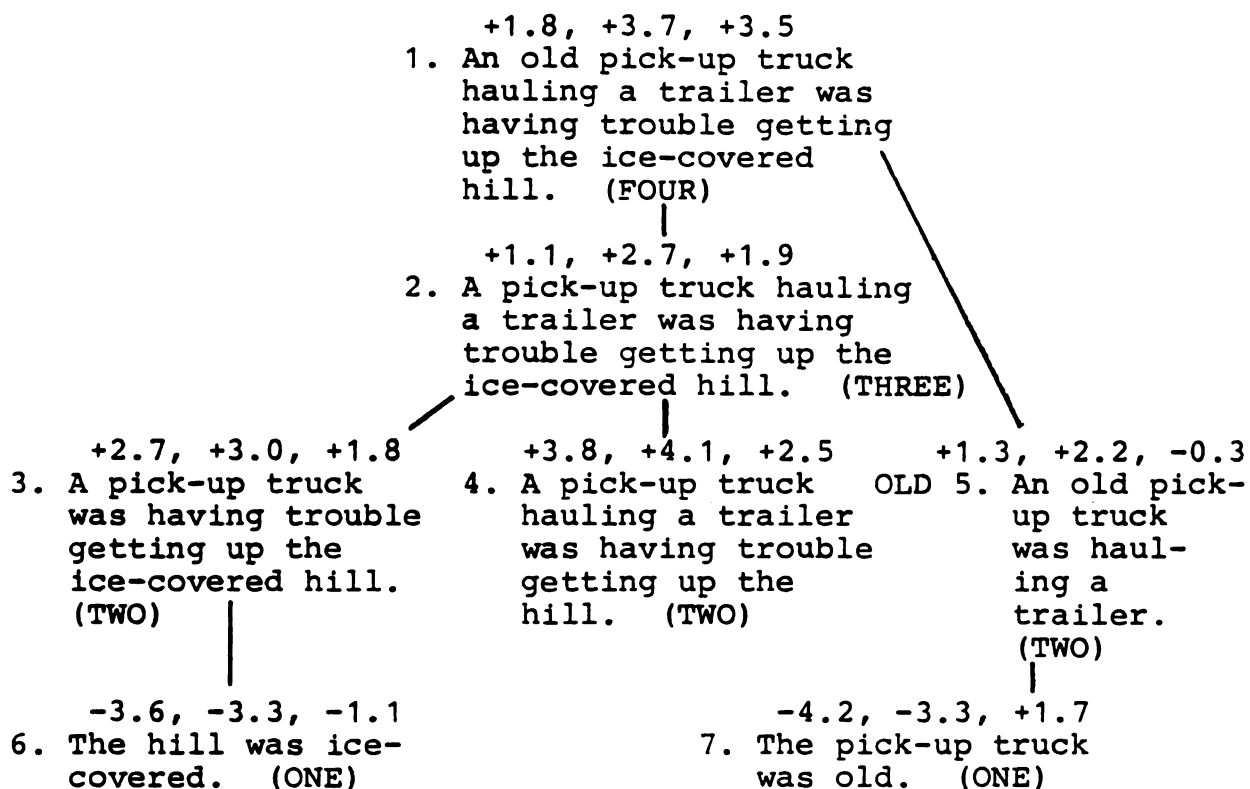
-4.2, -3.3, +1.7
7. The pick-up truck was old. (ONE)

Predictions

1 > 2 > 3 > 6   2 > 4   1 > 5 > 7

Note. The numerical values above each sentence represent the mean recognition ratings for (a) the native group, (b) the German group, (c) the mixed group. Sentence 5 in this idea set is OLD; the remaining sentences are NEW.
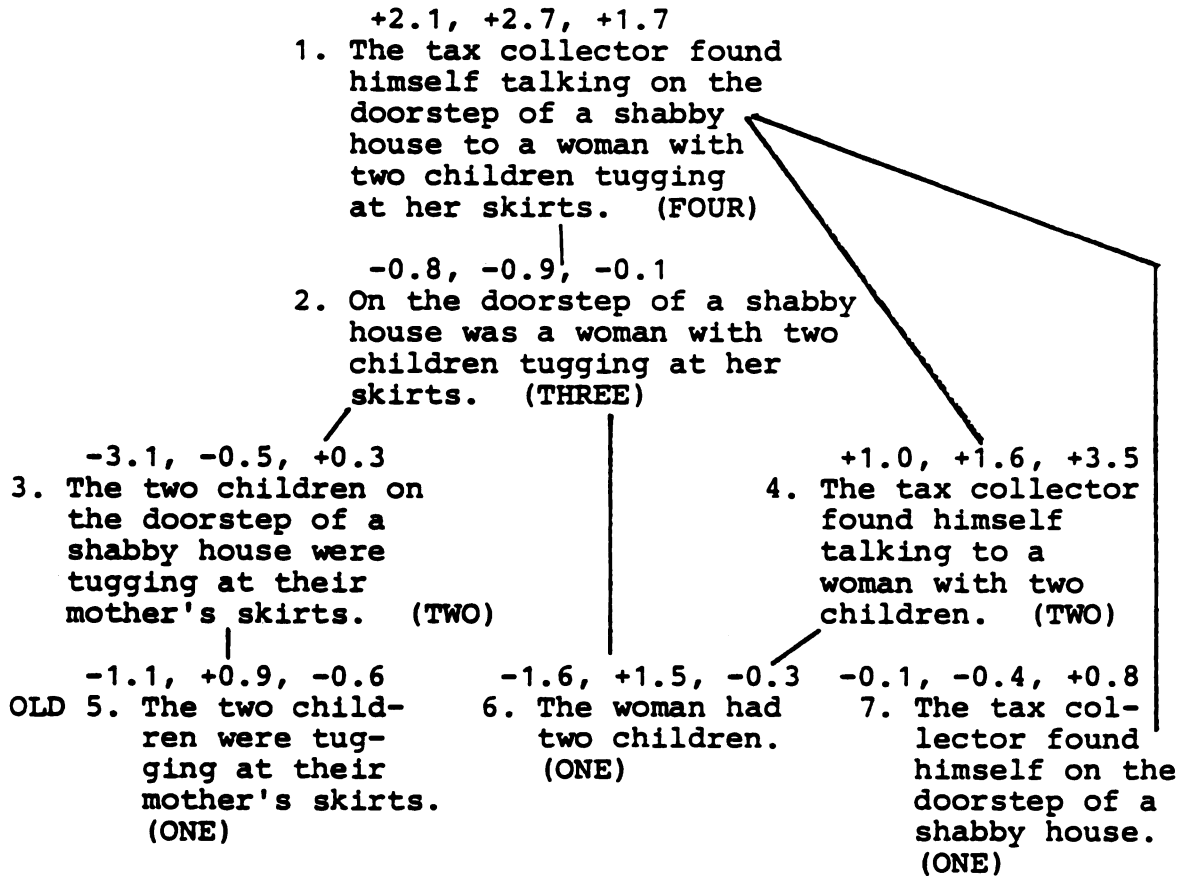
Idea Set D

+2.1, +2.7, +1.7
1. The tax collector found
himself talking on the
doorstep of a shabby
house to a woman with
two children tugging
at her skirts. (FOUR)

-0.8, -0.9, -0.1
2. On the doorstep of a shabby
house was a woman with two
children tugging at her
skirts. (THREE)

-3.1, -0.5, +0.3
3. The two children on
the doorstep of a
shabby house were
tugging at their
mother's skirts. (TWO)

+1.0, +1.6, +3.5
4. The tax collector
found himself
talking to a
woman with two
children. (TWO)

-1.1, +0.9, -0.6
OLD 5. The two child-
ren were tug-
ging at their
mother's skirts.
(ONE)

-1.6, +1.5, -0.3
6. The woman had
two children.
(ONE)

-0.1, -0.4, +0.8
7. The tax col-
lector found
himself on the
doorstep of a
shabby house.
(ONE)

Predictions

1 > 2 > 3 > 5   1 > 4 > 6   2 > 6   1 > 7

<u>Note</u>. The numerical values above each sentence represent
the mean recognition ratings for (a) the native group,
(b) the German group, (c) the mixed group. Sentence
5 in this idea set is OLD; the remaining sentences
are NEW.

LIST OF REFERENCES

LIST OF REFERENCES

Bartlett, F. C. Remembering. Cambridge, Mass.: Cambridge
    University Press, 1932.

Black, J. B., Turner, T. J., & Bower, G. H. Point of view
    in narrative comprehension, memory, and production.
    Journal of Verbal Learning and Verbal Behavior, 1979,
    18, 187-198.

Bormuth, John. Factor validity of cloze tests as measures
    of reading comprehension. Reading Research
    Quarterly, 1969, 4(3), 358-365.

Bransford, J. D., & Franks, J. J. The abstraction of
    linguistic ideas. Cognitive Psychology, 1971, 2,
    331-350.

Carpenter, Patricia A., & Just, Marcel. Integrative
    processes in comprehension. In David LaBerge & S. Jay
    Samuels (Eds.), Basic processes in reading: Perception
    and comprehension. Hillsdale, N. J.: Erlbaum
    Associates, 1977.

Carroll, John B. Defining language comprehension: Some
    speculations. In John B. Carroll & Roy Freedle (Eds.),
    Language comprehension and the acquisition of knowledge.
    Washington, D.C.: Winston, 1972.

Chafe, Wallace L. Discourse structure and human knowledge.
    In John B. Carroll & Roy O. Freedle (Eds.), Language
    comprehension and the acquisition of knowledge.
    Washington, D.C.: Winston, 1972.

Chihara, T., Oller, J., Weaver, K., & Chavez-Oller, M. A.
    Are cloze items sensitive to discourse constraints?
    Language Learning, 1977, 27, 63-73.

Cirilo, R. K., & Foss, D. J. Text structure and reading
    time for sentences. Journal of Verbal Learning and
    Verbal Behavior, 1980, 19, 96-109.

Clarke, Mark A., & Burdell, Linda. Shades of meaning:
    Syntactic and semantic parameters of cloze test
    responses. Paper presented at the Eleventh Annual
    TESOL Convention, Miami Beach, April 1977.

150

Crymes, R., James, G., Smith, L., & Taylor, H.
Developing fluency in English: With sentence-combining
practice in nominalization. Englewood Cliffs, N.J.:
Prentice-Hall, 1974.

Cziko, Gary A. Differences in first- and second-
language reading: The use of syntactic, semantic and
discourse constraints. The Canadian Modern Language
Review, 1978, 34(3), 473-489.

de Villiers, Peter A. Imagery and theme in recall of
connected discourse. Journal of Experimental Psychology,
1974, 103, 263-268.

Devine, Joanne. Developmental patterns in native and
non-native reading acquisition. Paper presented at
the Thirteenth Annual TESOL Convention, Boston,
March 1979.

Garrett, M. F., Bever, T.G., & Fodor, J. A. The active
use of grammar in speech perception. Perception and
Psychophysics, 1966, 1, 30-32.

Goodman, Kenneth S. The psycholinguistic nature of the
reading process. Detroit: Wayne State University
Press, 1967.

Goodman, Kenneth S., & Burke, Carolyn. The Goodman taxonomy
of reading miscues. In David Allen & Dorothy Watson (Eds.),
Findings of research in miscue analysis: Classroom
implications. Urbana, Ill.: ERIC-NCTE, 1976.

Goodman, Yetta, & Burke, Carolyn. Reading miscue inventory:
Procedure for diagnosis and evaluation. New York:
Macmillan, 1972.

Halliday, M. A. K. Notes on transitivity and theme in
English: II. Journal of Linguistics, 1967, 3, 199-244.

Hanzeli, Victor. Cloze tests in French as a foreign
language: Error analysis. In Eugene Brière & Frances
Hinfotis (Eds.), Concepts in language testing: Some
recent studies. Washinton, D.C.: TESOL, 1979.

Haviland, Susan E., & Clarke, Herbert H. What's new?
Acquiring new information as a process in comprehension.
Journal of Verbal Learning and Verbal Behavior, 1974,
13, 512-521.

Jones, Randall, & Spolsky, Bernard (Eds.). Testing
language proficiency. Arlington, Va.: Center for
Applied Linguistics, 1975

Kintsch, Walter. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum Associates, 1974.

Kintsch, W., & Greene, E. The role of culture-specific schemata in the comprehension and recall of stories. Discourse Processes, 1978, 1, 1-13.

Kintsch, W., & Kintsch, E. H. The role of schemata in text comprehension. International Journal of Psycholinguistics, 1978, 5(2), 17-29.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. Comprehension and recall of text as a function of content variables. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 196-214.

Kintsch, W., & Monk, D. Storage of complex information in memory: Some implications of the speed with which inferences can be made. Jounal of Experimental Psychology, 1972, 94, 25-32.

Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Krulee, G., Fairweather, P., & Bergquist, S. Organizing factors in the comprehension and recall of connected discourse. Journal of Psycholinguistic Research, 1979, 8(2), 141-163.

Lapkin, S., & Swain, M. The use of English and French cloze tests in a bilingual program evaluation: Validity and error analysis. Language Learning, 1977, 27, 279-314.

Lesgold, Alan. M. Pronominalization: A device for unifying sentences in memory. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 316-323.

Lesgold, A. M., Roth, S. F., & Curtis, M. E. Foregrounding effects in discourse comprehension. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 291-308.

Macnamara, John. Comparative studies of reading and problem solving in two languages. TESOL Quarterly, 1970, 4, 107-116.

Mehler, J., Segui, J., Pittet, M., & Barrière, M. Strategies for sentence perception. Journal of Psycholinguistic Research, 1978, 7(1), 3-16.

Miller, George A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 227-228.

Mistler-Lachman, Janet L. Depth of comprehension and sentence memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 98-106.

Morley, Joan. Improving aural comprehension. Ann Arbor: University of Michigan Press, 1972.

Morley, Joan. Listening dictation. Ann Arbor: University of Michigan Press, 1976.

Oller, John W. Assessing competence in ESL reading. TESOL Quarterly, 1972a, 6, 313-323.

Oller, John W. Scoring methods and difficulty levels for cloze tests of ESL proficiency. Modern Language Journal, 1972b, 56(3), 151-158.

Oller, J., & Conrad, C. The cloze technique and ESL proficiency. Language Learning, 1971, 21, 183-196.

Page, William D. The post oral reading cloze test: New link between oral reading and comprehension. Journal of Reading Behavior, 1975, 7(4), 383-389.

Plaister, Ted. Developing listening comprehension for ESL students. Englewood Cliffs, N.J.: Prentice-Hall, 1976.

Rivers, Wilga M. Listening comprehension. In Kenneth Croft (Ed.), Readings on English as a second language for teachers and teacher trainees. Cambridge: Winthrop, 1972.

Rumelhart, David. Understanding and summarizing brief stories. In David LaBerge & S. Jay Samuels (Eds.), Basic processes in reading: Perception and comprehension. Hillsdale, N. J.: Erlbaum Associates, 1977.

Sachs, J. D. S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 1967, 2, 437-442.

Smith, Frank. Understanding reading. New York: Holt, Rinehart and Winston, 1971.

So, Wu Yi. Bridging the Gap: Between learning English
    in the classroom and using English in lecture halls
    (Vols. 1 & 2). Lansing, Mich.: Michigan State Univer-
    sity Press, 1979.

Taylor, W. L. Cloze procedure: A new tool for measuring
    readability. Journalism Quarterly, 1953, 30, 415-433.