

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB RD., ANN ARBOR, MI 48106

8212195

Harris, Martin Steven

**THE ROLE OF SENTENCE CONTEXT IN THE PERCEPTION AND
PRODUCTION OF INTERNAL OPEN JUNCTURE**

City University of New York

Ph.D. 1982

**University
Microfilms
International** 300 N. Zeeb Road, Ann Arbor, MI 48106

THE ROLE OF SENTENCE CONTEXT IN THE PERCEPTION
AND PRODUCTION OF INTERNAL OPEN JUNCTURE

by

MARTIN S. HARRIS

A dissertation submitted to the Graduate Faculty
in Speech and Hearing Sciences in partial fulfillment
of the requirements for the degree of Doctor of
Philosophy, The City University of New York.

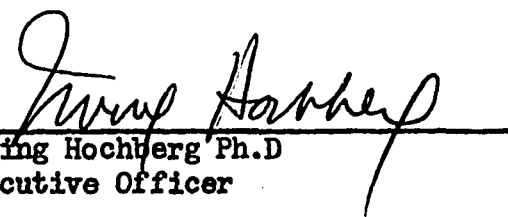
1982

This manuscript has been read and accepted for the Graduate faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

January 22, 1982
date


Harry Levitt Ph.D
Chairman of Examining Committee

January 22, 1982
date


Irving Hochberg Ph.D
Executive Officer

Katherine S. Harris Ph.D

Harry Levitt Ph.D

Lawrence Raphael Ph.D
Supervisory Committee

The City University of New York

Acknowledgments

I would like to thank my adviser Dr. Harry Levitt for his continuous understanding and patience during the course of this project. Without his constant good natured prodding I am sure this thesis would not have been completed. In addition I would like to acknowledge the efforts of the rest of my dissertation committee, Dr. Katherine Harris and Dr. Lawrence Raphael. I know that they spent many hours going over numerous preliminary drafts and I sincerely appreciate their thoughtful suggestions.

It is important for me to acknowledge the staff of the Acoustics Research Department of Bell Laboratories and especially Dr. Noriko Umeda for allowing me to pursue this project while I was a consultant. The time, equipment, and advice were all invaluable..

I would also like to offer special thanks to my uncle Mr. Harry Eider for teaching me many years ago, that one must never give up on a worthwhile goal no matter how many obstacles are placed on the path to success.

This thesis is dedicated to three very important people, my mother, my father and my wife. I suppose they drove me to complete this dissertation by simply having complete faith in my ability. There were many moments when I doubted myself, but I do not think that they ever did. I thank them for this faith and I hope to remain worthy of it in the future.

TABLE OF CONTENTS

<u>CHAPTER I - BACKGROUND REVIEW AND OUTLINE OF STUDY</u>	1
Prosodic features in perception.....	2
Acoustic characteristics and perception of word boundary in internal open juncture phrases.....	12
Contextual influences on perception and production.....	23
Rationale and outline of experiments.....	33
<u>CHAPTER II - EXPERIMENT I: EFFECTS OF CONTEXT ON INTELLIGIBILITY</u>	36
Method.....	37
Results.....	43
<u>CHAPTER III- EXPERIMENT II: ROLE OF SENTENCE CONTEXT ON PERCEPTION</u>	74
Overview.....	74
Method.....	75
Results.....	81
<u>CHAPTER IV - EXPERIMENT III: PERCEPTION AND PRODUCTION IN ISOLATION AND CONTEXT</u>	112
Method.....	114
Results.....	119
<u>CHAPTER V - DISCUSSION</u>	133
Experiment I.....	133
Experiment II.....	140
Experiment III	146
Summary and Conclusion.....	160
<u>BIBLIOGRAPHY</u>	162

LIST OF TABLES

1. Sentence Contexts From Which Phrases Were Excised for the Experimental Phrases of Experiment I.....	40
2. Percent Correct Per Phrase in Both Parts of Experiment I...	44
3. Mean Ratings by Fifty Listeners of "Commonness" of Phrases Used in Experiment I.....	44
4. Total Duration of Utterance in Milliseconds	49
5. Duration of Phoneme Preceding the Junctural Consonant.....	50
6. Duration in Milliseconds of Junctural Consonant.....	50
7. Percentage of Total Duration Occupied by Pre-junctural Phonemes.....	52
8. Percentage of Total Duration Occupied by Junctural Phoneme.	52
9. Relative Rate of Fall of Fundamental Frequency After Release of Consonant.....	60
10. Fundamental Frequency Immediately Before Closure and After Release for all Phrases in all Conditions.....	61
11. Peak Frequency of Each Phrase Before the Closure and After the Release.....	61
12. Presence or absence of Glottal Stop and Aspiration in Junctural Portion of all Phrases and Corresponding Percent Correct in Listening Task.....	64
13. All Utterances Used in Experiment II.....	78
14. Percent Correct for Each Phrase in Part I.....	84
15. Mean Ratings of "Commonness" of Phrases in Experiment II....	86
16. Percent Correct for Each Phrase in Part III.....	86
17. Percent Correct for Each Phrase in Part III.....	87
18. Spectrographic Characteristics of Junctural Consonant in Phrases in Correct and Anomalous Sentences and in Isolation.....	90

19.	Spectrographic Characteristics of Initial and Final Consonant Phrases in Correct Sentences.....	91
20.	Peak Fundamental Frequency of the First Word and the Second Word for the Word Initial and Word Final Consonant Conditions.....	101
21.	Fundamental Frequency at the end of the First Word and at the Beginning of the Second for the Same Phrases in correct and Anomalous Contexts.....	103
22.	Fundamental Frequency at the End of the First Word and at the beginning of the Second.....	104
23.	Peak Fundamental Frequency of the First Word and the Second Word of the Phrases that Appeared in Correct and Anomalous Utterances.....	105
24.	Mean Duration of Phrases and Components on Correct, Anomalous and Isolated Conditions. Percent Correct Responses in the Listening test are Included.....	110
25.	All Sentences Used in all Parts of Experiment III.....	116
26.	Percent Correct Identification for all Conditions of Experiment III	121
27.	Overall Percent Correct Judgments for the Four Main Conditions in Isolation.....	124
28.	Duration of Junctural Consonant and Preceding Phoneme in all Conditions of Experiment III.....	128
29.	Matrix Showing Significant Factors in Analyses of Variance for Isolated and Context Conditions.....	147
30.	Percent Correct Listening Judgments as a Function of Consonant Type Over all Three Experiments.....	156

LIST OF FIGURES

1. Spectrograms Showing How Duration was Measured in the Phrase "Great Eye"	48
2. Fundamental Frequency and Intensity Plots for Juncture Pairs in Isolated and Excised Conditions.....	54
3. Fundamental Frequency and Intensity Plots for Juncture Pairs in isolated and Excised Conditions.....	56
4. Fundamental Frequency and Intensity Plots for Juncture Pairs in Isolated and Excised Conditions.....	58
5. Spectrograms of "Grey Tie" in Isolated and Excised Conditions	65
6. Spectrograms of "Great Eye" in Isolated and Excised Conditions.....	66
7. Spectrograms of "Grey Day" in Isolated and Excised Conditions.....	67
8. Spectrograms of "Grade A" in Isolated and Excised Conditions.....	68
9. Spectrograms of "Plum Pie" in Isolated and Excised Conditions.....	69
10. Spectrograms of "Plump Eye" in Isolated and Excised Conditions.....	70
11. Percent Correct Responses in all Parts of Experiment II....	82
12. Spectrograms of the Phrase "Grey Day" in Three Conditions..	92
13. Spectrograms of the Phrase "Grade A" in Two Conditions.....	93
14. Spectrograms of the Phrase "Grey Tape" in Three Conditions.....	94
15. Spectrograms of the Phrase "Great Ape" in Two Conditions...	95
16. Fundamental Frequency Contours of the Phrase "Grey Day" and "Grade A" in a number of Conditions.....	97
17. Fundamental Frequency Contours of the Phrases "Grey Tape" and "Great Ape" in a number of Conditions.....	98
18. Fundamental Frequency Contours of the Phrases "It Swings" and "Its Wings" in a Number of Conditions.....	99

19. Overall Phrase Duration in Different Contexts	107
20. Duration of Phoneme Preceding Junctural Consonant in Different Conditions.	108
21. Duration of Junctural Consonant in Different Conditions.....	109
22. Percent Correct Responses in all Sections of Part I of Experiment III.....	120
23. Interaction Effect of Position and Type of Consonant for "Isolated Correct Condition".....	127

CHAPTER I

BACKGROUND REVIEW AND OUTLINE OF STUDY

This investigation was concerned with factors involved in the perception of internal open juncture phrases. In particular, the study sought to determine the relative importance of the acoustic cues and contextual information on boundary assignments in this juncture condition.

The review is divided into four major sections. The first section deals with the prosodic features of English, their acoustic characteristics and their perception. The second section is concerned with the characteristics of syllable and word boundary that apply to the internal open juncture condition. The third section presents a review of literature concerning the influence of context on perception and production. The fourth section is a discussion of the interaction of acoustic cues and contextual information in the perception of internal open juncture. This discussion is formulated in terms of the specific problem investigated in the experiments.

Prosodic Features in Perception

Prosodic features, as conventionally described by American linguists, refer to the characteristics of intonation, intensity and duration. According to Lehiste (1970) prosodic features are to be restricted to the condition in which these characteristics extend over more than one segment. The importance, in perception, of acoustic parameters such as fundamental frequency, intensity, and duration of segments of an utterance are well documented (Klatt, 1976; Lea, 1973; Lehiste, 1970; and Lieberman, 1967). Fry (1955) investigated the effects of duration and intensity on the perception of stress. He had several speakers utter disyllables (in sentences) "in which a change of function from noun to verb is commonly associated with a shift of stress from the first to the second syllable." (p.765) For example, some of the words were "object", "subject", "permit", etc. From spectrograms, duration and intensity measures of the segments corresponding to the speech sounds were made. In all segments there was a difference in vowel duration ratio between the noun and verb form. Intensity

ratios, in general, did not show clearly marked differences. To obtain information on the relative effectiveness of duration and intensity as stress cues he synthesized different versions of some of these disyllables. In these versions duration and intensity ratios of the first vowel to the second vowel were varied over a range, which, based on the speaker's production measurements, was viewed as an area of uncertainty. He then presented these synthesized disyllables in synthesized carrier sentences to a large group of listeners who were asked to state where the accent appeared in the critical word. He found that "in words where duration and intensity were operating in the same direction there was excellent agreement between the subjects, that is to say when the vowel was long and of high intensity listeners agreed that the vowel was strongly stressed, when it was short and of low intensity, it was judged as weakly stressed." (p.767) When duration and intensity were examined separately it was found that duration ratio had a greater effect on stress judgment than intensity.

In his 1958 work Fry extended the experiments to include the role of fundamental frequency variation in

stress judgment. The experiment contained two hypotheses, first, it was hypothesized that if two syllables differed in fundamental frequency, the syllable with the higher frequency will be judged as stressed. In the second portion he investigated the relationship, if any, of the fundamental frequency and the duration ratio to the judgment of stress. In this experiment synthesized disyllables were presented in isolation. The results, for an increase in fundamental frequency showed that changes tended to produce an all-or-none effect. That is, if the fundamental frequency contour was incremented by more than 5 Hz on a given syllable, that syllable was perceived as stressed. The magnitude of the frequency change did not affect the perception of stress beyond this value. The effect of changing duration ratios on noun and verb judgments was also clearly demonstrated. In an additional experiment, designed to determine the effect of sentence intonation on stress judgment Fry found that "sentence intonation is an over-riding factor in determining the perception of stress and that in this sense the fundamental frequency cue may outweigh the duration cue." (p.151). It is important to point out that the utilization of synthetic stimuli

in which the cues may be drastically simplified may not completely represent natural speech conditions.

Lieberman (1960) also examined fundamental frequency, envelope amplitude, and duration in natural speech. He examined pairs of syllables in which one syllable was perceived as stressed. The stressed syllable was found to have higher fundamental frequency than the unstressed syllable of the same utterance 90% of the time. It had a higher peak envelope amplitude 87% of the time and a longer duration 66% of the time. Data were also obtained for the same syllable in both stressed and unstressed conditions. In comparing each stressed syllable with its unstressed version, Lieberman found that the F_0 was higher 72% of the time, the envelope amplitude was higher in 90% of the cases and the duration longer in 70% of the cases. These data indicate that of the 3 variables considered, fundamental frequency is correlated most closely with stress, which is consistent with Fry's findings. Lieberman's results differed from Fry's in terms of the role of envelope amplitude and duration. Fry found duration to be more important than intensity. Lieberman found just the opposite. This difference may be due to the use of synthetic

speech by Fry and the use of natural speech by Lieberman. In addition, Lieberman's findings were correlational whereas Fry's observations were causal.

According to Bolinger (1958), who used both natural and synthetic speech, stress is determined by "pitch prominence," a rapid and wide departure from the smooth contour. He found duration to be a covariable with pitch but intensity did not play a critical role in stress judgments. Morton and Jassem (1965) used synthetic nonsense syllables and varied duration, pitch, and intensity systematically. The stimuli were constructed to have the same vowel in both syllables i.e. /sisi/ and /sasa/. Listeners were asked to mark which syllable they thought was stressed. Changes in fundamental frequency yielded much greater effects than changes in either intensity or duration. A raised fundamental was found to be more efficient than a lowered one in altering the perception of stress. They also found that the all or none effect observed by Fry was true for their data. That is, changes of 25 to 58% in F_0 were equally good in producing the stress changes observed.

It must be noted that in studies that use isolated words or syllables the major disadvantage is

that the simplicity of the stimuli do not allow the prosodic features that operate over the entire sentence to be important in stress perception (Huggins, 1972). If we consider more complicated material then the role of prosodic features should be more clearly observable.

A number of investigators have pointed out the value of prosodic cues in the perception of connected speech. Kozhevnikov and Chistovich (1965) presented narrow band filtered speech that was unintelligible. They asked subjects to identify the phrases. Subjects could not do this, but their ability to perceive correct word accent was almost perfect. In other words they were very poor at phrase intelligibility but because the information concerning pitch and timing was transmitted they were very good at accent location. It is apparently very difficult to destroy prosodic cues in running speech by filtering.

Bansel (1966) presented evidence from a study of Indian speakers of English suggesting that subjects rely on prosodic cues more than on segmental cues when these two types of cues are in conflict. In lists of English words spoken by Indians many words are spoken

with the incorrect stress pattern. Native speakers of English asked to identify these words match the pattern of stress as opposed to segmental cues. For example, for the word "character" the speakers produced $[kə'rʌktə]$ instead of $[k'æ rə ktə]$ and the listeners produced responses like "director", "collector" or "detractor". In these words the stress pattern is identical and since correct phonemes occur in many responses it is suggested that the listener is discarding segmental cues when the prosodic cues are in conflict with them.

Indications that the stress pattern of a word is important to its identity also comes from a paper by Brown and McNeill (1966). They prompted their subjects until a target word was on the tip of their tongues but yet not recallable. At this time each subject filled out a questionnaire concerning the "target-word". Of the words they thought sounded like the "target-word" and containing the same number of syllables, 74% had the same stress pattern as the target word. It is suggested that the stress pattern of a word is included in the way the word is represented in memory. It seems likely, therefore, that the stress pattern would be an important cue to the perception of the word.

Although the above studies provide evidence of the importance of prosodic cues in perception, it must be remembered that in two of these studies the stimulus material was either not English or, at best, distorted English. The Kozhevnikov and Chistovich study was of Russian where the use of stress is markedly different than in English. In the Bansel study, incorrect use of stress was employed by having non English speakers produce the utterances. In addition, non-native speakers may very well distort segmental characteristics which would also confound the results.

Huggins (1978) investigated the effects of changing durational relationships among characteristics of stressed and unstressed syllables on intelligibility of synthetic speech. He adjusted the durations of stressed syllables so that they would be considerably shortened and lengthened unstressed (short) syllables so that they would be longer than normal. This incorrect timing information caused the intelligibility to drop to about one-third of its normal value. The implication is that durational information is particularly important for intelligibility. It is important to bear in mind, however, that Huggins

did not simply remove durational information, but rather provided misleading durational information. If all syllables had the same duration, rather than reversing the duration ratio between stressed and unstressed syllables, then the effect on intelligibility might not have been so dramatic.

Jonn and Howarth (1968) showed the importance of temporal aspects of prosody in their research with the deaf. They made recordings of informal teaching sessions during which a teacher selected a spontaneous utterance of a deaf child and spent a few minutes repeating the remark fluently and having the child attempt to imitate that temporal pattern. This short time spent on each phrase produced 56% improvement in word intelligibility and a 200% improvement in the number of "sentence patterns" heard correctly. While adjustment in prosody presumably helped increase the intelligibility it is not known how much of the improvement may have been due to improved segmental characteristics resulting from the training.

It has also been demonstrated that when segmental and durational information is unchanged but the intonation contour is incorrect, intelligibility scores are considerably lowered. Wingfield and Klein (1971)

used sentences sharing a common five word sequence but having syntactic boundaries at different places within the string. For example: "To avoid attempts to influence voting, machines were installed." and "Due to our mayor's influence, voting machines were installed." From this pair they made prosodically anomalous sentences by splicing each sequence (influence voting machines were installed) with the other sentence's context. For example: "Due to our mayor's influence voting, machines were installed." Listeners were required to write what they heard. The errors in the prosodically anomalous condition were significantly greater than in the normal condition.

Acoustic characteristics and perception of word
boundary in internal open juncture phrases

The perception of boundaries can be discussed at a number of levels. There are accounts of boundaries between sentences, between words, between syllables, between morphemes and between phonemes. There is little agreement as to what constitutes a boundary signal in speech for each of these conditions. Witness the lack of success in artificial recognition systems for running speech (Shockey & Reddy, 1974; Pisoni & Sawusch, 1974). There is no doubt that the acoustic characteristics discussed in the previous section function as boundary markers in at least some, if not all of these different boundary conditions. For instance, Fujimura and Lovins (1977), in a review of factors defining syllable boundaries, noted the importance of extrinsic allophonic variation. There are distinct differences in acoustic characteristics between a phoneme in the syllable initial and syllable final position. These acoustic differences have been shown to be effective boundary markers under certain conditions (Christie, 1974; Davidsen-Nielsen, 1974). Umeda and Coker (1974) and Coker and Umeda (1975) have

documented some of the acoustic differences between initial and final allophones of various consonants. For instance, nasals, liquids, voiced stops and voiced fricatives are lower in intensity in initial position than they are in final position; and voiceless stops and voiceless fricatives are higher in initial position than they are in final position. Greater length of syllable initial consonants, and glottalization of syllable initial vowels have also been observed (Leniste, 1960).

A revealing approach to the analysis of syllable boundaries is to examine the change in acoustic characteristics when a phoneme is shifted from the end of one word to the beginning of the next, thereby producing a shift in the location of the word boundary. This boundary condition is known as internal open juncture. Changing the location of the juncture provides an efficient paradigm for investigating the characteristics of word boundaries.

The first major experiment designed to determine the acoustic cues relevant to the perception of internal open juncture was that of Bolinger and Gerstman (1957). They investigated word boundary in the phrases "lighthouse-keeper" and light-housekeeper.

Bolinger and Gerstman hypothesized that one way in which connectedness between syllables is signaled is by "disjuncture" which they defined as separation of syllable centers. The acoustic gaps in "lighthouse Keeper" versus "light housekeeper" which occur because of the voiceless consonants were experimentally lengthened and shortened. For example, the gap between light and house could be increased while the gap between house and keeper was reduced and vice versa. The modified stimuli were presented to nine subjects for judgment as either lighthouse - keeper or light housekeeper. The results showed that the stimuli were judged in accordance with the durational change of the silent portions. Specifically, if the phrase lighthouse-keeper were spoken but the silent portion between light and house was lengthened (by splicing tape) and the portion between "house" and "keeper" was shortened, listeners judged the phrase to be "light housekeeper" 80% of the time. They concluded that listeners apparently used the relative temporal separation as the cue to juncture.

The first detailed examination of acoustic cues relevant to the perception of internal open juncture using a variety of phrases was performed by Lehiste

(1960). Lehiste defined internal open juncture as "the boundary between two bounded sequences". She did not consider internal open juncture to be a phoneme. Examples of minimal pairs that differ only in the location of internal open juncture are: "grey-day - grade A" and "plum pie - Plump eye. Lehiste refers to the bounded sequence as "a phonetically manifested higher level unit in the phonological structure of English". In her 1960 study Lehiste presented these open juncture phrases in listening tests. The results showed that juncture was quite easily recognized. However, it should be noted that the phrases were spoken and presented to the listener in isolation, and consequently, the contextual cues typical of real speech, were not available. This factor is considered in greater detail later.

Lehiste (1960) also performed spectrographic analyses of the phrase pairs. An examination of the measurements presented in her monograph, shows that the duration of the junctural consonant is greater in the word initial as opposed to the word final position. It should be noted that Lehiste measured only 35% of the phrases in the sample. The implicit suggestion was, of course, that the listener used the

durational differences to recognize the phrases correctly.

In 1965 Lehiste expanded her study to languages other than English. She identified a number of boundary signals; such as laryngealization, breathy phonation, insertion of a glottal stop, modification of patterns of fundamental frequency, intensity, and duration. For example, postjunctural allophones are longer than pre junctural allophones, aspiration is present in some prejunctural conditions and not in postjunctural. These signals are by no means universal to all the languages she studied and no single feature could be found to be common to all manifestations of this type of word boundary (Garding, 1967). As noted earlier the subjects in Lehiste's studies heard phrases only in isolation, an experimental condition that emphasizes the role of segmental cues.

Garding (1967) studied the acoustic, articulatory, and perceptual aspects of internal open juncture in the context of meaningful sentences in the Swedish language. She placed eighty pairs of meaningful phrases, having contrastive juncture placement, in a sentence frame and had two speakers produce them at two different rates. A set of fifty pairs of meaning-

less phrases, also of the internal open juncture variety, were put in a carrier frame and produced at two rates of speech. Third she took a set of nine pairs of the meaningful utterances and put them in a narrative read at three rates. The meaningful and nonmeaningful material was subjected to acoustic analysis as well as perceptual testing.

On the basis of the listening tests each of the phrases was assigned a misinterpretation score. These scores indicated that three factors influenced identification: 1, the segmental context, 2, the stress pattern and, 3, the speech rate. It was found that segmental context played an important role and that postjunctural phonemes contain the most important cues. The postjunctural phonemes were ranked according to their effectiveness as markers of juncture. Garding interpreted this rank ordering as follows; "juncture is easily recognized when the phonemes display a special feature (aspiration, glottal onset) in postjunctural position and fail to display it in prejunctural position" In the nonsense material the intervocalic consonants and at least one of the neighboring vowels, usually both, were the same in corresponding sequences of the sense material. The

rank order is not the same in the sense and nonsense materials. The glottal onset retains its effectiveness as an initial cue but aspiration loses some of its power in the nonsense material. A secondary finding of importance for the current study is that in a nonsense phrase having an intervocalic /s/ where it would be possible to assign /s/ to either the first word or the second, the first is usually preferred.

In the acoustical analysis, which was performed by measuring spectrograms, Garding found a number of consistent changes when juncture location shifted from $V_1 + CV_2$ to $V_1C + V_2$ (V_1 refers to the first vowel, the + sign refers to the boundary, C refers to the consonant and V_2 refers to the second vowel). These changes occurred regardless of the phonemes involved. First of all, when the juncture location shifts so that the word initial consonant of the second word is moved to become the word final consonant of the first word, overall duration of the phrase is changed, usually because of a closure of the glottis before V_2 . There are changes in the overall intensity and the fundamental frequency. The juncture location is combined with an interval of low intensity corresponding to the closure of the glottis. This period is longer for V_1C

+V₂ than for V₁ + CV₂. Garding defined juncture as a marked syllable boundary in a phrase. In normal speech juncture occurs at word and morpheme boundaries between consecutive stressed vowels. The articulation of juncture compared to the articulation of an ordinary syllable boundary can be described as follows: "The speech organs slow down and move toward a neutral position. When a consonant follows juncture, the oral closure or constriction is typically prolonged, when a vowel follows, the airstream is checked by glottal closure or constriction." It is possible, however, that these changes occur only under certain circumstances, for instance, in a clearly articulated reading of lists. It is also possible that the cues used for distinguishing between initial and final consonants are speaker specific.

Two recent studies of juncture perception have been reported by Nakatani and Dukes (1977) and Dukes and Nakatani (1978). In their 1977 paper Nakatani and Dukes used synthesized speech as the stimulus material. They systematically changed portions of speech in the regions surrounding a juncture using a speech synthesis technique developed by Olive and Nakatani (1974). The speech was processed so as to

produce isolated phrases taken from different parent phrases. For instance, the phrase pair "known ocean-no notion" was segmented into four parts, the beginning sounding like "no" with some of its end missing, and the last part like "ocean" with some of the beginning cut off. The two middle portions represented the transition between the end portion and the juncture phoneme /n/. The first of these portions represented the change from the end of the first portion "no", to the middle of the /n/. The second represented the change from the middle of the junctural consonant to the beginning of the final portion "ocean".

In the perception experiment listeners heard hybrid phrases presented in random order. Their task was to write down what they heard. From the manner in which listeners perceived juncture Nakatani and Dukes were able to tell how much influence each portion had on perception. Nakatani and Dukes identified a number of features that were correlated with juncture perception. Specifically, 1, Juncture is perceived when a glottal stop is followed by laryngealization and the second word begins with a vowel. 2, The allophones of /l/ and /r/ differ in syllable final and syllable ini-

tial positions and appear to be juncture cues. Syllable initial /l/ is light and produced much like a glide, syllable final /l/ is dark and produced as a semivowel. Similarly, syllable - initial /r/ is a retroflexed glide, whereas syllable-final /r/ is a syllabic-retroflexed semivowel. Nakatani and Dukes also noted that although there were systematic changes in the duration of the first portion of the phrase, this was not believed to be a strong cue to juncture.

Nakatani and Dukes concluded from these results that the most important juncture cues are qualitative cues, i.e. cues that occur in only one form of the juncture (a burst, silence, frication etc.). Quantitative cues that occur in both forms of the juncture, such as differences in intensity etc. were not believed to be important juncture cues.

Nakatani and Dukes argued that listeners probably paid little attention to the weaker duration cues because strong segmental cues were present. While it may be that qualitative cues are important cues to juncture in the isolated phrases carefully articulated by readers, it is not at all clear what happens if these cues are changed in any way, as they very well might be in running speech. Perhaps more attention is

played to durational cues when the other cues are not so strong. Also one might ask what happens when speakers do not produce qualitative cues such as glottal stops (Umeda, 1978).

In the 1978 study, Dukes and Nakatani examined internal open juncture using phrases that were spoken in a sentence context and presented to listeners in isolation. The listeners were asked to indicate in a forced choice task which phrase they heard. They found that subjects could recognize the correct phrase a high percentage of the time. The percentage correct varied considerably depending on the type of phrase. For example, perception was better for word - initial voiceless consonants (93% correct) than for their voiced cognates (77% correct). Spectrograms were examined in order to determine what cues might supply the necessary junctural information. Cues such as aspiration, glottal stops, and laryngealization were shown to be important. For instance, presence of a glottal stop correlated highly with juncture perception in a phrase whose second word begins with a stressed vowel. Unfortunately Dukes and Nakatani did not compare the perceptual results of their phrases presented in isolation to perception of the same phrases in some context, i.e. a sentence frame.

In none of the studies that have been reviewed thus far, has the effect of sentence context been examined. In the Dukes and Nakatani study, even though the acoustic cues were enough to yield reasonably high percentages, this does not necessarily mean that these boundary cues are used by the listener when the phrases actually appear in running speech. The current experiments addressed the issue of whether the quantitative or qualitative cues shown to be important in the Nakatani and Dukes papers play a significant role in the perception of continuous speech that is dominated by contextual information.

Contextual influences on perception and production

It has been demonstrated by a number of researchers (Miller, Heise and Lichten (1951), Pollack and Pickett (1963), and Lieberman (1963)) that syntactic and semantic characteristics of an utterance can affect listener's perception. Although there are many forms of context, the most relevant to the present research is sentence context (the environment that a word has in a sentence).

One of the early experiments concerned with the influence of sentence context on intelligibility was performed by Miller, Heise and Lichten (1951). They performed articulation tests on words spoken in a background of noise and they found that words spoken in sentences were intelligible at much lower signal to noise ratios than the same words in isolation. The authors argued that in a sentence the range of possible alternatives for each word is limited by the context of the sentence. It is a fairly easy task to choose the correct item because of the constraints placed on the alternatives as a function of the syntactic and semantic characteristics of the sentence.

O'Neill (1957) replicated the Miller et al. study using many more words presented in the sentence context and in isolation. Intelligibility scores for words in context were consistently higher than for the segmented words.

An inherent problem in the methodology of the Miller, Heise and Lichten study was pointed out by Stowe, Harris and Hampton (1963). They demonstrated that because Miller et al used noise over the whole sentence the amount of context that was received by the listener was a function of signal to noise ratio.

As a result the effects of context and signal to noise ratio are confounded. Stowe et al. attempted to design an experiment that separated the effect of context from that of S/N ratio. They varied the number of words of context and presented the context words visually without noise. Context words were defined as those immediately preceding the test word. The number of context words was 0, 2, 3, 5, or 10. that preceded the test word. The test word was always a monosyllable and always in the sentence final position. The s/n ratio was -6 dB.

Stowe, Harris and Hampton found that average intelligibility of the test word decreased with decreasing number of words of context. They noticed that there was a large difference in difficulty of the various test words as a function of the "frequency effect" i.e. the less common words in the set were more difficult.

The frequency effect has been described in detail by Howes (1957) for words spoken in white noise; by Solomon and Postman (1952) for words presented visually; and by Savin (1963), who attempted to delineate the factors responsible for the effect. According to Savin the word frequency effect is the ten-

dency for common words to be perceived correctly at lower speech-to-noise ratios than uncommon words.

Pollack, Rubenstein and Decker (1959), in an investigation of the intelligibility of known and unknown message sets, pointed out that when the test items are unknown an important determinant of the intelligibility of a word is the frequency of occurrence in relation to the frequencies of the words that might be confused with it. This study was also done with the stimulus material presented in noise.

In 1963 Pickett and Pollack demonstrated that intelligibility of a word in an excerpt of fluent speech depended upon the excerpted duration. That is, the longer the excerpted duration the greater is the intelligibility. The principle is that the greater the contextual information available to the listener, the higher the percent intelligibility.

In 1964 Pollack and Pickett suggested that the intelligibility might have increased also because longer excerpts "provided a clearer auditory picture of the talkers communication act" (p.79). Here they were talking about "auditory context" which is made up of prosodic cues, such as voice pitch and stress. They designed an experiment which showed that intel-

ligibility of the initial word of an utterance is improved as more and more successive words of the utterance are heard even though every word but the initial word was known in advance by the listener.

There are context effects in production and in perception of speech. The studies reviewed thus far were confined mainly to measures of intelligibility as a function of amount of sentence context, masking of the stimuli with noise and the effects of word-frequency.

Many other investigators have pointed out the influence of context on various aspects of the overall perception process (Abrams and Bever, 1969; Smith and Goodenough, 1971; Triesman and Tuxworth, 1974; Triesman, 1968). Three studies that are particularly relevant to the current investigation are those of Lieberman (1963) Miller and Isard (1963) and Kalikow, Stevens, and Elliott (1977). These studies focused on the specific effects of sentence context on both the perception and production of speech. In the first Lieberman (1963) attempted to demonstrate that "the semantic and grammatical contexts of a sentence affect both the production and perception of speech in a consistent and complementary fashion" (p. 175). Lieber-

man investigated the possible changes in production in different contexts, by using the concept of redundancy. He defined redundant words as those in common maxims such as the word "nine" in "a stitch in time saves nine". A non-redundant counterpart would be "nine" in "please say the word nine".

Both redundant and non-redundant words were excised from sentences and presented in isolation to groups of listeners. These listening tests were made at a signal-to-noise ratio of +4 db using broad band masking noise. The non-redundant words were found to be more intelligible than the redundant words. For instance, in the non-redundant condition the average intelligibility score for all words was 59 percent correct. For the redundant conditions the score was 24 percent correct. Under a 30 db S/N almost all words regardless of redundancy were equally intelligible to the listener. It appears, according to Lieberman, that a speaker may alter the production of a word based on the context of the word in the sentence.

Although Lieberman measured the duration and the relative peak amplitude of the excised words his analysis of acoustic-cue changes in the redundant and

non-redundant words was incomplete in that differences in the fundamental frequency contours between the two conditions were not analyzed. Questions such as what changes the acoustic stimulus undergoes in various grammatical and semantic speaking conditions still remain unanswered.

Recently a test of everyday speech reception was developed (Kalikow, Stevens, and Elliott, 1977) which included the use of linguistic information. Basically the test items consisted of sentences presented in noise. The listener was required to fill in the last word of the sentence. These words were all monosyllabic nouns. In fifty percent of the cases the final word was highly predictable on the basis of the context of the sentence. In the other 50 percent the word was of low predictability. For normal hearing listeners the intelligibility score was significantly higher for the high predictability items, (i.e. 79% versus 42%) at almost all signal to noise ratios tested. There may be a confounding in the Kalikow et.al. study due to the possibility that a speaker may produce words more carefully and perhaps more intelligibly in an ambiguous context. If this were the case than the difference in intelligibility between high

and low predictable endings would be even greater if this confounding were removed.

If contextual information and redundancy influence both perception and production, then it might be true that if the sentence context is somehow made anomalous, (not conforming to linguistic rules of English) the perception should be hampered by an amount corresponding in some way to the degree of the anomaly unless acoustic cues are reinforced in an extraordinary way. Miller and Isard investigated this possibility using a perception and production paradigm. They systematically violated certain linguistic rules and tested subjects production and perception under these conditions. The stimuli that were used were ungrammatical strings of words, grammatical sentences which obey both semantic and syntactic rules, and anomalous sentences which were grammatically but not semantically correct. The subjects were asked to shadow (i.e. repeat out loud immediately) what the speaker was saying. They counted as correct a sentence in which all content words were shadowed correctly.

The sentence score for totally grammatical sentences was 89.4%. For anomalous sentences the score

was 81.1 percent. For ungrammatical strings of words the score was 61.1 percent. Individual word scores were much higher in all three conditions. The general conclusion was that linguistic rules of grammatical and semantic nature have measurable effects on our ability to hear sentences.

They employed a reading rate of approximately two words per second, a rate quite a bit slower than a normal speaking rate. This rate might have affected performance in a number of ways. For instance, slowing the rate might have changed the articulatory parameters somewhat in production and therefore the listeners might have received a greater amount of acoustic information yielding higher scores than would have been the case with normal speaking rates. Lastly, that the use of shadowing as the task variable may not be representative of natural speech perception. When shadowing perhaps the listener does not have to comprehend the content of the message. In addition the act of shadowing (speaking while listening) introduces a confounding variable.

One of the few studies on the effect of context on perception of internal open juncture is that of O'Connor and Tooley (1964). They used internal open

juncture phrases and measured how often the communication process fails when the only difference in phrases were junctural. Readers recorded junctural phrases in meaningful sentences. Unlike the Lehiste 1960 study, in which the phrases were spoken in isolation, in this study the phrases were produced in context. The phrases were then excised from the sentences and presented, in isolation, to groups of listeners using a two-alternative forced-choice technique. In this situation subjects misidentified approximately 33% of the phrases. In Lehiste's study, subjects were very good at identifying the internal open juncture phrases, misidentifying less than 15% of the phrases. It appears that if the phrases are not spoken in isolation (but are heard in isolation) they are not as easy to recognize. Although they did not systematically vary meaningfulness and phonological probability they did perform a post hoc analysis of their phrases and the corresponding percent correct identifications. This analysis led them to conclude that semantic and phonological factors might be important for the perception of internal open juncture.

Rationale and outline of experiments

The literature reviewed shows that there are substantial differences between word initial and word final consonants in internal open juncture phrases. Although it has been established that sentence context influences the production and perception of various types of speech stimuli, the effects of such context on internal open juncture phrases have not been systematically examined. The members of the phrase pairs used to study internal open juncture are, of necessity acoustically similar. It has therefore been assumed by investigators of internal open juncture that specification of the acoustic boundary cues would be sufficient to understand the perceptual differentiation of these phrases. However, since the boundary cues of the phrase pair may be affected by whether or not the phrase is in isolation, the premise of the previous investigations may be faulty. While the boundary cues of these isolated phrases may in fact be sufficient to distinguish between the member of the pairs it is also possible that when these phrases are spoken in context the cues themselves may be altered or diminished. If this is the case, then it could be

postulated that the phrases would become more difficult to correctly identify. In addition, it might also be true that the listener, under these conditions, uses the sentence context to determine which phrase was spoken. It is important to consider the sentence context in light of the fact that most phrases in a natural speech setting are not heard in isolation. It was the purpose of this research to investigate the role of sentence context on both the perception and production of internal open juncture.

Three experiments were performed. The first experiment was designed to determine whether there are differences in identifiability between internal open juncture phrases produced in isolation and produced in context but heard in isolation. Differences in the acoustic characteristics of the phrases between the two conditions were examined.

The second experiment was designed to investigate the effect of sentence context on the perception of internal open juncture. This experiment employed grammatically and semantically anomalous sentences to demonstrate the role of context. Acoustical analyses of the stimuli in correct and anomalous sentences were performed in order to determine if the same phrase was produced the same way in different contexts.

The third experiment examined specific hypotheses, developed from the results of the first two experiments, regarding the relative importance of contextual and acoustic cues in the perception of internal open juncture phrases.

CHAPTER II

EXPERIMENT I: EFFECTS OF CONTEXT ON INTELLIGIBILITY

This experiment was designed to determine whether there are differences in intelligibility between internal open juncture phrases produced in isolation and the same phrases excised out of meaningful sentences. It was hypothesized that subjects hearing internal open juncture phrases produced and presented in isolation would show higher identifiability scores than those who heard the same phrases excised out of the sentences and presented in isolation. It was further hypothesized that underlying these intelligibility differences there would be differences in the acoustic characteristics of the phrases spoken in different conditions.

Test materials were prepared by having a male reader read lists of phrases as well as phrases embedded in context. Three pairs of phrases, all with stop consonants at the juncture, were used. The experiment consisted of two parts. In part one listeners judged phrases that were spoken in isolation and in part two phrases that were spoken in sentences,

excised from the sentences and presented in isolation. In addition all the phrases were rated for "commonness" by a separate group of listeners in order to check for possible biases resulting from one member of a pair being more unusual or uncommon than the other.

The acoustic measures used to analyse the phrases were those demonstrated, in prior research, to be relevant to the definition of word and syllable boundary. Segment duration, fundamental frequency and intensity were measured. In addition, characteristics such as presence or absence of glottal stops, aspiration, and laryngealization were also examined.

Method

One male reader, approximately 45 years of age, generated the stimuli for the experiment. The reader was highly experienced at oral reading and was not aware of the purpose of the experiment. He was seated in an IAC sound booth and a high quality microphone was placed six inches from his mouth. He was given a list of thirty sentences to read and was instructed to read them in as natural a manner as he could. When he finished he was given a list of thirty isolated inter-

nal open juncture phrases. He was instructed to read these in as natural a manner as possible. Both readings were recorded on audio tape using a high quality Ampex tape recorder. The levels were monitored on a VU meter by the experimenter.

The experiment consisted of two parts. For part one, the phrases, recorded in list form, were rerecorded in a random order with all phrases repeated four times. Of the 15 phrase pairs recorded, three were selected for the experiment. The phrase pairs chosen were: "Grey day" - "Grade A"; "grey tie" - "great eye" and "plum pie" - "plump eye". The sentence contexts are reproduced in Table 1. The phrases always occupied a medial position in the sentence and the preceding and following word was the same for each pair member. These pairs were selected because previous researchers have shown that internal open juncture pairs with stop consonants (both voiced and unvoiced) at the juncture were quite easy for listeners to identify correctly. It was decided that three phrase pairs would be sufficient to demonstrate the differences between isolated phrases and phrases spoken in sentence context. The other twelve phrase pairs were included as buffers. That is, they were

used so the listeners would not establish a "set" for responding. Responses to these buffer phrases were not used in the data analysis.

For part two of experiment I, the phrases spoken in the sentences were excised from the sentences using an end point detecting program on a digital computer (Rabiner & Sambur, 1975). These segmented phrases were then rerecorded in random order (each appearing four times). The three experimental phrase pairs were the same as those in part one.

Forty college students including both males and females, served as the listeners in this experiment. All subjects were unaware as to the purpose of the experiment. They were native speakers of English and all reported that they were normal hearing and had no history of hearing problems. Sixteen participated in the first part of the experiment and twenty-four different subjects participated in the second part. The difference in number of subjects between the two parts of the experiment was due to the relative availability of subjects at the time of testing.

In both parts of the experiment the listeners were placed in an acoustically treated classroom (the

Table 1

Sentence contexts from which phrases were excised for
the experimental phrases of Experiment I

1. The egg was number one grade A in quality.
2. It was one grey day in town today.
3. He had a great eye for the ladies.
4. He wore a grey tie for three days straight.
5. She was so heavy that she even had plump eyes and fat lips.
6. The restaurant had plum pies and blueberry pies.

room was used as a language laboratory and was acoustically isolated by acoustical tile on the walls and the ceiling). The listener was seated in the center of the room facing a KLH loudspeaker from a distance of 15 feet. The frequency response of the loudspeaker was flat within

4 db from 30-8000 Hz, was used. The sound pressure level measured at the listener's ear was 75 db SPL. All listeners reported that they could hear the stimuli clearly.

In both parts the listeners were instructed to circle the phrases they heard in a two alternative forced choice paradigm. The two alternatives were the members of the phrase pair, i.e., "grey tie" or "great eye". The percentage of correct responses for each test phrase was tabulated.

In addition to the experiment a test was performed in which fifty college students (none of whom served as subjects) were asked to rate all the phrases in terms of a dimension labelled "commonness". The rating was on a scale of 1 to 5 where 1 was most common and 5 was most uncommon. The results of this rating were examined in relation to the listening test results obtained.

The percent correct for each phrase was tabulated. The percent correct for each part was compared using a t-test.

In addition overall and component durations were measured in the isolated and excised conditions. Fundamental frequency contours and intensity measures were made. In addition, spectrographic analyses of the phrases were also conducted.

Results

Perceptual measurements: Table 2 shows the percentage of correct identifications for both parts of experiment I. The overall identification score for the excised phrases was 81 percent correct, and for the isolated phrases 97 percent correct. A t-test was performed between the percent correct for the isolated and for the excised phrases. The difference was statistically significant [$t(38) = 5.46$ $p < .01$]. Table 2 shows that most of the errors occurred for those excised phrases in which the junctural consonant was in the word final position. For instance "plump eye" was recognized as "plum pie". For these phrases the listeners were correct only 65 percent of the time on average.

It should be noted that the results between listeners were highly consistent. For the isolated stimuli the mean number of correct responses was 23.31 with a standard deviation of 1.01. For the excised phrases the mean number of correct responses was 19.38 with a standard deviation of 2.76. The standard deviation for isolated phrases is very small due to

Table 2

Percent correct per phrase in both parts of Experiment I

<u>Word initial consonant</u>	<u>Excised</u>	<u>Isolated</u>	<u>Word final consonant</u>	<u>Excised</u>	<u>Isolated</u>
Plum Pie	100	98	Plump Eye	72	95
Grey Day	93	99	Grade A	68	95
Grey Tie	95	99	Great Eye	55	96
Average	96	99	Average	65	95

Table 3

Mean ratings by 50 listeners of "commonness" of
phrases used in Experiment I

<u>Phrase</u>	<u>Mean rating</u>	<u>Phrase</u>	<u>Mean rating</u>
Plum Pie	4.5	Plump Eye	4.9
Grey Day	2.3	Grade A	1.6
Grey Tie	3.0	Great Eye	3.2

1.0 - most common

5.0 - least common

the degree of difficulty of the task. The phrases were so discriminable that there was almost perfect performance. For the excised phrases the performance was somewhat poorer and so greater variability in response was observed. The fact that the standard deviations are both so small may also be attributed to the forced choice nature of the listening task.

Table 3 shows the judgments of "commonness" for each of the phrases. Each row of the table corresponds to a different pair. The columns identify the phrases and the mean ratings of "commonness" obtained for that phrase. The difference between members of a pair are not large enough or consistent enough to account for the differences found in perceptual recognition. The only substantial difference in commonness is between "grey day" and "grade A". "Grade A" was judged more common and yet there were far more errors for this phrase than for "grey day". There were, however, considerable differences between phrase pairs. For instance, the phrase "grade A" had a mean commonness judgment of 1.6 and for the phrase "plump eye" the mean commonness judgment was 4.9.

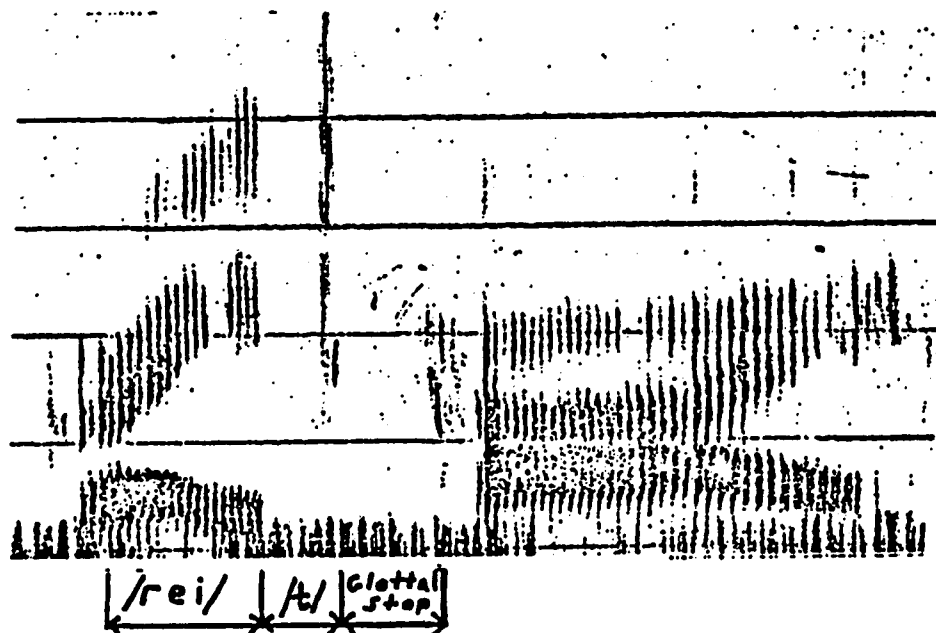
Acoustical measurements: Although specific hypotheses were not made with regard to acoustic cues, it was expected that both types of phrases would show similar acoustic characteristics and that the word initial consonant phrases would have different acoustic information from the word final consonant phrases. With these expectations in mind various acoustic measures were made.

The following duration measures were made from wide band spectrograms of the phrases in both conditions: 1) total duration of phrase; 2) duration of phoneme preceding junctural consonant; and 3) duration of junctural consonant. The junctural consonant was always a plosive, in which the duration was measured from the beginning of the closure to the release burst. The aspiration during the devoiced segment was not included. This definition was used for both the word initial and word final condition. The duration of the phoneme preceding the junctural phoneme was measured from the onset of regular glottal vibration of the transition to the beginning of the closure of the junctural consonant. For "m" in "plum pie-plump eye" the nasal resonances as seen on the spectrograms constituted the measured duration of that phoneme.

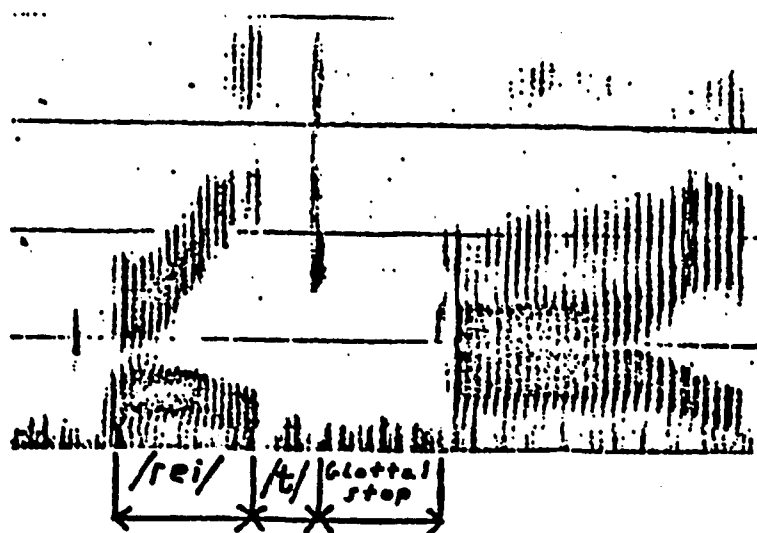
Because the segmentation of /r/ from the vowel in two of the pairs is ambiguous the measurement is specified as /rei/. Figure 1 is an example of the consonant and vowel measurements. The fact that the beginning of the /rei/ measurement was taken at the same point for both the isolated and excised phrases insures comparability between conditions.

Table 4 shows the total duration, in msec., of the utterance in the word initial consonant condition and the word final consonant condition for both the isolated version and the excised version. The isolated versions were found to be of greater duration than the phrases produced in context. It also may be seen that in both the isolated and excised versions, word initial consonant phrases are longer than their word final counterparts.

Table 5 shows the measured durations of the phoneme preceding the junctural consonant. The duration of this phoneme was found to be consistently longer in the word initial consonant condition. The mean duration in the word initial consonant condition in isolation is 266msec while in the word final condition it is 146msec. The phoneme duration is also consistently shorter in the excised version in every



Isolated



Excised

Fig. 1. Spectrograms showing how duration was measured in the phrase "Great Eye".

Table 4

Total duration of utterance in milliseconds

<u>Word initial consonant</u>	<u>Isolated</u>	<u>Excised</u>	<u>Word final consonant</u>	<u>Isolated</u>	<u>Excised</u>
Grey Day	870	630	Grade A	800	450
Grey Tie	880	700	Great Eye	830	640
Plum Pie	960	720	Plump Eye	870	620

Table 5

Duration of phoneme preceding the junctural consonant

<u>Word initial consonant</u>	<u>Isolated</u>	<u>Excised</u>	<u>Word final consonant</u>	<u>Isolated</u>	<u>Excised</u>
Grey Day	300	240	Grade A	220	200
Grey Tie	300	210	Great Eye	180	150
Plum Pie	200	125	Plump Eyes	40	40
Average	266	192	Average	146	130

Table 6

Duration in milliseconds of junctural consonant

<u>Word initial consonant</u>	<u>Isolated</u>	<u>Excised</u>	<u>Word final consonant</u>	<u>Isolated</u>	<u>Excised</u>
Grey Day	105	100	Grade A	90	45
Grey Tie	135	90	Great Eye	65	55
Plum Pie	75	45	Plump Eye	75	40
Average	105	78	Average	77	47

case. The mean duration for the excised version (across both word initial and word final conditions) is 161msec and for the isolated version it is 206msec.

Table 6 shows the measured durations of the junctural consonants. The junctural consonants, the stops /d/ /t/ and /p/ , are almost all longer in the initial position. The mean duration of the word initial consonants is 105ms. while for the word final it is 77ms. The junctural consonants are also longer in the isolated condition than in the excised condition.

Table 7 shows the percentage of the total phrase duration taken up by the prejunctural phoneme for both the isolated and the excised versions. It may be seen that the relative duration of the prejunctural phoneme is not very different for the two versions. The difference exceeds 4 percent in only the phrase "grade A".

Table 8 shows the proportion of the total phrase duration taken up by the junctural consonant for the excised and the isolated versions. In none of the six phrases is the relative duration very different for the two versions. The largest difference observed is 4 percent for the "grey day" phrase.

Table 7

Percentage of total duration occupied by pre-junctural phoneme

<u>Word initial consonant</u>	<u>Isolated</u>	<u>Excised</u>	<u>Word final consonant</u>	<u>Isolated</u>	<u>Excised</u>
Grey Day	34	38	Grade A	28	44
Grey Tie	34	30	Great Eye	22	23
Plum Pie	21	17	Plump Eye	04	05

Table 8

Percentage of total duration occupied by junctural phoneme

<u>Word initial consonant</u>	<u>Isolated</u>	<u>Excised</u>	<u>Word final consonant</u>	<u>Isolated</u>	<u>Excised</u>
Grey Day	12	16	Grade A	11	10
Grey Tie	15	13	Great Eye	08	09
Plum Pie	10	10	Plump Eye	09	08

Tables 7 and 8 show that the decrease in the duration of the phrase when it is produced in context is spread roughly evenly over the segments making up that phrase. Therefore the duration ratio change from the isolated to the excised versions cannot explain the decrease in percent correct between the two versions.

The fundamental frequency contours (Fo) of the phrases were obtained using an Fo - extraction algorithm designed by Rabiner and Gold (1975). Also obtained were short term intensity contours (I-contour). The plots were generated on a Data General Eclipse s/200 computer. The Fo and I- contours for each phrase pair are shown in Figures 2-4. Each figure corresponds to a different phrase pair. Both the Fo and I - contours were arranged with both members of a phrase pair set aligned in time at the onset of the voicing of the first vowel and placed on the same graph. This allowed visual comparisons of how much the two phrases differed from each other in the isolated and the excised conditions.

The upper left graph in Figure 2 shows the "plum pie - plump eye" pair in isolation. The upper right in Figure 2 shows the Fo contour for the same phrase

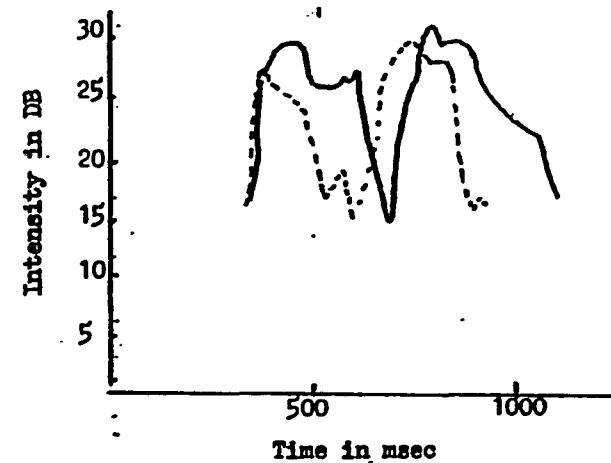
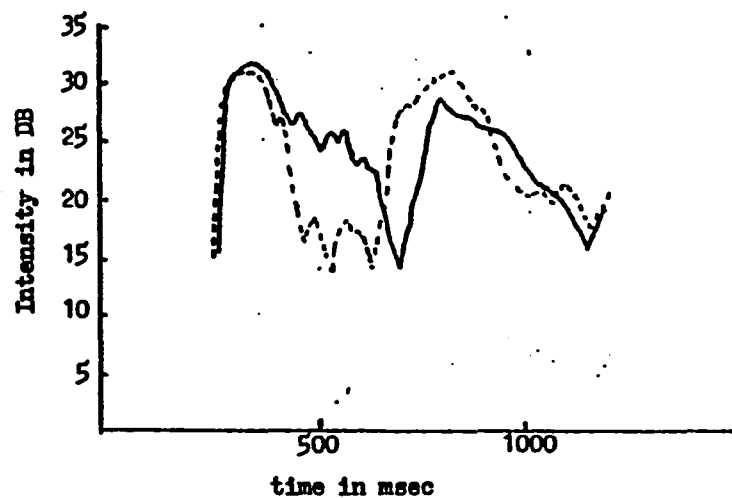
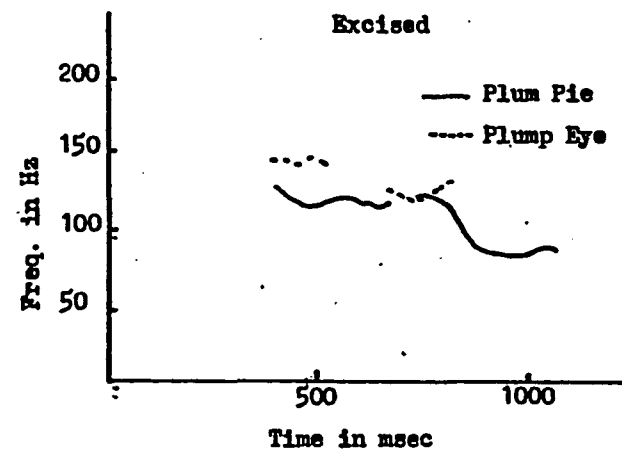
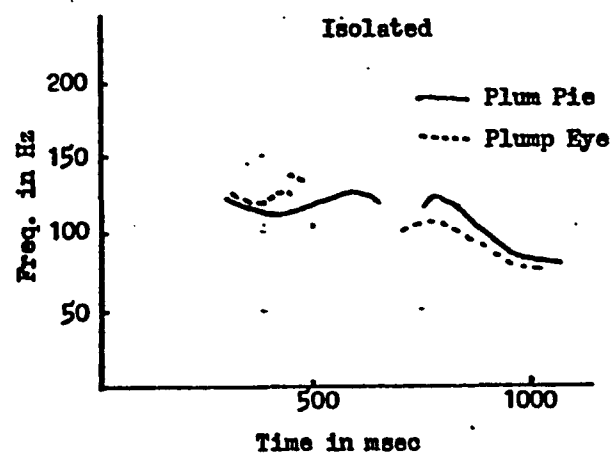


Fig. 2. F_0 and intensity plots for juncture pairs in isolated and excised conditions.

pair excised from a meaningful sentence. It should be noted that all of the perceptual errors for the "plum pie - plump eye" pair occurred for the "plump eye" phrase. Simple observation of Figure 2 shows that the form of the Fo contour is different for the two members of this phrase pair. In the isolated version of "plump eye", immediately after the break in the contour the Fo rises and then gradually decreases. This rise is not observed in the excised version of the phrase. For the "plum pie" phrase both the isolated and the excised versions show a rise in Fo immediately after the break in the contour at the vowel portion of "pie".

Figure 3 shows the Fo and I-contours for the phrase pair "grey day - grade A". The Fo contours for the phrase pair in isolation look quite similar. However, the Fo contours for the pair excised from the sentence are quite different from one another. The excised "grade A" phrase is virtually flat except for a deep dip before the word boundary and has only a very short blank portion which is the glottal stop. The flatness of the contour is a reflection of the sentence context that surrounded it as compared to the isolated version which is rich in Fo variation from

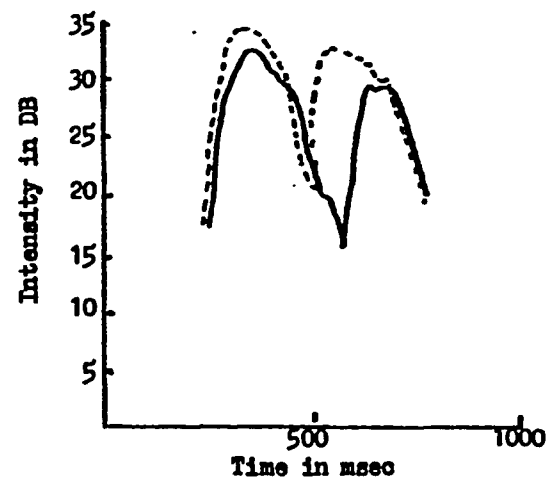
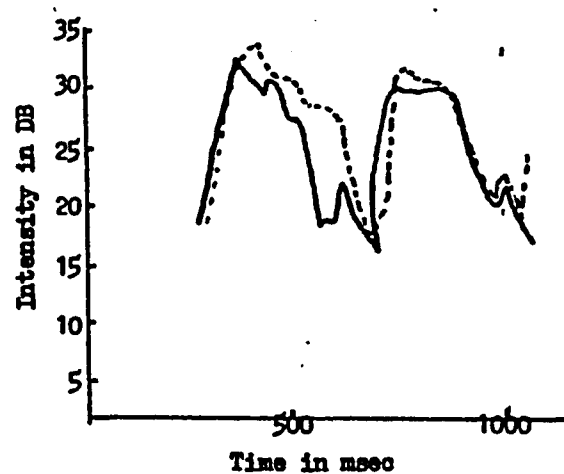
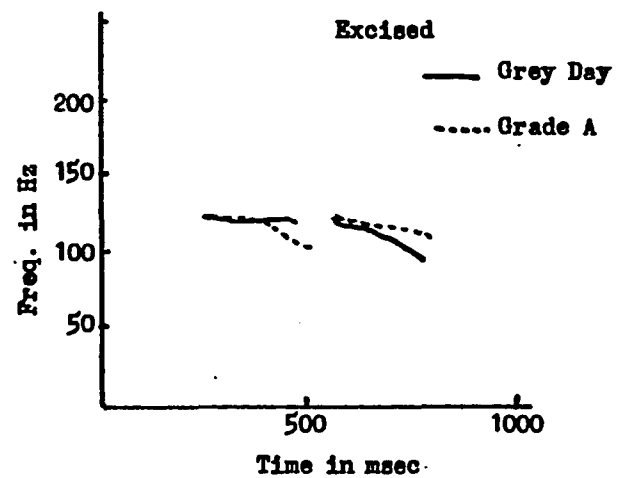
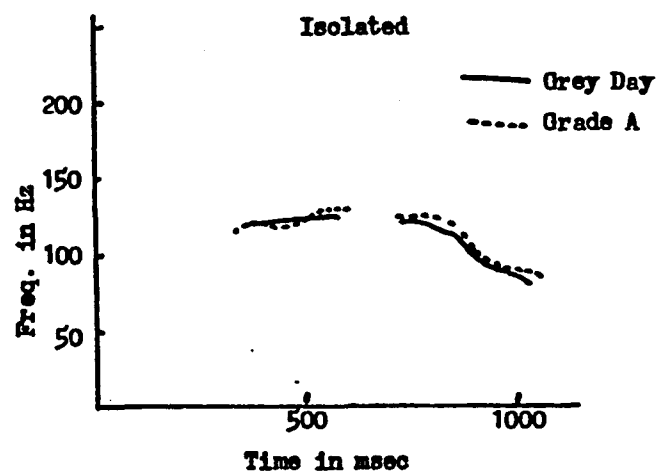


Fig. 3. F_0 and intensity plots for juncture pairs in isolated and excised conditions.

beginning to end. Other than the flatness of contour and the duration of the glottal stop there appears to be no difference between the contours. That is, the vowel /ei/ begins at its maximum frequency in both cases and the point just before the stop in both curves is at the same frequency as the beginning of the following vowel.

Figure 4 shows the F_0 and I - contours for the "grey tie - great eye" phrase pair. For the "grey tie" phrase the point before the closure of the stop is considerably lower in frequency than the beginning frequency of the following vowel. This is true in both the isolated and the excised versions, although the frequency difference is much greater in the isolated version. For the "great eye" phrase the frequency after the release is equal to the frequency before the closure for the isolated version and slightly lower in the excised version. Approaching the closure, both versions of "great eye" decrease in frequency at the same rate. After the release, however, the situation changes. For the isolated version the vowel begins at its maximum frequency and remains there before decreasing.

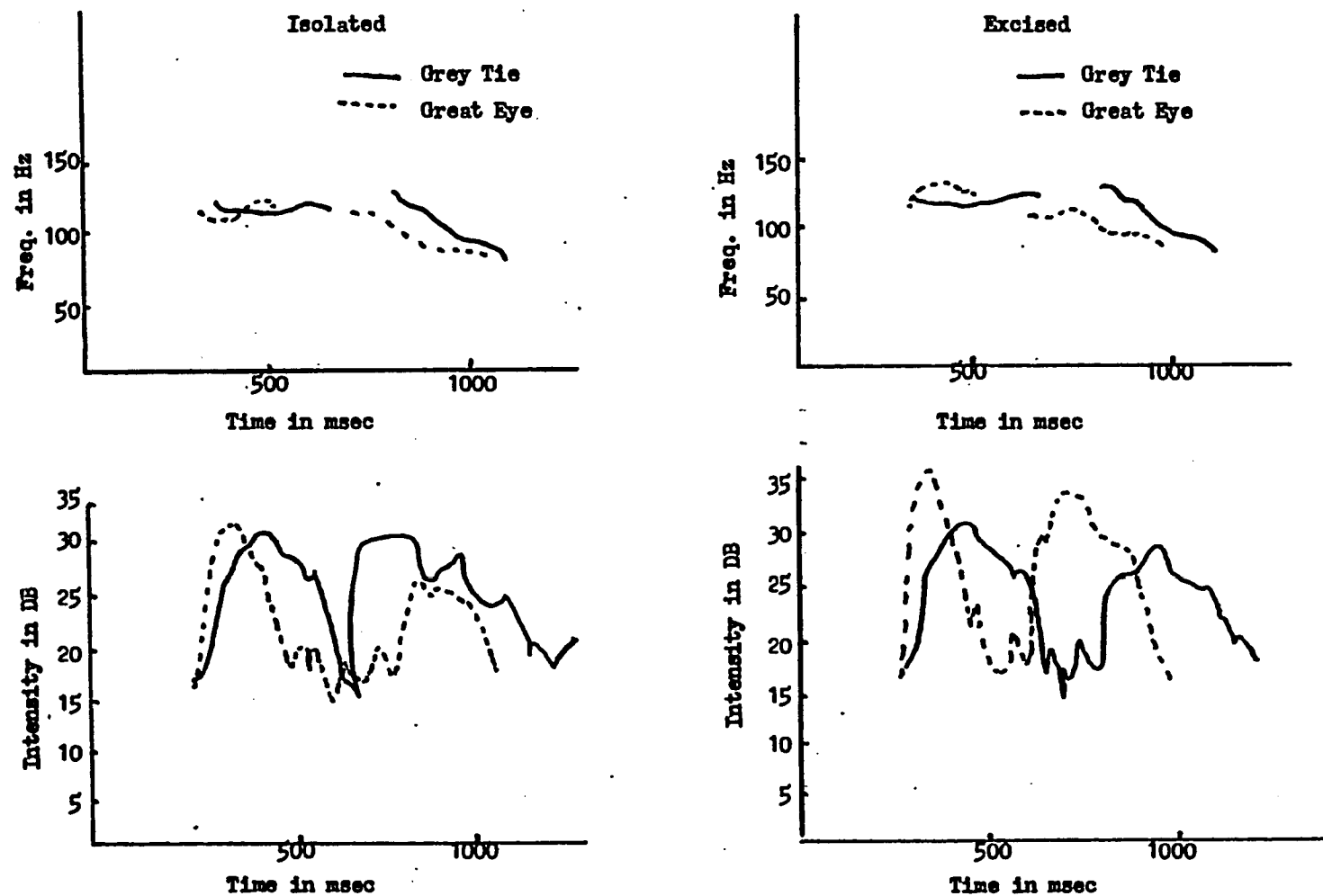


Fig. 4. F₀ and intensity plots for juncture pairs in isolated and excised conditions.

The I-contours in the lower portions of Figures 2-4 correspond to the Fo plots above them in the time dimension. These plots were useful in helping to interpret the Fo plots. They clearly indicate a pattern of falling intensity before the juncture and maximal intensity on the vowel following the juncture (Garding, 1967). Although there are some specific form differences between the phrase pairs these are not consistent enough to differentiate the boundary conditions under investigation.

Although no specific hypotheses were made regarding differences in Fo contour between the word initial consonant and word final consonant conditions a number of characteristics were examined. These characteristics are tabulated in Tables 9-11. Table 9 shows the only consistent difference gleaned from the Fo contours. This table shows the relative rate of fall of Fo after the release of the consonant for all pairs in both conditions. Relative rate of fall is determined by measuring the slope of the Fo contour during the 100 msec from the release of the junctural consonant. For all phrases the rate of fall is greater in the word initial consonant condition. This is true for both the excised and isolated phrases.

Table 9

Relative rate of fall of fundamental frequency
after release of consonant

<u>Isolated</u>	<u>Excised</u>
Plum pie > Plump eye	Plum pie > Plump eye
Grey day > Grade A	Grey day > Grade A
Grey tie > Great eye	Grey tie > Great eye

Table 10

Fundamental frequency immediately before closure and after release
for all phrases in all conditions

<u>Phrase</u>	<u>Fo before closure</u>		<u>Phrase</u>	<u>Fo after release</u>	
	<u>Isolated</u>	<u>Excised</u>		<u>Isolated</u>	<u>Excised</u>
Plum Pie	125	125	Plum Pie	125	125
Plump Eye	140	140	Plump Eye	100	125
Grey Day	120	100	Grey Day	120	120
Grade A	120	120	Grade A	120	120
Grey Tie	110	115	Grey Tie	125	120
Great Eye	110	110	Great Eye	110	100

Table 11

Peak Frequency of each phrase before the closure and after release

<u>Phrase</u>	<u>Before closure</u>		<u>Phrase</u>	<u>After release</u>	
	<u>Isolated</u>	<u>Excised</u>		<u>Isolated</u>	<u>Excised</u>
Plum Pie	130	130	Plum Pie	125	125
Plump Eye	145	150	Plump Eye	110	135
Grey Day	125	120	Grey Day	120	115
Grade A	125	120	Grade A	120	115
Grey Tie	120	115	Grey Tie	130	120
Great Eye	125	125	Great Eye	115	105

Table 10 shows the F_0 immediately before closure and after release for each phrase in each condition. The relationships are not consistent between phrases. It is therefore necessary to examine each phrase pair individually. The F_0 before closure for "plum pie" is identical to the F_0 after release of the consonant. For "plump eye" the F_0 is lower after the release than it is at the closure. These relationships hold for both the isolated and excised conditions. For "grey day" the F_0 before closure is equal to that after release in the isolated condition but lower at the closure in the excised condition. In "grade A" the F_0 is the same both before closure and after release in both conditions. For "grey tie" there is an increase in F_0 after the release in both conditions. "Great eye" shows a drop only after release in the excised condition.

Table 11 shows the peak F_0 for each phrase both before closure and after release of the consonant. For all phrases in the isolated condition except "grey tie" there is a lower peak frequency after the release than before the closure. The "grey tie" phrase shows a higher peak F_0 after the release. The same relationships exist in the excised phrase condition.

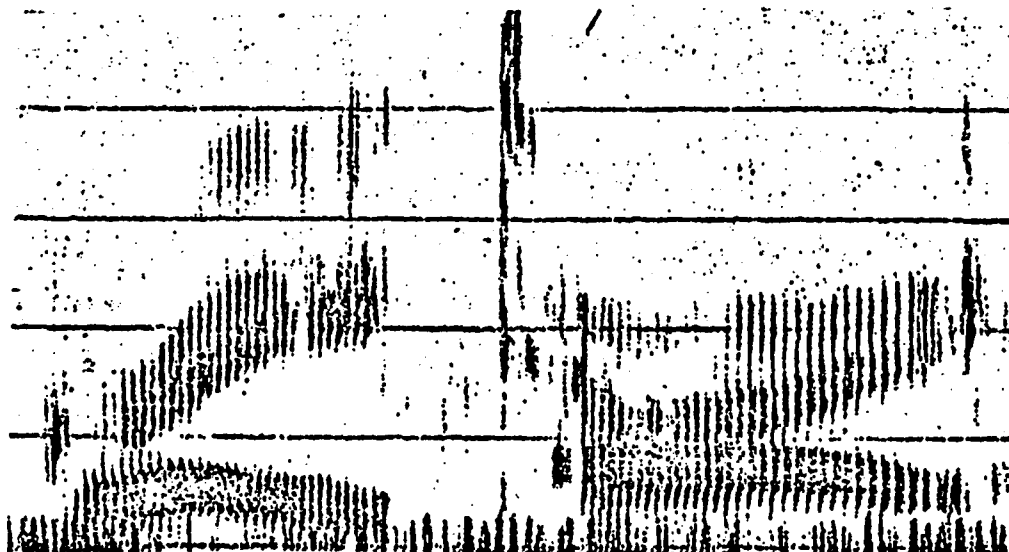
Glottal characteristics: Spectrographic analysis has been shown to provide information on vocal tract behavior. In light of the findings of Dukes and Nakatani (1978) it is necessary to examine the cues concerning glottal condition that may be associated with juncture perception. Figures 5-10 show the spectrograms for all the phrases in isolated and excised conditions. Looking at these figures it is obvious that the duration differences between the isolated and excised versions are considerable. These differences have been discussed earlier in this section. In addition, presence or absence of such conditions as glottal stops, aspiration, laryngealization provides information concerning word boundary. A glottal stop is considered present if between the release of the consonant and regular vocal cord vibration of the following vowel there a blank portion seen on the spectrogram. Aspiration is defined as a period of "noise" on the spectrogram immediately after the release of the consonant. In all cases in which a glottal stop is observed on a spectrogram a trained linguist agreed that a glottal stop could be heard.

Table 12 shows these glottal conditions and includes the corresponding listening task results for

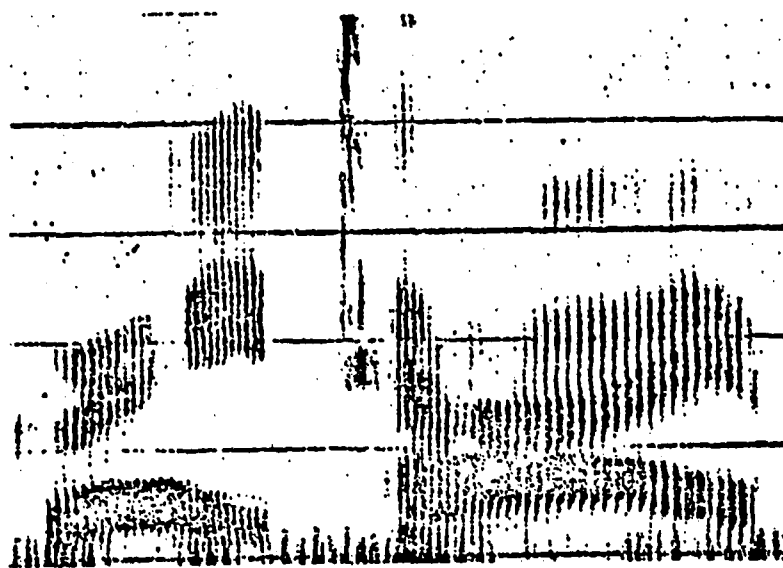
Table 12

Presence or absence of glottal stop and aspiration in junctural
portion of all phrases and corresponding percent correct
in listening task

	Isolated			Excised		
	stop	Asp.	% Corr.	stop	Asp.	% Corr.
Grey Tie	-	+	98	-	+	95
Great Eye	+	-	98	+	-	55
Grey Day	-	-	98	-	-	93
Grade A	+	-	100	-	-	68
Plum pie	-	+	100	-	+	100
Plump Eye	+	-	88	+	-	78

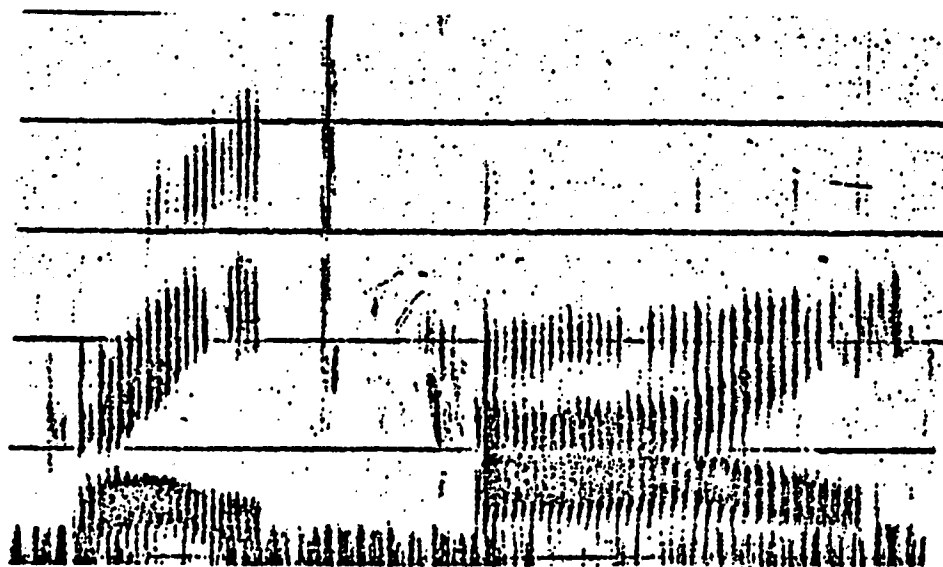


Isolated

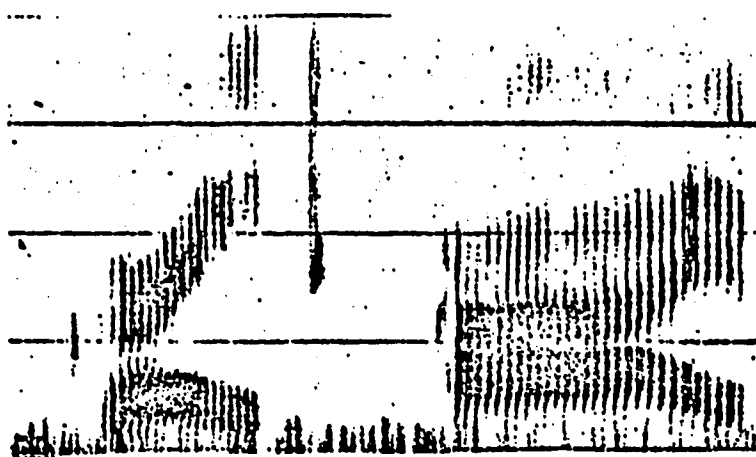


Excised

Fig. 5. Spectrograms of "Grey Tie" in isolated and excised conditions.

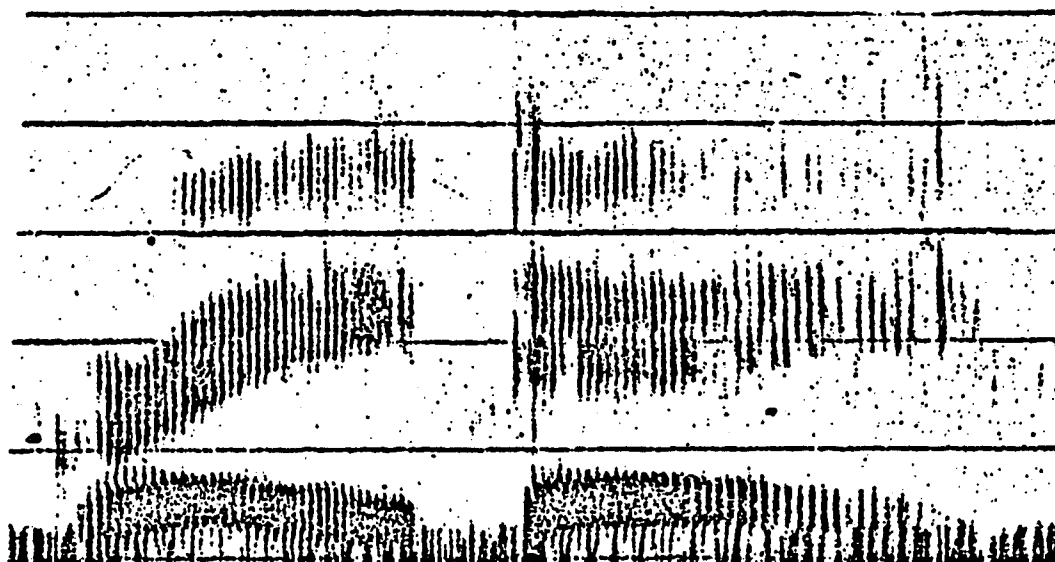


Isolated

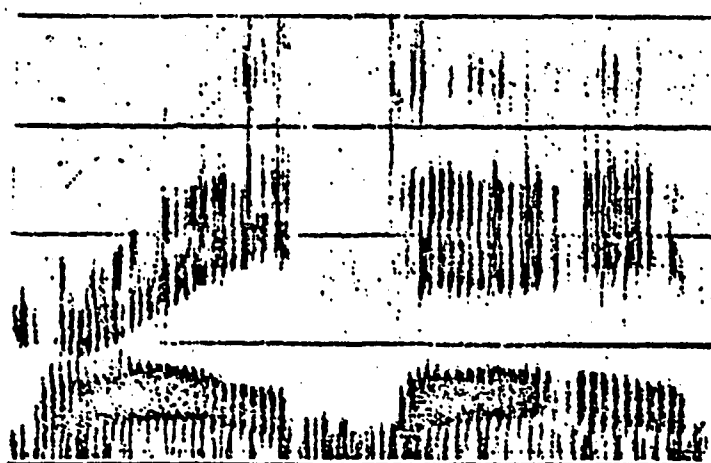


Excised

Fig. 6. Spectrograms of "Great Eye" in isolated and excised conditions.

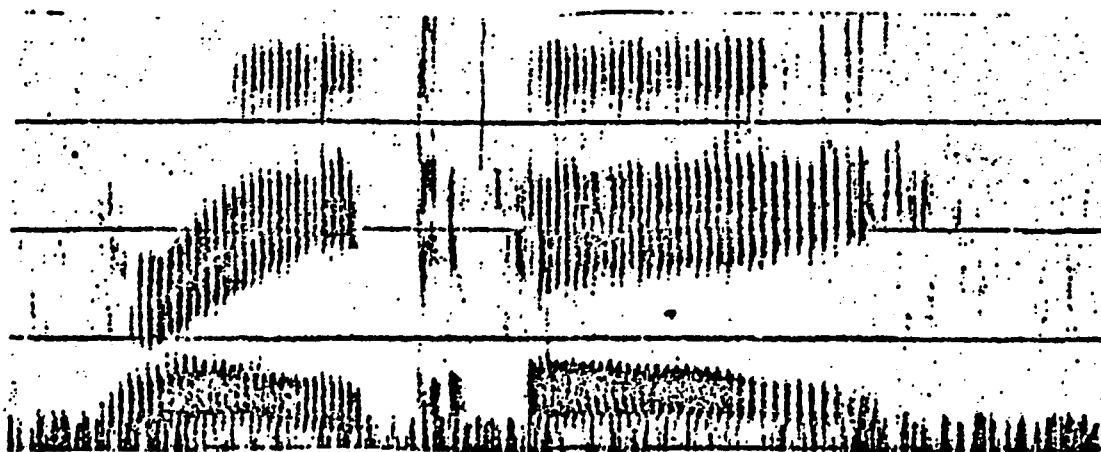


Isolated

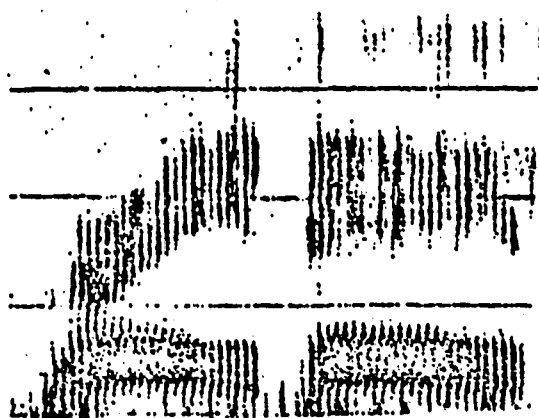


Excised

Fig. 7. Spectrograms of "Grey Day" in isolated and excised conditions.

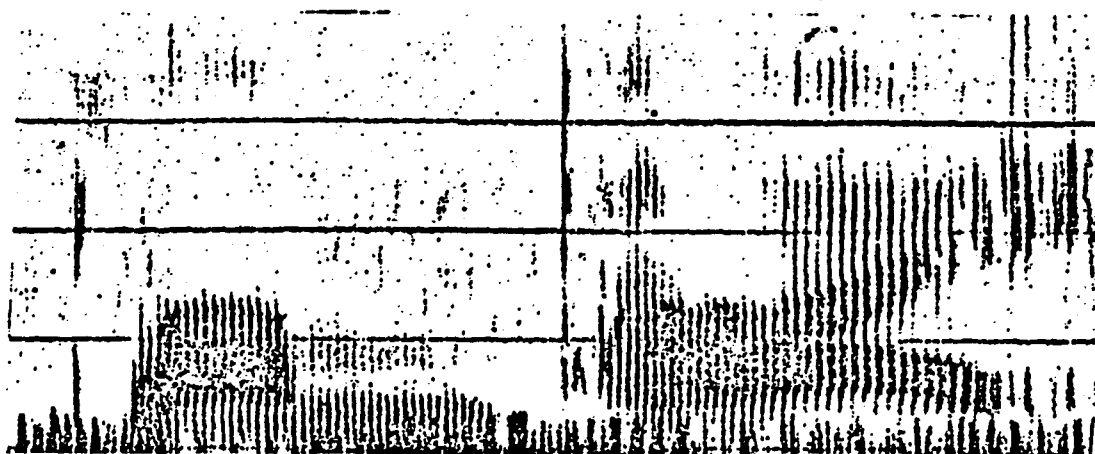


Isolated

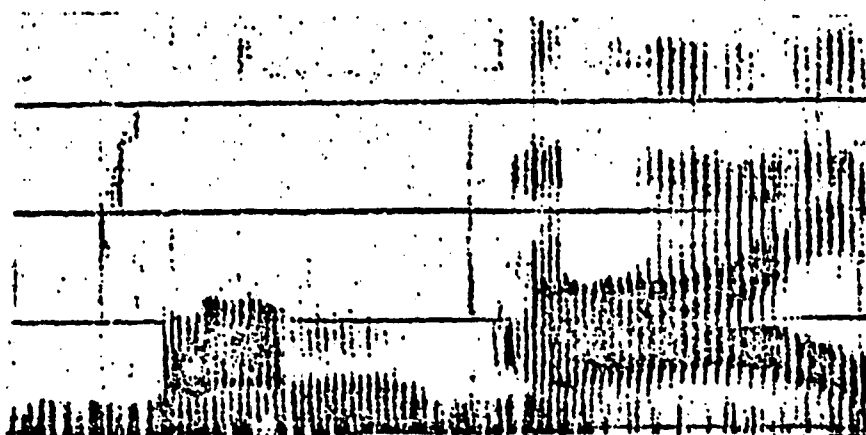


Excised

Fig. 8. Spectrograms of "Grade A" in isolated and excised conditions.

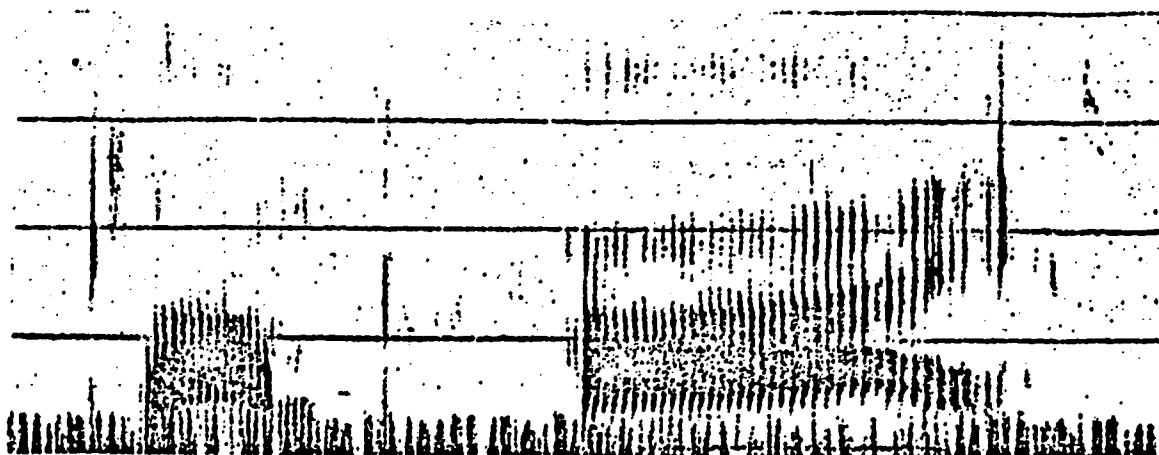


Isolated

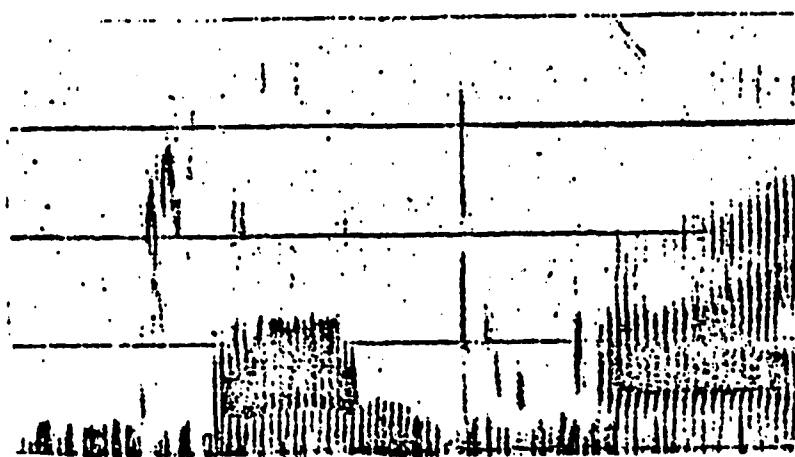


Excised

Fig. 9. Spectrograms of "Plum Pie" in isolated and excised conditions.



Isolated



Excised

Fig. 10. Spectrograms of "Plump Eye" in isolated and excised conditions.

each phrase. For the "grey tie - great eye" pair and the "plum pie - plump eye" pair the presence or absence of glottal stops and aspiration does not directly relate to the difference in correct identification of the isolated and excised conditions. The "grey day" - "grade A" pair presents rather different results and these are discussed below.

Looking at figure 5 it is clear that in the isolated version of grey tie the aspiration after the release is longer than it is in the excised version. Other than that there seem to be no other potential qualitative cue differences between the isolated and excised versions. It should be noted that the correct identification in the perception for "grey tie" (Table 2) was very high in both conditions. Looking at the spectrograms for "great eye" (Fig. 6) in the two conditions reveals no difference, other than the overall duration difference. Both versions show a clear glottal stop. The qualitative measures provide no clue as to what could have accounted for the difference in perception scores for the excised and the isolated versions of "great eye". The difference between "grey tie" and "great eye" is that the "great eye" has the glottal stop and the grey tie does not

but shows some aspiration. This is true for both versions.

Figure 7 shows the spectrograms for "grey day" in the isolated and excised conditions. The burst portion of the /d/ seems a good deal more intense in the isolated version and the closure period is longer in the isolated version. The spectrograms in Figure 8 are of "grade A" in both versions. There is laryngealization after the burst and a glottal stop in the isolated version. The excised version however, has no glottal stop and no laryngealization. Again, aside from the duration differences, these are the only disparities between the spectrograms. There was a high percentage of errors in recognition of the grade A phrase in the excised condition. The "grade A" spectrogram in the excised condition looks very similar to the "grey day" spectrogram in the excised condition. The lack of a glottal stop may account for the relatively poor performance in the recognition of the "grade A" phrase in the excised condition.

The spectrograms for "plum pie" are seen in Figure 9. aside from the durational differences the only difference seems to be that the burst and aspira-

tion are more intense in the isolated version. Figure 10 shows the spectrograms for "plump eye". The excised version is a bit difficult to interpret. There is probably a glottal stop after the burst. In the isolated version the glottal stop is very clear.

The results of this study clearly indicate that there are differences in perception and production of stop consonant internal open juncture phrases in isolated and excised conditions. The results also indicate that neither the relative duration ratios, the intensity contours nor the F_0 contour differences seem to affect the discrimination. However, it seems likely that the articulatory cues do affect the discrimination process. The assumption is that the articulation of the phrase (primarily the word boundary cues) is not as carefully controlled in contextual conditions as it is in isolation because the context carries much of the recognition burden.

Chapter III

EXPERIMENT II: ROLE OF SENTENCE CONTEXT IN PERCEPTION

Overview

Based on the results of experiment I it is clear that the production of internal open juncture phrases is different when the phrases are spoken in context as opposed to when they are spoken in isolation. Experiment II was performed to examine that contextual effect in more detail. The perception of internal open juncture phrases in sentences having varying degrees of linguistic rule violation was measured.

To accomplish this readers read sentences containing internal open juncture phrases. Some of the sentences were grammatically and semantically anomalous when the wrong pair member was read. Others were only semantically anomalous,

and still others were completely correct regardless of which pair member was used.

There were three parts to this experiment. In Part 1 it was hypothesized that when the sentences were semantically and grammatically anomalous that listeners would make a large number of errors on those

utterances that are not actual sentences. That is, they would substitute the pair member that made the utterance into a sentence, even though the acoustic stimulus provided opposing cues.

It was further hypothesized that in part 2 (in which the utterances were only semantically anomalous) the same sort of error pattern would occur, but to a lesser degree than in part one. Lastly it was hypothesized that when the sentences were semantically and syntactically correct with either member of a phrase pair, there would be a large number of errors due to the fact that the context offered no help in deciding which phrase was correct.

Method

In this experiment more phrase pairs were used than in Experiment I in order to demonstrate the contextual effects across a wider range of phonetic boundary conditions. That is, in addition to the stops, fricatives, nasals and liquids were all represented. The acoustic analysis here was performed to indicate that the boundary cues were present in the incorrect readings as well as in the correct readings. The same

acoustic measurements used in experiment I were made in this experiment.

In the first part of this experiment the utterances consisted of two types, the first of which were complete sentences containing internal open juncture phrases. The sentences of the second were identical to the first except that they contained the other member of the internal open juncture pair, thus rendering them semantically and grammatically anomalous. There were six pairs of utterances. Examples of the two types of utterance are: "The eggs were grade A extra large" and "the eggs were grey day extra large." In the second part of the experiment the utterances consisted of two types, the first of which was complete sentences containing internal open juncture phrases. The second was identical to the first except that they contained the other member of the internal open juncture pair thus rendering them semantically anomalous while grammatically correct. Examples of the two types of utterance are: "The zoo was full of great apes and other primates"; "the zoo was full of grey tapes and other primates." There were six sentences of each type.

In part 3 the stimuli also consisted of internal open juncture phrases of two types. In this case, however, both phrases were in grammatically and semantically correct sentences, "I asked for more ice but my waiter didn't hear me." "I asked for more rice but my waiter didn't hear me." There were six sentences of each type. A complete list of all sentences and sentence like utterances is shown in Table 13. This table also shows in which conditions each utterance was used.

Three male readers ranging in age from 40-60 generated the stimuli used in this experiment. All three could be regarded as professional readers because they were volunteers for the "reading for the blind" program. None of the readers was aware of the purpose of the experiment. Each reader read exactly the same material. The use of more than one reader allowed for random checks to be made with regard to effects of specific speaker characteristics. Each reader was placed in an IAC sound booth and a high quality microphone was placed six inches in front of his lips. Each reader was given lists of sentences and sentence-like utterances and instructed to read

All utterances used in Experiment II

Grammatically and
Semantically anomalous

1. He said that if ice cream he would kill me.
2. The eggs were grey day extra large.
3. He just didn't see Mabel to lift more than fifty pounds.
4. Did you seem able and Norman in the street.
5. If it keep sticking it will drive me crazy.
6. The bird flapped it swings in a most unusual manner.

Correct counterpart

- He said that if I scream he would kill me.
- The eggs were grade A extra large.
- He just didn't seem able to lift more than fifty pounds.
- Did you see Mabel and Norman in the street.
- If it keeps ticking it will drive me crazy.
- The bird flapped its wings in a most unusual manner.

Grammatically correct
Semantically anomalous

1. The zoo was full of grey tapes and other primates.
2. I hoped they had plump eye on the menu.
3. The chinese food came with more ice than we possibly could have eaten.
4. The workers went on strike so that the management's nude eel would not be implemented.
5. Being a bee feeder, he felt out of place in a dairy restaurant.
6. He wished he could seem able again because she was so beautiful.

Correct counterpart

- The zoo was full of great apes and other primates.
- I hoped they had plum pie on the menu.
- The chinese food came with more rice than we possibly could have eaten.
- The workers went on strike so that the management's new deal would not be implemented.
- Being a beef eater he felt out of place in a dairy restaurant.
- He wished he could see Mabel again because she was so beautiful.

Either correct

1. He saw a number of different kinds of grey tapes in the magazine.
2. He saw a number of different kinds of great apes in the magazine.
3. I asked for more rice but my waiter didn't hear me.
4. I asked for more ice but my waiter didn't hear me.
5. It was a unique experience to meet a nice man in this day and age.
6. It was a unique experience to meet an ice man in this day and age.
7. No one had ever seen this new deal before.
8. No one had ever seen this nude eel before.
9. No one knew that plump eyes really turned him on.
10. No one knew that plum pies really turned him on.
11. He was the only bee feeder in the tribe.
12. He was the only beef eater in the tribe.

exactly what was written in as natural a manner as possible. They were permitted to practice the utterances before the recording was made. Sentences with anomalous structure were read two weeks prior to the correct versions of the sentences. In addition, at a subsequent date the three readers read each phrase in isolation for comparison purposes. All lists of sentences and sentence-like utterances were recorded on audio tape using a high quality Ampex tape recorder (model # PR-10). The recording level was monitored on a VU meter by the experimenter. From the master tapes, three stimulus tapes were created, one for each part of the experiment. Each tape contained the recordings of all three speakers for that particular part of the experiment.

Seventy-five college freshmen, both male and female, served as listeners in this experiment. All were native English speakers with no known hearing difficulties (as determined by questionnaire). None of the listeners had phonetics training and none were aware of the purpose of the experiment. Twenty-five different listeners, randomly selected from the 75 total listeners participated in each part of the experiment.

Each group of 25 listeners was placed in the same classroom used in Experiment I. The tape was played on a high quality Ampex (model # Pr 10) tape recorder and through a KLH loudspeaker. The measured frequency response was 4db from 30-8000 Hz. The sound level measured at the ear of a listener in the center of the room was approximately 75 db SPL. All listeners reported that they could clearly hear the stimuli.

In each part, the listeners were instructed as follows: "You will hear a number of utterances on this tape. Simply write down exactly the words that you hear in each utterance. You will have 15 seconds between utterances to write your answers." The listeners were not told that some of the utterances were syntactically or semantically anomalous and that others were not.

A response was considered correct if the open juncture phrase spoken by the reader was written. If the other member of the phrase pair, or anything else was written the response was scored as incorrect. The percent correct judgments were tabulated for the "correct" and "incorrect" utterances for each part of the experiment.

In addition to the experiment a test was performed in which fifty college students (none of whom served as subjects) were asked to rate all phrases in terms of a dimension labelled "commonness". The rating was on a 1 to 5 scale where 1 was most common and 5 was most uncommon. The results of this rating were examined in relation to the listening test results obtained in both experiments. This was done to insure that neither phrase of a pair was more likely to be written simply due to the fact that it might be a more "common" phrase.

Results

Perceptual measurements: Results of all parts of Experiment II are shown in Figure 11. Histograms depicting the percent correct judgements by all subjects for all speakers are shown. The data were combined across speakers because there were no significant differences in listeners responses as a function of speaker differences. The histograms labelled I are from Part I. They show that the grammatically and semantically correct versions of the sentences are

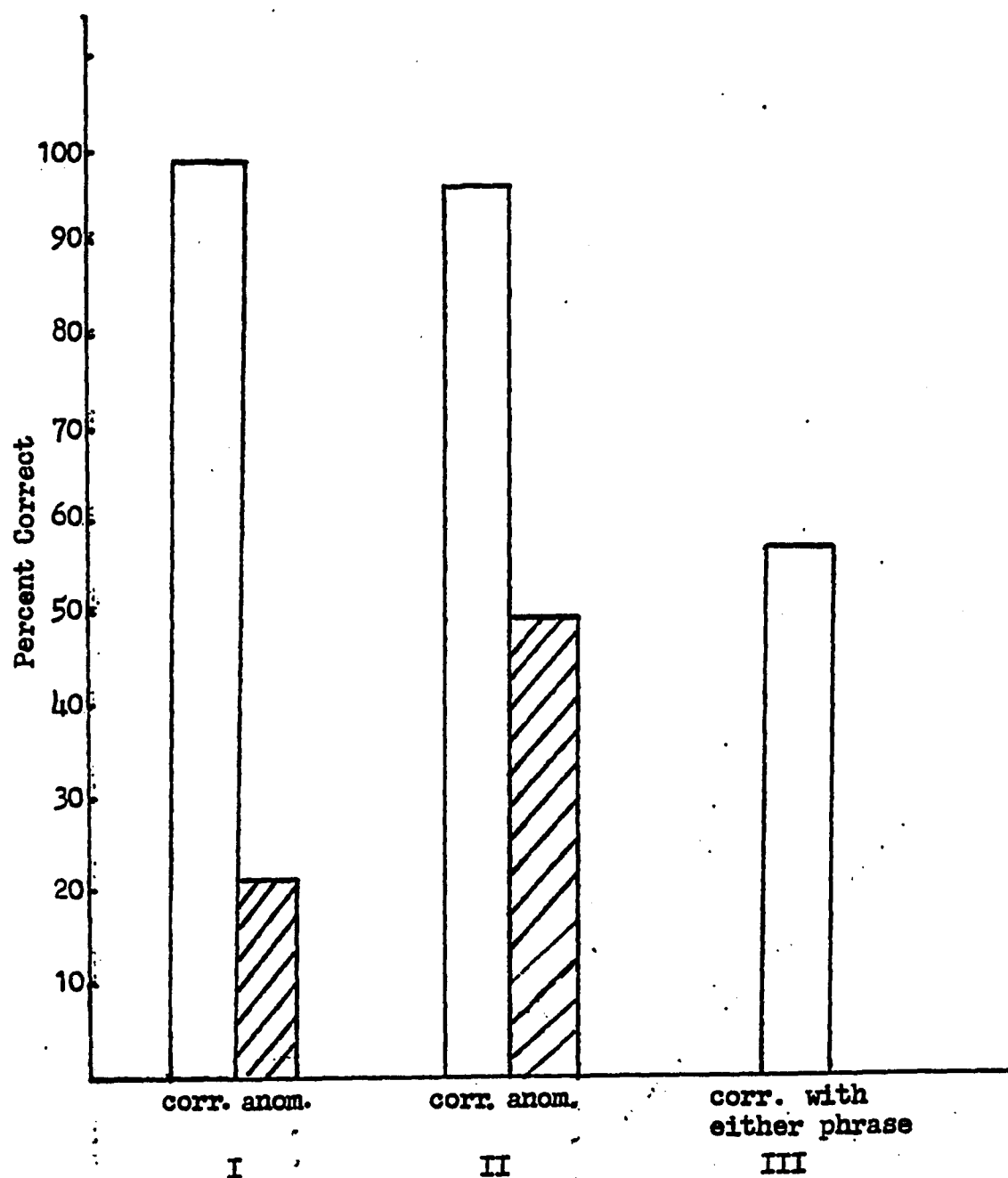


Fig. 11. Percent correct responses in all parts of Experiment II

reported accurately 97.5 % of the time (open bar). When the grammatically and semantically incorrect version of the sentence is presented (the one spoken with the contrasting phrase) the subjects reverse that version and report the one that makes sense. For example, if the spoken utterance was "He said that if ice cream he will kill me", the listener wrote "I scream". The rate of correct responses to the grammatically and semantically anomalous sentence is 18% (the hatched bar).

The individual differences between listener's responses were very small. The mean correct number of responses was 21.14 with a standard deviation of 2.14. Similarly small standard deviations were found for parts II and III of this experiment.

The results of Part 1 are shown in Table 14. The phrases on the left side of the table are those that were read in correct context. Those on the right were read in anomalous context. It is shown that the sentences with the correct phrase were all identified a great deal more often than those with the incorrect phrase. This was true for every phrase.

It should be noted that the responses scored as "incorrect" were always the other member of the phrase

Table 14

Percent correct for each phrase in Part I

<u>Sentences with correct phrase</u>		<u>Sentences with anomalous phrase</u>	
Keeps ticking.	97%	Keep sticking	46%
Its wings	100%	It swings	14%
I scream	98%	Ice cream	13%
Grade A	100%	Grey day	0%
Seem able	73%	See Mabel	21%
See Mabel	95%	Seem able	16%

pair. In only two cases did a listener make an error that was not one of the members of a phrase pair.

Table 15 shows the judgments of "commonness" for each of the phrases. Each row of the table corresponds to a different phrase pair. The columns of the table identify the phrases and the mean judgments of "commonness" obtained for that phrase. Comparing the percentages correct in Table 14 with the degree of "commonness" of the phrases shown in Table 15, it is clear that the bias associated with phrase familiarity could not account for the listening test result. In other words the phrase that is more "common" does not consistently yield higher percent correct judgments. The results for part 2 of the experiment (semantically anomalous sentences) are shown in histograms labelled II in Figure 11. Here, the correct sentences are judged correctly 95% of the time (clear bar). The semantically incorrect utterances were judged correctly only 45% of the time (hatched bar). That is, the subjects reversed the phrases to produce a meaningful sentence 55% of the time. The percent correct for all sentences in Part II is shown in Table 16. The results of Part III may be seen in the histogram labelled "III" in Figure 11 and in Table 17.

Table 15

Mean ratings of "commonness" of phrases used in Experiment II

<u>Phrase</u>	<u>Mean rating</u>	<u>Phrase</u>	<u>Mean rating</u>
See Mabel	2.8	Seem Able	2.3
Keep Sticking	2.7	Keeps Ticking	2.3
It Swings	2.2	Its Wings	2.6
I Scream	1.6	Ice Cream	1.0
More Rice	2.7	More Ice	1.5
Bee Feeder	4.2	Beef Eater	3.5
New Deal	2.2	Nude Eel	5.0
Grey Tapes	4.3	Great Apes	3.7
A Nice Man	1.5	An Ice Man	2.1

1.0 - most common

5.0 - least common

Table 16

Percent correct for each phrase in Part II

<u>Sentences with Correct phrase</u>		<u>Sentences with Anomalous phrase</u>	
More Rice	100	More Ice	84
Plum Pie	100	Plump Eye	23
Beef Eater	95	Bee Feeder	04
See Mabel	93	Seem Able	52
New Deal	99	Nude Eel	25
Great Ape	99	Grey Tape	80

Table 17

Percent correct for each phrase in Part III

<u>Sentences with Correct phrase</u>		<u>Sentences with Anomalous phrase</u>	
A Nice Man	94	An Ice Man	24
More Rice	100	More Ice	64
Plum Pies	100	Plump Eyes	22
Bee Feeder	04	Beef Eater	99
Grey Tape	94	Great Ape	78
New Deal	100	Nude Eel	10

In these sentences either member of the phrase pair set would have made correct English sentences. It was found that the listeners were incorrect in phrase choice quite often. For the sentences where the consonant was in the word-initial position there were 18% errors. When the consonant was word final the error rate was 50%. Although the percentages are different, it is clear that the pattern is the same as in Experiment I. That is, the greatest number of errors are in the sentences in which the reader spoke the word final consonant phrase.

Acoustic measurements: We examined a number of acoustic characteristics with the following two points in mind. First, comparisons were made between the word initial and word final consonant phrases in all conditions in order to determine what differences exist between the members of a phrase pair. Second, comparisons were made between the various productions of the same phrase, that is, the phrase spoken in correct and anomalous contexts. These were necessary in order to be reasonably sure that the reader read the phrase he was instructed to read in all conditions. The production analyses consisted of systematic comparisons of duration, fundamental fre-

quency contours, and glottal characteristics as measured from wide band spectrograms. Because the measurement of acoustic characteristics for the three speakers were very similar, the acoustic data for only one speaker is presented here. This particular speaker was chosen at random from the three. Comparison of the acoustic findings in Experiment II with those of Experiment I, where applicable, are made in the discussion section.

Glottal characteristics. Examination of spectrograms of the phrases in various conditions revealed a consistent similarity between the same phrase spoken in isolation, in a grammatically and semantically correct sentence and in utterance in which the grammar and/or semantics is inappropriate due to the internal open juncture phrase spoken. A number of characteristics were examined. Presence or absence of a glottal stop, presence or absence and duration of aspiration, strength and duration of a burst were some of the important characteristics examined. Table 18 shows, for a number of the phrases, these characteristics in the correct sentences, anomalous sentences, and in isolation. The glottal characteristics were the same across all three columns of Table 18.

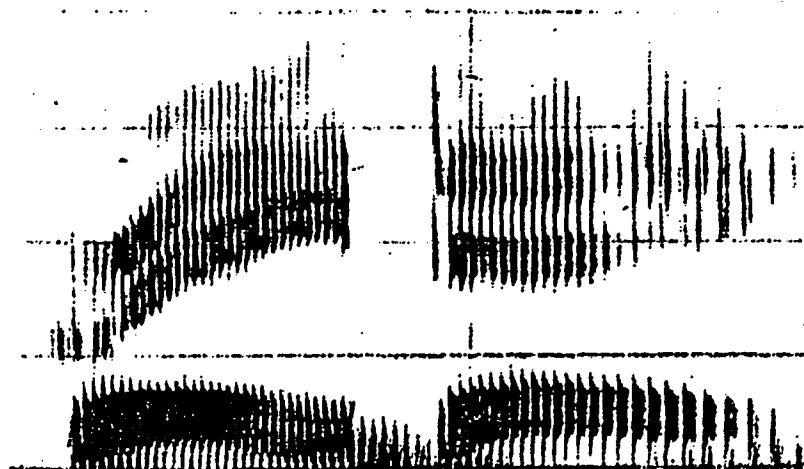
Table 18

Spectrographic characteristics of junctural consonant in phrases
in correct and in anomalous sentences and in isolation

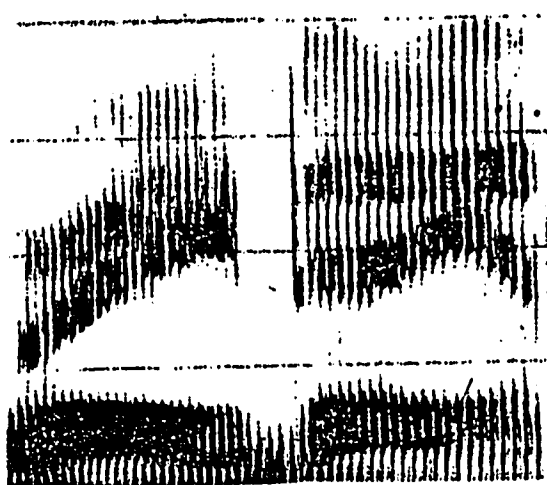
<u>Phrase</u>	<u>Characteristics in all three conditions</u>
Grey Tape	No glottal stop aspiration strong burst
Bee Feeder	Strong frication to high frequency
Grey Day	No glottal stop no aspiration
Keep Sticking	Short burst and aspiration
It Swings	/w/ is devoiced strong /s/
Plump Eyes	Glottal stop no aspiration
More Ice	High third formant mild glottalization
Nude Eel	Glottal stop no aspiration
Ice Cream	Long aspiration

Table 19Spectrographic characteristics of initial and final consonant phrases in correct sentences.

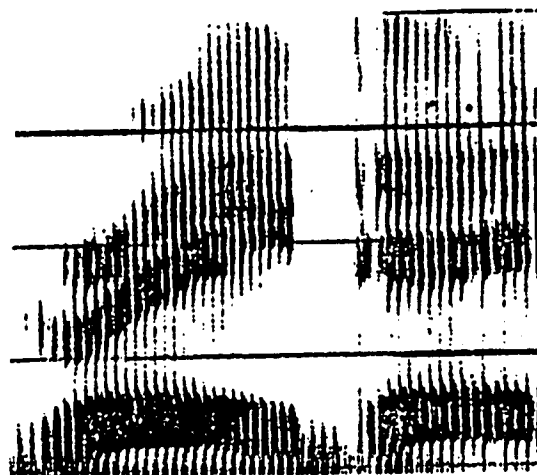
<u>Word initial consonant</u>	<u>Characteristics</u>	<u>Word final consonant</u>	<u>Characteristics</u>
grey day	No glottal stop No aspiration F_2 stays high before/d/	grade A	Glottal stop No aspiration F_2 goes down before/d/
grey tapes	No glottal stop, aspiration Strong burst	great apes	Glottal stop No aspiration No burst
plum pie	Burst Strong aspiration	plump eyes	Glottal stop No aspiration
bee feeder	Strong frication To high frequency	beef eater	Preceding vowel Frication, frication Not as high as bee-feeder
it swings	/w/ is devoiced, Strong /s/	its wings	/w/ is not devoiced
new deal	Burst No aspiration No glottal stop	nude eel	Glottal stop
keep sticking	Burst and aspiration Both short	keeps ticking	Strong burst and Long aspiration
more rice	Low third formant	more ice	High third formant Mild glottalization
I scream	No aspiration No glottal stop	ice cream	Long aspiration No glottal stop



Isolated

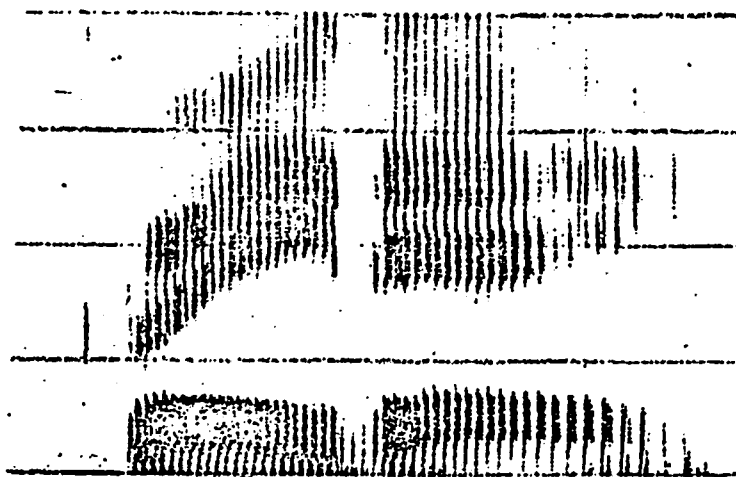


Correct Sentence

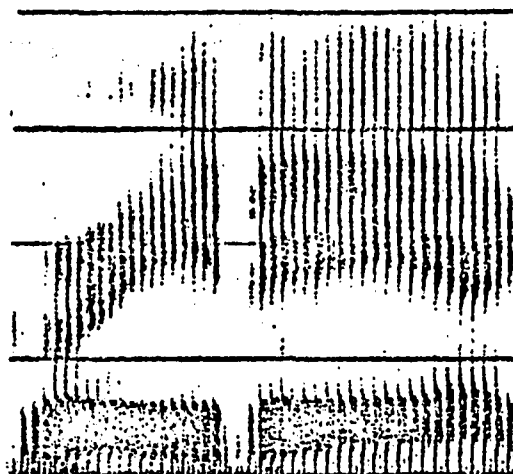


Anomalous Sentence

Fig. 12. Spectrograms of the phrase "Grey Day" in three conditions

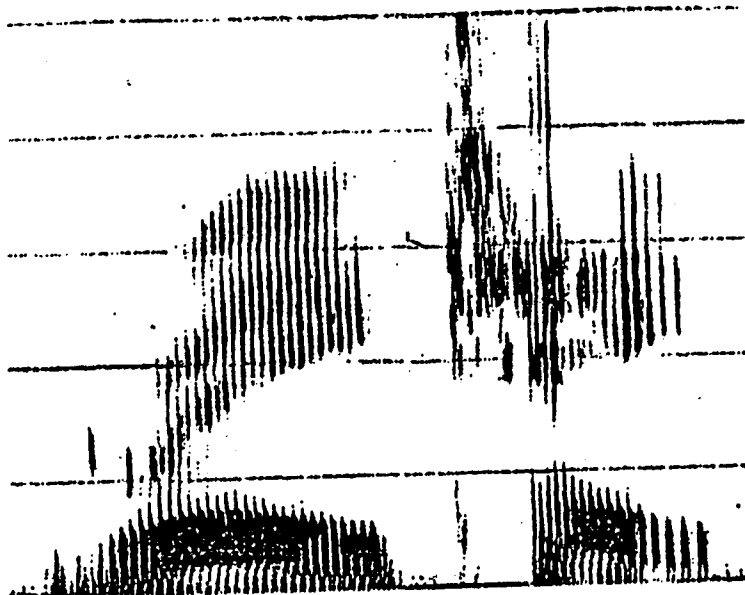


Isolated

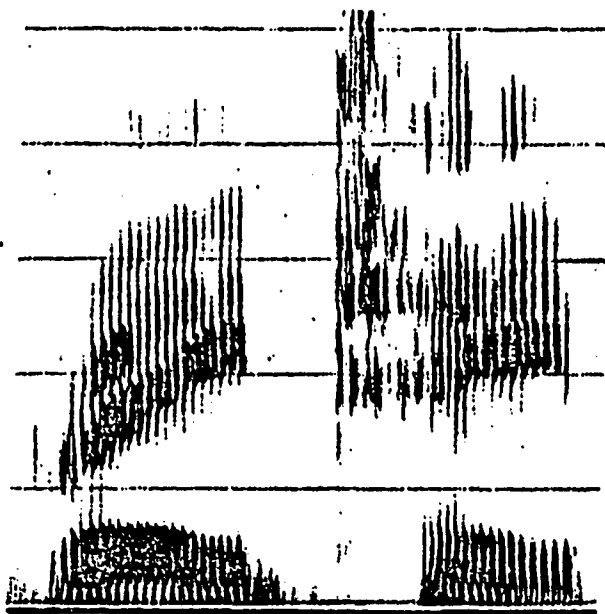


Correct Sentence

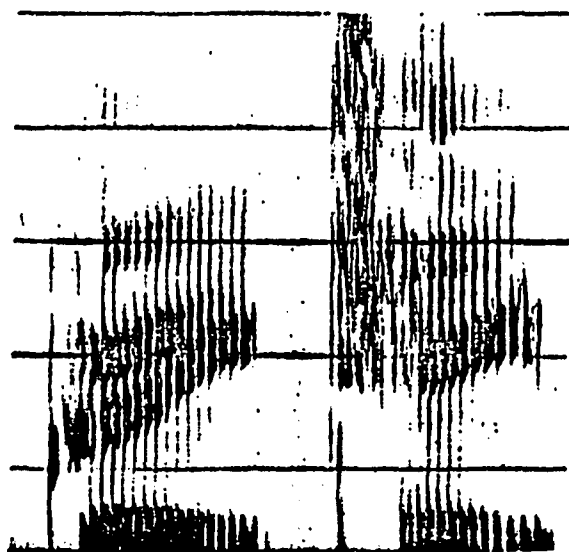
Fig. 13. Spectrograms of the phrase "Grade A" in two conditions.



Isolated

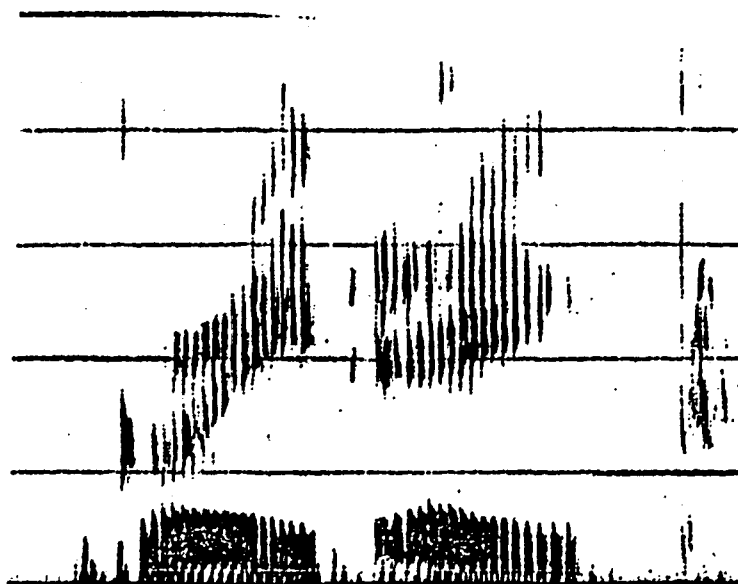


Correct Sentence

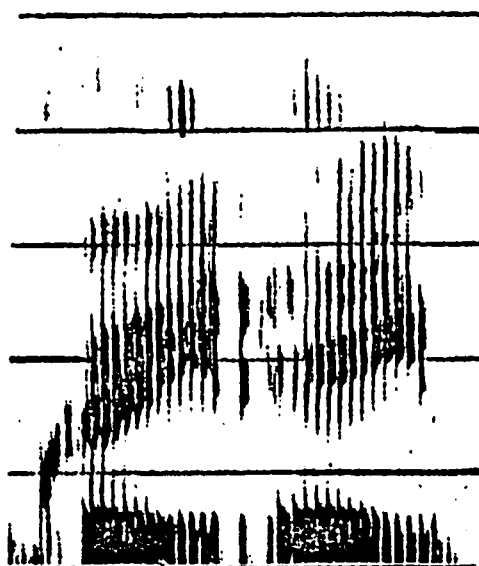


Anomalous Sentence

Fig. 14. Spectrograms of the phrase "Grey Tape" in three conditions.



Isolated



Correct Sentence

Fig. 15. Spectrograms of the phrase "Great Ape" in two conditions.

For example, the phrase "grey tape" has no glottal stop, a strong burst and aspiration in all three conditions of the table. Figures 12-15 are examples of spectrograms of some of the representative phrases.

Table 19 shows spectrographic characteristics of both members of phrase pairs in correct sentences. This table shows the word initial consonant phrases and their characteristics on the left side and the word - final consonant phrases on the right side. Differences between word-initial and word final phrases were observed for every phrase examined. For instance, looking at the second row of Table 19, the phrase "grey tape" has no glottal stop but there is some aspiration. Its counterpart "great ape", has a glottal stop and no aspiration. Figures 14 and 15 show the spectrograms of this phrase pair.

Comparing spectrograms in this manner shows that the glottal characteristics, at least, are consistent within a phrase across conditions (correct, anomalous) and that the two members of a phrase pair are different as far as these characteristics are concerned.

Fundamental frequency. Figures 16 through 18 show the F_0 contours for three typical phrase pairs.

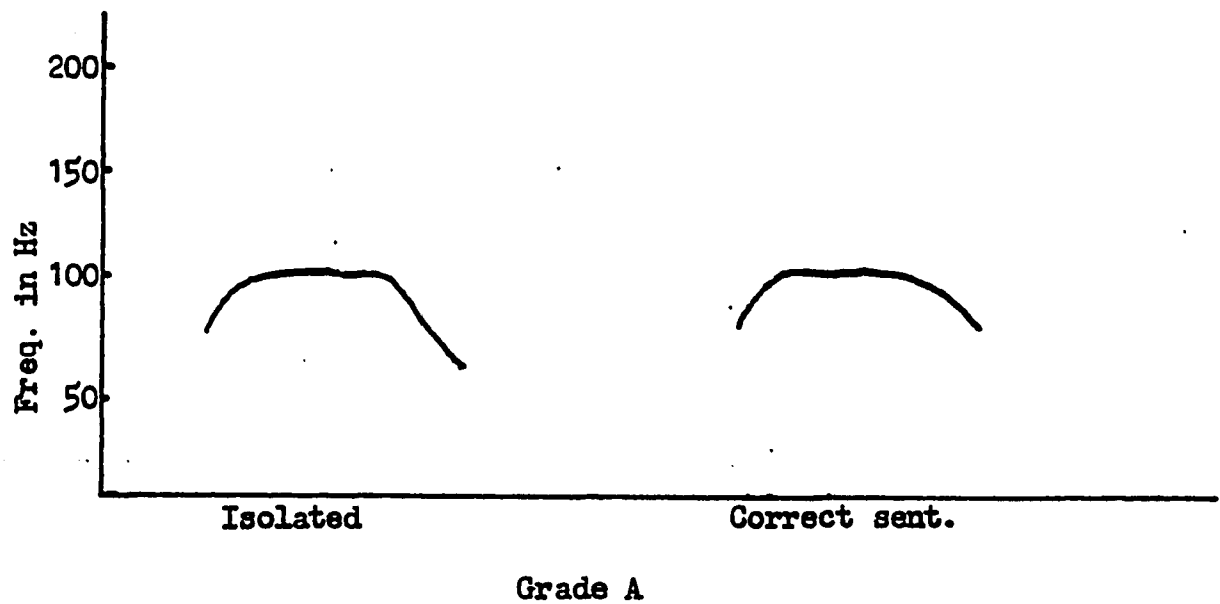
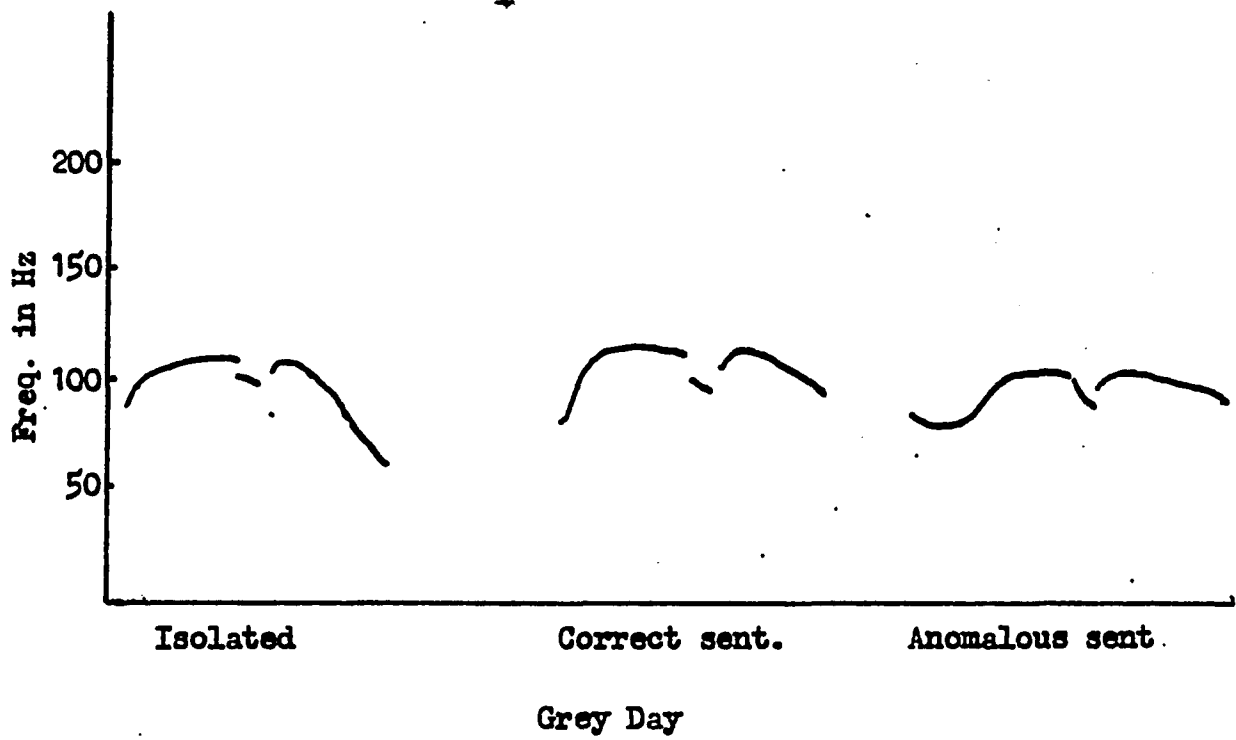


Fig. 16. F₀ contours of the phrases "Grey Day" and "Grade A" in a number of conditions.

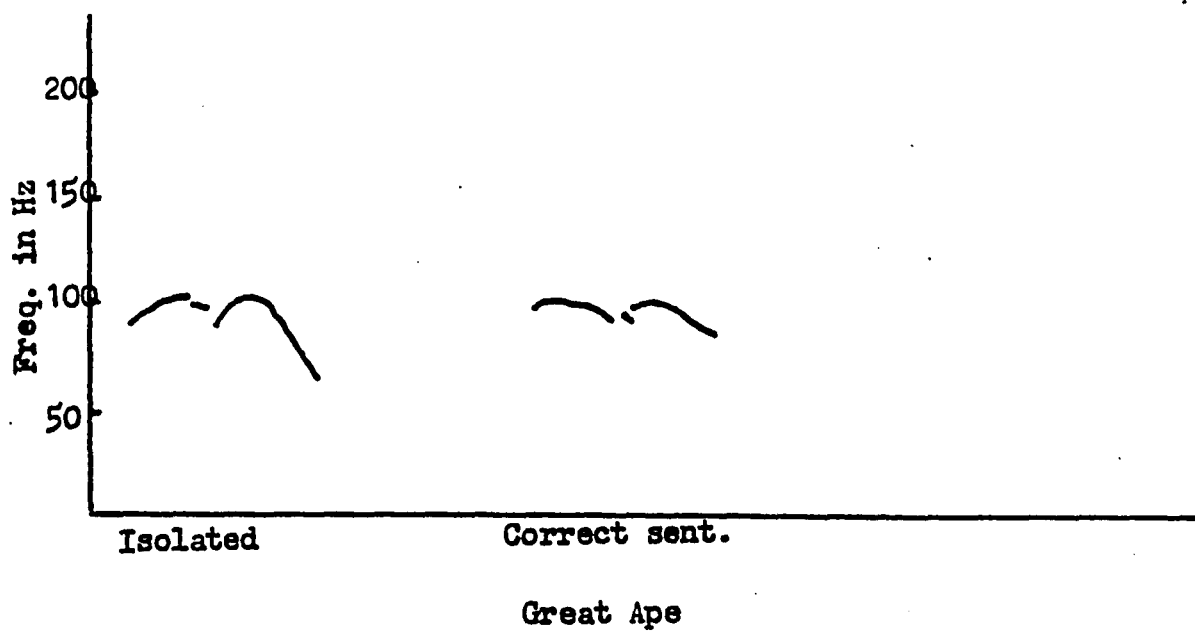
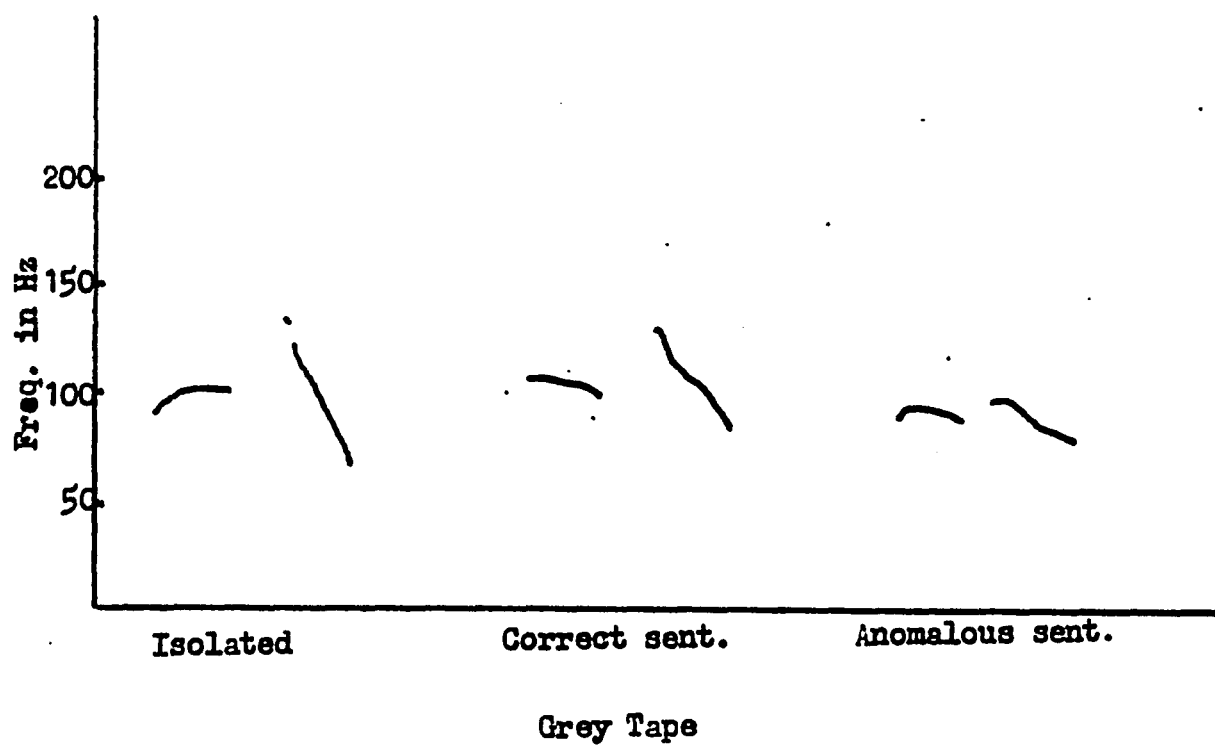


Figure 17. F0 contours of the phrases "Grey Tape" and "Great Ape" in a number of conditions.

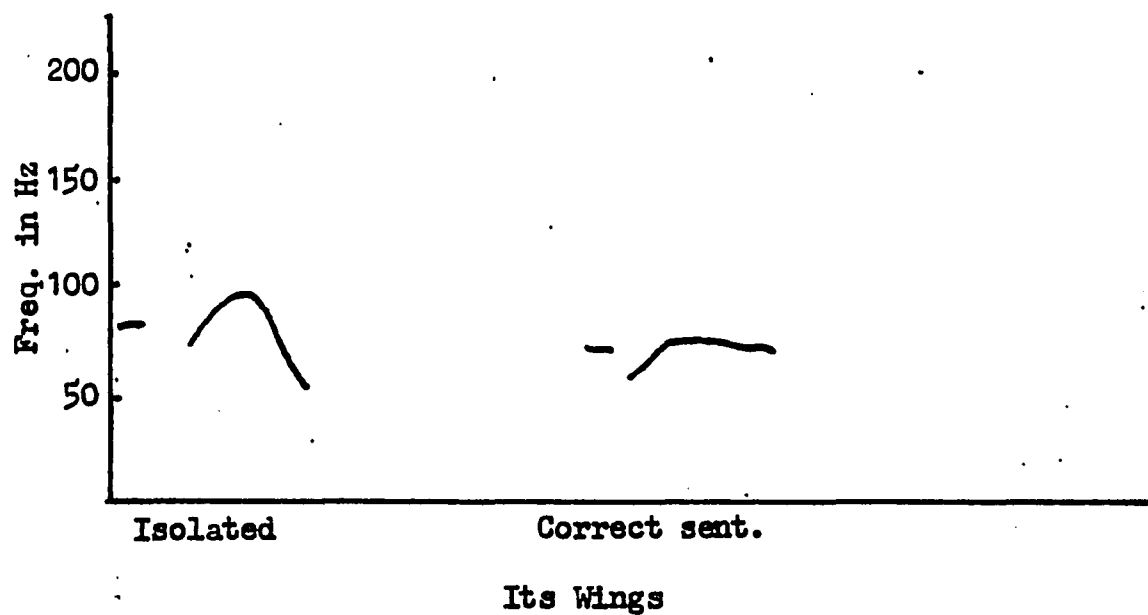
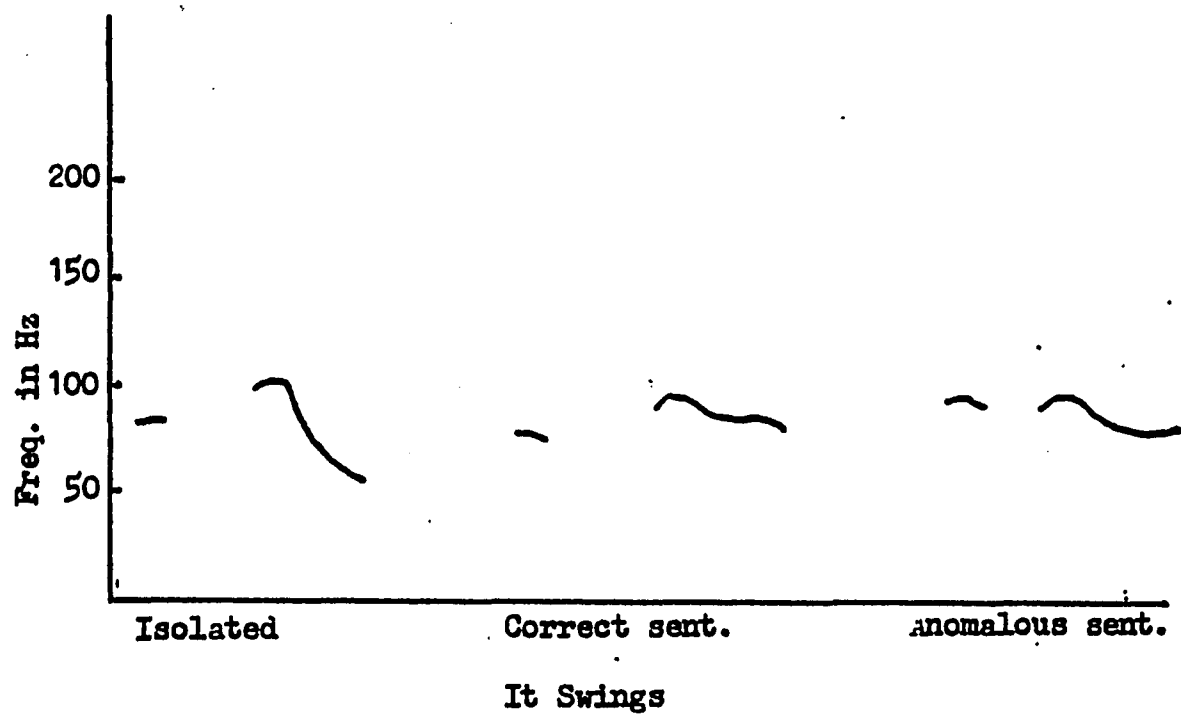


Fig. 18. F₀ contours of the phrases "It Swings" and "Its Wings" in a number of conditions.

Tables 20 through 23 show measurements of specific characteristics from the contours of phrases in all conditions. Phrases from all three parts of the experiment were analyzed. with the exception of three pairs. These were "a nice man"-"an ice man", "see Mabel"-"seem able", and "more rice"-"more ice". They were excluded because the Fo contours were not clear enough to permit measurement of the characteristics chosen for analysis. The characteristics measured include measures of peak fundamental frequency and Fo at the end of the first word and the beginning of the second word of the phrases of one speaker. Table 20 shows the peak Fo of the first word and the second word of word initial and word final consonant phrases.. Looking at the direction of the change in the word initial consonant phrases it is seen that the peak Fo increases from the first word to the second in 5 phrases, decreases in one and remains the same in two. For the word final condition the peak Fo increases in three, decreases in one and remains the same in four phrases. A t-test performed to test the significance of difference between the changes in peak Fo in the two conditions indicated that there were no significant differences ($t(7) = < 1$).

Table 20

Peak fundamental frequency of the first word and the second word for the word initial and word final consonant conditions

<u>Word initial consonant</u>			<u>Word final consonant</u>		
	<u>Peak F₀</u>	<u>Peak F₀</u>		<u>Peak F₀</u>	<u>Peak F₀</u>
	<u>First Word</u>	<u>Second word</u>		<u>First word</u>	<u>Second word</u>
I scream	100	115	Ice cream	105	120
grey day	100	100	grade A	95	95
it swings	100	110	its wings	95	100
grey tapes	95	120	great apes	100	100
plum pies	85	110	plump eyes	105	105
new deal	90	100	nude eel	100	105
bee feeder	125	95	beef eater	110	95
keep sticking	130	140	keeps ticking	115	120

Table 21 shows the peak Fo of the first word and the second word of the phrases that appeared in correct and incorrect utterances. There were no significant differences between the first word and the second word for the correct versus the incorrect usage of the same phrase.

Table 22 shows the Fo at the end of the first word and the beginning of the second word for the word initial and word final consonant phrases. The differences between the Fo at the end of the first word and the beginning of the second were computed for the word initial consonant condition and the word final consonant condition. Comparison of the frequency differences indicates no significant difference between the word initial consonant and the word final consonant phrases.

Table 23 shows the Fo at the end of the first word and the beginning of the second word for the same phrase read in a correct and an incorrect utterance. There is no significant difference on this measure, between the phrase used in the correct or incorrect context.

Duration. Measurements were made from wide band spectrograms, of the total duration of the phrase, the

Table 21

Peak fundamental frequency of the first word and the second word of the phrases that appeared in correct and incorrect utterances

<u>Phrase</u>	<u>Correct</u>		<u>Incorrect</u>	
	<u>First Word</u>	<u>Second Word</u>	<u>First Word</u>	<u>Second Word</u>
grey tape	95	120	95	115
bee feeder	125	95	120	95
grey day	100	100	100	100
keep sticking	130	140	135	150
it swings	100	110	100	100
plump eyes	105	105	110	105
nude eel	100	105	95	100
ice cream	105	120	110	125

Table 22

Fundamental frequency at the end of the first word and at the beginning of the second

	<u>Word Initial</u>			<u>Word Final</u>	
	<u>End of</u> <u>First Word</u>	<u>Beg. of</u> <u>Second</u>		<u>End of</u> <u>First Word</u>	<u>Beg. of</u> <u>Second</u>
I scream	105	105	Ice cream	100	110
grey day	95	90	grade A	95	95
it swings	90	100	its wings	80	80
grey tapes	100	115	great apes	100	105
plum pies	80	100	plump eyes	100	80
new deal	80	90	nude eel	75	75
bee feeder	70	65	beef eater	110	95
keep sticking	120	125	keeps ticking	115	110

Table 23

Fundamental frequency at the end of the first word and at the beginning of the second for the same phrases in correct and incorrect contexts

Phrase	Correct		Incorrect	
	End of First Word	Beg. of Second	End of First Word	Beg. of Second
grey tape	100	115	105	110
bee feeder	70	60	75	60
grey day	95	90	90	90
keep sticking	120	125	120	120
it swings	90	100	90	100
plump eyes	90	85	90	80
nude eel	75	75	85	85
ice cream	100	110	105	115

phoneme preceding the junctural consonant and the junctural consonant itself. Measures were taken from isolated conditions, the phrase in a correct sentence and in an incorrect utterance. Figure 19 shows that the mean duration of the phrases is greatest in isolation, gets shorter in a correct sentence and increases slightly in an incorrect utterance. The word initial consonant phrase is always slightly longer than the word final consonant phrase.

Figure 20 shows the duration of the phoneme preceding the junctural consonant. It is shorter in the sentence than in the isolated reading and there is no change between a correct sentence and an incorrect utterance. In Figure 21 the duration of the junctural consonant itself is plotted. There is an increase in the duration of the word initial consonant in incorrect utterances. Table 24 shows that this durational emphasis of the initial consonant does not increase accuracy of report by listeners.

The results of this experiment indicate that when context is available to the listener he determines what he heard largely on the basis of that context. This was demonstrated by the fact that the listeners made decisions on the basis of the context more fre-

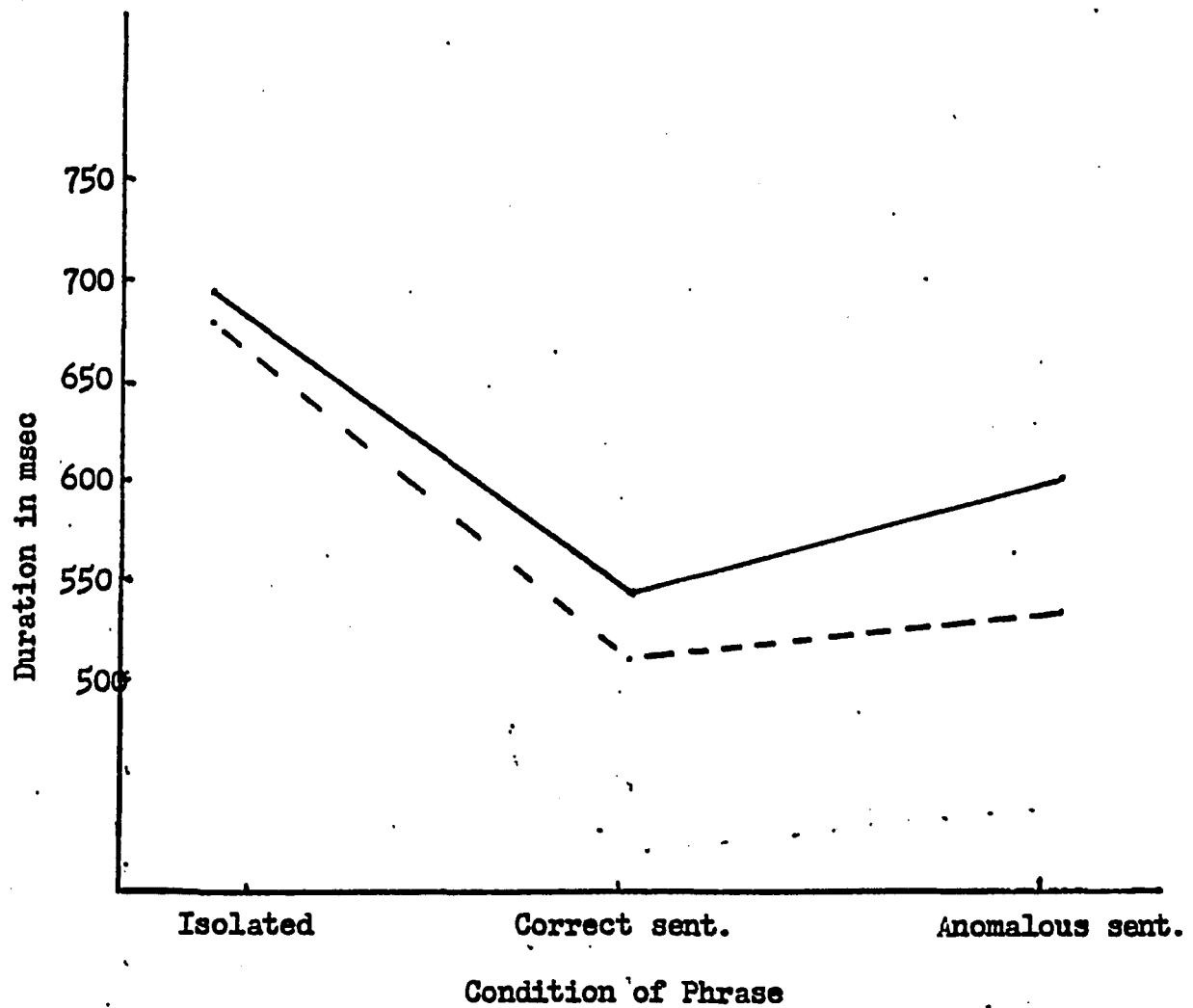


Fig. 19. Overall phrase duration in different contexts.

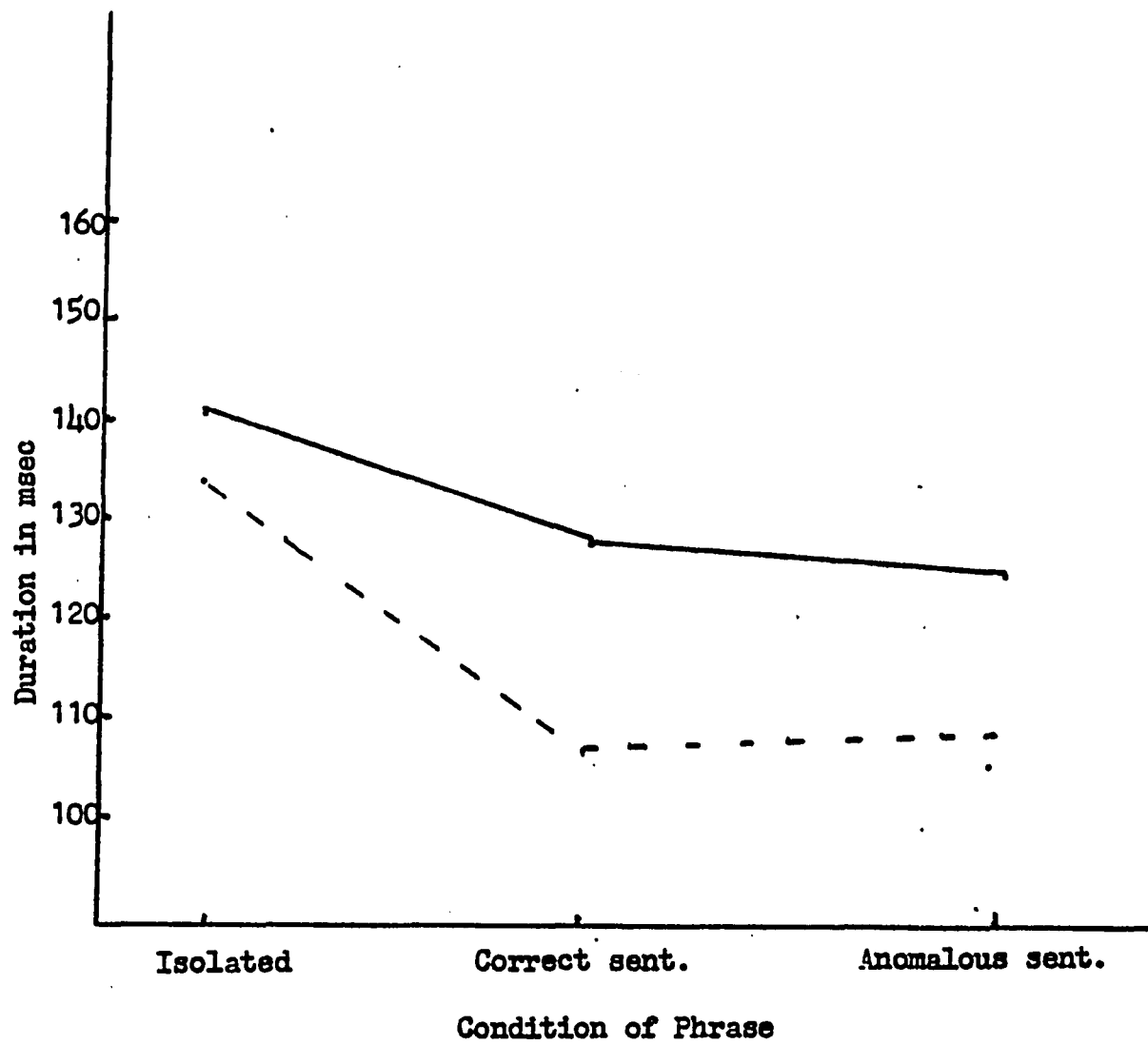


Fig. 20. Duration of phoneme preceding junctural consonant in different contexts.

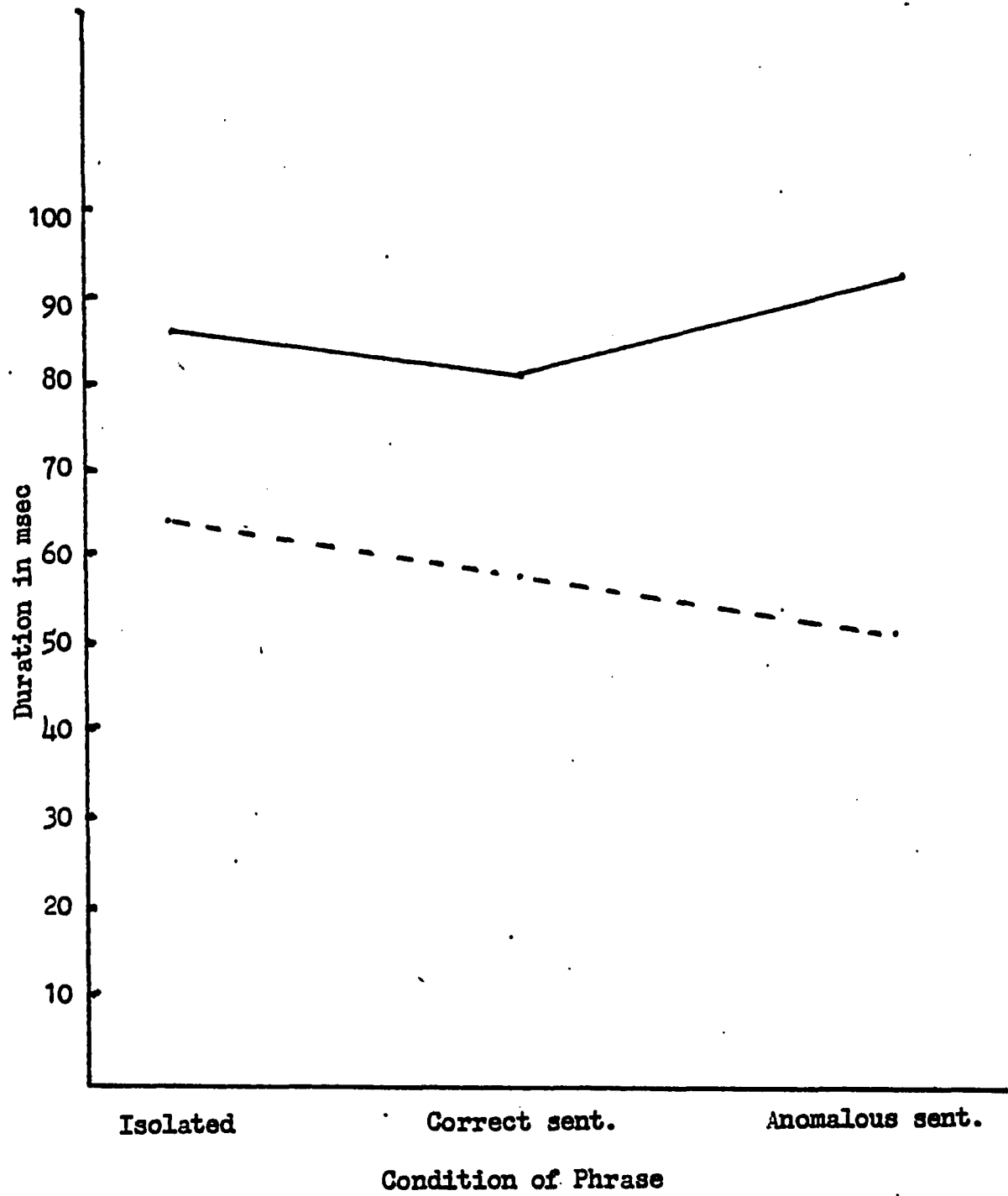


Fig. 21. Duration of junctural consonant in different contexts.

Table 24

Mean duration of phrases and components in correct, incorrect and isolated conditions. Percent correct responses in the listening test are included

Duration of	Word Initial Condition					Word Final Condition				
	<u>Isol.</u>	<u>Corr.</u>	<u>% Corr.</u>	<u>Anom.</u>	<u>% Corr.</u>	<u>Isol.</u>	<u>Corr.</u>	<u>% Corr.</u>	<u>Anom.</u>	<u>% Corr.</u>
Overall	690	550	90	600	27.5	680	520	88	550	35
Preceding phoneme	144	122	90	123	27.5	135	108	88	110	35
Junctural Consonant	87	80	90	94	27.5	75	67	88	48	35

quently when the phrase was totally (grammatically and semantically) incorrect than when it was only semantically incorrect. When the context provided no differentiating information the phrases were judged based on the acoustic cues to boundary. This is apparent because here, as in experiment I, the word final consonant condition yielded a great many more errors than the word initial consonant condition. Unfortunately the same phrases were not used in all three parts of the experiment and thus there was some confounding between the effect of context and between-phrase differences that may have occurred. Experiment III was designed to remedy this problem and also to allow for an evaluation of the consonant placement effect.

CHAPTER IV

EXPERIMENT III: PERCEPTION AND PRODUCTION IN ISOLATION AND CONTEXT

The results of Experiments I and II provide information on the role of contextual information in the production and perception of word boundary in internal open juncture phrases. As a consequence of these results certain hypotheses can be formulated regarding the factors that affect the recognition of the boundary in internal open juncture. Experiment III addressed these hypotheses by combining elements of the first two experiments into a larger experiment of balanced design in which the specific effects of context on boundary perception and production could be measured through a listening test paradigm. In this experiment the reader produced the phrases in anomalous and correct contexts as in Experiment II. However, the same set of phrase pairs appeared in all three experimental conditions (semantically and syntactically correct; semantically anomalous, syntactically correct; semantically and syntactically anomalous). In this way confounding between ex-

perimental conditions and interphrase differences were avoided. In addition, the phrases were constructed so as to include a representative set of consonants in the junctural position. In the second part of the experiment all the phrases were excised from the recordings and presented in isolation to another group of listeners for identification. In this way the effect of context on the production of contrasting phrase pairs could be examined.

It was hypothesized that the subjects would make most errors when the phrases were grammatically and semantically anomalous and somewhat fewer errors when they were only semantically anomalous. It was further hypothesized that the errors when both phrases were correct would be a function of the acoustical characteristics only because the context provided no cue to phrase recognition. In addition, on the basis of results of experiment II, a significant difference in recognition of word-initial consonant phrases and word-final consonant phrases was expected. Specifically it was expected that the word initial consonant phrases would be more accurately identified than their word final counterparts. Based on previous research by Dukes and Nakatani cited earlier it was

also hypothesized that phrases with stop consonants (/d, /t/) would be more often correctly identified than those with the consonants (/m/ & /s/). In Part 2 it was expected that the phrases, in isolation, would be identified by acoustic cue and therefore the word initial consonant phrases would have higher scores than the word finals.

Method

Four pairs of internal open juncture phrases were used. A representative set of consonants was used in the junctural position. The phrase pairs were: I see - Icy; Free Danny - Freed Annie; Grey Tape - Great Ape; See Mabel - Seem Able. The junctural consonants, /t/, /d/, /m/, and /s/ include three of the major consonant classes. Constraints on the overall size of the experiment and the difficulty of finding appropriate phrase pairs prevented the inclusion of other basic consonant types such as the voiced fricatives and glides.

Each phrase was placed in three different sentence frames. In the first set of frames the phrase was grammatically and semantically anomalous half of the time and completely correct in the other half of

the frames. In the second group of frames the phrase was grammatically correct and semantically anomalous in half of the frames and completely correct in the other half. In the third group of frames the phrases were completely correct in every sentence. All of the sentences are reproduced in Table 25.

For the second part of the experiment all of the phrases from every sentence (anomalous and correct) were excised from the recordings for presentation in isolation.

As in Experiment I, one male reader, highly practiced in oral reading, read all of the sentences and sentence - like utterances. He was permitted to practice reading the utterances until he was confident that he could read them naturally even though some did not make sense. He read all the anomalous sentences (interspersed among many buffer sentences) first. Two weeks later he read all the correct sentences. The recordings were made in an IAC sound booth with a high quality microphone placed six inches from the speaker's lips. A high quality Ampex tape recorder (model PR-10) was used to make the recordings. Recording level was monitored by the experimenter with a VU meter.

Table 25

All sentences used in all parts of Experiment III

Grammatically Semantically
anomalous

1. Did icy fish at the store
2. The I see water was blue and cold
3. Perhaps they will Freed Annie in the morning
4. They have already Free Danny from jail
5. The evolutionary period was grey tape anthropoid man
6. They recorded on great ape extra strength
7. He did not see Mabel to lift the weight
8. Did you seem able and Norman in the street

Correct counterpart

- Did I see fish at the store
- The icy water was blue and cold
- Perhaps they will free Danny in the morning
- They have already freed Annie from jail
- The evolutionary period was great ape anthropoid man
- They recorded on grey tape extra strength
- He did not seem able to lift the weight
- Did you see Mabel and Norman in the street

Grammatically correct
Semantically anomalous

1. It is cold and I see without earmuffs
2. Did icy fish talk
3. I have freed Annie yogurt in the cafeteria
4. They have free Danny for money
5. The zoo was full of grey tapes and other primates
6. The recording was on great ape which is stronger
7. He wished he could seem able again because she was so beautiful
8. He tried to see Mabel but everyone knew he wasn't

Correct counterpart

- It is cold and icy without earmuffs
- Did I see fish talk
- I have free Danny yogurt in the cafeteria
- They have freed Annie for money
- The zoo was full of great apes and other primates
- The recording was on grey tape which is stronger
- He wished he could see Mabel again because she was so beautiful
- He tried to seem able but everyone knew he wasn't

Either Correct

1. On a cold day will I see Mary freeze
2. On a cold day will icy Mary freeze
3. They have freed Annie for lunch
4. They have free Danny for lunch
5. There were pictures of grey tapes in the brochure
6. There were pictures of great apes in the brochure
7. He did not see Mabel at first
8. He did not seem able at first

For part 1 of the experiment three tapes were constructed, one for each condition. One tape contained the grammatically and semantically anomalous utterances and their correct counterparts randomly placed among many buffer sentences. A fifteen second pause occurred after each utterance. The second tape contained the semantically anomalous utterances and their correct counterparts as well as the buffer sentences. Tape three contained the sentences that were both semantically and grammatically correct with either phrase member.

For part 2 of the experiment all the excised phrases were placed in a randomly selected order and recorded among buffer phrases for presentation in isolation. There was a four second interval between phrases.

The subjects were 128 college freshmen enrolled in introductory Psychology courses. All subjects were native speakers of English and none had any known hearing difficulty. Thirty two subjects served in each condition of part 1 of the experiment and the same number served in part 2. They were randomly assigned as to the experimental conditions.

Each group of 32 listeners was seated in an acoustically treated classroom (described earlier). The tapes were played through a high quality Ampex tape recorder into a KLH loudspeaker. The frequency response was 4 db from 30-8000 Hz. The sound level measured at the ear of a listener in the center of the room was approximately 75 db SPL.

In each section of part 1 the listeners were instructed as follows: "You will hear a number of utterances on this tape. Simply write down exactly the words that you hear in each utterance. You will have fifteen seconds between utterances to write your answers." The listeners were not told that some of the utterances were semantically and/or syntactically anomalous and that others were not. A response was considered correct if the open juncture phrase spoken by the reader was written. Any other response was considered incorrect. The percent correct judgments for each utterance were tabulated.

In part 2 thirty two listeners were instructed to simply write down the phrase they heard on the tape. The same scoring procedure and method of tabulation as for Part 1 was used.

Results

Results of all sections of Part 1 are depicted graphically in Figure 22 and in tabular form in Table 26. In Figure 22, histograms show the percent correct judgments by all subjects. The histogram at the left shows that the grammatically and semantically correct versions of the sentence were reported correctly 91% of the time (open bar). For the sentences that were both grammatically and semantically anomalous the percentage of correct responses was only 11% (hatched bar). The two histograms in the middle show the results for section II. The open bar histogram shows that in the semantically correct version the percent correct was 86.5. In the semantically anomalous condition the percent correct (hatched bar) was 26. The histogram on the right of the figure shows the percent correct in the condition in which either phrase was acceptable. Here the listeners were correct 63 percent of the time. Two analyses of variance were performed on these data. One analysis was done for the two anomalous conditions and the other analysis for their correct counterparts. Since the variance of an estimated proportion varies as a function of the

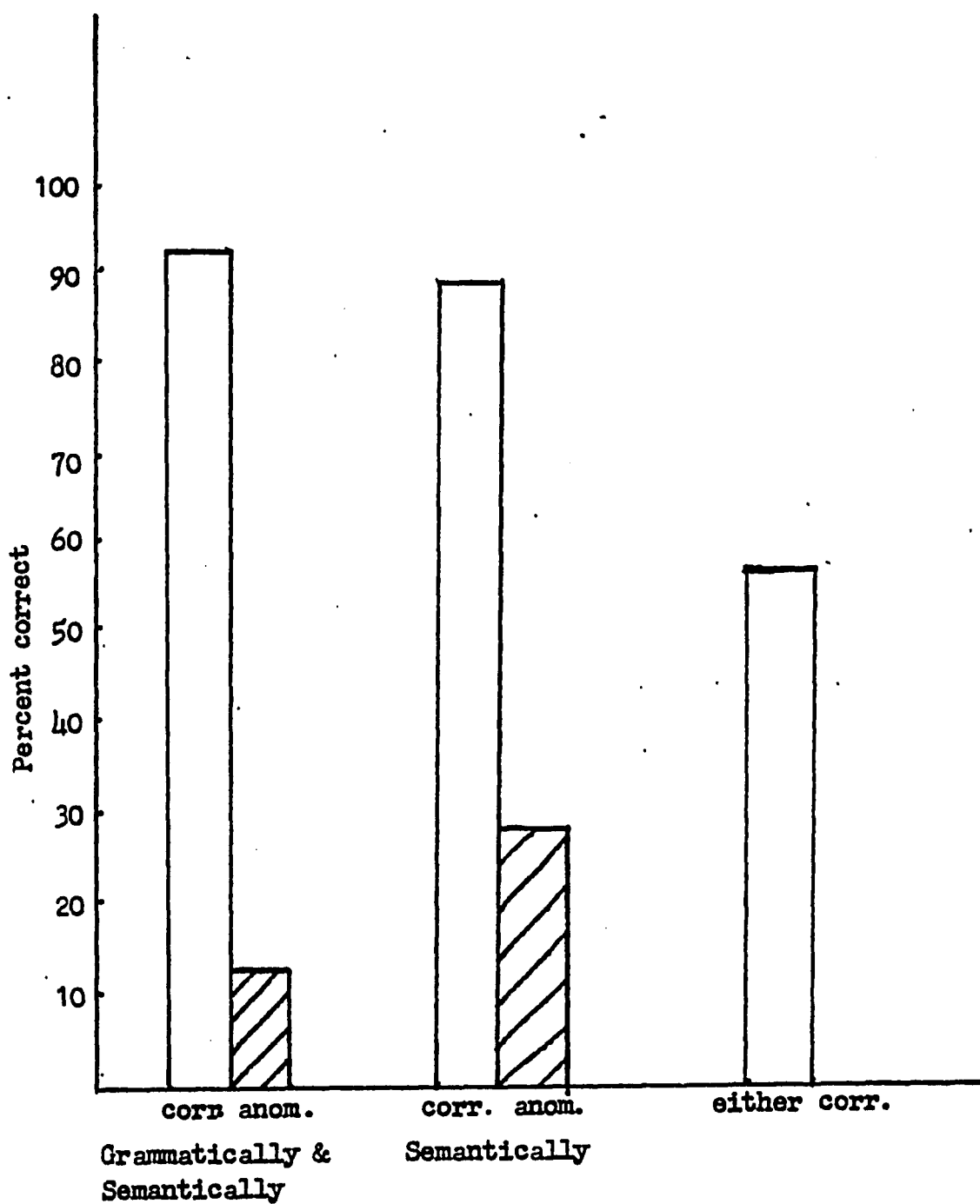


Fig. 22. Percent correct responses in all sections of part I of Experiment III.

Table 26

Percent correct identification for all conditions of Experiment III

<u>Phrase</u>	<u>Semantic & Syntactic</u>				<u>Semantic only</u>				<u>Either Correct</u>	
	<u>Anomalous</u>		<u>Correct</u>		<u>Anomalous</u>		<u>Correct</u>		<u>Context</u>	<u>Excised</u>
	<u>Context</u>	<u>Excised</u>	<u>Context</u>	<u>Excised</u>	<u>Context</u>	<u>Excised</u>	<u>Context</u>	<u>Excised</u>		
Icy	0	93.5	97	73.5	16	81	94	65.5	22	78
I see	6	31	81	43.5	6	75	88	74	87	71
Freed Annie	9.5	87	97	72	19	75	62.5	58.5	37.5	76
Free Danny	9.5	87	100	92	25	88	94	83.6	91	97
Great Apes	34	97	78.5	90	31.5	64	94	60	51.5	43
Grey Tapes	9.5	97	97	100	40.5	100	94	97	91	100
Seem Able	6.5	38.5	81	44	16	56	72	54	31.5	53.5
See Mabel	0	59	97	62.5	53.5	63	94	53.5	72	72.5

proportion, the data were subjected to an arc-sine transform in order to stabilize the error variance prior to the analysis of variance. The factors in both analyses were: Sentence form- grammatically & semantically anomalous - & semantically anomalous, Consonant position- consonant after the word boundary and consonant before the boundary; and Consonant type - /m/ /s/ - /d/ /t/; resulting in a 2x2x2 design. The consonants were grouped as /m/ /s/ and /d/ /t/ because there was only one observation per consonant for the listeners. Therefore it was necessary to form some pooling in order to obtain the within cells variance needed for any analysis. The choice of how the consonants would be grouped was decided upon based on the fact that the stops shared the same place of articulation. The pooling of /s/ and /m/ was done to achieve balance in the analysis. Because we wanted to show that internal open juncture phrases with stop consonants were easier to identify than phrases with other types of consonants it was felt that the pooling of /s/ and /m/ was justifiable.

In the analysis for the anomalous sentences only, the difference between grammatically and semantically anomalous sentences and semantically anomalous sen-

tences was significant ($F(1,8) = 6.61$ $p < .05$]. Neither of the other two factors was significant.

For the analysis of the same factors but in the all correct situation none of the three factors significantly influenced the percent correct. The scores were all quite high, the overall percentages correct were cited above.

In section III where either member of the phrase pair was both grammatically and semantically correct an analysis of variance comparable to those done above would not have been possible because no incorrect condition existed. In this condition when the consonant was word initial the percent was very high, 88%, but when it was word final the percent correct was only 38%.

The results for part 2 of experiment III where all phrases were excised out of all conditions and presented in isolation, may be seen in Table 28. First of all the percent correct for all four conditions varied only from 68% to 75% correct. Because the phrases were excised and because of the results obtained for excised phrases of experiment I these percentages were considered quite reasonable. The percent correct for Experiment I excised phrases was

Table 27

Overall percent correct judgments for the four main
conditions in isolation

Gram. & Sem. <u>Anomalous</u>	<u>Correct</u>
73.40	73.1
Semantically <u>Anomalous</u>	<u>Correct</u>
76.0	68.25

81 and that for the correct phrases in this experiment was only 70. This difference may be accounted for by the fact that phrases having other than stop consonants at the juncture are included in this experiment. Taking only the stops from this experiment the rate of correct responses is 82 percent which is almost identical to the result for Experiment I.

Two analyses of variance analogous to those of Part 1, were performed on the data for these conditions.

The first analysis of variance, for the phrase pairs excised from semantically and/or syntactically anomalous sentences, showed that the only significant factor was that of consonant type ($F(1,8) = 12.56$ $p < .01$). Phrase pairs involving stop consonants (/d/ and /t/) in the junctural position were identified correctly more often than those pairs involving the consonants /m/ and /s/. The second analysis of variance, for phrase pairs excised from semantically and syntactically correct sentences showed both the consonant type and the consonant position were statistically significant factors ($F(1,8) = 17.71$ $p < .01$ and $F(1,8) = 5.87$ $p < .05$, respectively). In addition, a statistically significant interaction between these two factors was also obtained ($F(1,8) = 6.82$ $p < .05$).

Figure 23 shows these two main effects and how they interact. Consonant type is plotted on the abscissa. The ordinate shows the percentage of correct responses. Two curves are shown, one for the effect of consonant before the boundary (B1), the other for the effect of consonant after the boundary (B2). The two curves intersect, indicating an interaction between the two variables. For /s/ and /m/ the scores are slightly higher for consonants before the boundary, for the stops the score for consonants after the boundary is much higher. Note that both curves have a positive slope indicating that, for either consonant position, scores for the stop consonants are higher than those for the others.

In addition to the perceptual results, spectrograms of every phrase were made and duration measures were taken from these in a manner similar to the first two experiments. The duration measures are shown in Table 28. This table shows the duration of the junctural consonant and the phoneme that precedes it for all phrases in all conditions. The presence or absence of glottal stops and aspiration periods are also reported where applicable as in the previous experiments. Because these acoustic characteristics are

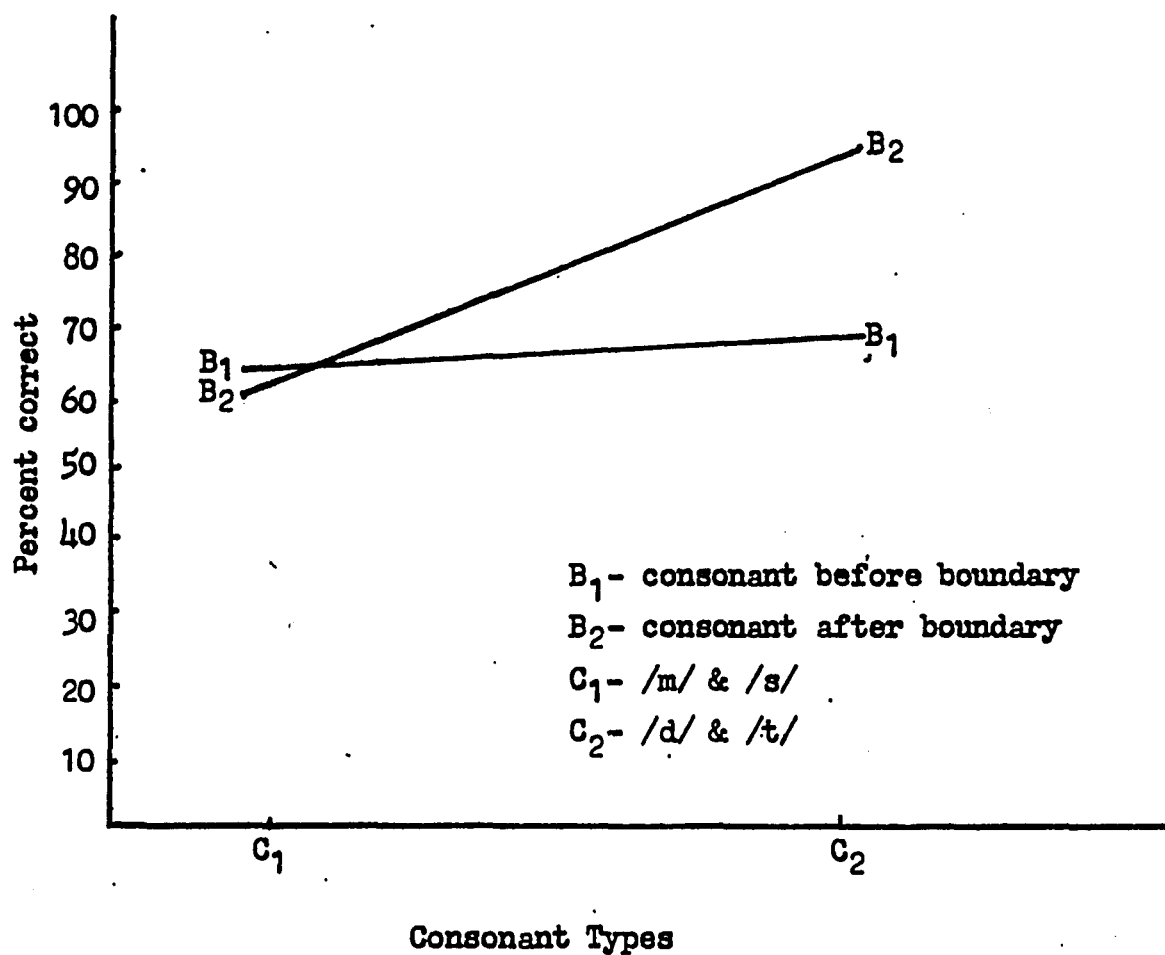


Fig. 23. Interaction effect of position and type of consonant for the "isolated correct" condition.

Table 28

Duration of junctural consonant and preceding phoneme in all conditions of Experiment III

<u>Phrase</u>	<u>Semantic & Syntactic</u>				<u>Syntactic only</u>				<u>Either Correct</u>	
	<u>Anomalous</u>		<u>Correct</u>		<u>Anomalous</u>		<u>Correct</u>			
	<u>P.P.</u>	<u>J.C.</u>	<u>P.P.</u>	<u>J.C.</u>	<u>P.P.</u>	<u>J.C.</u>	<u>P.P.</u>	<u>J.C.</u>	<u>P.P.</u>	<u>J.C.</u>
Icy	140	110	140	130	160	110	140	100	115	110
I see	115	140	100	130	120	130	140	140	90	130
Freed Annie	120	35	90	30	135	35	125	35	130	45
Free Danny	100	90	130	80	150	80	130	80	140	85
Great Apes	130	45	125	40	150	55	125	65	115	50
Grey Tapes	130	85	130	90	135	80	130	90	135	85
Seen Able	70	65	70	70	80	70	80	80	75	75
See Mabel	65	80	70	85	80	80	60	110	75	90

only applicable to half of the phrases used in this experiment, these results were not presented in tabular form.

The duration of the preceding phoneme or the junctural consonant cannot account for the poor performance of "I see" in the excised conditions for the syntactically and semantically anomalous versions or the correct counterparts. In the anomalous as well as in the correct versions the durations do not deviate significantly from those in other conditions.

Another interesting finding as observed from Table 26, was that the percent correct for "great ape" in the anomalous context condition was considerably higher than for all other phrases in this condition. Examination of the duration of the junctural consonant and the preceding phoneme for this phrase shows nothing that might account for this finding. The durations are quite consistent with those found for other tokens of this phrase, and are quite different from those found for the tokens of "grey tape". This phrase also yields low percentages correct for the excised versions of the semantically anomalous condition, its correct counterpart and the excised version of the "either correct" condition. Examining the

duration figures shows nothing that would lead to the conclusion that the durations are somehow responsible for these lower percentages. In every case the duration of the junctural consonant is considerably shorter than it is for every token of the "grey tape" phrase. In addition there is a glottal stop in every version of "great ape" and a long aspiration period in every version of "grey tape". With these consistent differences we might expect that the identification scores for the "great ape" phrases in excised conditions would be quite a bit higher. However, examining the data for the "great eye" excised phrase (another voiceless stop condition) of Experiment I shows that the percent correct identification is also quite low and consistent with the "great ape" data shown here. The anomaly here may be the high scores for the "great ape" excised phrase in the semantically and syntactically anomalous condition and its counterpart. There is nothing in the acoustics examined here, of these two tokens to account for the superior identification scores. The durations are comparable to other tokens of the same phrase and there is a glottal stop of similar duration to those found for other tokens. The only other version of "great ape" that

shows an unexpected percent correct identification is in the correct context version of the syntactically and semantically anomalous condition. Here the percent correct was only 78.5%. A considerably higher value should have been obtained based on the other correct context scores. Because the acoustic cues for this version were again similar to those in other versions the reasons for the poor performance must be sought elsewhere. It is possible that because the sentence context was rather complicated the listeners were not able to correctly process the semantic content.

The phrases "See Mabel" and "seem able" in general, have the lowest correct identification scores in this experiment. For these phrases the duration data may shed considerable light. The differences between the duration of the preceding phoneme and the junctural consonant of "see Mabel" and "seem able" are not very great. In other words duration cues might not differ enough to make differentiation of the phrase pairs in excised conditions easy for a listener. Here the context is the all important determiner and it is shown that when the context is correct and unambiguous the identification scores, at least for the word initial condition, are quite high.

There are also particularly low scores for the phrase "Freed Annie" in the correct version of the semantically anomalous condition. This is true for both the context and excised versions. Examination of the duration data does not help to explain why this difficulty occurred. The phrase is very similar in duration, to other versions and is quite different from the "Free Danny" phrase in the length of the junctional consonant. In addition "Freed Annie" does contain a glottal stop and the consonant is released without aspiration. There is, therefore, no reason, from the acoustics examined here, for this token to yield such low percentages of correct identification.

CHAPTER V

DISCUSSION

Experiment I

The results of experiment I demonstrate that a listener has some difficulty in recognizing boundary location in internal open juncture phrases when the phrases are excised from sentence context and presented in isolation. In condition I there were virtually no errors. That is, when produced and heard in isolation, these stimuli were reported with almost perfect accuracy. In condition II, however, when the phrases were excised and presented in isolation a substantial error rate was observed. This result is consistent with that of Dukes and Nakatani (1977) in that both studies showed a substantial error rate for the excised phrases.

In the excised condition there were no obvious syntactic or semantic contextual cues and in the isolated condition there could not have been any contextual cues. Yet there were considerably more errors made in the excised condition. It is necessary to ex-

amine what production factors are responsible for this difference in listening task results. This discussion centers on the production differences that exist between the two conditions.

Because there were only three phrase pairs used here, results are discussed mainly on a between condition basis. That is, this investigation was concerned primarily with cue differences between the two conditions, not with an extensive evaluation of cues to juncture, as was the study of Dukes and Nakatani.

The first item that requires some discussion is the selection of only one speaker to produce the stimuli. Although he was a highly practiced oral reader who had been utilized as a speaker in many such studies, it is possible that speaker cues introduced were not sufficient to provide listeners with information to make boundary decisions. The only way we can determine if the cues were adequate is to examine the scores. The listeners performance was so high that we must assume that the speaker uttered the intended phrases. In addition, in Experiment II three speakers were used and no significant differences as a function of speaker were observed. This fact provides further justification for the use of one trained speaker in this type of investigation.

In the current investigation there was a mean error of only 4% for the stops in word initial position. For the stops in the same position Dukes and Nakatani showed an error rate of 10%. Considering the speaker and condition differences in the two experiments these differences are not mutually inconsistent. In addition, because of the relatively low error rates in both experiments we focused on those conditions where error rates are higher on average, i.e. on the phrases in which there were vowels in word initial position. Here Dukes and Nakatani found 16% errors and the current experiment showed an average error rate of 35%. Because Dukes and Nakatani did not break down their results into specific vowels and because they used many more phrases it is impossible to determine why their percentage of errors was so much lower than those found here.

Dukes and Nakatani claimed that a glottal stop at the onset of a vowel provided the strongest cue for juncture in their experiments. Examination of Table 12 shows that there is a glottal stop present in all of the vowel initial phrases in the isolated condition and in two of the three phrases in the excised condition. It should be noted that the duration of the

glottal stop was greater for the isolated phrases than it was for the excised phrases. Figure 6 shows this duration difference for the phrase "great eye". Despite the presence of the stop, the excised vowel-initial phrases showed substantially lower percentages of correct scores than the consonant initial phrases. The percent correct was not consistently higher for those phrases in which a glottal stop was present. If the glottal stop were a necessary and sufficient cue for perception of juncture then the vowel initial phrases in the excised condition that had a glottal stop should have been equal in percent correct to the same phrases in the isolated condition. Because the score for the excised phrases was so much lower it must be concluded that the difficulty is due to something other than the presence or absence of a glottal stop at vowel onset.

Leniste (1960) and Dukes and Nakatani (1978) both claimed that aspiration on the word initial voiceless stop consonant is a cue to juncture. If this is the case, then it would be expected that the error rate would be quite low for voiceless plosives as the junctural consonant since in the prejunctional position there is no aspiration which contrasts with the post-

junctural position in which the consonant is released with aspiration. This study showed that for both excised and isolated conditions scores for the word initial voiceless stop phrases were quite high.

Although Table 12 shows that when aspiration is present in the isolated condition it is also present in the excised condition, it should be noted that the aspiration time is longer in the isolated condition. Figure 5 shows this for the phrase "grey tie". The possible significance of this duration difference in aspiration time is discussed below.

Fundamental frequency contours do not appear to provide enough consistent information for consideration of differences between contours to be important cues to juncture. The only totally consistent finding was that the rate of fall of the F_0 on the second word was always greater for the word initial consonant condition. It is interesting, however, that examination of the actual contours as presented in figures 2-4 did reveal that there are considerable differences in general form of the contours between the word initial and word final consonant conditions for some of the phrases. The differences that are seen in these figures on visual inspection, do not appear to be

greater in the isolated condition than in the excised condition. In fact the differences look greater in the excised condition. If Fo contour were a cue to juncture, the assumption is that a greater difference in contours would be more discriminable to a listener than a lesser difference. Therefore there should have been fewer errors in the excised condition. However, almost all errors were in the excised condition, not in the isolated condition. It, therefore, must be concluded that the Fo contour per se did not serve to distinguish the boundary position in these phrases.

The data are much more consistent with respect to duration. The prejunctional and junctional phonemes for all three phrase pairs were longer in the word initial consonant condition. This difference in absolute duration is, however, not reflected when the relative duration of the prejunctional and junctional phonemes are examined. It is therefore unlikely that the duration differences themselves account for the differentiation between isolated and excised versions.

As indicated above the absolute presence or absence of such cues as glottal stops and aspiration exist in both the isolated conditions (where there are almost no listening errors) and the excised condition

(where there are a considerable number of errors). It is therefore necessary to look elsewhere in order to account for the differences in listening scores between the two conditions. One possibility is in the duration of the cues. As mentioned earlier the glottal stops and the aspiration periods were always longer in the isolated conditions. Perhaps the listener requires cues to juncture to be of a certain minimal duration in order to be able to perform this type of identification without error. However, when a sentence context is present even minimal duration cues are sufficient because the listener relies mainly on the cues provided by the context to locate the word boundaries correctly. Verification of this supposition was the main purpose of Experiment II.

Experiment II

In experiment II the results clearly show that the semantic and grammatical characteristics of a sentence strongly influence the frequency of correct identifications. It must be reiterated at this point that listeners were not instructed about the correctness or incorrectness of the utterance in terms of grammar and semantics. The fact that they were aware of the possibility of an anomalous utterance is obvious from the results obtained. That is, the spoken, yet anomalous phrase was reported 18% of the time in Part I and 45% of the time in Part II. In other words some of the time the listeners did write the spoken phrase that was anomalous. If they were operating on the assumption that all the utterances were correct sentences this would not have happened.

The influence of the sentence context over acoustic cues to juncture in perception is noted by the difference in error rate for Parts I and II. The utterances that were both grammatically and semantically anomalous yielded the highest error rate, and those that were only semantically anomalous yielded a somewhat lower error rate. It appears that the

listener takes into account the nature of the entire utterance when making a decision about what he heard. In other words when the phrase spoken made the utterance completely unacceptable to the structure of the language (Part I) the listeners were very strongly influenced and thus reported the phrase that was correct in English even though (as will be pointed out) the acoustic cues contraindicated that phrase. In Part II, where the utterance spoken was only semantically anomalous, (it was acceptable grammatically) the listener still reported the phrase that would make the utterance semantically correct 55 percent of the time. However, this percentage was not nearly as high as it was for Part I (82)%).

Comparing the results for Part III with those of Experiment I is quite informative. First of all, the largest percent of errors was in the phrases where the consonant was pre-junctural. This was exactly the finding in Experiment I. In the excised phrases of Experiment I the listeners made errors because the context was removed leaving them with only acoustic cues which in some cases (as discussed in the previous section) are not necessarily sufficient. In Part III of this experiment, while there was a context

provided, it offered no help in choosing the correct phrase because either phrase could have been correct. Again, the listener had to depend on the acoustic cues which often proved to be insufficient for correct identification. This happened more often when the second word of the phrase began with a vowel. Again this is in agreement with the findings of Dukes and Nakatani (1978).

For two phrase pairs ("bee feeder - Beef eater" and "new deal - nude eel") the data were not consistent with the pattern obtained for the majority of the phrase pairs considered. Specifically, the percentages of 4 and 10 for "bee feeder" and "nude eel" were extremely low. Even though it was expected that "nude eel" would have a lower percentage than "new deal" based on the position of the consonant, the magnitude of the difference should be examined more closely. A possible explanation is that the phrase "nude eel", and to a lesser extent, "bee feeder" are not very common. As shown in Table 15, the phrase "nude eel" was rated 5 on a five point scale where 5 was the "least common". The difference between "new deal" and "nude eel" was greater than for any other phrase pair. Although the difference between "bee feeder" and "beef

eater" is not so large it is true that "bee feeder" was rated as one of the "least common" phrases of all those used in the experiment.

In order to test the assumption that the results obtained in this experiment were due to context effects in perception and not to context effects in production, the detailed acoustic analyses were conducted. All of these analyses clearly showed that the phrase spoken in isolation, in correct context and in anomalous context shared the same production features. That is, the speakers produced the phrases in a similar fashion in the different conditions. The similarities are clearly seen in Table 18.

The acoustic characteristics in the isolated and excised phrases in Experiment I show similar patterns. For instance, comparison of Table 12 with Table 19 where applicable (plosives only) shows that the characteristics are the same in both Experiments I and II.

For the F_0 contours a comparison of Tables 10 and 11 of Experiment I with Tables 20 and 22 of Experiment II (again for plosive phrases only) shows a lack of consistency, indicating considerable variability in F_0 contours between the two experiments. For example, in

the excised phrase "plump eye" in Experiment I the peak frequency before closure is 15 Hz higher than it is after the release (Table 11). In Experiment II (Table 20) the peak frequency is the same (105 Hz) both before closure and after release. This type of inconsistency is seen many times in the F_0 data both within and across experiments. Such inconsistency makes reliable interpretations of this aspect of acoustic information highly dubious.

Along with the glottal characteristics, the duration results are clearly the most reliable and for this experiment are similar to the results of Experiment I. That is, the duration of the phrase and its junctural components were greater when the phrase was spoken in isolation. However, when the phrase was in an Anomalous context (Parts I and II combined) the duration sometimes increased. This was especially true of the duration of the junctural consonant in post -junctural position (figure 21). It is possible that the reader, when confronted with this type of situation, tried to produce the phrases in anomalous context as carefully as possible. The extra care in production was manifested by the increased consonant duration in the anomalous conditions. However, despite

this increased duration, listeners still made many incorrect judgments of the anomalous phrases. If increased duration of the junctural consonant was a sufficient cue for juncture discrimination in phrases produced in isolation and this increased duration was manifest in the anomalous context production of the phrases then the phrases should have been identified correctly. In addition, if, as according to Dukes and Nakatani, the acoustic cues are sufficient for juncture discrimination then how are we able to account for the poor performance of the listeners in the anomalous conditions? The most probable explanation in the framework of this experiment is that the production cues to juncture are either not perceived by the listener or are subjugated to a secondary role in the decision making process. It is concluded that for the case of internal open juncture phrases, the sentence context in which the phrase appears has the primary role in the identification process, even when the phrases produced are anomalous to the sentence context.

This conclusion was verified in the framework of Experiment III in which all phrases appeared in all

possible conditions so that comparisons between perception and production results could be made without the possible confounding that could have occurred in Experiment II.

Experiment III

The results of experiment III indicate that the hypotheses based on the results of the first two experiments are essentially correct. To be specific it was shown that the grammatically and semantically anomalous phrases were reported correctly only 10% of the time. For this condition in Experiment II with different phrases the score was 18% correct. When the phrases were grammatically correct but semantically unlikely the percent correct was 28 in this experiment and 45 in experiment II. Therefore in both experiments it is clear that the more anomalous the phrase the more it will be incorrectly identified.

Examination of the data for this experiment points out, quite clearly and in a quantitative manner, the role of context in perception of internal open juncture phrases. The matrix in Table 29 helps to illustrate this point by showing the significant factors in the four analyses of variance. In the context condition

Table 29

Matrix showing significant factors in analyses of variance
for isolated and contextual conditions

<u>Correct</u>	<u>Isolated</u>	<u>Context</u>
	consonant position consonant type	nothing
<u>Anomalous</u>	consonant type	degree of anomaly of context

where the phrases were correct the analysis of variance indicated that none of the three variables affected the means. This occurred because the scores for all the phrases were very high (87.5 % correct). There were few errors in these conditions because both the context and the acoustics provided information to the listeners as to the appropriate identification of the phrase. In the anomalous context condition, the only significant difference found was between the two degrees of "anomaly". In other words the grammatically and semantically anomalous phrases were heard correctly more infrequently (10% correct) than the semantically unlikely phrases (28.25% correct). The available acoustic cues did not appear to be used by the listeners. In other words, it made very little difference what type of consonant or whether it was before or after the word boundary. The listeners simply reported phrases on the basis of the context in which the phrase appeared.

Turning next to the isolated phrases which were excised from the correct contexts, it can be seen that the identifications were significantly affected by type of consonant and location of the consonant. Here there is, of course, no context in the listening and

the judgments must be made on the basis of the acoustic information in the stimuli themselves. As shown in Table 28 listeners make a considerable number of errors in these conditions. /d/ and /t/ phrases were easier (83% correct) than /s/ and /m/ phrases (60% correct) to identify and phrases in which the consonants were word initial were easier (78.5% correct) to identify than those in which they were word final (65% correct). For the isolated condition in which the phrases were excised from anomalous contexts the only significant factor was that /d/ and /t/ were identified correctly more often (86% correct) than /s/ and /m/ (59% correct). It was expected that the location of the consonant would also have been significant. However, it appears that the lack of significance could be attributed to a very low score for the phrase "I see" from the grammatically and semantically anomalous condition. The percent correct for this phrase was so low that it pulled down the mean correct for the word initial consonants enough so that there was no difference between that and the word final consonants.

The phrase pair I see - icy is somewhat of a problem due to the lack of parallel structure between

it and the other phrase pairs used. It was chosen because it was possible to fit it into all the contexts necessary to meet the conditions of a balanced design. No other phrase pair investigated appeared to be able to fill this necessary condition.

Table 30 presents the overall listening results across the three experiments. This table shows, for each major consonant type, the percent correct judgments by the listener. The table is organized with the consonant type, in the excised conditions, yielding the highest percentage correct in listening scores at the top and continuing in descending order. For instance the voiceless initial stops have the highest listening scores for the excised conditions. The descending order was considered only for the excised phrases because in context the phrases were judged on the basis of other than phonetic cues. This table was prepared in order to compare the behavior of the phrases across the three experiments, where possible. Because all phrases did not appear in every experiment the comparisons can not be complete. However, some useful information is found for those cases in which the phrases crossed over experiments and for the cases in which different phrases with the

same consonant or consonant type appeared across experiments. The consonant types will be examined in descending order of accuracy of identification by the listeners for the excised conditions.

For the voiceless stops in initial position we see very consistent performance for all phrases over all experiments. The identification scores for every phrase, regardless of condition, are 94% or better. This is, of course, with the exception of the anomalous context condition where a significant discrepancy exists between Experiments II and III for the phrase "grey tape". In Experiment II the anomalous version of "Grey tape" was identified correctly 80% of the time but in Experiment III it was identified only 25% of the time on the average. The duration of the junctural consonant in Experiment II was 70 msec. and in Experiment III it was 85 msec. and 80 msec. for the two anomalous versions. It is possible, although unlikely, that this difference played a role in the identification difference since the longer consonants yielded the lower identification scores. Possibly, sentence differences or speaker differences not taken into account across experiments are responsible for the discrepancy. In addition it

must be pointed out that the fact that voiceless stops in the initial position yield the highest identification scores, is in agreement with the findings of Dukes and Nakatani (1978).

The second highest listening scores occurred for the voiced stop phrases in initial position. Their behavior is highly consistent across the conditions except those in which the phrases appeared in anomalous context. That is, the listeners identified these phrases a very high percentage of the time in Experiments II and III. In the anomalous conditions these phrases were identified quite infrequently, as was hypothesized. Dukes and Nakatani found that voiced initial stops were much less often identified correctly. This difference between our results and theirs is probably due to phrase differences and differences in experimental conditions.

In the word final voiceless fricatives the scores for correct sentences were very high, 97% for "beef eater" and 95.5% for "Icy". In the anomalous context condition both "Ice cream" and "Icy" yielded very low scores as expected. However, in the excised conditions of Experiment III the scores were the third highest of any type of consonant.

For the voiced stops in final position there were large differences between phrases and between experiments. These phrases, spoken in isolation (Experiment I) and correct context (Experiments II and III) were identified quite often (an average of 92% if we do not include "nude eel"). However, in excised condition they were identified correctly much less often (average of 71%). The anomalous conditions for these phrases occurred only in Experiment III and were described earlier.

For the voiceless stops in the final position the listener's performance was quite different as a function of Experiment and condition. The phrases read in isolation (Experiment I) yielded very high identification scores but the excised versions and the context versions showed considerable variability. The low scores for "Plump eye" in Experiment II are no doubt a function of the likelihood of that phrase in context. "Great ape" in Experiments II and III yielded almost identical scores for correct context conditions. The average scores for the excised phrases across the experiments was 70%. Dukes and Nakatani (1978) considered all conditions in which the consonant was word final and the vowel began the second word as one type.

They found that when vowels were in initial position the average percent correct was 84%. Taking the last three conditions reported here, which were all vowel initial conditions, the average percent correct was only 72%. However, the overall position of these vowel initial phrases in these experiments and in the Dukes and Nakatani work is very similar.

The results for the word initial nasal consonant represented by "see Mabel" across Experiments II and III shows considerable consistency between experiments. That is, the scores for "see Mabel" in correct context are high in both experiments and are quite low and similar in anomalous contexts.

For the initial voiceless fricatives the differences in listening results are largely a function of the differences in phrases. "Bee feeder" is almost never correctly identified but "I scream" was identified 98% of the time in a correct sentence in Experiment II. "I see" was correctly identified 84.5% of the time in a correct sentence but very infrequently in an anomalous sentence, much like "bee feeder".

Finally, for the word final nasal consonant represented by "seem able" in correct context the

identification scores were 73 and 61 for Experiments II and III respectively. For the anomalous versions the scores were 16 and 11 % respectively for the two experiments.

Examination of Table 30 shows that, by and large, phrases and phrase types yield relatively comparable identification scores across experiments. This is especially true for the voiced and voiceless stops in the initial position and the nasal in initial and final position. The differences that exist are most surely attributable to the fact that three separate experiments were performed with somewhat varying conditions and with different speakers and some different phrases in each experiment.

The overall conclusions drawn from these findings would best be demonstrated by looking at the context effects on production and perception in the three experiments. Looking first at production it is clear from the first two experiments that the acoustic cues to juncture, as specified in the literature (Nakatani & Dukes, 1978; and Lehiste, 1960), are present in the production of the phrases in isolation and in sentence context. The only reliable difference between the phrases spoken in the two conditions were differences

Table 30

Percent correct listening judgments as a function of consonant type over all three experiments

<u>Consonant Type</u>	<u>Experiment I</u>		<u>Experiment II</u>		<u>Experiment III</u>			
	<u>Excised</u>	<u>Isolated</u>	<u>Correct Sent</u>	<u>Anomalous sentence</u>	<u>Corr. sent.</u>	<u>Anom. sent.</u>	<u>Excised correct</u>	<u>Excised Anomalous</u>
<u>Voiceless stops initial</u>								
Plum pie	100	98	100					
Grey tie	95	99						
Grey tape			94	80	95	25	99	99
<u>Voiced stops -init.</u>								
New Deal			99	18				
Free Danny					95	17	90	88
<u>Voiceless fricatives - Final</u>								
Beef eater			97					
Ice cream				13				
Icy					95	8	70	87
<u>Voiced stops - final</u>								
Grade A	68	95	100					
Nude eel			10					
Freed Annie					80	14	65	81
<u>Voiceless stops final</u>								
Plump eye	72	95	22	23				
Great eye	55	96						
Great ape			88		86	33	75	80
<u>Initial nasal</u>								
See Mabel			94	21	88	27	63	56
<u>Voiceless fricatives init..</u>								
Bee feeder			04	04				
I scream			98					
I see					85	06	59	53
<u>Final nasal</u>								
Seem able					61	11	50	47

in duration of some of these cues i.e. glottal stops and aspiration. Specifically, the cues were longer in the isolated conditions than in the sentence context conditions. In isolation the production of the phrase is considered to be optimal with regard to articulatory parameters. This would account for the extended duration of the junctural cues. When the phrase is produced in context the articulation is by no means optimal. The junctural cues may be present but they are of shorter duration.

Experiment III shows the effects of this less precise articulation on the perception of these phrases in context and removed from the context. If the production cues, in the phrases spoken in context, were sufficient to differentiate pair members (as they apparently are when the phrases are spoken in isolation) then the phrases should have been recognized correctly almost 100% of the time when they were removed from the context and presented in isolation. The results showed a considerable error rate for all the phrases heard in isolation. This must be interpreted as evidence that the acoustic cues in production are only totally sufficient for accurate recognition when the phrases are spoken in isolation.

The results of Experiment III also demonstrate that when the phrase appeared in a completely correct context it was almost never misidentified when the context was not ambiguous as to which phrase belonged. That is, the acoustic cues that appear to be, by themselves, insufficient for perfect identification are totally sufficient when the sentence context adds syntactic and semantic cues as well. When there is ambiguity as for the condition in which either phrase was correct in the context the error rate was high because the context could not be used as a factor in the decision making process.

These results also demonstrate the relative importance of the acoustic cues and the contextual cues. This is evidenced by the findings that when the phrases are presented in contexts that are grammatically and semantically consistent with one phrase it is that phrase that will be heard the majority of the time even if it is actually the other member of the phrase - pair that was produced. In other words, when the acoustic cues were contradictory to the contextual information, the listener apparently considered primarily the contextual cues in his decision.

In order to generalize from these results one must take into account the effects of different speakers, their individual styles, the phrases and conditions under which they were spoken and the linguistic background of the listeners. Considering these variables the results are remarkably consistent even though the speakers were different in the three experiments and the same phrases did not appear in all paradigms.

SUMMARY AND CONCLUSION

Experiment I showed that phrases spoken in isolation and presented in isolation are very easy to identify. However the same phrases spoken in a sentence and presented in isolation cause a certain amount of difficulty. Primarily, word final consonant phrases were more difficult than word initials.

Experiments II & III showed that when the context is varied such that the phrases are anomalous in certain sentences, the listener appears to be using the contextual cues even though the acoustic cues to juncture are present in the phrases. It was shown that when context was correct for the phrase almost no errors were made, and that when context was present but provided no information as to which phrase was spoken the listener depended on the acoustics and in so doing made a considerable number of errors.

The acoustic analyses in Experiments I & II identified specific acoustic contrasts between word initial and word final consonant phrases and Experiment III specified the importance of some of those differences in the recognition of internal open juncture phrases.

The results of the three experiments which each used different internal open juncture phrases and different speakers were quite consistent with one another.

The results of Experiment III seem to be consistent with the idea of a trading relationship that exists between acoustic cues and contextual cues provided by a sentence structure. For internal open juncture phrases it appears that the acoustic cues delimit the choice to one of two possibilities and the context in which the phrase appears is used in making the final decision.

Bibliography

- Abrams, k. and Bever, T.G. Syntactic structure modifies attention during speech perception and recognition. Quarterly Journal of Experimental Psychology, 1969, 21 280-290.
- Bansal, R.K. The intelligibility of Indian English: Measurements of the intelligibility of connected speech, and sentence and word material presented to listeners of different nationalities. 1966, Ph.D thesis London University.
- Brown, R. and McNeill, D. The tip of the tongue phenomenon. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 325-337.
- Bolinger, D.L. and Gerstman, L.J. Disjuncture as a cue to constructs. Word, 1957, 13, 246-255.
- Bolinger, D.L. A theory of pitch accent in English, Word, 1958, 14, 109-149.
- Christie, W.M. Some cues for syllable juncture perception in English. Journal of the Acoustical Society of America. 1974, 55, 819-821.
- Coker, C. and Umeda, N. The importance of spectral detail in initial-final contrasts of voiced stops. Journal of Phonetics, 1975, 3, 63-68.
- Davidson-Nielsen, N. Syllabification in English words with medial sp, st, sk. Journal of Phonetics, 1974, 2, 15-45.
- Dukes, K.D. and Nakatani, L.H. Acoustic and phonetic correlates of word juncture perception. 1978, Unpublished memorandum.
- Fry, D.B. Duration and intensity as physical correlates of stress. Journal of the Acoustical Society of America, 1955, 27, 765.
- Fry, D.B. Experiments in the perception of stress. Language and Speech, 1958, 1, 126-152.

- Fujimura, O. and Lovins, J.B. Syllables as concatenative phonetic units. 1977, Unpublished memorandum.
- Garding, E. Internal juncture in Swedish. Lund: C. W. K. Gleerup, 1967.
- Howes, D. On the relation between the intelligibility and frequency of occurrence of English words. Journal of the Acoustical Society of America, 1957, 29, 296-305.
- Huggins, A.W.F. On the perception of temporal phenomenon in speech. Journal of the Acoustical Society of America, 1972, 51, 1279-1290.
- Huggins, A.W.F. Speech timing and intelligibility. Attention and Performance VII, Hillsdale, N.J.: Lawrence Erlbaum, 1978.
- John, J.E. and Howarth, J.N. The effects of time distortion on the intelligibility of deaf children's speech. Language and Speech, 1965, 8, 127-134.
- Kalikow, D.N., Stevens, K.N. and Elliott, L.L. Development of a test of speech intelligibility using sentence materials with controlled predictability. Journal of the Acoustical Society of America, 1977, 61, 1337-1351.
- Klatt, D.H. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America, 1976, 59, 1208-1221.
- Kozhvnikov, V.A. and Chistovich, L.A. Speech: Articulation and perception. Translated from Russian. (Washington, D.C.: U.S. Department of Commerce Clearinghouse for Federal Scientific and Technical Information) 1965.
- Lea, W.A. Segmental and suprasegmental influences on fundamental frequency contours. In L. Hyman (Ed.) Consonant types and tone. University of Southern California Occasional Papers in Linguistics, 1973, No. 1.
- Lehiste, I. An acoustic-phonetic study of internal open juncture. Phonetica, 1960, 5, Supplement.

- Leniste, I. Juncture, Proceedings of the 5th congress on phonetics, 1965, 171-200.
- Lehiste, I. Suprasegmentals. Cambridge: MIT press, 1970.
- Lieberman, P. Some acoustic correlates of word stress in American English. Journal of the Acoustical Society of America, 1960, 32, 451-454.
- Lieberman, P. Some effects of semantic and grammatical context on the production and perception of speech. Language and Speech, 1963, 6, 172-179.
- Lieberman, P. Intonation perception and language. Cambridge Massachusetts: MIT Press, 1967.
- Miller, G.A., Heise, G.A. and Lichten, W. The intelligibility of speech as a function of the context of the test materials. Journal of the Acoustical Society of America, 1951, 41, 329-355.
- Miller, G.A. and Isard, S. Some perceptual consequences of linguistic rules. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 217-228.
- Morton, J. and Jassem, W. Acoustic correlates of stress. Language and Speech, 1965, 8, 159-181.
- Nakatani, L.H. and Dukes, K.D. Locus of segmental cues for word juncture. Journal of the Acoustical Society of America, 1977, 62, 714-719.
- O'Connor, J.D. and Tooley, O.M. The perceptibility of certain word boundaries. In honour of Daniel Jones, D. Abercrombie, (Ed.), London: Longmans Ltd. 1964.
- Olive, J.P. and Nakatani, L.H. Rule synthesis of speech by word concatenation: A first step. Journal of the Acoustical Society of America. 1974, 55, 660-665.
- O'Neill, J.J. Recognition of intelligibility test materials. Journal of Speech and Hearing Research, 1957, 1, 262-267.
- Pickett, J.M. and Pollack, I. The intelligibility of excerpts from fluent speech: Effects of rate of

utterance and duration of excerpt. Language and Speech, 1963, 6, 151-164.

Pisoni, D.B. and Sawusch, J.K. Some stages of processing in speech perception. In A. Cohen and S.G. Nooteboom (Eds.), Structure and process in speech perception, New York: Springer-Verlag, 1975.

Pollack, I., Rubenstein, H., and Decker, L. Intelligibility of known and unknown message sets. Journal of the Acoustical Society of America, 1959, 31, 273-279.

Pollack, I. and Pickett, J.M. The intelligibility of excerpts from conservation. Language and Speech, 1963, 6, 165-172.

Pollack, I. and Pickett, J.M. Intelligibility of excerpts from fluent speech: Auditory versus structural context. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 79-84.

Rabiner, L.R. and Gold, B. Parallel processing techniques for estimating pitch periods of speech in the time domain. Journal of the Acoustical Society of America, 1969, 46, 442-448.

Savin, H.B. Word frequency effect and errors in the perception of speech. Journal of the Acoustical Society of America, 1963, 35, 200-206.

Shockey, L. and Reddy, R. Quantitative analysis of speech perception: Results from transcription of connected speech from unfamiliar languages. Paper presented at the Speech Communication Seminar, Stockholm, Sweden, August 1-3, 1974.

Solomon, R.L. and Postman, L. Frequency of usage as a determinant of recognition thresholds for words. Journal of Experimental Psychology, 1952, 43, 195-201.

Smith, F. and Goodenough, C. Effects of context, intonation and voice on the reaction time to sentences. Language and Speech, 1971, 14, 241-250.

Stowe, A.N., Harris, W.P. and Hampton, D.B. Signal and context components of word recognition behavior. Journal of the Acoustical Society of America, 1963, 35, 639-644.

Triesman, A.M. Contextual cues in selective listening. Quarterly Journal of Experimental Psychology, 1968, 20, 242-248.

Triesman, A.M. and Tuxworth, J. Immediate and delayed recall of sentences after perceptual processing at different levels. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 38-44.

Umeda, N. and Coker, C. Allophonic variations in American English. Journal of Phonetics, 1974, 2, 1-5.

Umeda, N. Occurrence of glottal stops in fluent speech. Journal of the Acoustical Society of America, 1978, 64, 88-94.

Wingfield, A. and Klein, J.F. Syntactic structures and acoustic pattern in speech perception. Perception and Psychophysics, 1971, 9, 23-25.