This is to certify that the

thesis entitled

Cross Cultural Study on Pause Characteristics
in On-going Speech by
U.S.A. Students and Japanese Students

presented by

Hirotaka Nakasone

has been accepted towards fulfillment
of the requirements for

Master of Arts degree in Audiology and Speech Sciences

Major professor
Dr. Oscar Tosi

Date May 15, 1979

CROSS CULTURAL STUDY ON PAUSE CHARACTERISTICS IN ON-GOING SPEECH BY U.S.A. STUDENTS AND JAPANESE STUDENTS

BY

Hirotaka Nakasone

A THESIS

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

MASTER OF ARTS

Department of Audiology and Speech Sciences

1979

ABSTRACT

CROSS CULTURAL STUDY ON PAUSE CHARACTERISTICS
IN ON-GOING SPEECH BY
U.S.A. STUDENTS AND JAPANESE STUDENTS

By

Hirotaka Nakasone

This study was concerned with the measurement of the pause durations in on going speech uttered by the two language groups, U.S.A. speakers and Japanese speakers. Comparison of the pause distribution patterns was done by classifying the pause continuum into five categories as follow: juncture pauses, phrasal pauses, pauses between words, pauses between syllables, and pauses within syllables. Two different speech modes were used: reading of a short passage and 2 minutes spontaneous speech. 20 subjects (ten males and ten females) from each group participated in the study. Pauses were defined in terms of the acoustical point of view as proposed by Tosi (1974). An electronic pause detector, pausimeter, and associated instruments were applied for the detection and measurements of the pauses. Analysis of the data provided the following findings: (1) Japanese speakers had the longer mean pause durations between and within syllable than U.S.A. speaker did, but no differences were found in the longer pause categories (i.e.,

juncture, phrasal and between-words pauses). (2) In general, males had longer pauses than females within a language group. (3) As a general trend, the difference of the pause durations in longer speech units between Japanese males and females was greater than the difference between U.S.A. counterparts.

Accepted by the faculty of the Department of Audiology and Speech Sciences, College of Communication Arts, Michigan State University, in partial fulfillment of the requirements for the Master of Arts Degree.

Thesis Committee Chairman

Oscar I. Tosi, Ph.D.

ACKNOWLEDGEMENT

I sincerely express my gratitude to my thesis director, Dr. Oscar I. Tosi, for his ideas, effort and support in instrumentation, and encouragement throughout this project. I am indebted to the members of thesis committee, Dr. Leo V. Deal, for his detailed corrections of the draft to the shaping up of this thesis, and Mr. Paul A. Cooke, for his analytical reviews of the draft.

Special thanks are due Ms. Susan J. Mirowski for her painstaking transcriptions of the materials and for her help during the measurement procedure. Her help was invaluable.

Finally, I thank my wife, Aiko, for her support and assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Pauses in on-going speech have been investigated by many researchers from various points of view, by different approaches, and with different definitions applied. Tiffany and Carrel (1977) proposed a classification of types of pauses. One type of pause was termed as "unfilled pause" which was considered to be a genuine interval of silence. Anohter type was called "filled pause," which was further devided into "adventitious fills," "repetition fills," and "steady fills." Another way of viewing types of pauses is to look from the speaker's side, i.e., a pause can be purposefully inserted by a speaker during the speech, or a pause can be a manifestation of uncertainty about what and how to say the next string of phonemes. The former types of pauses are commonly given a name of "syntactic" or "grammatical" pause; the latter is called a "hesitation"pause.

Reviewing previous literature in a studying pauses revealed that there are primarily three types of approaches. One approach has been taken by psycholinguists who studied pauses as one of the significant measures in relation to cognitive process. This approach is represented by studies by Goldman-Eisler (1951, 1952, 1958, 1961, 1968 and 1972), Levin, Silverman, and Ford (1967), Taylor (1969), Kowal,

O'Connell, and Sabin (1975), and Levin and Silverman (1965). According to this approach, pauses are related to the temporal factor in the cognitive process of an individual's speech production. Goldman-Eisler (1951) stated that pause distributional pattern and certain functions in various speech situations were related to personality characters of individuals. Further, Goldman-Eisler (1968) posed that certain ratios of signal and pause tended to be characteristic of individuals and neither were influenced by different speech context nor changed according to different types of listeners. Levin, Silverman, and Ford (1964) summarized that pauses were inversely related to the automaticity of the cognitive process. In relation to language experience of the speaker, Kowal, O'Connell, and Sabin (1975) reported that less language experienced children had greater frequency and duration of pauses in their speech. Taylor (1969) stated that pause latency and hesitation reflect the amount of time of conceptualizing for the sentences to be produced. Levin and Silverman (1965) attempted to classify pauses into 15 types of hesitation variables and three fluency variables, but they differentiated silent pauses from hesitation pauses. Without specifying particular duration values, Levin and Silverman speculated that the very brief pause, the time for taking a breath, could serve an obvious and necessary physiological function. Its increase above a normal base level could be a sign of anxiety; and if the pause were longer, it may indicate that some cognitive organization is taking

place. Another type of approach is seen in linguistic studies in which pauses were studied in relation to syntactic or grammatical structure of a sentence in a spoken language. Fodor, Bever, and Carrett (1974) implied that some pauses may be syntactically conditioned and not reflecting language formulation process. This implication is partially in agreement with the preceeding study result by Hawkins (1971), who reported that most pauses under study were related to specific pauses occurring at syntactic structures. Rochester and Gill (1973) shared the same result with Hawkins, i.e., pauses under study were related to specific syntactic structures. Syntactic pauses occurring at syntactic juncture were compared to hesitation pauses from a perceptual aspect by Boomer and Dittmann (1962). Boomer and Dittmann found that the perceptual threshold for the duration of pauses was higher than for duration of hesitation pauses. In other words, a group of listeners tended to hear more pauses of shorter duration if they occurred at unexpected ungrammatical speech units in a sentence than pauses of the same duration if they occurred at expected grammatical junctures in a sentence. By the use of paired comparisons-psychophysical method these authors established the threshold for hesitation pause to be about 200 msec and for juncture pauses to be somewhere between 500 and 1,000 msec, thus concluding that any arbitrary definition of speech pauses in terms of duration alone would violate certain under-lying linguistic and psychological realties.

In both types of approaches, psycholinguistic and linguistic studies, different assumptions were made with reference to the duration of pauses in on going speech under study. For example, Levin and Silverman (1965) timed pauses by a stop watch and considered only those pauses lasting longer than one second. Goldman-Eisler, by the use of a chronograph in most of her pause studies cited previously, set 250 msec as the minimum pause duration for her study by stating that pauses shorter than 250 msec were related to the normal process of speech production and not related to the internal speech formulation process.

The third approach has been employed primarily by acoustic phoneticians, speech scientists, or speech pathologists to study pauses in running speech with emphasis on the different durational characteristics of pauses as variables and with less concern with pauses as a function of hesitation or as to where in a running speech pauses occurred. Love and Jeffress (1971) attempted to study brief pauses in the fluent speech of stutterers and nonstutterers by using a seriesof electronic equipment including a self-devised counter, called a speech-pause counter. They concluded that stutterers had a significantly greater number of pauses of 150 to 250 msec in length. They discarded pauses below 88 msec on an assumption that pauses of such duration were attributes of the movements of the vocal folds. Pauses of 10 to 25 msec long were also eliminated from the study for a different reason: these were a function of low-energy

fricatives and stop consonants in speech. With the advent of an electronic pause detecting device, such as a pausimeter (Tosi, 1965), it was made possible to detect pauses as short as 5 msec. In reviewing the literature relating to pauses, Tosi (1974) pointed out a chaotic status of studies of pauses because of the lack of common agreement of definition of pauses. Consequently, he proposed the following operational definition of pause:

> Pause is flow of acoustic energy of which the relative amplitudes remain below a predetermined value of a parameter called "pause-maximum amplitude, Lp," provided the duration of such levels of amplitude is more than a predetermined amount of time, indicated by another parameter called "pause-minimum duration, Tp." The parameter Lp is expressed as a percentage or dB ratio with respect to the average peak amplitudes (pressure or voltage) of the rectified waves of the sample of sound analyzed. The parameter Tp is expressed in milliseconds (Tosi, 1974).

Tosi further proposed that the above operational definition of pause should complement rather than replace or conflict with any linguistic, psychological, phonetic, or musical concept of pause and should provide a standard definition of pause to make comparisons possible between different studies of pauses. Indeed, this operational definition of pause from a speech acoustic view allowed an experimental design to measure the lowest perceptual threshold for hearing pauses, which is about 5 msec as suggested by Licklider (1962). Along with this fact, another important aspect of utilization of the operational definition of pauses in the study of speech lies in the fact that it may extend the study of pauses to include so-called voice onset time (VOT),

which is associated with plosive cognates and defined as the time from which the point of the pressure trace begins its greatest drop to the point of the first glottal pulse. VOT is known to be one of major cues which distinguish the voiced-unvoiced plosive cognates and is commonly referred to as the time lag between the moment of peak pressure for plosive consonant and onset of glottal vibration associated with the following vowel (Agnello and McGlone, 1970). The acoustical energy during this time interval can be analogous to "pause," or "articulatory pause," if Tosi's definition of pauses is applied. Studies dealing with VOT from both perceptual and productive aspects are numerous with human subjects (Zlatin, 1974, Lisker, 1975, Lisker and Abramson, 1964, 1967, 1970, and William, 1977) as well as with animal subjects (Kuhl and Miller, 1978). A general agreement among these investigators relating to VOT values is that minimum duration value for VOT (or pause) is somewhere around 5-10 msec; the lowest range was 15-20 msec, and the highest was 100-120 msec. Lisker and Abramson (1964) studied 11 different languages to compare VOT values and found that VOT values varied from language to language. Since the detectable range of pauses is as short as 5 msec if electronic equipment is used, any pauses of this range should be given attention too. Study of pauses by use of such an instrument was coined as "Pausometry." Several pausometric studies were then conducted. Black, Tosi, Singh and Takefuta (1966) studied pause duration characteristics of four different language groups

of native American English, Indian, Hindi, and Japanese speakers. This study resulted in similar distributions across four different language groups when each language group was speaking in its own native language as well as in English. Significant difference was found in the median duration of pauses between two levels of proficiency in English as rated by aural comprehension of English. Tosi and Lashbrook (1969) investigated pauses of 12 male subjects at two significant moments of their circadian rhythm: early in the morning and late afternoon. They reported no statistical differences among pauses. The average median duration of pauses in this study was 200 msec. Physiological status change in the individual induced by the administration of a drug (philocybin) was also studied by Tosi, Fischer, and Rocky (1968). This study yielded statistically significant differences in the median duration values of pauses from a group of subjects with average median duration of 50 msec, which is almost four times as short as pause durations in normal speech condition. However, in their study, all pauses were lumped in the measurement; the nature or types of pauses was not considered. In addition, it was noted by Tosi (1974) that previous studies using a pausimeter described above were concerned with obtaining pause measures as a function of "articulatory" pauses.

Statement of Problem

In spite of vigorous attempts to study the nature of pauses in a speech, there is much confusion about the definition of pauses, especially relating to the duration of pauses or the minimum duration of pauses to be established for investigation. There has been an attempt to provide a uniform definition of pauses from a speech acoustic point of view with primary intention to make comparisons possible across existing study fields of study (Tosi, 1974). To date, however, pausometric studies which applied Tosi's definition did not reflect types of pauses with reference to the nature, function, or locations of pauses in a running speech (Black, Tosi, Singh, and Takefuta, 1966, Tosi, Fischer, and Rocky, 1968, Tosi and Lashbrook, 1969, Love and Jeffress, 1971). A recent study (Deputy, 1978) was concerned with hesitation and articulatory pauses in two groups of children, one with normal speech production and the other with two or more consistent misarticulations; types of pauses were inferred by the different duration of pauses, by arbitrarily setting 10-50 msec for articulatory pauses, 51-25- msec for articulatory and hesitation pauses mixed, and 251-3000 msec for hesitation pauses only. Deputy noted that there was possibly overlapping of articulatory and hesitation pauses in the 51-250 msec category. In deed, the distinction between pauses caused by speaker's hesitation and pauses occurring at syntactic juncture appeared to be practically impossible. Kowal, O'Connell, and Sabin (1975) implied that it is

possible that pauses for linguistic processing and cognitive processing would overlap. From a perceptual point of view, it has also been shown by Boomer and Dittmann (1965) that shorter hesitation pauses were more readily detectable than juncture pauses by a group of listeners. Both researchers cited above concluded that arbitrarily classifying pauses into either hesitation pauses or juncture pauses solely based upon the duration may not be acceptable.

Purpose of the Study

The general purpose of the present study was to measure the duration of pauses distributed in the speech of U.S.A. speakers and Japanese speakers and to compare the pause distributional patterns by breaking down the pause distributional continuum into five speech categories. The operational definition of pauses suggested by Tosi (1974) was used. A designation of "hesitation factor" to any pauses was avoided and filled pauses were not considered to be "pauses" in this study. It was a logical choice when applying Tosi's operational definition of pause along with an electronic pause detector, a pausimeter. The primary reason for employing this design was that the author of this study was particularly interested in the rather brief durations of pauses in U.S.A. speakers and Japanese speakers, as well as in overall pause distributional characteristics across the two different languages, English and Japanese.

Classification of pauses in on-going speech into five

speech categories was not free from problems. The nature of this problem is clearly stated by Goldman-Eisler (1972), who expressed doubt as to whether the stream of spontaneous speech can meaningfully be divided in accordance with grammatical structures. Hawkins (1971) also addressed problems to be overcome in describing the location of pauses in grammatical terms.

This five categories of pauses according to positions where they occurred in reading and in spontaneous speech were as follows:

(1) Juncture pauses (JNP): any pauses occurring at the juncture of syntactic clauses and pauses between sentences.

(2) Phrasal pauses (PHP): any pauses occurring within a clause or sentence, but occurring in a phrasal unit. Phrase was defined as a group of two or more words, yet without constituting a sentence, i.e., not including a verb.

(3) Between-word pauses (BWP): any pauses occurring between two words, or between a word and the following phrase. Word was defined as a minimum unit of spoken language which was bearing grammatical rules and function. The compound word was treated as a single word in this study.

(4) Between-syllable pauses (BSP): all pauses occurring between syllables but within a word except for pauses equivalent to VOT (voice onset time; the

time lag between the moment of peak pressure for plosive consonant and onset of glottal vibration associated with the following vowel).

(5) Within-syllable pauses (WSP): any pauses occurring within a syllable (equivalent to VOT). Pauses categorized in (4) and (5) were later considered as Articulatory pauses.

The specific purposes of the present study were to answer the following questions:

(1) Is there any difference between the native American English speakers (Us) and the native Japanese speakers (Js) in terms of duration of pauses;

a. when Js are speaking in English?

b. when Js are speaking in Japanese?

(2) If there is, what category of pause is different from the other

a. in larger units of a speech structure, such as at sentence or clause juncture (JNP) or in phrasal unit (PHP),

b. in a smaller unit of speech structure as in a word unit (BWP), or

c. in rather brief segments of speech, such as in between-syllable (BSP) or in within-syllable (WSP).

(3) Is there any difference between male speakers and female speakers in terms of duration of pauses in each category:

a. within the same language group?

b. across the two language groups?

II. EXPERIMENTAL PROCEDURE

Subjects

20 native U.S.A. students (10 males and 10 females) and 20 Japanese students (10 males and 10 females) enrolled at Michigan State University participated as subjects, ages ranging from 18 to 31 years for both language groups. All the subjects were free from pathological conditions of speech and from speech defects. All the subjects were volunteers for this study.

Originally, prior to the measurement of pauses, the Js group was divided into two proficiency levels rated in terms of overall Englishness in the reading of a short passage in English by a panel of seven judges consisting of graduate students in speech pathology. The result of this rating was a clear cut line between male and female speakers; nine males fell in the lower ranks, nine females in the upper ranks, and two in the marginal ranks. Because of this result, and in terms of equalizing the number of subjects in sub-groups (by sex), it was believed to be appropriate that the classification was changed from the proficiency level to the two different sexes.

Recording

Speech samples of the subjects were recorded in a sound isolated room (IAC: Industrial Acoustics Company, Inc.) by a high-sensitive dynamic microphone (Unidyne III, model 545) with the average microphone-to-mouth distance of about 15 cm into a high fidelity tape recorder (Voice Identification, Inc., model 700). The input level of each speaker was controlled in order to be equal for each speaker by monitoring a VU meter attached to the tape recorder. Communication between the speaker and the experimenter was conducted through the microphone used for the recording, from speaker to the experimenter, and through a pair of headphones attached throughout the speech production. This circuit served two purposes: one for communication between the speaker and the experimenter and the other for the speaker to monitor his own intensity level of speech production while engaged in reading and in spontaneous speech production. Prior to the production of the speech samples, each speaker had an opportunity to get acquainted with the reading materials, and each was asked to rehearse in the recording booth while the intensity level was being checked. The speaker was informed of the proper speech intensity level and advised to monitor it by listening to his or her own voices through the headphones. The intensity level of the feedback source through the headphones was adjusted to the most comfortable level for each speaker. Spontaneous speech was timed by a stop watch; and as two minutes passed, the recording was

stopped by the experimenter.

Materials

For both U.S.A. and Japanese language group, a passage "If you throw a stone" was prepared (Appendix A). A translated version of the passage was prepared by the author of this study for the Japanese language group only (Appendix B). Speakers of both language groups were given an opportunity to get acquainted with the passage. The Japanese group read the original English passage for about five to ten minutes before entering the booth. U.S.A. subjects read the material at least once before entering the recording booth. The Japanese group also read the Japanese version aloud once before entering the recording booth. A topic to be used for spontaneous speech was left to the speakers, both for Japanese speakers in English and in Japanese and for U.S.A. speakers in English. The speakers were told that they had to maintain the speech for two minutes with as little interruption as possible. For those speakers who could not pick up the topic for spontaneous speech, several choices of pictures were given. After two to three minutes were allowed for the speaker to form an idea to talk about, the subject began speaking. When two minutes had passed recording was stopped. Recordings were made of one speaker at a time; each language group took about a week, totaling about two weeks for the entire recording sessions. The instructions given as to the rate of reading and the level of speech

production were as follows:

> When you read the passage in English (or in Japanese only for Japanese speakers), please keep the following two things in your mind. (1) Read the passage at your most comfortable reading rate and loudness level. (2) Imagine that you have to pass informations in the passage as accurately as possible to your listener. For your two minutes speech, the topic is left to you. You will talk in such that you are talking to your friends, to a group of classmates, or to your members of family. Should you see a difficulty to select a topic you want to talk, you may look at picture cards prepared in the folder placed in front of you to start with your speech. You do not have to describe the picture, but use these pictures in aiding you to form an idea of your speech. I will give you a sign to start your speech after you have decided what to talk about.

In the course of recording the speaker, recording conditions were monitored by the author through the headphones attached to the main tape recorder. When the speaker was stuck in the middle of reading or spontaneous speech because of nervousness or when the speaker made excessive noise by gestures around the microphone, he was informed of such activity and the recording was repeated.

Instrumentation

Figure 1 is a block diagram of instrumental design with a pausimeter and associated equipment for detection and measurement of pauses in speech. Since the explanation of the circuitory designs and functions of this pausimeter has been made elsewhere (Tosi, 1974, Deputy, 1978), only a brief description referring to this equipment will be made.

This equipment is composed of three major circuits, referred to as Block I, Block II, and Block III. Block I consists of an amplifier, an attenuater, and a rectifier.

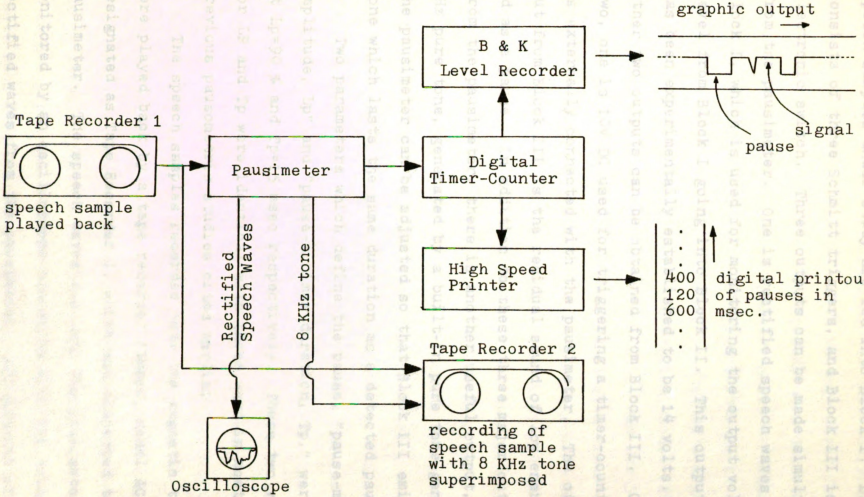


Figure 1. Block diagram for pause detection.

Speech samples under study are fed into Block I. Block II consists of three Schmitt triggers, and Block III is an electronic switch. Three outputs can be made simultaneously from the pausimeter. One is rectified speech waves from Block I, which is used for monitoring the output voltage level from Block I going into Block II. This output level has been experimentally established to be 14 volts. The other two outputs can be obtained from Block III. Of the two, one is 10v DC used for triggering a timer-counter which is externally connected with the pausimeter. The other output from Block III is the residual sound of low energy detected as a pause. In addition to these three major outputs from the pausimeter, there is another useful output, an 8 KHz pure tone, generated by a built-in pure tone generator. The pausimeter can be adjusted so that Block III emits this tone which lasts the same duration as a detected pause.

Two parameters which define the pause, "pause-maximum amplitude, Lp" and "pause-minimum duration, Tp," were set at Lp=90 % and Tp=15 msec respectively. These two values for LP and Tp were identical with those used in most of the previous pausometric studies cited earlier.

The speech samples recorded onto the magnetic tapes were played back by a tape recorder (Ampex, model AG 600), designated as Tape Recorder 1, which was connected to the pausimeter. The speech waves fed into the pausimeter were monitored by an oscilloscope connected with the outlet of rectified waves from the pausimeter. Adjustments of the

attenuator and amplifier were made for each individual speech samples so that the average output voltage level from rectifier remained at 14 volts. When pauses were detected, the pausemeter triggered a timer-counter (Hewlett-Packard, model 5326 B) for digital displays of pause durations in msec. These displays were simultaneously printed by a high-speed printer (Mohawk Data Sciences Corporation, model 1200). A level recorder (Bruel & Kjar, model 2305) was fed 5v signals from the event marker on the rear panel of a timer-counter. The durations of DC signals from this marker were in synchrony with the durations of pauses as displayed on the timer-counter and also with the printed output from high-speed printer. Another tape recorder (Magnecord, model 1022), designated as Tape Recorder 2, was prepared to record speech samples directly coming out from Tape Recorder 1 and to record the 8 KHz pure tone signals generated by a built-in pure tone generator in the pausimeter. The pausimeter was adjusted so that it generated 8 KHz pure tones when pauses were detected. The tone lasted the same duration as that of the detected pause. The original speech sample was recorded on channel 1, and the 8 KHz tones were recorded on channel 2 of the Tape Recorder 2. This recording of the original speech samples with 8 KHz tones superimposed was later used for determining exact locations of pauses ocurring in running speech. The level recorder was operated at 10 cm/sec paper speed and cue words from on going speech as played back were written on the running graphic paper by

another experimenter, a native American English speaker having experiences in phonetics and majoring in Speech Sciences. When the speech samples under process were spoken in Japanese, the author wrote the cue words.

Through the set of instrumentation described above, three outputs for each speech sample were produced: (1) graphic displays on the level recorder on which X-axis was time scale, and on Y-axis 5v DC for existing signal and 0v DC for pauses; (2) digital outputs in msec by a high-speed printer; and (3) the speech samples re-recorded with 8 KHz pure tone superimposed.

Reliability of Instruments

Reliability of the pausimeter and connected instruments was checked prior to the operation. One of the recorded speech samples was played back five times repeatedly through the set of instruments with the specifications of two parameters (Lp=90 %; Tp=15 msec). The average error rate was computed on the basis of the average deviation of duration of pauses in per cent between the first and the second, the third, the fourth, and the fifth run. It was found that the average error rate of durations of pauses in msec was less than five per cent. The same procedure was repeated every 30 minutes during the operation of the instruments to check the error rate. The error rate throughout the entire operation remained quite constant.

Determination of Location of Pauses

In order to determine the location of pauses within the speech sample analyzed and to assign correctly the measured duration of pauses, based on the three outputs described above, the following steps were followed.

Step 1. Phonetic transcripts were made for all spontaneous speech samples (Example, see Figure 2. c), but only one transcript was made for reading samples, one in English (Appendix C) and one in Japanese (Appendix D). The phonetic transcripts for English were prepared by a native English speaker who had experience in phonetic transcription and who was majoring in Speech Sciences. Transcripts for Japanese speech samples were prepared by this author.

Step 2. Digital outputs indicating durations of pauses printed on a strip of a printer paper (Figure 2. b) were matched and transferred to the zero signal regions on the level recorder graphic output (Figure 2. a).

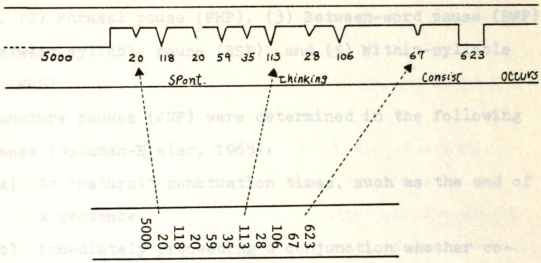
Step 3. The re-recorded speech samples with the 8 KHz pure tone superimposed were played back at a reduced rate of tape speed, four times slower than the original recording rate. While listening to the recorded speech sample at the reduced rate, the experimenter read through the transcript of the speech samples. When an 8 KHz tone was heard, a check (indicated by a horizontal line drawn above the transcript in Figure 2.

c) was made in the corresponding segment of the transcript. This checking was made carefully, by first listening to the tape at original rate of speed then at the rate of half of the original and, finally, at the slowest rate -- one-fourth rate of the original.

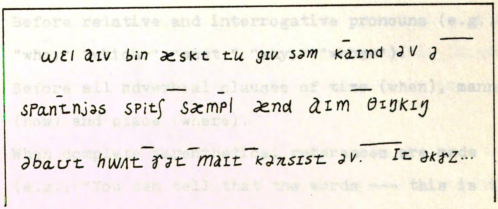
Step 4. Each digit indicating the duration of a pause which had been transferred to the level recorder graphic output prepared in Step 2 was then transferred to the transcripts prepared in Step 3. Cue words written on the graphic paper approsimately every five or six words during the operation of the instruments served to locate either the ending or the beginning of a sentence or a group of words in the transcript.

Figure 2(a-c) is a diagram showing the procedure with three outputs from a spontaneous speech sample.

a. (The graphic output from the level recorder with cue words written. The base line or dipped regions indicate the pauses detected. The digits on this graph were transfered from output b in the sequence.)



b. (The printer output. Each number indicates the duration of pause in msec.)



c. (The pausimeter output consisted of 8 KHz tones superimposed onto speech samples. When the tones were heard during the playback, they were marked by horizental lines on top of speech sound transcribed indicating the position of the pause occurred.)

Figure 2(a-c). Pause outputs from the Level Recorder(a), the Printer(b), and the Pausimeter(c).

Categorization of Pauses

The pause and its duration entered in the transcript were classified into five categories: (1) Juncture pause (JNP), (2) Phrasal pause (PHP), (3) Between-word pause (BWP), (4) Between-syllable pause (BSP), and (5) Within-syllable pause (WSP).

1. Juncture pauses (JNP) were determined in the following cases (Goldman-Eisler, 1968):
   (a) At "natural" punctuation times, such as the end of a sentence.
   (b) Immediately preceeding a conjunction whether co-ordinating (e.g., "and," "but," "neither," "therefore") or subordinating (e.g., "if," "when," "while," "as," "because").
   (c) Before relative and interrogative pronouns (e.g., "who," "which," "what," "why," "whose").
   (d) Before all adverbial clauses of time (when), manner (how) and place (where).
   (e) When complete parenthetical references are made (e.g., "You can tell that the words --- this is the phonetician speaking --- the words are not sincere").
2. Phrasal pauses (PHP) were determined in the following cases:
   (a) When a pause occurred at the end of a phrase and the smallest unit of a phrase consisted of two words. For example, pauses in "I will go / to the meeting" or "The house / was yellow" are considered

to be PHP.

(b) When the repetition of pauses in two or more words units occurred.

3. Between-word pauses (BWP) were determined in the following cases:

(a) When pauses occurred after a single word, or a single compound word.

(b) When pauses occurred in the middle of a phrase.

4. Between-syllable pauses (BSP) were determined in the following cases (This category includes all the pauses within a word except those equivalent to so-called VOT):

(a) When pauses occurred within a word, regardless whether a word was mono- or multisyllabic words, or in a compound word.

(b) When pauses occurred within a consonant cluster in a word, such as "g/ret" (great), "spar/kl" (sparcle).

5. Within-syllable pauses (WSP) were determined in the following cases:

(a) When pauses occurred within a syllable, such as in plosive phoneme + vowel, "p/arti" (party), "t/ɛIkIŋ" (taking).

Analysis of the Data

The pause duration measured in msec was the dependent variable in this study. Three independent variables were two native speaker groups, two speech modes, and two different sexes. Further, the pause durations in the conditions

described above were measured in five pause categories.

The mean duration of pauses for each speaker in a given condition in each pause category was computed. Then, mean pause durations according to sex groups were computed and tabulated in tables. From these tables bar graphs were also plotted. Finally, the data were analyzed by two sets of four three-way analysis of variance (ANOVA) . The analysis was conducted by use of IBM 6500 at Michigan State University. One set of four ANOVA was conducted for the comparison between the group of U.S.A. English speakers (Us) and the group of the native Japanese speakers (Js) utterring in English, in two speech modes, reading and spontaneous speech, and by two sexes. Another set of ANOVA was used for the comparison between Us utterring in English and Js in Japanese in the same conditions. Out of five categories, JNP, PHP, BWP, and BSP were treated by ANOVA. But WSP was discarded from the ANOVA and treated separately because of the rare incidence of pauses of its category in Js.

III. RESULTS

The results of this study will be discussed in three sections. In the first section, a general view of the results will be presented. In the second section, the results of the pause durations in running speech when the both speaker groups were speaking in English for each condition will be presented. The third section will be related to the results of the comparisons of pause durations of Js performing the task in their own native language, Japanese, with those of Us in English.

General Interpretation of the Results

Figure 3 is a graphic representation of the grouped mean pause durations of Us, Js in English, and Js in Japanese in reading speech mode, discriminated by sex, and in the five pause categories: JNP, PHP, BWP, BSP, and WSP. Figure 4 illustrates the same results as Figure 3 except speech mode was spontaneous speech. These two figures were prepared from the summarized tables of the grouped mean pause durations (Appendix E). By visual inspections it was found that the pause durations in spontaneous speech mode were constantly longer than those in reading mode regardless the sex and the group difference. The differences of pause durations, however, for the categories of BSP and WSP in

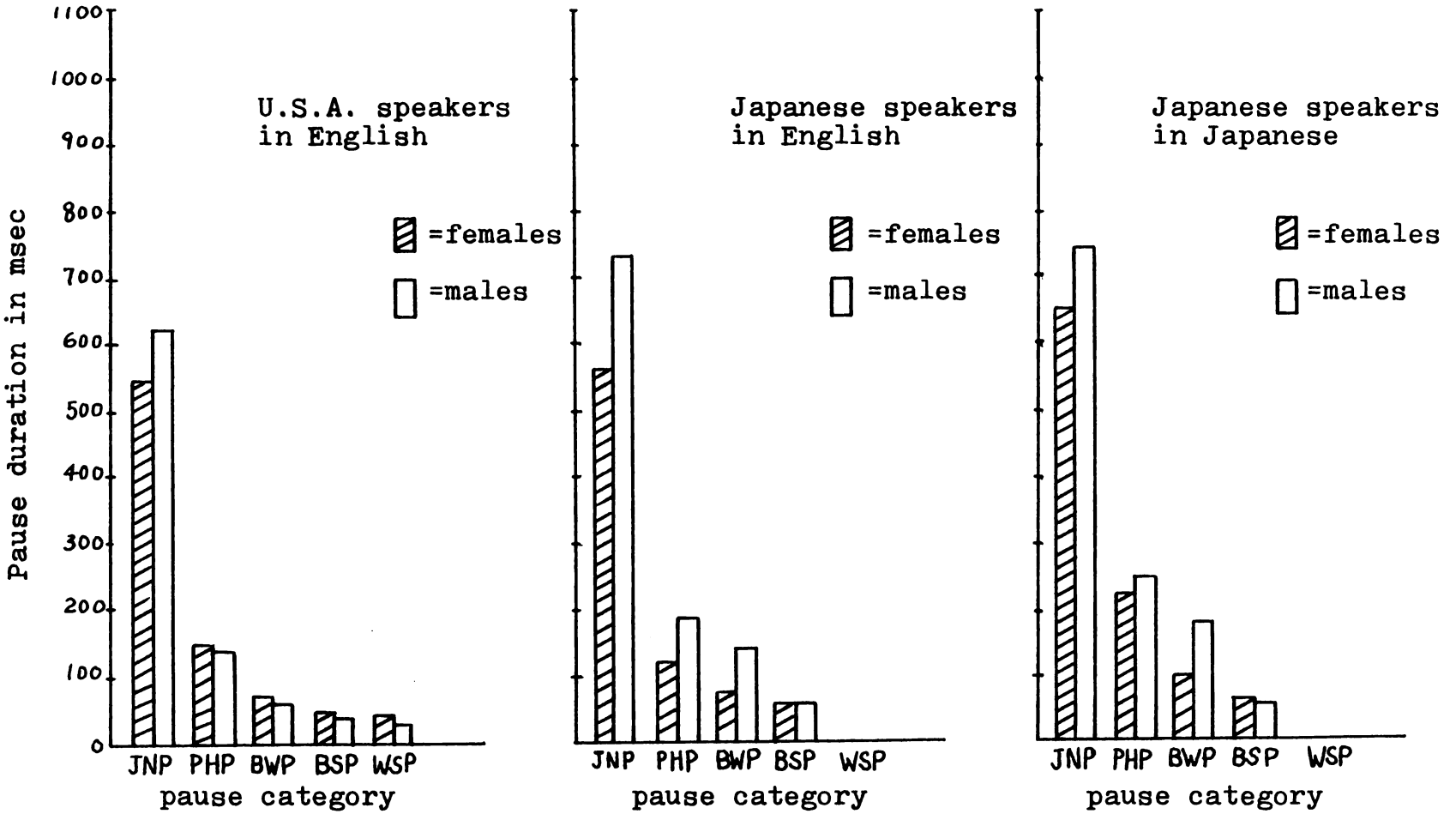


Figure 3. Graphs of grouped mean pause durations of five categories by U.S.A. speakers in English and by Japanese speakers in English and in Japanese in the task of reading of a short passage, discriminated by sex. (Five categories are: JNP=pauses occurring at grammatical juncture; PHP=pauses occurring in phrasal unit; BWP=pauses occurring in the word unit; BSP=pauses occurring between syllables but within a word; and WSP=pauses occurring within a syllable.)

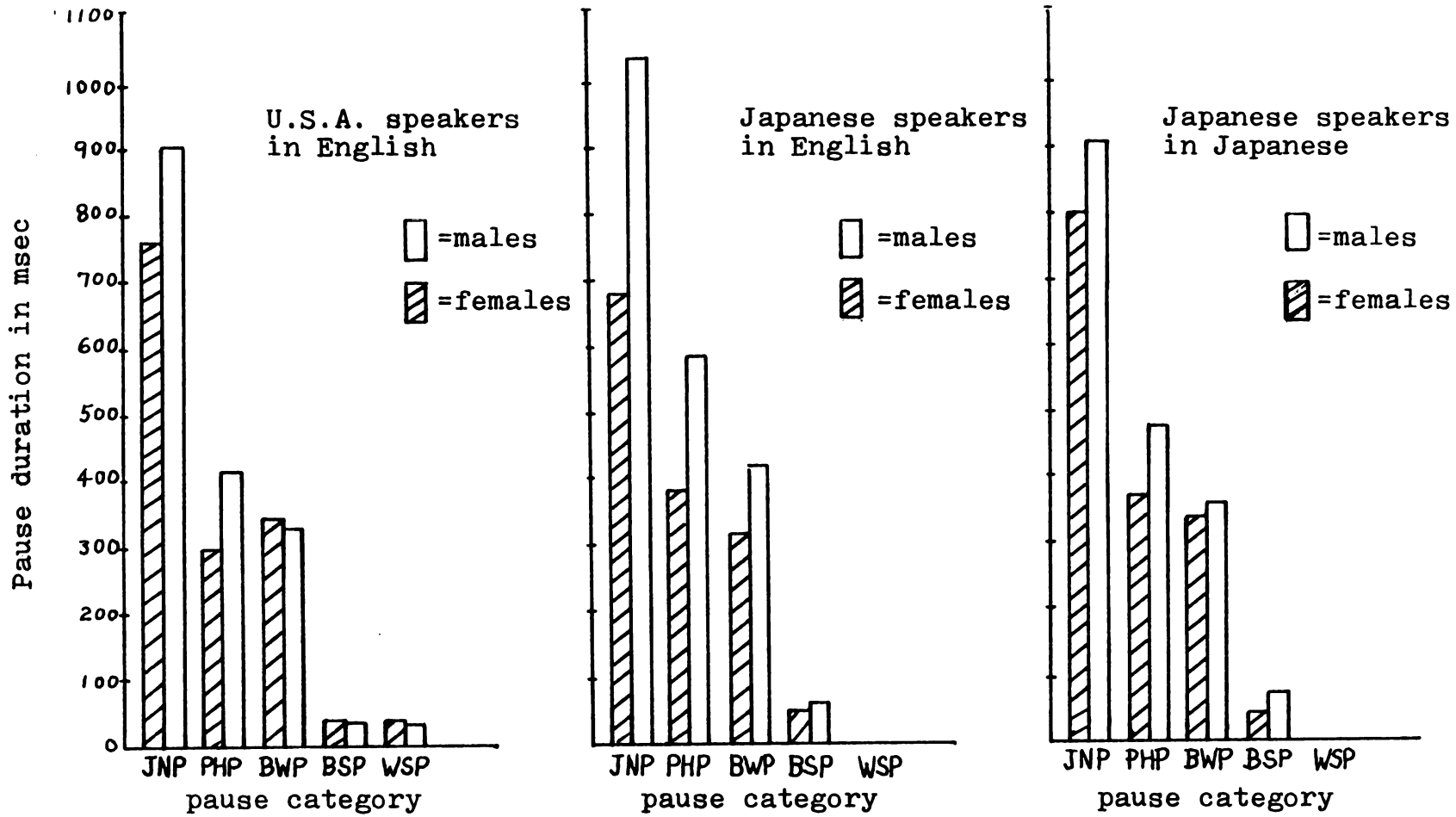


Figure 4. Graphs of grouped mean pause durations of five categories by U.S.A. speakers in English and by Japanese speakers in English and Japanese uttering spontaneous speech, discriminated by sex. (Five categories are: JNP=pauses occurring at grammatical juncture; PHP=pauses occurring in phrasal unit; BWP=pauses occurring in the word unit; BSP=pauses occurring between syllables but within a word; and WSP=pauses occurring within a syllable.)

respect to speech modes diminished to negligiable amount. The mean pause duration in reading in the category of JNP for two speaker groups ranged from 539.0 msec (Us females) to 734.1 msec (Js males in Japanese). In spontaneous speech the pause durations in JNP were relatively longer than those in reading mode and ranged from 682.0 msec (Js females) to 1035 msec (Js male in English). The next longest pauses were found in the category of PHP, ranging from 120.6 msec (Js females in English) to 244.7 msec (Js males in English), then followed by pauses categorized as BWP, of which mean pause durations ranged from 60.3 msec (Us males) to 180.4 msec (Js males in Japanese). Pause durations in BSP in the reading task Us yielded relatively shorter values (48.8 msec for females and 42.1 msec for males) than Js both in English and in Japanese. In the last category, WSP, pauses were totally missing in Js, whereas Us males and females produced 41.6 and 31.3 msec, respectively. The data discussed above were then subjected to the two sets of analysis of variance. The first analysis of variance was conducted for the purpose of statistical analysis of the pause durations in the four pause categories by Us and Js both performing in English language, and discriminated by sex (WSP was discarded from the analysis because of absence of pauses of this category in Js). The second analysis of variance was conducted for same purpose except that Js were performing the tasks in their own native language.

Comparison Between U.S.A. and Japanese Speaking English

A. JNP: Table 1 shows the summary of ANOVA of the duration of JNP in msec by 20 Us and Js in English language, discriminated by sex, and in two speech modes, reading and spontaneous speech. It was found that pause durations in JNP (pauses occurring at the end of a sentence and at syntactic junctures) were not different for the Us and Js groups. Main effects of sex and speech mode were found to be statistically significant, ($p=0.01$, and $p=0.01$, respectively). Neither two-factor interaction nor three-factor interaction was significant.

B. PHP: Table 2 represents the summary of ANOVA of the duration of pauses in PHP (pauses occurring between phrases) by Us and Js in the same experimental condition as in JNP. Again, there was no significant difference in the mean pause duration for the two language groups. And quite similarly, sex and speech mode main effects were significant ($p=0.05$, and $p=0.001$, respectively). No two or three-factor interaction was statistically significant.

C. BWP: Table 3 shows the summary of ANOVA of duration of pauses in BWP (pauses occurring in the word unit) by Us and Js in the same experimental conditions. The difference of the mean BWP between Us and Js, here at word level, was not significant. In this category of pause, only the main effect of mode was found to be significant at the lefel of 0.001. No interaction effect was significant.

D. BSP: Table 4 shows the summary of ANOVA of the

Table 1. Summary of analysis of variance of the mean duration of JNP by U.S.A. and Japanese speakers in English.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	38874.153	1	38874.153	.531
Sex (B)	699099.528	1	699099.528	9.548 **
Mode (C)	1093880.151	1	1093880.151	14.940 ***
Two-way Interaction				
A x B	108847.635	1	108847.635	1.487
B x C	9050.385	1	9050.385	.124
C x A	74865.966	1	74865.966	1.022
Three-way Interaction				
A x B x C	19659.585	1	19659.585	.269
Residue	5271816.911	72	73219.679	

** Significant at p=0.01
*** Significant at p=0.001

Table 2. Summary of analysis of variance of the mean duration of PHP by U.S.A. and Japanese speakers in English.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	21813.013	1	21813.013	2.267
Sex (B)	77066.113	1	77066.113	8.008 **
Mode (C)	1087644.800	1	1087644.800	113.019 ***
Two-way Interaction				
A x B	123.008	1	123.008	.013
B x C	4721.664	1	4721.664	.491
C x A	11592.112	1	11592.112	1.205
Three-way Interaction				
A x B x C	19207.602	1	19207.602	1.996
Residue	692898.818	72	9623.595	

** Significant at p=0.01
*** Significant at p=0.001

Table 3. Summary of analysis of variance of the mean duration of BWP by U.S.A. and Japanese speakers in English.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	33472.562	1	33472.562	3.003
Sex (B)	30811.250	1	30811.250	2.791
Mode (C)	1424391.938	1	1424391.938	129.045 ***
Two-way Interaction				
A x B	45849.888	1	45849.888	4.154
B x C	963.272	1	963.272	.087
C x A	207.368	1	207.368	.019
Three-way Interaction				
A x B x C	1748.450	1	1748.450	.158
Residue	794731.060	72	11037.931	

*** Significant at p=0.001

Table 4. Summary of analysis of variance of the mean duration of BSP by U.S.A. and Japanese speakers in English.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	4687.922	1	4687.922	6.512 *
Sex (B)	991.232	1	991.232	1.377
Mode (C)	.480	1	.480	.001
Two-way Interaction				
A x B	2984.124	1	2984.124	4.145
B x C	1054.152	1	1054.152	1.464
C x A	357.858	1	357.858	.497
Three-way Interaction				
A x B x C	7.081	1	7.081	.010
Residue	51830.610	72	719.870	

* Significant at p=0.05

duration of pauses by Us and Js in BSP (pauses occurring between syllables but within a word). In this category of pauses, as seen in Table 4, only group effect was statistically significantly different (p=0.01). No interaction effect was significant. The duration of pauses was quite similar for both sexes, in two speech modes, and within a group.

E. WSP: Pauses of this category was not included in ANOVA because its incidence with Js was rarely observed when they are speaking or reading in English.

Comparison Between U.S.A. and Japanese Speaking Their Native Language

A. JNP: Table 5 represents the results of ANOVA of the mean durations in JNP for the two native language groups, Us in English and Js in Japanese, discriminated by sex, and in two speech modes, reading and spontaneous speech. According to the analysis, there was no significant difference in this category between Us group in English and Js in Japanese. The main effect of sex in this category on the mean duration of pause was not significant, and two speech modes produced a quite obvious statistical difference (p=0.001). Two-factor, or three-factor interaction was not significant.

B. PHP: Table 6 represents the result of ANOVA of the mean duration of pauses in PHP for the two native language groups, discriminated by sex, in two speech modes, reading and spontaneous speech. The mean duration of PHP for Us

was significantly different from that of Js ($p=0.01$). Sex effect and speech mode effect were both significant ($p=0.05$, and $p=0.001$, respectively). There was no interaction effect.

C. BWP: Table 7 shows the result of ANOVA of mean duration of pause in BWP for the two native language groups, discriminated by sex, in two speech modes, reading and spontaneous speech. In this category of pauses, no factor was found to be significantly different except the main effect of speech mode ($p=0.001$).

D. BSP: Table 8 is the result of ANOVA of the mean duration of BSP in msec for the two native language groups, discriminated by sex, in two speech modes, reading and spontaneous speech. In this category of pauses there was only one main effect of the native language group (significant at $p=0.001$). Language group-sex interaction was significant ($p=0.05$).

Table 5. Summary of analysis of variance of the mean duration of JNP by U.S.A. speakers in English and Japanese speakers in Japanese.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	73041.741	1	73041.741	1.007
Sex (B)	231243.265	1	231243.265	3.187
Mode (C)	870925.980	1	870925.980	12.002 ***
Two-way Interaction				
A x B	641.278	1	641.278	.009
B x C	43175.278	1	43175.278	.595
C x A	7459.453	1	7459.453	.103
Three-way Interaction				
A x B x C	2212.356	1	2212.356	.30
Residue	5224777.643	72	72566.356	

*** Significant at p=0.001

Table 6. Summary of analysis of variance of the mean duration of PHP by U.S.A. speakers in English and Japanese speakers in Japanese.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	137116.800	1	137116.800	9.913 **
Sex (B)	79279.232	1	79279.232	5.732 *
Mode (C)	820935.200	1	820935.200	59.352
Two-way Interaction				
A x B	50.881	1	50.881	.004
B x C	4642.104	1	4642.104	.336
C x A	45668.124	1	45668.124	3.302
Three-way Interaction				
A x B x C	1059.968	1	1059.968	.077
Residue	995881.378	72	13831.686	

* Significant at p=0.05
** Significant at p=0.01
*** Significant at p=0.001

Table 7. Summary of analysis of variance of the mean duration of BWP by U.S.A. speakers in English and Japanese speakers in Japanese.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	22271.138	1	22271.138	2.161
Sex (B)	10769.440	1	10769.440	1.045
Mode (C)	1240020.000	1	1240020.000	120.300 ***
Two-way Interaction				
A x B	20269.345	1	20269.345	1.966
B x C	12310.722	1	12310.722	1.194
C x A	6643.012	1	6643.012	.644
Three-way Interaction				
A x B x C	2925.781	1	2925.781	.284
Residue	742154.622	72	10307.704	

*** Significant at p=0.001

Table 8. Summary of analysis of variance of the mean duration of BSP by U.S.A. speakers in English and Japanese speakers in Japanese.

Source of Variation	SS	DF	MS	F
Main effects				
Group (A)	7091.378	1	7091.378	14.052 ***
Sex (B)	304.981	1	304.981	.604
Mode (C)	590.785	1	590.785	1.171
Two-way Interaction				
A x B	4712.450	1	4712.450	9.338 **
B x C	55.778	1	55.778	.111
C x A	1416.245	1	1416.245	2.806
Three-way Interaction				
A x B x C	456.968	1	456.968	.906
Residue	36333.892	72	504.637	

** Significant at p=0.01
*** Significant at p=0.001

IV. DISCUSSION AND CONCLUSION

General Discussion of Results

The purpose of this study was to measure the duration of pauses distributed in different speech units and to compare the measured pauses between the two different native language groups by applying the operational definition of pauses suggested by Tosi (1974). 20 native American English speakers (Us) and 20 native Japanese speakers (Js) participated in this study as subjects. Further, the pause duration was measured and compared as dependent variable of three different factors. The first factor was related to two different native language groups, Us and Js. The second factor was associated with sex. The third factor was two different speech modes, reading and spontaneous speech.

Although the Us groups were measured of their pauses only in their native language, American English, the Js group performed the tasks in English as well as in their native language, Japanese. Therefore, there were actually three groups of subjects. The first group consisted of Js speaking in English, the second group consisted of Us in English, and the third group, Js speaking in their own native language, thus creating two comparison sets. The first set of comparisons was made between two groups, Us and Js

both speaking in English language; and the second set of comparisons was made between Us speaking in English and Js speaking in Japanese. Male speakers and female speakers in each group were treated separately. Each speaker made two types of speech samples: reading a passage lasting on the average of 45 seconds and speaking for two-minutes spontaneous topic. Reading was included for the purpose of establishing a base line to be compared with actual on-going speech. Then measurements of pauses produced by each speaker in the conditions stated above were made according to where they actually occurred.

In general, this study revealed that Us and Js as groups did not differ in terms of the mean durations of JNP and PHP as long as both groups were performing in their own native language. When Js were performing in English, however, they produced relatively longer mean pause durations in PHP than Us did. An indication of this result is that Us and Js shared the same pause durational values in greater speech units when they spoke in their own mother language. In view of the duration of pauses in this category, the study by Boomer and Dittmann (1962) is worth of mentioning. In their study it was reported that the juncture pauses were around 500 to 1000 msec, the range which the pause durations in JNP in the present study fell in.

As to pauses in a smaller unit of speech structure as in a word unit, BWP, both speaker groups were not significantly different, whether Js were performing the task in

English or in Japanese. Differences in pause durations for two groups became somewhat distinct in brief segments of speech as in BSP and WSP. Duration values of pauses categorized as BSP (pauses occurring between syllables but within a word except for pauses equivalent to VOT) were constantly longer than those of Us's whether in English or in Japanese and whether in reading or in spontaneous speech. Although the differences of pause durations of this category between two groups were relatively small, the average difference in the order of 10 msec, the statistical analysis showed a significant difference. Js had greater BSP values, at the level of 0.05 for the comparison between Js in English and Us, and of 0.001 for the comparison between Js in Japanese and Us. A special notion about the pauses of this category is that neither a factor of speech modes, reading or spontaneous speech, nor a factor of sex had significant effect upon the pause duration while the rest of the categories, JNP, PHP, and BWP did. To explain what possibly accounted for this difference of BSP values for Js and Us requires more research. The best assumption, however, is that it could be attributable to a certain phonological system present in the Japanese language and was carried over into the articulatory activity in English. Among many possible unknown accounts, one candidate is a phonetic phenomenon associated with so-called "soku-on" (Fukui, 1978). This phenomenon "soku-on" is defined as an extended duration of silent interval inserted by Japanese speakers between

the vowel and the following plosive consonant in CVC context. According to the author's informal observation, this type of silent interval, "soku-on," occurred in Js's speech samples frequently and detected as pauses by the pausimeter. Since they occurred between syllables and within a word, they were all categorized as BSP.

Pauses categorized as WSP yielded a contrasting result between the two speaker groups. WSP was defined for the purpose of this study to be a pause detected pausometrically within a syllable. Specifically, the pauses of this category were considered to be equivalent to so-called VOT. The pausimeter rarely detected WSP in Js's speech productions neither in English nor in Japanese. The overall average WSP values for Us in the present study were 36.7 msec and 38.4 msec in reading and in spontaneous speech, respectively. The above figures, when considered to be equivalent to VOT, roughly agreed with various studies (Zlatin, 1974, Lisker, 1975, Lisker and Abramson, 1964, 1967, 1970 and William, 1977).

From the fact that Japanese phonology contains plosive cognates such as /p/ and /b/, /t/ and /d/, and /k/ and /g/ and from the previous research evidences that VOT value is one of the important cues in discrimination of the voiced plosives from the unvoiced plosives, it would be reasonable to assume that VOT of some duration must be accompanied in the production of plosives by Japanese speakers. Therefore, the absence of pauses categorized as WSP, or equivalently voice onset time in Japanese speakers were probably caused

by too high setting of the parameter, Tp (the pause-minimum duration), thus escaping the detection by the pausimeter. This parameter was kept constant with Tp=15 msec for all the speech samples processed. If this parameter, Tp, had been set lower value for Js, then WSP might have been detected but in the duration range shorter than 15 msec.

Sex Effects on Pause Duration

When the pause duration was viewed from the point of the sex effect, it was found that the mean durations of pauses by males were longer than females in JNP and PHP for both speech modes when Js were performing in English. In addition, comparing native languages, males had longer PHP during spontaneous speech. Again, as in the case of the mode, sex difference had no influence upon the pause durations in the categories of BSP and WSP. Although it was not revealed by statistical analysis, a general trend of pause durations as a function of sex in the two speaker groups was that the differences of the mean pause durations between Js males and females in the categories of JNP, PHP, and BWP both in English and in the native language were greater than the difference between Us males and females in the same pause categories. An indication drawn from the above result is that when the pauses are to be investigated in the longer speech units, such as juncture pauses and phrasal pauses used in this study, care should be exercized about treatment of sex differences with Japanese subjects, but not so with

the case of U.S.A. subjects. In order to see whether the reading rate of speech samples were in any way associated with the effects of sex difference, a table of the average reading time required for the subjects to complete the reading task was prepared (Appendix F). The average reading time for each pair of sexes in each group was treated by two-tailed t-tests. None of pairs of comparisons resulted in significant difference at 0.05 level. Nevertheless, the different trend for two language groups was clear. The average reading rate of Us males was slightly shorter than that of Us females, though Js females had the shorter average reading rate than Js males did.

Speech Mode Effect on Pause Duration

The subjects were constantly producing the greater mean pause durations during spontaneous speech as compared to reading in the pause categories of JNP, PHP, and BWP for Us and for Js performing both in Japanese and in English. The mode effect, however, disappeared in the pause category of BSP and WSP. Other factors held constant, the sole difference the mode effect produced was dividing line between three longer speech units, JNP, PHP, and BWP and two brief units, BWP and BSP. In other words, the mean pause durations in BSP and WSP in reading and in spontaneous speech were identical within Us and Js (both in English and in Japanese), regardless of the sex difference. Since the pauses in BSP and WSP were measured within the word in this

study, the above finding can be related to the remark by Hawkins (1971) who considered that the word was the smallest unit where a hesitation pause can occur, not within the word. Further, if the well-established notion about hesitation pauses are related to the cognitive process in speech formulation, then it leads to a conclusion that the pauses categorized as BSP and WSP in this study are not related to hesitation phenomena in speech process but rather limited to articulatory levels of speech production.

Implication for Further Research

Implicit in the results of this study is that within a language group there is a dividing line among the five categories of pause, viz., pauses in brief segments of speech and those in longer segments such as juncture pauses, phrasal pauses, and the pauses occurring in the word unit. Since that dividing line was found to be somewhere around 40 to 65 msec (general mean pause durations in BSP), it would be worth lowering the minimum pause duration to the values stated above, provided an adequate instrument is available.

With reference to the pause category of WSP (equivalent to VOT), it would be a wise exercise to study the VOT values by a certain language speaker group under study prior to the establishment of the parameter, pause-minimum duration, to prevent brief pauses from escaping the detection by the pausimeter.

Although no attention was given to hesitation phenomena

associated with pause durations in the present study, there is a suggestive notion on the phenomena which emerged from the measurement design applied in this study.

With an assumption that the reading task involves only articulatory and syntactic pauses but does not require as much pause duration for cognitive process as spontaneous speech does, then the extra amount of the pause duration values in the spontaneous speech may be considered as pause duration due to hesitation. By subtraction, then, the pause durations due to hesitation occurring in a particular speech unit in spontaneous speech can be separated from the those in the same speech unit in reading. If the separations of the pauses into syntactic, hesitation, and articulatory pauses according to where in running speech they occur, the research would be more fruitful inasmuch as the purpose is to study pauses as a measure of cognitive process involved in speech formulation.

Conclusion

The present study showed that the mean pause durations in BSP and WSP, i.e., pauses occurring between and within-syllables, differed for the two speaker groups, U.S.A. speakers (in English) and Japanese speakers (both in English and in Japanese). The Japanese speakers had longer mean pause durations than U.S.A. speakers did in the former category while they produced no pauses in the latter category. Further, the pause durations of these two categories were found to be homogeneous between the sexes and in the two

speech modes, reading and spontaneous speech, within each language group.

In contrast, the mean pause duration in longer speech units (i.e., juncture, phrasal, and between-word) did not differ significantly between the two groups. However, in these three longer speech units, spontaneous speech mode yielded the greater mean pause durations than reading mode did across the two groups, whereas the different sexes in each group produced the pause durations in different degrees. In other words, the difference of pause durations between Japanese males and females utterring in English and in Japanese was greater than the difference between the U.S.A. males and females.

A conclusion implied by the results stated above is that the difference of the pause durations between the two language groups was associated with articulatory process of speech production and that the variability of the mean pause durations within the language group emerges in speech units as short as the word unit.

APPENDICES

APPENDIX A

A READING PASSAGE IN ENGLISH

If you throw a stone into a quiet pond, you see small waves start out from the place where the stone fell into the water. These waves spread steadily outward in circles. A chip of wood floating on the water will bob up and down as the waves pass. This happens because motion is handed on from the stone to the chip by means of the waves. After the set of waves has gone by, the surface of the water is again quiet and the chip is still.

The waves on the pond give you a picture in slow motion of what happens when you hear a sudden noise. An exploding firecracker disturbs the air around it, just as the stone disturbed the water. Waves move out through the air in all directions. When these waves hit your ear, they pass some of the motion on to your eardrum, and you hear the sudden sound.

APPENDIX B

A READING PASSAGE IN JAPANESE

小石を静かな池の中に投げるとそれが落ちたあたりからいくつもの小さな波が外側に向って走っていくのが見られます。波は着実な速度で広がり続けます。水面に浮かぶ板切れは波が通り過ぎる都度に上下に揺れるでしょう。その理由は小石によって生じた運動が波により板切れへと転移されるからなのです。そして波が消え去ってしまうと，池の水も元の静かなたたずまいに戻り，板切れもピタと動きを止めます。

この池の波はあなたが突然音を聞くときの様子をスローモーションで鮮明に物語っているのです。花火の爆発音はまさに小石が水を乱すごとく周囲の空気を乱し，その中をあらゆる方向へと散らばって行きます。この空気の波があなたの耳に届く時，波は鼓膜に伝えられ，それが音として聞こえるのです。

APPENDIX C

PHONETIC TRANSCRIPT FROM ENGLISH PASSAGE

ɪf ju θro ə ston ɪntu ə kwaɪət pɔnd, ju si smɔl weɪvz stɑrt aut frɔm ðə pleɪs hwɛr ðə ston fɛl ɪntu ðə watɚ. ðɪz weɪvz spred stɛdɪ autwɚd ɪn sɝkɪlz. ə tʃɪp ɛv wʊd flotɪŋ ɔn ðə watɚ wɪl bɔb ʌp ənd daun æz ðə weɪvz pæs. ðɪs hæpnz bɪkɔz moʃən ɪz hændɪd ɔn frɑm ðə ston tu ðə tʃɪp baɪ mɪnz ɛv ðə weɪvz. æftɚ ðə sɛt ɛv weɪvz hæz gon baɪ, ðə sɝfɪs ɛv ðə watɚ ɪz əgeɪn kwaɪət ænd ðə tʃɪp ɪz stɪl.

ðə weɪvz ɔn ðə pɔnd gɪv ju ə pɪkʃɚ ɪn slo moʃən ɛv hwɑt hæpnz hwɛn ju hɪr ə sʌdn nɔɪz. æn ɪksplodɪŋ fɑɪrkrækɚ dɪstɝbz ðə ɛr ɑraund ɪt, dʒʌst əz ðə ston dɪstɝbd ðə watɚ. weɪvz muv aut θru ðə ɛr ɪn ɔl dɪrɛkʃənz. hwɛn ðɪz weɪvz hɪt jur ɪr, ðɛɪ pæs səm ɛv ðə moʃən ɔn tu jur ɪrdrʌm ənd ju hɪr ðə sʌdn saund.

APPENDIX D

PHONETIC TRANSCRIPT FROM JAPANESE PASSAGE

koiʃiwo sizukana ikeno nakani nageruto, sorega otʃita atari kara ikutsumono teisana nami ga sotogawani muka'te haʃi'te ikunoga miraremasu. nami wa tʃakuʒitsuna sokudode hirogari tsuzukemasu. suimen ni ukabu itagiraewa nami ga tori sugiru tsudoni dʒjogeni jureru deʃo. sono rijuwa koiʃini jo'te ʃodʒita undoga nami ni jori itakire eto ten'i. sareru kara nanodesu. soʃite nami ga kiesa'te ʃimauto, ikeno mizumo motono ʃizukana tatazumaini modori, itagiremo pitato ugoki wo tomemasu.

kono ikeno nami wa anataga totsuzen otowo kikutoki no josuwo sio moʃonre senmeini monogata'te irunodesu. hanabino bakuhatsuonwa masani kojiʃiga mizuwo mirasugotoku ʃuino kuki wo midaʃi, sopo nakawo arajuru hokoeto tʃiraba'te juki masu. kono kuki no nami ga anatano mimini todoku toki nami wa komakuni tsutaerare, sorega ototo ʃite kikoerunodesu.

APPENDIX E

TABLES OF THE GROUPED MEAN AND S.D. OF THE DURATION OF PAUSES IN MSEC BY U.S.A. SPEAKERS AND JAPANESE SPEAKERS IN ENGLISH, AND JAPANESE SPEAKERS IN JAPANESE.

		U.S.A. speakers in English					
		female		male		female and male	
		mean	S.D.	mean	S.D.	mean	S.D.
JNP	Read.	539.0	126.0	622.4	91.1	580.7	115.3
	Spon.	764.3	235.5	907.4	308.1	835.8	276.3
PHP	Read.	146.1	69.7	141.0	67.9	143.4	66.8
	Spon.	294.3	103.6	413.9	126.8	354.1	128.3
BWP	Read.	69.7	28.3	60.3	18.5	64.7	23.4
	Spon.	342.7	155.0	327.9	144.6	335.3	146.1
BSP	Read.	48.8	33.9	42.1	17.0	42.9	30.9
	Spon.	47.4	24.8	39.8	12.3	44.0	20.1
WSp	Read.	41.6	21.0	31.3	4.6	36.7	11.4
	Spon.	40.1	15.7	36.3	12.3	38.4	13.7

		Japanese speakers in English					
		female		male		female and male	
		mean	S.D.	mean	S.D.	mean	S.D.
JNP	Read.	561.9	140.4	730.1	242.3	646.1	211.5
	Spon.	682.0	327.2	1035.3	474.8	748.6	346.5
PHP	Read.	120.6	31.0	187.2	54.5	153.9	55.0
	Spon.	376.1	165.6	579.5	235.6	382.5	156.6
BWP	Read.	72.1	26.1	146.6	31.8	109.4	47.6
	Spon.	319.4	95.2	419.1	176.3	369.3	147.1
BSP	Read.	54.9	27.7	55.3	28.4	55.1	27.3
	Spon.	47.4	19.2	52.7	27.1	50.2	23.2
WSP	Read.	0.0	0.0	0.0	0.0	0.0	0.0
	Spon.	0.0	0.0	0.0	0.0	0.0	0.0

APPENDIX E (continued)

		Japanese speakers in Japanese					
		female		male		female and male	
		mean	S.D.	mean	S.D.	mean	S.D.
JNP	Read.	647.9	182.1	734.1	189.5	687.6	182.9
	Spon.	794.5	426.1	905.1	398.2	849.8	405.4
PHP	Read.	223.8	76.8	244.7	94.1	234.3	84.3
	Spon.	370.7	170.2	472.6	170.5	421.6	173.8
BWP	Read.	96.3	42.7	180.4	89.2	140.8	81.5
	Spon.	331.5	103.0	356.2	109.5	343.9	104.2
BSP	Read.	53.9	15.7	52.1	9.9	53.0	12.8
	Spon.	44.5	10.8	69.1	25.9	56.8	23.1
WSP	Read.	0.0	0.0	0.0	0.0	0.0	0.0
	Spon.	0.0	0.0	0.0	0.0	0.0	0.0

APPENDIX F

A TABLE OF THE GROUPED MEAN AND S.D. OF THE READING TIME IN SECONDS FOR A SHORT PASSAGE IN ENGLISH BY U.S.A. SPEAKERS AND JAPANESE SPEAKERS AND IN JAPANESE BY JAPANESE SPEAKERS.

	U.S.A. speakers in English		Japanese speakers in English		Japanese speakers in Japanese	
	male	female	male	female	male	female
Mean	45.98	50.27	64.29	54.68	60.86	53.47
S.D.	3.14	5.11	7.00	6.09	5.41	4.88

BIBLIOGRAPHY

BIBLIOGRAPHY

Agnello, J.G., and R.E. McGlone. Distinguishing features of /p/ and /b/. Journal of the Acoustical Society of America, July, 1970. (Abstract)

Black, J., Tosi, O., Singh, S., and Takefuta, T. A study of pauses in oral reading of one's native language and English. Language and Speech, 1965, 9, 237-241.

Boomer, D.S. Hesitation and grammatical encoding. Language and Speech, 1965, 8, 148-158.

Boomer, D.S., and Dittmann, A.T. Hesitation pauses and juncture pauses in speech. Language and Speech, 1962, 5, 215-220.

Deputy, P.N. Pausimetric analysis of speech production in children with consistent misarticulations versus children with no consistent misarticulations. Unpublished Doctoral Dissertation, Michigan State University, 1978.

Fodor, J.A., Bever, T.G., and Carrett, M.F. The Psychology of Language. New York: McGraw-Hill, 1974.

Fukui, S. Perception for the Japanese stop consonants with reduced and extended durations. Bulletin of the Phonetic Society of Japanese. 1978, 159, 9-12.

Goldman-Eisler, F. The measurement of the time sequences conversational behavior. British Journal of Psychology, 1951, 42, 355-361.

Goldman-Eisler, F. Individual differences between interviewers and their effects on interviewees' conversational behavior. Journal of Mental Science, 1952, 98, 660-671.

Goldman-Eisler, F. The distributions of pause durations in speech. Language and Speech, 1961, 4, 232-237.

Goldman-Eisler, F. The predictability of words in context and length of pauses in speech. Language and Speech, 1958, 1, 226-231.

Goldman-Eisler, F. Psycholinguistics: Experiments in Spontaneous Speech. London: Academic Press, 1968.

Goldman-Eisler, F. Pauses, Clauses, sentences. Language and Speech, 1972, 15, 103-113.

Hawkins, P.R. The syntactic location of hesitation pauses. Language and Speech, 1971, 14, 277-288.

Kowal, S., O'Connell, D.C., and Sabin, E.J. Development of temporal patterning and vocal hesitations in spontaneous narratives. Journal of Psycholinguistic Research, 1967, 4, 195-206.

Kuhl, P.K., and Miller, J.D. Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. Journal of the Acoustical Society of America, 1978, 63, 905-917.

Levin, H., and Silverman, I. Hesitation phenomena in children's speech. Language and Speech, 1965, 8, 67-85.

Levin, H., Silverman, I., and Ford, B.L. Hesitations in children's speech during explanation and description. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 560-564.

Licklider, J. Basic correlates of the auditory stimulus, in Handbook of Experimental Psychology, S.S. Stevens, Ed., (J. Wiley & Sons, New York, 1962), Chap.25, p.1204.

Lisker, L. Is it VOT or a first-formant transition detector? Journal of the Acoustical Society of America, 1975,57, 1547-1551.

Lisker, L., and Abramson, A.S. A cross-language study of voicing in initial stops: acoustical measurements. Word, 1964, 20, 384-422.

Love, L.R., and Jeffress, L.A. Identification of brief pauses in the fluent speech of stutterers and nonstutterers. Journal of Speech and Hearing Research, 1971, 14, 229-210.

Rochester, S.R., and Gill, J. Production of complex sentences in monologues and dialogues. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 203-210.

Taylor, I. Content and structure in sentence production. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 170-175.

Tiffany, W.R., and Carrell, F. Phonetics: Theory and Application. (2nd ed.). New York: McGraw-Hill, 1977.

Tosi, O. A Method for Acoustic Segmentation of Continuous Sound into Pauses and Signals and Measurement of Segment Durations. Unpublished Doctoral Dissertation, Ohio State University, 1965.

Tosi, O. Pausometry: Measurement of a low levels of acoustic energy. In World Papers in Phonetics, Phonetics Society of Japan, Tokyo The Phonetic Society of Japan, 1974. 129-144.

Tosi, O., Fischer, R., and Rockey, M. Quantitative measurement of time contraction induced by philocybin. Arzneimettel Forschung, 1968, 18, 535-537.

Tosi, O., and Lashbrook, W. Pauses and Circadian Rhythm. Paper presented at the 78th Meeting of the Acoustical Society of America, San Diego, California, 1969.

Williams, L. The perception of stop consonant voicing by Spanish-English bilinguals. Perception and Psychophysics, 1977, 21, 289-297.

Zlatin, M.A. Voicing contrast: Perceptual and productive voice onset time characteristics of adults. Journal of the Acoustic Society of America, 1974, 56, 981-994.