

INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

8713765

Hnath-Chisolm, Theresa Ellen

EFFECTS OF QUANTIZATION ON THE UTILITY OF VOICE FUNDAMENTAL
FREQUENCY AS AN ACOUSTIC SUPPLEMENT TO SPEECHREADING

City University of New York

PH.D. 1987

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

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark .

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received _____
16. Other _____

University
Microfilms
International

EFFECTS OF QUANTIZATION ON THE UTILITY
OF VOICE FUNDAMENTAL FREQUENCY AS AN
ACOUSTIC SUPPLEMENT TO SPEECHREADING

by

THERESA HNATH-CHISOLM

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.

1987

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.

April 21, 1987
Date

Arthur Boothroyd
ARTHUR BOOTHROYD, Ph.D.
Chair of Examining Committee

April 21, 1987
Date

Irving Hochberg
IRVING HOCHBERG, Ph.D.
Executive Officer

IRVING HOCHBERG, Ph.D.

HARRY LEVITT, Ph.D.

MARK WEISS, Ph.D.

Supervisory Committee

JAMES M. PICKETT, Ph.D.

External Examiner

Abstract

THE EFFECTS OF QUANTIZATION
ON THE UTILITY OF VOICE FUNDAMENTAL FREQUENCY
AS AN ACOUSTIC SPEECHREADING SUPPLEMENT
By

Theresa Hnath-Chisolm

Adviser: Professor Arthur Boothroyd

Fundamental voice frequency (F_0) has been shown to be a highly effective supplement to speechreading when presented by hearing. When the same information is presented via the sense of touch, however, it provides only a small improvement of speechreading performance. The purpose of this study was to test the hypothesis that the poor performance of tactile F_0 supplements is the result of degradation of fundamental frequency contours. This degradation results either from the inherently poor frequency resolution of the skin in single channel, temporal displays, or from the quantization that results from the use of discrete transducers, in multichannel, spatial displays.

Data were obtained for the recognition of words presented in sentence length stimuli by 12 normal adult subjects by speechreading, with and without auditory supplements. There were 5 supplements: (a) full F_0 contour, (b) the F_0 contour quantized at one, four,

eight, and twelve levels. Performance under all supplemented speechreading conditions was found to be significantly better than performance by speechreading alone. The most effective supplement was the full F_0 contour. Quantization at twelve levels resulted in a significant decrease of effectiveness. Supplemented speechreading performance diminished, in a monotonic fashion, as the number of quantization levels was further decreased. It was also found that sentence length was a significant factor determinant of performance in all conditions except speechreading + full F_0 .

In the main experiment, the full F_0 contour consisted of the output of an electroglottograph. In order to generate the quantized signals, however, it was necessary to extract and measure F_0 , and to regenerate sine wave copies. In a supplementary experiment, it was found that the processing itself resulted in a marked loss of effectiveness, even before quantization was deliberately introduced.

The principal conclusion is that spatial tactile F_0 displays should provide more than 8 levels of quantization if they are to be maximally effective.

ACKNOWLEDGEMENTS

It would not have been possible for me to complete this work without the support and encouragement of many people. I am very thankful to all the friends, fellow students, and family members who were willing to listen to my complaints and provide me with much needed words of encouragement.

If I attempted to mention each one by name these acknowledgements would be longer than the dissertation. There are a few people, however, to whom I owe special thanks. I wish to thank Liat Kishon-Rabin, Larry Medwetsky and Teresa Schwander, who tolerated me these last two years. I am very grateful to Eudy Yeung for his technical assistance, as well as for his friendship. I am grateful to my family for understanding my absence on many occasions: my mother, Alda Namias; my father, David Hnath; my grandmother, Margaret Seckler, my sisters, Janet Namias and Megan Hnath, and my in-laws, the Chisolms and the Jones.

The two people I have shared every step in the doctoral program with can never be thanked enough for all that they have done: Debbie Lynn and Laurie Hanin. I want to especially thank Laurie Hanin for being the most understanding co-worker with whom I have ever had the privilege to work.

I wish to express my gratitude and appreciation for all the time and good advice that my committee

members, Dr. Irving Hochberg, Dr. Harry Levitt, and Dr. Mark Weiss, gave to me. I also thank Dr. James Pickett for serving as external examiner. The most difficult person to thank is the chairman of my committee, Dr. Arthur Boothroyd. Words can never adequately express the gratitude I have towards him. Not only has he been my mentor, he has become, and I believe will always remain, a very special friend.

Last, but certainly not least, I want to thank my husband, Richard, who married me in the middle of all the insanity. He always believed that I would finish, even when I had my doubts. I hope that in our future years I can give to him all that he has given me.

TABLE OF CONTENTS

CHAPTER	PAGE
1 Introduction...	1
2 Review of the Literature.....	6
Part 1: Importance of access to voice fundamental frequency during speechreading.....	8
Part 2: Ability of the profoundly deaf to use auditory Fo information to enhance speechreading performance.....	19
Part 3: Transmitting fundamental frequency via the sense of touch.....	27
Physiological and psychophysical data relevant to the tactile transmission of voice Fo.....	27
Encoding voice Fo into tactile stimulus patterns.....	32
The perception of speech via tactile displays of voice Fo.....	35
Part 4: Quantization of voice fundamental frequency.....	42
3 Main Experiment.....	45
Purpose.....	45
Methodology.....	46
Preparation of test materials.....	46
1. Test materials.....	46
2. Initial recording of test material.....	48
3. Measurement of range of speaker's Fo.....	51
4. Quantization of the Fo contours.....	51
Experiment.....	57

	1. Instrumentation.....	57
	2. Subjects.....	59
	3. Procedure.....	59
	Results.....	64
	Preliminary discussion of results.....	77
4	Supplementary experiment.....	79
	Purpose.....	79
	Methodology.....	79
	Preparation of test material.....	79
	Experiment.....	80
	1. Instrumentation.....	80
	2. Subjects.....	80
	3. Procedure.....	82
	Results.....	85
	Preliminary discussion of results.....	90
5	Comparison of results of main and supplementary experiments.....	91
	Purpose.....	91
	Methodology.....	91
	Results.....	93
	Preliminary discussion.....	96
6	Effects of pitch extraction errors on performance.....	97
	Purpose.....	97
	Methodology.....	98
	Results	
	1. Two-way analysis of variance for a randomized design.....	101
	2. T-test on differences between matched-pairs of supplemented speechreading scores for error and non-error sentences.....	106
	Preliminary discussion.....	112

7	Discussion.....	113
8	Clinical implications, research implications, and future research.....	134

Appendices

A.	Topic related sentence sets.....	140
B.	Pitch quantization software.....	173
C.	Individual subject history.....	180
D.	Informal history questionnaire.....	181
E.	Instructions.....	182
F.	Raw data for sentence length analysis.....	183
G.	Evaluation of percentage of pitch extraction errors.....	186
	References.....	188

List of Tables

Table	Page
3.1	Randomization of sentence sets and groupings...62
3.2	Order of presentation of conditions and sentence set groups in the main experiment.....63
3.3	Word recognition scores as a function of test condition and replication for 12 subjects tested in the main experiment.....65
3.4	ANOVA summary table for the main effects of condition and replication and condition by replication interaction in the main experiment.....67
3.5	Results of the Student-Newman Keuls procedure applied to differences between condition means of the arcsine transforms of the raw data in the main experiment.....70
3.6	Mean percentage correct word recognition scores as a function of sentence length and condition.....73
3.7	ANOVA summary table for the effects of sentence length and condition.....74
4.1	Order of presentation of conditions for subjects in the supplementary experiment.....84
4.2	Word recognition scores as a function of condition obtained by 12 subjects in the supplementary experiment.....86
4.3	ANOVA summary table for the main effect of condition in the supplementary experiment.....88
4.4	Results of Student-Newman Keuls procedure applied to the means of the arcsine transformations of the raw data in the supplementary experiment...89
5.1	Percent reduction of error in supplemented speechreading conditions in the main and supplementary experiments.....94
6.1	Percentage words correctly recognized for each of the non-error sentences by each subject in each condition.....99

6.2	Percentage words correctly recognized for each of the error sentences by each subject in each condition.....	100
6.3	Percentage of words correctly recognized in sentences which did not contain pitch extraction errors and sentences that did contain pitch extraction errors after processing - data for ANOVA.....	102
6.4	ANOVA summary table for the main effects of condition and sentence type and the interaction between condition and sentence type.....	103
6.5	Illustrating the method used for arriving at an estimate of the effects of pitch extraction errors from which subject and sentence effects have been cancelled.....	107
6.6	Blocks of scores of error and non-error sentences under the processed and unprocessed Fo conditions.....	108
6.7	Differences between processed and unprocessed scores for matched pairs of error and non-error sentences.....	110
7.1	Methods and results of studies which investigated the perception of sentence length stimuli by speechreading alone and speechreading supplemented with acoustic Fo.....	115
7.2	Methods and results of studies which investigated the perception of sentence length stimuli by speechreading supplemented with an acoustic Fo signal and by speechreading supplemented with a constant frequency signal.....	120
7.3	Number of data points contributed by each of the 12 subjects tested in the supplementary experiment for each of the conditions evaluated in the post-hoc data analysis on the effects of pitch extraction errors on performance.....	127
B.1	Representative input voltages to IBM PC-XT and equivalent voltage input to VCO needed to replicate pitch contour.....	174
C.1	Individual subject history for subjects who participated in the main experiment.....	180
C.2	Individual subject history for subjects who participated in the supplementary experiment...	180

F.1	Percent correct word recognition as a function of sentence length by speechreading alone.....	183
F.2	Percent correct word recognition as a function of sentence length by speechreading + Fo quantized at one level.....	183
F.3	Percent correct word recognition as a function of sentence length by speechreading + Fo quantized at four levels.....	184
F.4	Percent correct word recognition as a function of sentence length by speechreading + Fo quantized at eight levels.....	184
F.5	Percent correct word recognition as a function of sentence length by speechreading + Fo quantized at twelve levels.....	185
F.6	Percent correct word recognition as a function of sentence length by speechreading + unprocessed Fo.....	185
G.1	Sentences analyzed to determine percentage of pitch extraction errors.....	187

List of Figures

Table	Page
1.1 Example of quantization of a Fo contour.....	2
3.1 Instrumentation for the initial recording of test material.....	49
3.2 Sample spectrograms of full audio and EGG signal.....	50
3.3 Instrumentation for preparation of test tapes..	54
3.4 Sample spectrograms of unprocessed Fo signal and Fo signal quantized at 12, 8, 4 and one level(s).....	56
3.5 Instrumentation for testing subjects.....	58
3.6 Mean performance as a function of condition and replication in the main experiment.....	66
3.7 Mean performance as a function of condition and sentence set presentation in the main experiment.....	68
3.8 Mean performance as a function of sentence length within conditions in the main experiment.....	75
3.9 Mean performance as a function of condition within sentence length in the main experiment.....	76
4.1 Sample spectrograms of unprocessed Fo signal and processed Fo signal.....	81
4.2 Mean performance as a function of condition in the supplementary experiment.....	87
5.1 Percent reduction of error due to supplementation in the main and supplementary experiment.....	95
6.1 Percent words correctly recognized for each sentence type in each of three conditions.....	104
6.2 Mean difference between matched-pairs of unprocessed & processed scores for two sentence types.....	111
7.1 Sample spectrograms illustrating the effects of a pitch extraction error and the effects of time sampling constraints sampling time.....	124

CHAPTER 1
INTRODUCTION

The research to be described is concerned with the role of auditorily presented voice fundamental frequency (F_0) contours as an aid to speechreading and the negative effects of degradation of these contours by quantization. Quantization is defined as the process of representing a continuous variable by a limited set of states. An illustration of what is meant by quantization is provided in Figure 1.1. It will be seen that the continuously varying F_0 contour is represented by a set of discrete frequencies, giving a "stepped" contour that only approximates the original. The effects of degradation caused by quantization on the utility of F_0 as a supplement to speechreading are as yet unknown. However, it is expected that important information will be lost as a result of quantization.

Previous research has shown that auditory presentation of fundamental frequency results in dramatic improvements in the speechreading performance of normal hearing subjects. With the addition of auditorily presented F_0 improvement in performance over speechreading alone has been found for the perception of consonants (Rosen, et al., 1979; Breeuwer and Plomp,

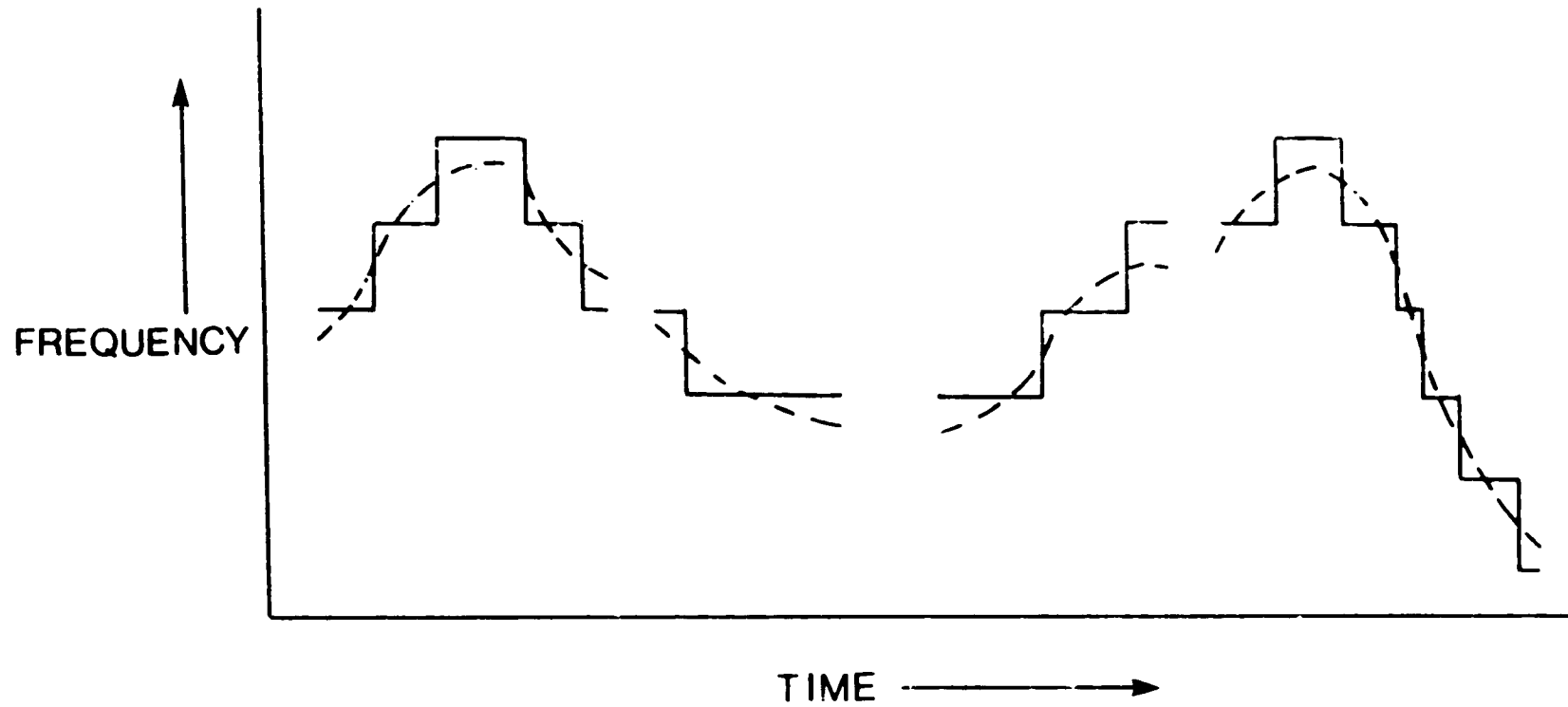


FIGURE 1.1: EXAMPLE OF THE EFFECTS OF QUANTIZATION. THE DOTTED LINE REPRESENTS THE UNPROCESSED F_0 CONTOUR. THE SOLID LINE REPRESENTS THE F_0 CONTOUR QUANTIZED AT 6 LEVELS.

1986); suprasegmental speech feature contrasts (Risberg and Lubker, 1978); sentence length stimuli (Risberg, 1974; Risberg and Lubker, 1978; Breeuwer and Plomp, 1986); and, connected discourse (Rosen, et al., 1981; Grant, et al., 1985).

Encouraged by these findings, researchers have tried to replicate Fo contours tactually for the profoundly deaf. A variety of display formats have been used to transform fundamental frequency into tactile patterns. It has, for example, been displayed as pulses/second in a single vibrator or as locus in a spatial array of vibrators (Rothenberg, et al., 1977; Grant, 1980; Boothroyd, 1984). Each of these displays generates a different stimulus configuration. The former generates temporal patterns while the latter generates spatiotemporal patterns. It has been suggested that spatiotemporal stimulus patterns are more appropriate than temporal patterns for the perceptual organization of the tactile system and thus should be used in tactile speechreading aids (Kirman, 1973). Spatiotemporal Fo displays have been used more than temporal ones in practical applications.

Although it might be feasible to build a continuous spatiotemporal display of Fo, engineering and mechanical constraints typically require that discrete output transducers be used in such a display.

When this is done, the normally continuous F_0 contour of a sentence is degraded as a result of frequency quantization.

Practical attempts to build spatiotemporal tactile displays of F_0 as aids to speechreading have been disappointing (Boothroyd and Hnath, 1986; Grant, et al., in press). Although, significant improvement in speechreading performance has been found, this improvement was much smaller than that reported for normal auditory presentation of F_0 .

The present study is based on the hypothesis that a major reason for the poor performance of existing spatiotemporal tactile F_0 displays is related to the quantization introduced by the transducer array. There are, however, other possible factors, the primary one being inherent limitations in the tactile system. In order to separate the effect of quantization from effects of the limitations of the tactile system, this study employed the auditory modality. The general goal was to measure the effects of quantization on the utility of a spatiotemporal display of voice fundamental frequency as a supplement to speechreading. There is, of course, no guarantee that an effective quantized acoustic signal would also be effective as a tactile supplement, but acoustic effectiveness may be considered a necessary but insufficient condition.

The specific goals of the study were:

1. To measure changes in speechreading performance resulting from supplementation by auditorily presented F_0 .
2. To measure changes in the utility of F_0 contours as a speechreading supplement resulting from the introduction of quantization.
3. To establish the smallest number of quantization levels for which there is no, or minimal, loss of speechreading assistance.

CHAPTER 2

REVIEW OF THE LITERATURE

This investigation is based on the observation that access to voice fundamental frequency (F_0) is important during speechreading. Research studies whose findings support this premise will be reviewed in the first section of this chapter.

The most common way to provide the hearing-impaired access to F_0 is through the use of amplification. There is evidence, however, that many profoundly deaf individuals cannot perceive variations in F_0 . Thus, even if fundamental frequency is made audible, some profoundly deaf individuals may not be able to use this acoustic information to supplement speechreading. The second section of this chapter addresses this issue.

Another way to provide the hearing-impaired access to fundamental frequency is through the use of tactile displays. Tactually presented F_0 does not appear to be as effective a speechreading supplement as does auditorily presented F_0 . The consequences of frequency quantization may be one of the reasons for the limited effectiveness of tactually presented F_0 . Research

studies whose findings are relevant to the tactile presentation of fundamental frequency, with special emphasis on their presentation through the spatiotemporal displays in which quantization occurs, will be reviewed in the third section.

There has been very little research in which the effects of quantizing fundamental frequency have been considered, and none of the research done has evaluated the effects of quantization on intelligibility. Such research, as is available, will be reviewed in the last section of this chapter.

Part 1: Importance of access to voice fundamental
frequency during speechreading

The results of research studies whose purposes have been to evaluate the ability of normal hearing subjects to perceive speech by speechreading alone and by speechreading supplemented with auditorily presented fundamental frequency provides the strongest evidence for the importance of access to this speech parameter during speechreading.

As evidence of the potential benefit of F_0 as a speechreading supplement, it has been pointed out that the addition of low-pass filtered speech to speechreading produces a dramatic improvement in speech perception (e.g. Sanders and Goodrich, 1971; Risberg, 1974). Low-pass filtered speech, however, may include several kinds of information in addition to F_0 . Examples are intensity variations and first formant frequency variations. In order to separate the contributions of F_0 from other factors, Risberg (1974) evaluated the ability of normal hearing subjects to identify known sentences under three conditions: (1) speechreading alone; (2) speechreading supplemented auditorily with an amplitude modulated constant frequency signal (i.e. providing information about the overall amplitude of the speech signal); and, (3) speechreading supplemented auditorily with an amplitude

modulated fundamental frequency signal (i.e. providing information about the overall amplitude of the speech signal and F_0 variations). Although both acoustic signals served to supplement speechreading, the signal which contained information on F_0 variations appeared to provide greater assistance. Risberg pointed out, however, that since a closed-set task was used to measure performance, the relevance of these results to normal speech communication, which is more akin to an open-set response task, is questionable.

Improvements in performance over speechreading alone on an open-set response task were found with the addition of an auditorily presented amplitude modulated F_0 signal by Risberg and Lubker (1978). Normal hearing subjects were required to answer simple questions (e.g. How much is $1 + 2$?) under three conditions: (1) speechreading alone; (2) speechreading + amplitude modulated constant frequency; and (3) speechreading + amplitude modulated + F_0). While little difference was found between performance in the speechreading alone and speechreading + amplitude modulated constant frequency condition, a gain of 40.6 percentage points over speechreading alone was found with the addition of the amplitude-modulated F_0 signal. The investigators noted that the small improvements over speechreading alone found with the addition of the amplitude modulated constant frequency signal were surprising,

since both this, and the amplitude modulated fundamental frequency signal provided the same segmental information (i.e. information about the durations and relative durations of voiced phonemes). They therefore conducted a second experiment to evaluate the exact role of access to F_0 during speechreading using the same conditions. They measured the ability to discriminate, in a closed-set task, three suprasegmental contrasts: (1) syllable stress; (2) phrase boundary; and, (3) word stress. Three new groups of normal hearing subjects were tested, each under only one condition. Scores obtained under both auditorily supplemented speechreading conditions were significantly better than those obtained by speechreading alone for all contrasts. Word stress was the only contrast for which a significant difference was found between performance under the speechreading + amplitude modulated constant frequency condition and that under the speechreading + amplitude modulated F_0 condition. Speechreading alone performance for the same contrasts was also measured in seven hearing-impaired subjects because the investigators believed that these subjects were more "experienced" speechreaders than the normal subjects. No significant differences were found between the scores of the hearing-impaired and those of the normal hearing subjects. Based on the results obtained in these experiments, the investigators concluded that the most

important supplementary information provided to the subjects during speechreading was F_0 . However, the data do not appear to warrant such a strong conclusion. Although, on the question task the best performance was obtained under the condition which provided information about F_0 variations, little difference was found between scores obtained on the discrimination task under this condition and the one that did not provide information about F_0 variations. In addition, the signal which contained information about F_0 variations also contained information about overall amplitude variations. Thus, the relative contribution of information about F_0 variations and about overall amplitude variations to the observed improvements in speechreading performance was not made clear .

Several investigators have found improvements in speechreading performance with the addition of an auditorily presented F_0 signal that did not contain any information about overall amplitude variations. For example, Breeuwer and Plomp (1986) measured the number of syllables correctly recognized by normal hearing subjects in sentence length stimuli by speechreading alone and by speechreading supplemented with an auditorily presented constant amplitude F_0 signal. A different set of 13 short (8 or 9 syllables) sentences was used to obtain scores for each trial in each condition. A significant difference was not found for

performance in the two conditions on the first trial. However, the mean percentage correct score obtained on the second trial in the auditorily supplemented condition (51.4%) was found to be significantly better than that obtained for speechreading alone (16.7%). Performance of the same subjects in two other auditorily supplemented speechreading conditions was also evaluated in this experiment. The acoustic signals were: (1) amplitude modulated F_0 ; and, (2) constant amplitude constant frequency (i.e providing information about voicing). While performance in both these supplemented conditions was found to be significantly better than speechreading alone, the addition of the constant amplitude constant frequency only resulted in only a very small improvement over speechreading alone. In addition, performance under the speechreading + amplitude modulated F_0 condition was found to be significantly better than that obtained under the speechreading + constant amplitude F_0 condition. These investigators conducted a second experiment in which the ability of normal hearing subjects to recognize both consonants and vowels was evaluated in three conditions: (1) speechreading alone; (2) speechreading + constant amplitude F_0 ; and, (3) speechreading + amplitude modulated F_0 . Consonants were presented in a /vowel-consonant-vowel/ nonsense syllable context. Vowels were presented in a /consonant-vowel-consonant/ nonsense syllable context.

The recognition of consonants under both supplemented conditions was found to be significantly better than that by speechreading alone. The scores obtained in the two supplemented conditions for the recognition of vowels were not, however, found to be significantly different from those obtained by speechreading alone. It was not determined if the differences between the scores obtained for either consonant or vowel recognition in the two supplemented conditions were significant.

Other investigators have evaluated the effects of supplementing speechreading with auditorily presented information about F_0 using the connected discourse tracking procedure (CDT) developed by deFilippo and Scott (1978). In this procedure subjects are required to speechread, and repeat verbatim, text read by a speaker. Performance is measured as the number of words per minute correctly perceived (i.e. the tracking rate). Changes in performance are commonly expressed in one of two ways: (1) as the percentage gain in improvement over performance in a baseline condition (usually speechreading alone); or, (2) as a percentage of the normal tracking rate (i.e. the rate obtained when subjects are allowed to "hear" unfiltered spoken text).

Using the connected discourse tracking procedure, Rosen, et al (1981) found improvements over

speechreading alone with the addition of auditorily presented Fo information. These investigators used a laryngograph to derive the Fo of the speaker. (A laryngograph or an electroglottograph measures changes in impedance across the vocal folds (Borden and Harris, 1980)). Pulses, triggered by the laryngograph signal, were low-pass filtered at 20k Hz and high-pass filtered at 4k Hz before presentation to normal hearing subjects. This was done so that the subjects still heard the pitch of the speaker's voice, but in a reduced form. The investigators believed that this signal provided information about Fo similar to that provided to deaf persons using a single channel cochlear implant designed to present Fo (Douek, et al., 1977). Percentage gains re: speechreading alone of 55-146% were obtained. These results appear to demonstrate that normal hearing subjects can use acoustic speech information presented in an unnatural way to supplement speechreading.

Grant, et al. (1985) also found that auditorily presented Fo information served as a significant supplement to speechreading. These investigators also used the connected discourse tracking procedure. Tracking data were obtained over 5 one-hour testing periods for three normal hearing subjects in three conditions: (1) speechreading alone; (2) speechreading + constant amplitude Fo; and (3) speechreading +

amplitude modulated Fo. Performance in both supplemented speechreading conditions was found to be significantly better than that obtained by speechreading alone. The results obtained for the speechreading + amplitude modulated Fo condition were found to be significantly better than those obtained in the speechreading + constant amplitude Fo condition. The investigators also obtained data over 5 one-hour testing sessions for one normal hearing subject under the condition of speechreading + voicing information. This subject had also participated in the first experiment. A mean score of 62.8% (re:a normal tracking rate of 108.3 wpm) was obtained. This score is very similar to the mean score obtained under the speechreading + constant amplitude Fo condition (67.9%). The investigators believed that the similarity of these results suggests that a signal which contains information about voicing is as effective an acoustic speechreading supplement as one which contains information about Fo variations. It should be noted that the subject's speechreading alone performance was not tested after training/testing in the speechreading + voicing condition. It is possible that her speechreading ability improved during this period. Thus, the design of the study makes it difficult to accept any conclusions as to the effectiveness of voicing information as an acoustic supplement.

Rosen, et al. (1979) found that signals which contained either voicing information or Fo information were equally effective acoustic supplements when nonsense syllable stimuli were used. The ability of normal hearing subjects to discriminate consonants presented in a vowel-consonant-vowel nonsense syllable context in three conditions: (1) speechreading alone; (2) speechreading + voicing; and, (3) speechreading + Fo. The acoustic signals did not contain any information about overall amplitude variations. Discrimination performance increased from 43.9% correct by speechreading alone to 71.9% with the addition of the voicing signal and 73.6% with the Fo signal. These results suggest that for the discrimination of consonants, an acoustic signal which only contains information about voicing provides the speechreader with essentially equivalent assistance to that which is provided by signals which contain information about Fo variations

The results of the research studies reviewed thus far demonstrate that normally hearing subjects, when presented with a speechreading task, perform much better when low-frequency acoustic cues are used as a supplement. Although Fo is not the only cue that can provide assistance, it is a valuable supplement when presented alone, i.e. as a constant amplitude pulse

train. In addition, the results of the Rosen, et al. (1981) study suggest that F_0 is a valuable acoustic supplement even when it is presented in an unnatural way. Data on the magnitude of the effect of supplementing speechreading with information about voicing are, however, somewhat inconclusive. Both Grant et al.'s (1985) and Rosen, et al.'s (1979) data suggest that it is a strong supplement, while the data of Breeuwer and Plomp (1986) suggest that this cue provides only a weak supplement. Further investigation into the effects of supplementing speechreading auditorily with information about variations in fundamental frequency and with that about voicing alone (i.e. constant frequency) would help to clarify this issue.

It will be noted in the foregoing review that the issue of utterance length has largely been ignored. Some investigators have used single word stimuli and others have used sentences of short or unspecified length. If as Lieberman (1967) suggests, a primary function of intonation, which is carried by F_0 , is to provide cues that listeners use to segment sentences for higher syntactic processing, then it might be expected that sentence length (and hence complexity) would be a critical variable. It is generally acknowledged that longer sentences are harder to speechread than shorter sentences. It might well be,

then, that Fo functions as a more effective supplement for longer sentences, but this issue has not been addressed.

Part 2: Ability of the profoundly deaf
to use auditory Fo information to enhance
speechreading performance

Profoundly deaf individuals are commonly defined as exhibiting 3-frequency pure tone averages greater than 90 dB. Although there are many degrees of hearing loss within this group, the audiograms of most profoundly deaf subjects show residual hearing below 500 Hz. Voice Fo is accessible from frequencies below 500 Hz. It might thus be predicted that the profoundly deaf should be able to use Fo information to enhance speechreading performance. However, caution must be taken in interpreting the results of the previously discussed studies with normal hearing subjects to the speechreading performance that might be expected if auditorily presented fundamental frequency information is made available to individuals with profound deafness, as these listeners often exhibit poor frequency discrimination.

The ability of profoundly deaf individuals to discriminate low frequency stimuli has been evaluated by several investigators (e.g. Butler and Albrite, 1956; DiCarlo, 1962; Gengel, 1969). Although the results obtained differ slightly depending on the psychophysical measurement technique used, it has, in general, been found that mean frequency difference

limens are larger than those obtained for normal hearing subjects at the same frequencies and sensation levels. It has also been found that the variation in discrimination ability for the same degree of hearing loss is very large. Thus, individual frequency discrimination ability cannot be predicted solely on the basis of degree of loss.

Poor frequency discrimination ability may result in an inability to use F₀ information to enhance speechreading performance. Since the relationship between acoustically aided speechreading performance and frequency discrimination ability has not been systematically explored, it is not known if this is the case. It has, however, been shown that many profoundly deaf individuals gain little or no assistance in speechreading from the use of amplification. For example, results obtained for the visual and audio-visual recognition of common words reveal small mean gains in performance (1-15 percentage points) with the addition of acoustic information (Numbers and Hudgins, 1948; Hopkins, 1953; Hudgins, 1954). Similar results for profoundly deaf children were obtained by Craig (1964) for the visual and auditory-visual perception of sentences. Improvements in performance of only 5.0-8.5 percentage points were found when audition was used to supplement speechreading. Erber (1972) presented evidence which suggested that these improvements were

primarily due to the presence of time-intensity information; that is, amplification failed to provide many of the profoundly deaf with spectral information. He concluded that the information received through the ears of the profoundly deaf who do not perceive spectral variations is similar to that which can be perceived vibrotactually. Indeed, several investigators have suggested that these profoundly deaf individuals do not possess true hearing, and they are actually "feelers" (Boothroyd and Cawkwell, 1970; Nober, 1970; Ericson & Risberg, 1977; and Erber, 1979). If this is the case, then the poor frequency resolution of the skin within the normal range of speech (Rothenberg, et al., 1977) might account for the minimal improvements in speechreading performance found with the addition of the amplified acoustic speech signal.

Although many profoundly deaf individuals do not perceive spectral variations in the amplified acoustic speech signal, it is not clear whether this is a result of a lack of auditory capacity or a lack of training. It has been shown that some of these individuals do have difficulty perceiving, auditorily, the linguistic contrasts signalled by variations in F_0 . For example, Stark and Levitt (1974) found that the performance of profoundly-deaf children on many tasks that required the recognition of several linguistic features which

may be contrasted by Fo (i.e. pause, question vs. statement and location of stress) was at chance levels. The subjects' task required making same/different judgements. The perception of pause was better than that of the other features and performance on question vs. statement items was the worst. Similar results were obtained by McGarr (1976) and by Gold (1978) for the identification of the same features. Pause was most often identified correctly. Stress was easier to identify when it occurred at the beginning rather than at the end of a sentence. The identification of questions was the most difficult. Although these results may be interpreted to indicate a lack of the capacity to perceive Fo variations, Stark and Levitt (1974) suggested that they might be the result of a lack of emphasis in training.

In support of Stark and Levitt's contention, Most (1985) found that there was a relationship between the performance of some 3-10 year old profoundly deaf subjects on tasks requiring the perception of synthetic intonation contours and the emphasis placed on auditory skills in the childrens' educational programs. Subjects were required to both imitate and discriminate speech stimuli which consisted of flat, rising, and falling intonation contours. Experienced listeners rated the imitations on a ten-point scale. The imitations of some of the children with hearing losses

within the range of 91-110 dB received very high scores, while the imitations of others were judged to be very poor. The imitations of children with losses greater than 110 dB were all rated as being very poor. The forced-choice test was only appropriate cognitively for subjects older than four years. While 50% of the children over four, with hearing losses in the 91-110 dB range did well on the forced-choice test, all those with hearing losses greater than 110 dB did poorly. The children in the 91-110 dB group who performed well on both the imitative and forced choice tasks were in training programs which emphasized auditory training, while those who did poorly were in programs which did not. Thus, there was a relationship between the degree of prior auditory training and the performance of children in the 91-110 dB range. Since none of the children with hearing losses greater than 110 dB did well on these tasks, the emphasis on auditory training did not appear to be a factor in their performance. (This last conclusion should be treated cautiously as only 4 of Most's subjects had hearing losses in excess of 110 dB).

The results of the research on the auditory perception of Fo by the profoundly deaf are by no means conclusive. The indications, however, are that some subjects, particularly those with losses in excess of 110 dB may be incapable of perceiving changes in Fo.

For those in the 90 to 110 dB range, many may have the potential for the perception of F_0 , but the exact role of training in determining performance remains unclear.

Sensory displays of F_0 have been developed to provide speechreading assistance to those profoundly deaf individuals who appear to lack the auditory capacity to perceive F_0 . There is a logical appeal to the use of sensory substitutes for the transmission of F_0 . There is, however, no guarantee that a modality other than hearing is capable of integrating the resulting patterns with visual speech information. The only evidence available in the literature comes from cochlear implant research. Fourcin, et al. (1984) reported improvement in speechreading performance for one post-lingually deaf subject with the use of the "cochlear implant" developed by Douek, et al (1977). This is the only cochlear implant designed to present only F_0 information. Performance was measured using the connected discourse tracking procedure (deFilippo and Scott, 1978). The subject's speechreading performance increased approximately 50% (re: speechreading alone) in the supplemented condition. The magnitude of this improvement is only slightly poorer than the poorest performance reported by Rosen, et al. (1979) for normal hearing subjects listening to "reduced" F_0 (Rosen, et al., 1979). It of interest to note that although the frequency discrimination ability

of seven subjects fitted with this cochlear implant was found to be much less acute than that found in normal listeners, positive results were obtained for the perception of question/statement contrasts and word stress by the same subjects (Fourcin, et al., 1979). The relevance of these findings to the issue of sensory substitution, however, is questionable since the cochlear implant does, in fact, tap into the auditory system.

Although there is a clear need for further research into the ability of the profoundly deaf to use natural or transformed Fo to supplement speechreading, some general conclusions can be drawn from the literature. First there are some profoundly deaf who can perceive Fo auditorily, with no assistance other than amplification. Their aided lipreading performance, however, suggests that they may not all be using Fo as an effective supplement. Second, for those who do not perceive (or are unable to use) Fo via amplification, a cochlear implant that transmits Fo may be a viable alternative.

Cochlear implants, however, are not the only option for the transmission of Fo. Moreover, this approach has the drawback of requiring the surgical implantation of extra- or intra-cochlear electrodes and is thus non-reversible. The use of tactile aids avoids

the problems associated with surgery. Moreover, in a recent comparative review of research into cochlear implants and tactile aids, it was found that the best implant performance was not substantially better than that of experienced multichannel tactile aid users (Pickett and McFarland, 1985). This finding, in conjunction with those demonstrating that Fo information provides an effective supplement to speechreading, suggests the need for research which provides information that can be used in the development of an effective tactile Fo display.

Part 3: Transmitting fundamental frequency
via the sense of touch

There have been many attempts to transmit speech information through the tactile sense. The results, which have been reviewed recently by Reed, et al. (1982), Sherrick (1984), and Pickett and McFarland (1985) have shown that tactile aids can serve as successful supplements to speechreading. The impetus for the present research comes from those studies dealing specifically with the tactile transmission of F_0 . These studies will therefore be reviewed here. Since F_0 has been coded both temporally, as rate of vibration, and spatiotemporally, as locus of vibration, in tactile displays, background physiological and psychophysical information relevant to temporal and spatiotemporal pattern perception will also be outlined.

Physiological and psychophysical data relevant to the tactile transmission of voice F_0

Two receptors are believed to mediate the perception of vibratory stimuli. These have been termed Pacinian corpuscles and Meissner's corpuscles (Geldard, 1972; Sherrick and Craig, 1982). The receptive field of the Pacinian corpuscle is usually large, with poorly defined borders and a single point of maximum sensitivity (Valbo and Johansson, 1978).

The response of such a receptor provides precise information about the time of occurrence of mechanical events, but only crude information about the place of occurrence (Sherick and Craig, 1982). In contrast, Meissner's corpuscles have larger receptive fields, but with more sharply defined borders, within which are several areas of peak sensitivity (Valbo and Johansson, 1978). These receptors are highly sensitive to transient stimuli and have very good spatial "tuning". Thus, they are believed to have the potential to register precise information about both timing and location (Sherrick and Craig, 1982).

Measurements of the absolute sensitivity of the skin to vibratory stimuli have shown that it is responsive to frequencies within the range of approximately 10 to 1000 Hz (e.g. Verillo, 1963; 1968; Talbot, et al., 1968). The exact range depends on several methodological factors, such as the size of the contactor, whether or not a static surround is used, and site of stimulation on the body. In general, it has been found that absolute sensitivity is fairly poor and independent of frequency between 10-50 Hz. Sensitivity increases for frequencies above 50 Hz, with the best sensitivity at 250 Hz, then rapidly decreases for higher frequencies. It is generally believed that Pacinian corpuscles mediate the response above 50 Hz. There is uncertainty about the receptor responsible

for the responses to stimuli between 10-50 Hz, but evidence points to Meissner's corpuscles (Merzenich and Harrington, 1969).

It will be seen from the foregoing, that the skin is sensitive to vibrating stimuli within the range of voice F_0 . The encoding of F_0 , however, either as frequency of vibration or as locus of vibration will involve the stimulation of two different receptor systems. These differ in terms of absolute sensitivity as a function of frequency, and they respond differently to spatial information. The exact way in which these differences will affect the perception of possible tactile transforms of F_0 is uncertain.

Although the skin is sensitive to frequencies within the range of voice F_0 , its frequency discrimination capacity is poor. In the most recent and most thorough investigation into tactile frequency discrimination, Rothenberg, et al (1977) obtained frequency difference limens for vibratory stimulation of the forearm. The results were influenced by both the type of stimulus (pulsed vs. sinusoids) and by the psychophysical method (constant stimuli vs. detection of frequency modulation). Results were also influenced by stimulus frequency. The best difference limen was 10% (roughly 1/7 octave) for pulsed stimuli at 20 Hz. As frequency increased, the difference limen increased,

reaching 15% at 100 Hz for pulsed stimuli (roughly 1/5 octave) and 25% at 100 Hz for sinusoids (roughly 1/3 octave). Above 100 Hz difference limens increased rapidly.

Traditional measures of tactile spatial resolution, utilizing static pressure applied simultaneously at two points, have shown that a distance of 2.5-4.0 mm is needed for discrimination at the fingertips. The fingertips are the most sensitive body site and at other sites the static two-point limen becomes larger (Weinstein, 1968). More recently, Johnson and Phillips (1981) have shown that when a two alternative forced-choice procedure is used, subjects can always discriminate two stimuli from one, even when the stimulators are side by side (occupying a total space of 1 mm). Static measures, however, are of questionable relevance to the perception of patterns that are varying over time. Békésy (1959), for example, showed that there were marked masking effects when two or more sites are stimulated simultaneously with vibratory stimuli. Even when stimulators are widely spaced over the body gross masking effects have been observed (Gilson, 1969).

It has been shown that subjects can detect at least a 1/4" shift in stimulus location in a time varying tactile pattern (Boothroyd, 1984; Boothroyd, et

al. 1986). This was found when the reference stimulus encompassed movements of zero to 2 1/4".

Tactile patterns which vary in both space and time give rise to sensations of movement. The spatial and temporal properties of these sensations, however, are not necessarily identical to those of the stimulus. Two striking examples of a lack of correspondance are the "rabbit" phenomenon and the "phi" phenomenon. The rabbit phenomenon has been described as a sensation of stimuli "hopping" between two stimulus locations. This sensation can be induced if tap stimuli are used and the number of taps delivered to each stimulation site remains small (less than 18). This sensation has been found with contactors as close as 2 cm and as far apart as 35 cm. In addition, it has been found that both the interstimulus onset interval (greater than 20 msec) and the number of contactors used were not critical in inducing this sensation (Geldard and Sherrick, 1972).

The "phi" phenomenon (also known as tactile apparent movement) has been described as a vibrating, continuous "gouging" of the skin between loci of stimulation (Sherrick and Rogers, 1966). To elicit this sensation vibratory stimuli are used. It has been shown that the clarity of this sensation is dependent on three factors: (1) stimulus duration; (2) interstimulus onset interval; and, (3) number of

stimulators. Using two stimulus sites, they found that as stimulus duration increased from 100 to 300 ms, a concurrent increase in interstimulus onset interval of approximately 125 to 250 ms was needed to produce good sensations of movement (Sherrick and Rogers, 1966). A similar relationship between stimulus duration and interstimulus onset interval was found for the use of two to four stimulus locations by Kirman (1974). In addition, he found that the optimal interstimulus onset interval needed to produce good sensation of movement decreased as the number of stimulators increased from two to four, regardless of stimulus duration. It has further been shown that the perception of tactile apparent movement is independent of stimulus magnitude, stimulus frequency, distance between stimulators, and body site and that it can be elicited with either vibrotactile or electrocutaneous stimuli. (see Sherrick, 1968 for review).

Encoding voice Fo frequency into tactile stimulus patterns

Based on the information just reviewed, researchers have suggested several ways of encoding Fo for the tactile sense. Rothenberg, et al (1977) have proposed the temporal encoding of Fo as rate of vibration. They based this suggestion on the frequency discrimination results they obtained with the warble

tone method. They believed that with the warble tone method, frequency stimuli were presented in a manner closer to the way they would be when a speech derived variable, such as F_0 , was encoded. The psychophysical results suggested that there are at least seven differentiable frequency steps within the range of 10 to 90 Hz on the forearm, and possibly ten at the more sensitive fingers. Thus, in order to effectively display F_0 to the tactile sense, a recoding of the range of voice F_0 (70-500 Hz) into the vibrotactile range of 10 to 90 Hz would be necessary.

Sherrick (1984) suggested that since Rothenberg, et al.'s (1977) results indicated that tactile discrimination of more than a few steps in frequency is difficult, researchers should combine frequency, intensity, and spatial dimensions to enlarge processing capacities. He cited studies by Taylor (1978) and Geldard, et al. (1981) which demonstrated gains in the number of bits of information transmitted when these dimensions are combined. Sherrick also suggested that redundantly encoding a speech parameter may result in an enhancement of stimulus pattern recognition. The issue of redundancy has been raised in a more general context by Reed, et al. (1982). They point out that redundancy operates at many levels in the speech perception process, its function being to reduce the probability of error when information loss occurs in

one or more sources. They further argue that the ultimate effectiveness of a speechreading aid depends on the subject's use of the information, rather than the available information, and in degraded conditions redundant encoding can be very beneficial. (It should be noted that Reed, et al. (1982) raised the issue of redundancy in the context of what information should be displayed in a sensory aid. They believe that sensory aids should encode cues which are visible as well as those that are invisible, as the redundant presentation of visible cues through a sensory speechreading aid could be of real value). If F_0 is redundantly encoded in both spatial and temporal dimensions, then patterns of movement will be generated. Kirman (1974) believes that such spatiotemporal patterns may be particularly appropriate for the perceptual organization of the tactile system.

The results of the Boothroyd (1984) and Boothroyd, et al. (1986) studies provide empirical evidence which suggests that F_0 can be perceived via spatiotemporal tactile transformations. There are, however, limits to the spatial resolution that can be provided by a fixed number of tactile transducers. These limits, in turn, impose limits on the frequency resolution that can be provided in a spatial display of F_0 . For example, if an 8-channel display is used to encode the frequency range of 70 to 500 Hz (i.e. approximately two and one-

half octaves), then each channel would represent a change of approximately one-third of an octave. Thus, an 8-channel spatial encoding of F_0 would provide essentially equivalent frequency resolution to the temporal encoding scheme suggested by Rothenberg, et al. (1977). By increasing the number of channels frequency resolution can, however, be improved in a spatial display. If a 16-, rather than 8-channel display is used, then each channel would represent a change of approximately one-sixth (rather than one-third) of an octave.

The perception of speech via tactile displays of voice

F_0

Rothenberg and Molitor (1979) demonstrated the feasibility of the tactile perception of word stress via several temporal recoding schemes. Seven linear and three logarithmic modes of converting the voice fundamental frequency range (70-500Hz) into the vibrotactile frequency range of 10 to 300 Hz were explored. A single channel vibrator was applied to the forearms of both normal hearing and deaf subjects. Results for all subjects demonstrated above chance performance with simple one- or two-octave reductions (i.e. transmission of every second or fourth glottal pulse). The investigators found that simple frequency division transformations resulted in a significant

ability to identify intonation patterns associated with moderate-to-strong patterns of English sentence stress. They did point out, however, that if subtler stress patterns are to be discriminated, other transformations would probably be necessary.

Plant and Risberg (1983) found that word stress and word syllable number and type (i.e. monosyllables, trochees, spondees, and trisyllables) could be perceived tactually via a temporally recoded one-octave reduction of the F_0 range of the speaker. Stimuli were presented through a hand-held vibrator to 10 normal hearing subjects. Mean percentage correct scores of 79% for the word stress task and 49% for the syllable number and type task were obtained. Both these scores were found to be above chance.

Grant, et al. (1986) demonstrated that the perception of several speech contrasts was possible via a tactile spatiotemporal transformation of F_0 . They investigated the perception of statement/question contrasts, phrase boundaries, and word stress. A ten-channel electrocutaneous display was used to present stimuli to the forearms of one deaf and one hearing subject. Both subjects obtained scores greater than 90% on all three tasks after a relatively small amount of training.

Boothroyd (1984) also obtained positive findings for the tactile perception of speech contrasts via a spatiotemporal transformation of Fo. An eight-channel vibrotactile display was used in the evaluation of the tactile perception of monotone vs. natural intonation contours, question/statement contrasts, and word stress. Stimuli were presented to normal hearing subjects at two body sites: the thenar eminence of the hand and the medial surface of the forearm. Although the average scores for both monotone/natural and question/statement contrasts were found to be above chance at both body sites, scores obtained at the thenar eminence were significantly better than those obtained at the forearm. Two of the eight subjects were able to identify word stress at or close to 100%. However, the group mean performance for this contrast was below chance. It was noted that the two subjects who exhibited superior performance on this task had prior experience with the tactile display (Boothroyd, 1986). The perception of the same three speech contrasts by one congenitally deaf subject (PTA=105 dB) via hearing aids alone and hearing aids + the spatiotemporal tactile transformation of Fo was also reported. In the hearing aid alone condition only performance on the word stress task was above chance. In the hearing aid + tactile stimulation condition, performance on all three tasks was above chance levels.

It can be concluded from these studies that speech contrasts signalled by variations in voice F_0 can be perceived through the tactile sense as long as appropriate temporal or spatiotemporal transforms are used.

The effects of supplementing speechreading with either temporal or spatiotemporal tactile transforms of voice fundamental frequency has not been extensively investigated. Only two reports appear in the literature (Boothroyd and Hnath, 1986; Grant, et al., 1986). These two research studies will be discussed in detail.

Boothroyd and Hnath (1986) measured the recognition of words in sentences as a function of brief training in three conditions: (1) speechreading alone; (2) speechreading plus a temporal encoding of voice fundamental frequency; (3) speechreading plus a spatiotemporal encoding of voice fundamental frequency. An eight channel vibrotactile F_0 display (described in detail by Boothroyd, 1985) was worn on the forearm for both aided conditions. The frequency range of the display was adjusted to encompass the speaker's fundamental frequency range during the spatiotemporal condition. During the temporal condition the same device was modified so that only one solenoid was activated. The solenoid vibrated at one-half of the

input fundamental frequency. Subjects were two normal hearing young adults. Masking noise was used to assure that the subjects could not "hear" the vibrations during the supplemented conditions. Data were collected over six sessions. Subjects were given brief (10 minute) periods of training in each condition prior to testing in all but the first session. Training was done using live voice and testing was conducted using videorecorded sentence sets. The same speaker was used in the presentation of both training and testing materials. The number of words correctly recognized in each test session by both subjects in the three conditions was calculated. Significant differences in performance were not found for the scores obtained under the three conditions in the first four sessions. Performance in the last two sessions was found to be significantly better with tactile supplements than without. A significant difference was not, however, found between the scores obtained with the temporal and spatiotemporal transformations of F_0 . The investigators believed that this was the result of the short amount of training time in each condition (less than one hour) and the mixing of conditions within sessions. Evidence of significant learning effect across sessions was found only in the condition in which spatiotemporal coding was used.

The effect of supplementing speechreading with a

spatiotemporal electrocutaneous display of fundamental frequency has been evaluated by Grant, et al., 1986. A ten-channel display, described in detail by Grant (1980) was used in this study. Data were obtained using the connected discourse tracking procedure (deFillippo and Scott, 1978). One normal hearing and one profoundly hearing-impaired adult served as subjects. Based on differences in overall speechreading ability between the two subjects different testing materials were used. Materials were presented by a different speaker to each subject. The average tracking rate for the normal hearing subject was 40% (re: normal tracking rate) by speechreading alone and 52% (re: normal tracking rate) for the supplemented condition. Tracking data were also reported for the normal hearing subject under the condition of speechreading supplemented with auditorily presented fundamental frequency information. The investigators compared the subject's performance in this condition with that measured in the speechreading alone and speechreading + tactually presented Fo information conditons. The average difference between the speechreading only and aided speechreading rates (re: normal tracking rates) were found to be 12 percentage points with tactile transmission and 36 percentage points with auditory transmission. Thus, much smaller improvements in performance were obtained with tactile supplementation. It was also found that

it took a longer time to achieve improvements in performance over speechreading alone with tactile supplementation. Since however, the investigators found no evidence of asymptotic performance for this subject, they believed that the limits of performance with tactile supplementation remained unresolved. They also suggested that the quantization of the F_0 , which resulted from spatiotemporal transformation, might have affected performance in the tactile condition.

The results of the Boothroyd and Hnath (1985) and the Grant, et al (1986) studies demonstrate that tactually encoded F_0 information serves as a significant, but small, supplement to speechreading.

Part 4: Quantization of voice fundamental frequency

Although auditorily presented voice fundamental frequency provides a powerful supplement to speechreading, tactually presented F_0 provides only a minimal supplement. There are several possible reasons for the failure of tactile displays of F_0 to provide the assistance to speechreading that would be predicted from the results of acoustic presentation. First, the tactile sense may be inherently unsuitable for processing F_0 or for integration with visually perceived speech patterns. It may also be the case that in the processing and recoding of F_0 for the sense of touch critical information is removed or distorted, thus rendering the resulting stimulus patterns ineffective. These factors may also combine to reduce the effectiveness of tactually presented F_0 .

The present study is concerned with one of the ways in which F_0 becomes degraded during presentation via a spatiotemporal tactile display, that is, degradation resulting from quantization into a fixed number of discrete levels. The effects of quantizing F_0 have been briefly studied in connection with work on bandwidth reduction in speech synthesis and transmission. Flannigan (1957) examined data on difference limens and F_0 range and concluded that quantization of F_0 contours into 80 steps should have

no audible effect on the quality of speech from a male talker. This does not, however, represent a lower limit for acceptability or intelligibility.

The question of just detectable levels of quantization was addressed empirically by Rosenberg, et al. (1971) using resynthesized speech. They found quantization at 4 bits (i.e. 16 steps) was just detectable to listeners. If it is assumed that non-detectable degradations have no effect on intelligibility, then these data suggest that quantization of fundamental frequency into 16 steps would provide the same intelligibility as that offered by continuous F_0 . It should be noted that this is 16 steps for a single talker. To cover the F_0 ranges of males, females, and children, it might require two or three times as many steps. These data do not indicate the maximum amount of F_0 quantization that could be introduced without measurably affecting intelligibility. Moreover, the relationship between the contribution of F_0 to auditory speech perception and visual speech perception is not clear.

It is reasoned that the issue of quantization can be investigated via the auditory modality. Thus, any effects due to quantization can be separated from those due to any limitations of the tactile sense. By determining the number of levels into which an acoustic

Fo signal can be quantized without reducing its' effectiveness as a speechreading supplement, the least acceptable channel number for a spatial tactile display of Fo can be inferred.

CHAPTER 3
MAIN EXPERIMENT

Purpose

This investigation dealt with the effectiveness of audible fundamental frequency (F_0) contours as supplements to speechreading. The specific purpose of the main experiment was to examine the negative effects of distorting those contours by quantization. The specific hypotheses of this experiment were:

(1) speechreading performance would improve significantly with the addition of auditorily presented F_0 information;

(2) there would be some "critical" level(s) of quantization at which supplemented speechreading performance would be equivalent to that obtained with continuous F_0 contours;

(3) supplemented speechreading performance would diminish, in a monotonic fashion, as the number of quantization levels decreased beyond the "critical" level(s);

(4) utterance length would be an important factor in evaluating the effects of supplementing speechreading with auditorily presented F_0 information.

Methodology

Preparation of test materials

1. Test materials.

Performance was measured as the number of words correctly perceived in sentence length stimuli. These sentences were "topic-related", that is, each sentence was about a common topic of which the subjects were made aware. Such stimuli provide phonetic, semantic, syntactic, and contextual information similar to that available in conversational speech. The use of a topic-related sentence test to measure speech perception performance was initially reported by Boothroyd (1981), who developed such a test for use with hearing-impaired adolescents. A similar test was also used in the previously described investigation into tactile speechreading supplements conducted by Boothroyd and Hnath (1986). However, neither of these topic-related sentence tests contained enough material for the purposes of this study, and it was first necessary to develop sufficient test material.

A total of 48 sets, of 12 topic-related sentences each, were written for use in this and other investigations at the City University of New York. Each sentence in a set is about one of the following 12 topics: food, family, work, clothes, homes, animals, sports & hobbies, weather, health, seasons & holidays,

money, and music. Topic order remains constant from set to set. Within each set there are 4 statements, 4 questions, and 4 commands. Length varies continuously from 3 through 14 words, giving a total of 102 words per set. Sentence type and length are counterbalanced across topic in the 48 sets.

An initial version of the 48 sentence sets was audiorecorded and three normal hearing adult listeners were asked to indicate with a yes/no response whether or not each sentence was likely to occur in an everyday conversation between adults. The criterion for accepting a sentence was that all three listeners had to agree on its acceptability. Only 24 (out of a total of 576) sentences were judged by at least one listener to be unlikely to occur. Only two of these 24 sentences were judged to be inappropriate by more than one listener. These sentences were rewritten and rerecorded. The same listeners were asked again to judge the acceptability of the new sentences. On the first presentation of the new sentences all three listeners agreed that they were appropriate to everyday conversation. The final 48 sentence sets are included in Appendix A.

2. Initial recording of test material.

The instrumentation used in the initial recording of the test material is illustrated in Figure 3.1. A 1/2" videorecorder (Panasonic AG6300) and a color video camera (Panasonic WV-3240) were used. A Sony KX-1901 color monitor was used to display the video image during recording. Two audio signals were recorded: (1) full audio signal from an Electrovoice 63A microphone situated 6" from the speaker's mouth; and, (2) an electroglottograph (EGG Fo) signal from a Synchronoice electroglottograph. Sample spectrograms of the same sentence recorded from the Electrovoice 63A microphone and the Synchronoice electroglottograph are shown in Figure 3.2. The EGG recording was made to provide a signal whose most intense component was voice Fo and whose spectrum showed little or no influence of supralaryngeal resonance or supralaryngeal sound sources. The speaker was a female doctoral candidate in Speech and Hearing Sciences with a northeastern American accent.

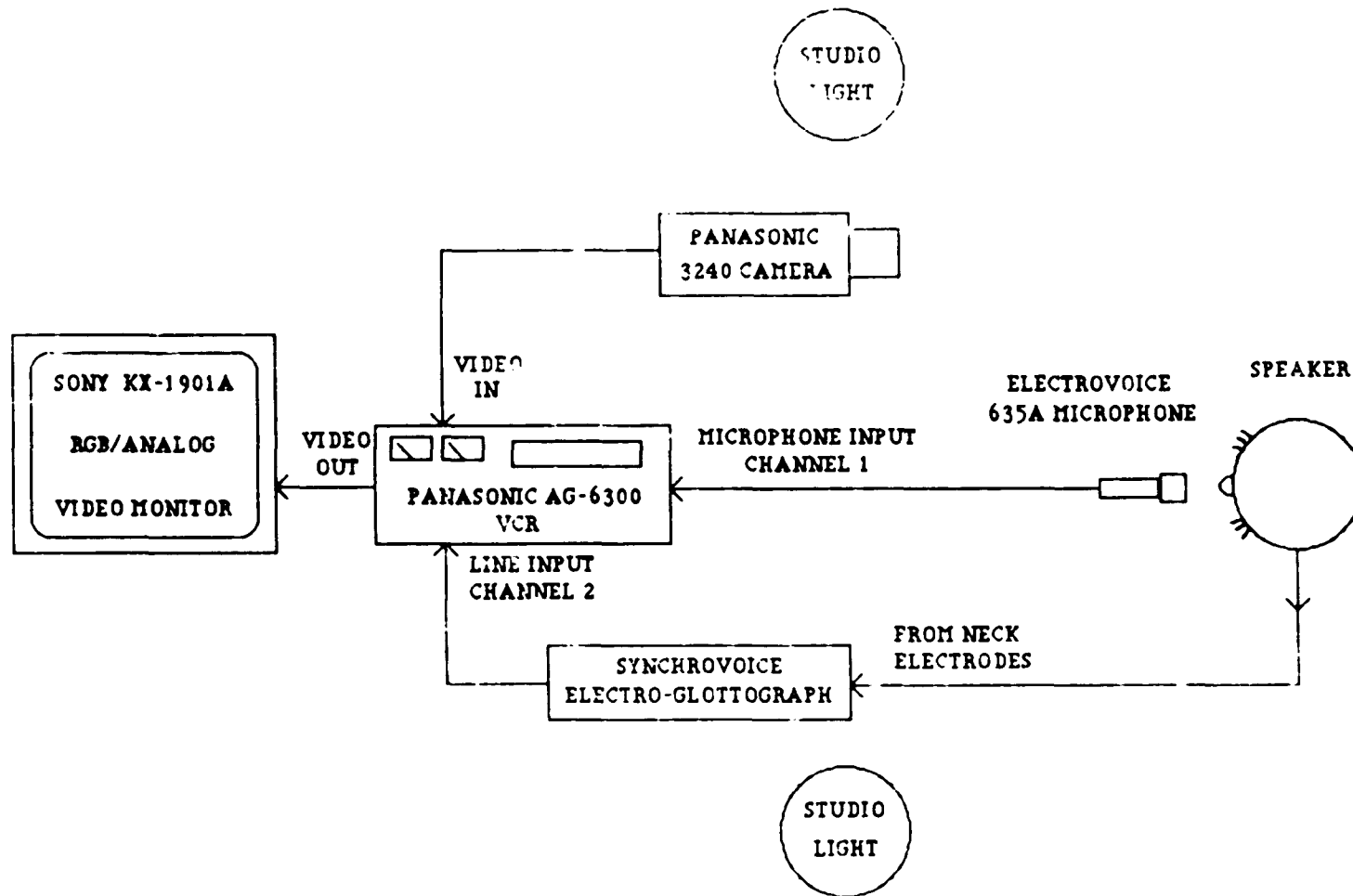


FIGURE 3.1 INSTRUMENTATION FOR THE INITIAL RECORDING OF TEST MATERIAL

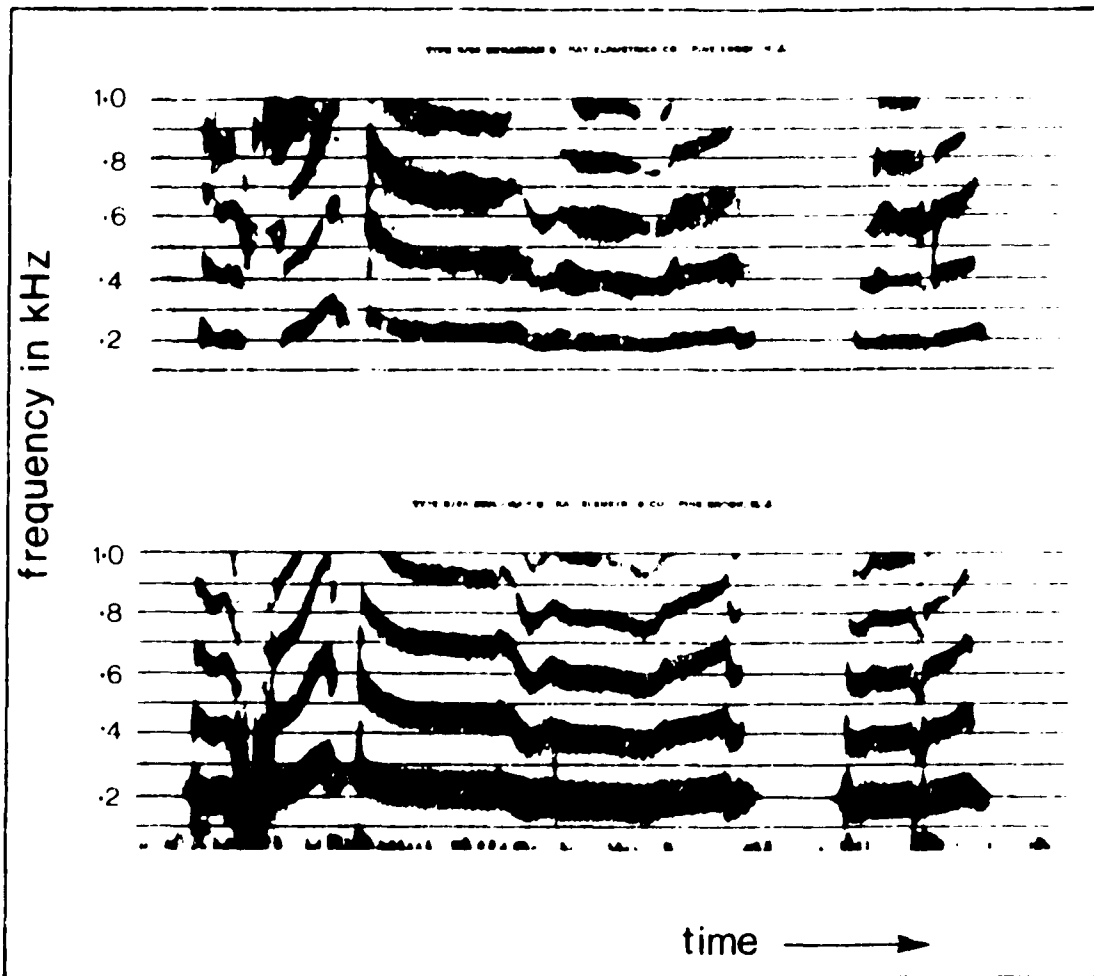


FIGURE 3.2: SAMPLE SPECTROGRAMS ILLUSTRATING THE SAME UTTERANCE RECORDED VIA THE ELECTROVOICE 635A MICROPHONE (UPPER) AND THE SYNCHROVOICE ELECTROGLOTTOGRAPH (LOWER).

3. Measurement of the range of the speaker's voice Fo.

The range of the speaker's Fo was measured using equipment and software available at the City University of New York Speech and Hearing Sciences Laboratories. The first step was to feed the recorded EGG signal to a pitch extractor (Visipitch Kay 6387). The pulse output of Visipitch was fed to a custom built interpulse interval timer which measured the time between pulses (Boothroyd and Jiang, 1985). The output of the interpulse interval timer was fed to the binary input of an ISAAC interface (Cyborg Isaac 91A). After conversion to frequency, graphic representations of the variations of the Fo of the speaker were generated using an AppleIIe computer and a software program written for the measurement of voice pitch (Yeung and Boothroyd, 1985). A random sample of six sentences were analyzed using this procedure. The lowest frequency measured was approximately 100 Hz and the highest frequency approximately 330 Hz.

4. Quantization of the Fo contours.

Four levels of frequency quantization were chosen for use in this study. These were 1 level, 4 levels, 8 levels and 12 levels. The coarsest of these, one level, represents the most extreme form of quantization. It corresponds to a simple voiced/voiceless indication. Quantization at four

levels was used because it was midway between one and eight levels. Eight levels were chosen because an 8-channel tactile display of F_0 had previously been used by the investigator in an experiment evaluating the effectiveness of tactile F_0 contours as a speechreading supplement (Boothroyd and Hnath, 1986). Thus, results obtained with quantization at 8 levels had implications for practical devices that were already in use. The finest degree of quantization of F_0 , into 12 levels, was used because it was the maximum that was achievable within constraints imposed by the speed of the computer program used for F_0 measurement, quantization, and regeneration. The time needed by this program for a single pitch sample was 40 msec. Thus, when the value of F_0 changed rapidly, the finite sampling rate produced F_0 quantization equivalent to approximately 12 levels.

Five test tapes were prepared from the original recording. Each of the test tapes contained one of the following acoustic signals:

- (1) EGG signal
- (2) F_0 contours quantized at 12 levels
- (3) F_0 contours quantized at 8 levels
- (4) F_0 contours quantized at 4 levels
- (5) F_0 contours quantized at 1 level.

The instrumentation used to prepare the test tapes

and to quantize the F_0 contours is illustrated in Figure 3.3. The video signal was directly dubbed and the EGG signal was first fed through a bandpass filter (Allison AL2ABR). The filter was set at a low frequency cutoff of 90Hz and a high frequency cutoff of 360 Hz. The filter settings were chosen to allow only the measured fundamental frequency range of the speaker to be passed. In dubbing the unprocessed F_0 test tape the output of the bandpass filter was fed directly into the audio input of the VHS recorder.

To quantize the fundamental frequency contours the output of the bandpass filter was first fed to a pitch extractor (Visipitch 6387). This provided a varying DC voltage output proportional to the fundamental frequency. This voltage served as input into an A/D converter (Cyborg Isaac D91A). After conversion, the digital representation of the DC voltage was transformed into the appropriate quantization levels using an IBM PCXT computer and the Pitch Quantization program included in Appendix B. The transformation into appropriate quantization levels involved the following steps:

- (1) The log of the digital representation of 90 Hz was subtracted from the log of the digital representation of the current value of F_0 (X) (i.e. F_0 (X) - 90);

- (2) (F_0 (X) - 90) was divided by the step size, S

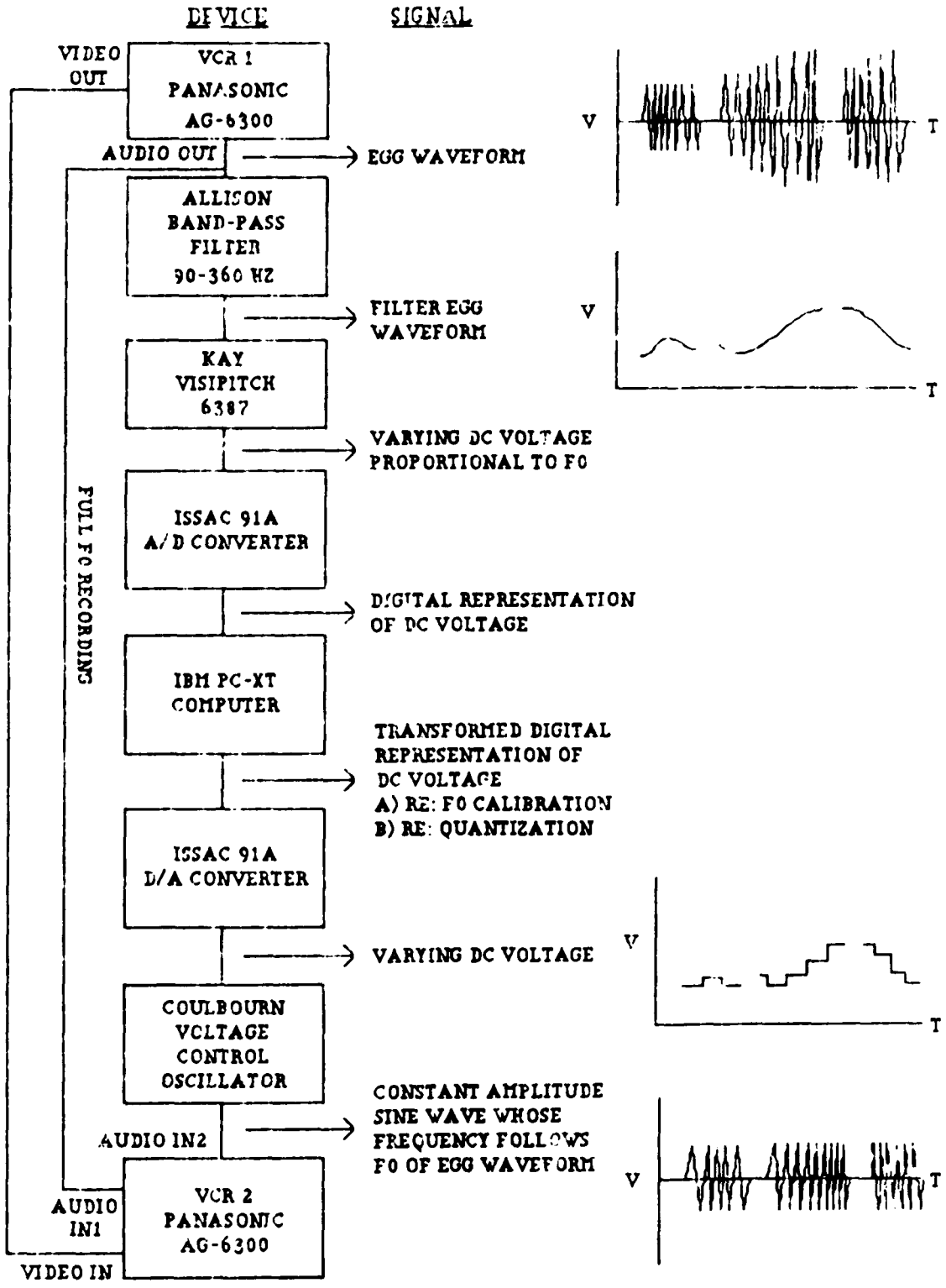


FIGURE 3.3 INSTRUMENTATION FOR PREPARATION OF TEST TAPES

(S depended on the desired number of quantization levels);

(3) The result of $(F_o(X) - 90)/S$ was then rounded off to the nearest integer;

(4) The integer of $(F_o(X) - 90)/S$ was multiplied by S and was added to $(90 + 1/2 S)$;

(5) Then the antilog of the integer of $((F_o(X) - 90)/S) \times S + (90 + 1/2 S)$ was taken to yield the quantized digital representation.

The quantized digital signals were converted to equivalent varying DC voltages by a D/A converter (Cyborg Isaac D91A) for input into a voltage controlled oscillator (Colbourn S24-05). The output of the VCO provided a constant amplitude sine wave whose frequency followed the F_o contour provided by the EGG waveform. This signal was recorded onto the audiotrack of the VHS recorder. Sample spectrograms showing the filtered unprocessed F_o signal and the signal quantized at 1, 4, 8, and 12 levels are shown in Figure 3.4A-E.

It should be noted that the choice of the processing technique was determined by both available equipment and the need to prepare the test tapes within a reasonable period of time. The technique, however, was not ideal. It resulted in: (1) the introduction of pitch extraction errors (a shortcoming of Visipitch); (2) time sampling constraints mentioned earlier; (3) loss of harmonic structure by virtue of use of sine

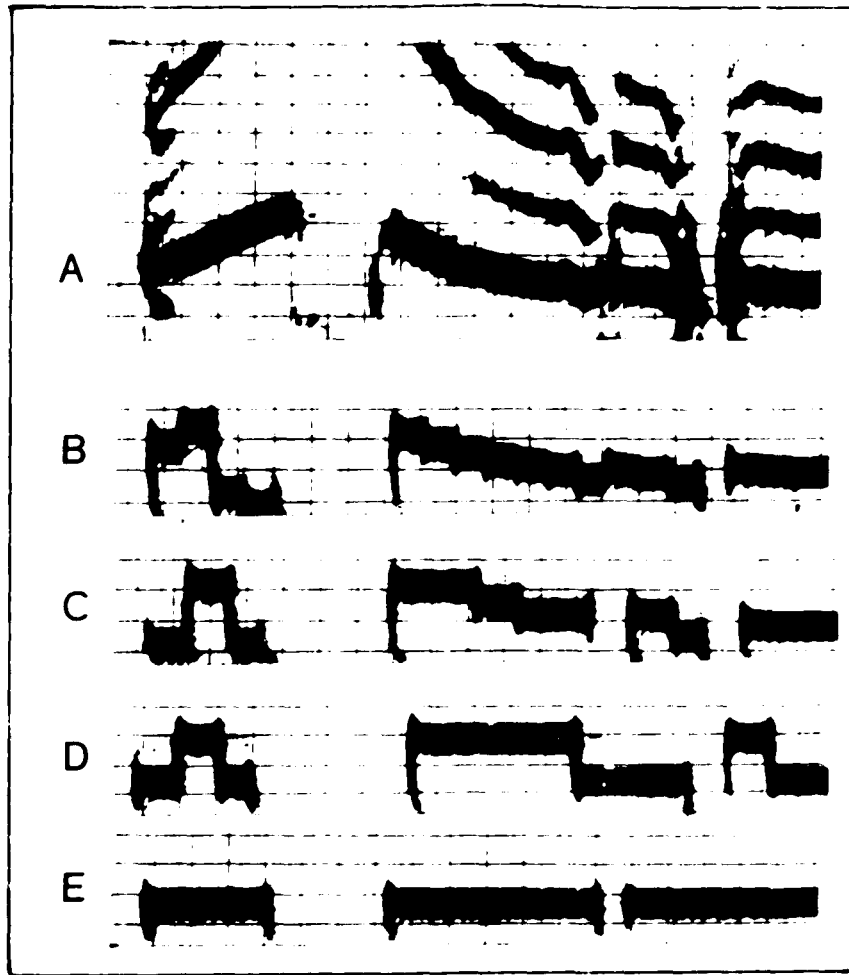


FIGURE 3.4: SAMPLE SPECTROGRAMS ILLUSTRATING (A) UNPROCESSED F_0 SIGNAL (B) 12 LEVELS OF QUANTIZATION (C) 8 LEVELS OF QUANTIZATION (D) 4 LEVELS OF QUANTIZATION (E) 1 LEVEL OF QUANTIZATION.

waves; and (4) loss of information related to amplitude variations that show in the EGG signal and result from variations in the area of vocal fold contact. Thus, it was realized that even without quantization, degradation of the F_0 contours would occur using this processing technique. It was, however, believed that frequency quantization would be the most important kind of degradation, and that these other forms of degradation introduced through the processing technique would be relatively unimportant.

Experiment

1. Instrumentation.

The instrumentation used in the experiment is illustrated in Figure 3.5. The test tapes were played on a 1/2 inch VHS recorder (Panasonic AG-6300). The video signal was displayed on a 19 inch color monitor (SONY KX1901A). The audio signal was fed to a low-pass filter set at 360 Hz for the purpose of reducing transient frequency splatter at points of frequency change in the quantized conditions. The signal was then amplified by a linear amplifier (Realistic SA 150) and fed to the subjects through TDH49 binaural headphones. An attenuator (Grason-Stadler 1293) provided adjustment of listening levels in 5 dB steps.

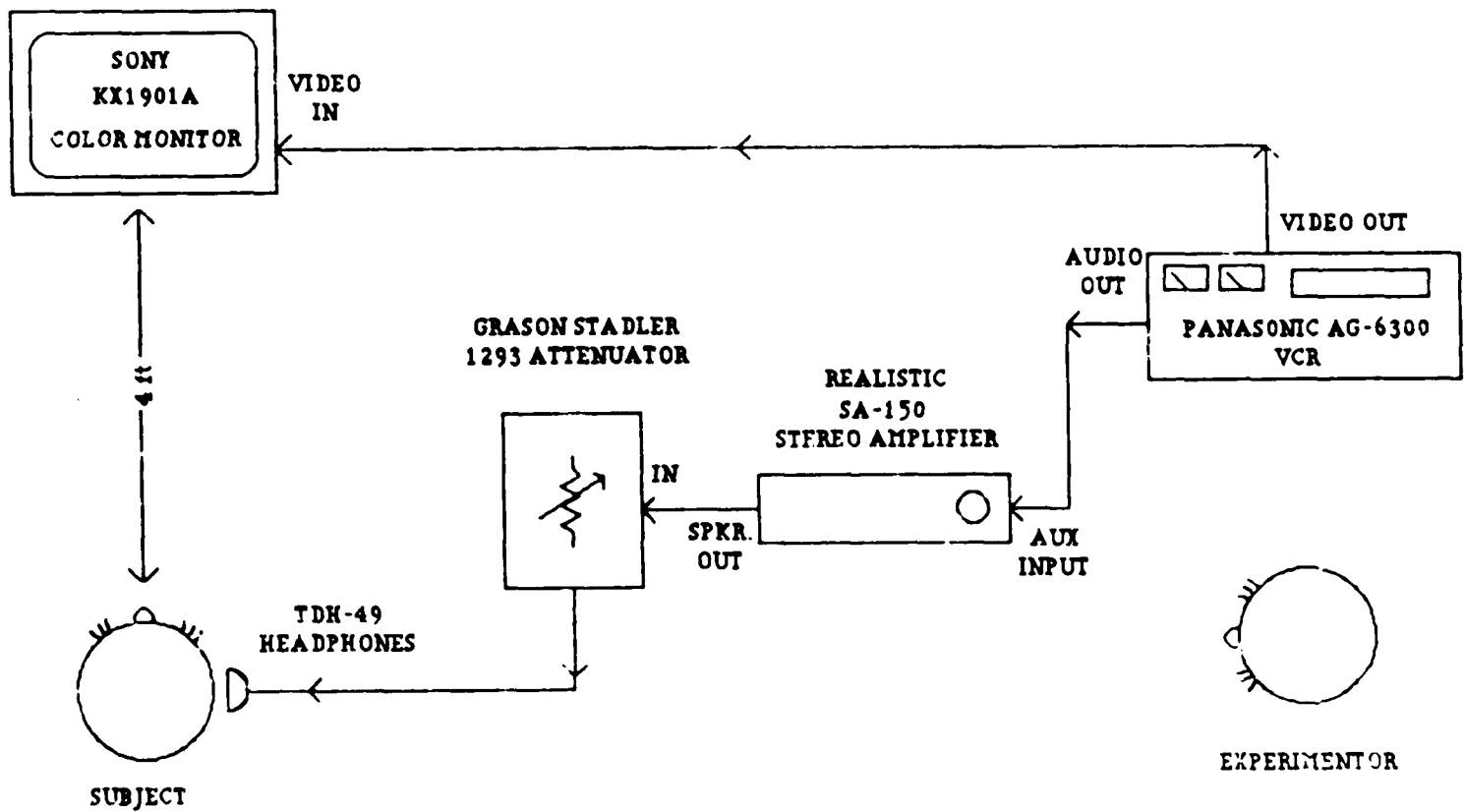


FIGURE 3.5 INSTRUMENTATION FOR TESTING SUBJECTS

2. Subjects.

Twelve native English speaking adults served as subjects. Individual subject information will be found in Appendix C. The subjects were between the ages of 24 and 40 years. Four were male and eight were female. Subjects were all from middle class backgrounds. All were college graduates. None of the subjects had any formal training in speechreading; six, however, had backgrounds in speech and hearing sciences. No subjects had a known history of hearing loss and all reportedly had normal or corrected normal vision. The information on language, hearing, and vision history was obtained by questionnaire prior to participation in the experiment (Appendix D).

3. Procedure.

To examine the effects of frequency quantization on the utility of F_0 as an acoustic speechreading supplement, performance was measured in the following conditions:

1. speechreading alone
(obtained by turning off the sound during presentation of one of the test tapes)
2. speechreading + binary voicing information
(i.e. quantization at 1 level)
3. speechreading + F_0 quantized at 4 levels
(or 2 levels/octave for the F_0 range of speaker used in this study)
4. speechreading + F_0 quantized at 8 levels
(or 4 levels/octave for the F_0 range of speaker used in this study)

5. speechreading + Fo quantized at 12 levels
(or 6 levels/octave for the Fo range
of speaker used in this study)
6. speechreading + unprocessed Fo contours
(obtained by using the EGG signal).

Subjects were tested individually in three, one and one-half hour sessions. During each session eight sentence sets were presented in each of two conditions. The first two sentence sets in each group of eight were presented to familiarize subjects with the task and to provide experience in a condition. The last six sentence sets were presented to obtain two scores, each based on the average of three sentence sets. The use of three sentence sets to obtain one score decreases error variance.

To determine the order of presentation of sentence sets one randomization of the 48 sets was made. This randomization was then divided into six groups of 8 sets each (see Table 3.1). Within each of the six sentence set groups, the order of presentation of sets was randomized for each subject. The order of presentation of conditions and sentence set groups was counterbalanced across subjects as shown in Table 3.2.

During testing the subjects were seated 4 feet from the monitor. The audio signal was adjusted for most comfortable listening prior to testing in each

auditorily supplemented condition. This level was used throughout testing in that condition.

Subjects were given written instructions prior to testing in each condition. The instructions are included in Appendix E. Subjects were informed of the topic of each sentence before viewing it. After viewing the sentence once, subjects were required to repeat as much of it as possible. The experimenter circled the words correctly perceived on printed answer sheets for subsequent scoring and analysis. Scores were the number of words correctly perceived in a group of three sentence sets. Thus, each score was based on 306 words (102 per set).

Table 3.1
Randomization of Sentence Sets and Groupings

GROUP A	GROUP B	GROUP C
40	45	9
44	8	22
39	28	16
15	36	37
13	27	30
42	3	5
12	7	1
14	31	46

GROUP D	GROUP E	GROUP F
6	41	34
38	20	19
47	11	43
26	25	29
10	18	4
2	35	33
17	32	23
21	48	24

Note: Each of the numbers corresponds to one of the sentence sets included in Appendix A.

Table 3.2
Order of Presentation of Conditions (C) and Sentence
 Set Groups (SS) for Subjects Tested in the Main
 Experiment

TEST SESSION:						
SUB- JECT	C/SS ¹	C/SS	C/SS ²	C/SS	C/SS ³	C/SS
1	I/A	II/B	III/C	IV/D	V/E	VI/F
2	III/E	IV/F	V/A	VI/B	I/C	II/D
3	V/C	VI/D	I/E	II/F	II/A	IV/B
4	II/A	I/B	IV/C	III/D	VI/E	V/F
5	IV/E	III/F	VI/B	V/A	II/C	I/D
6	VI/C	V/D	II/E	I/F	IV/A	III/B
7	I/B	II/A	III/D	IV/C	V/F	VI/E
8	III/F	IV/E	V/B	VI/A	I/D	II/C
9	V/D	VI/C	I/F	II/E	III/B	IV/A
10	II/B	I/A	IV/D	III/C	VI/F	V/E
11	IV/F	III/E	VI/A	V/B	II/D	I/C
12	VI/D	V/C	II/F	I/E	IV/B	III/A

Key to Conditions:

- I: Speechreading alone
- II: Speechreading + Fo Quantized at 1 Level
- III: Speechreading + Fo Quantized at 4 Levels
- IV: Speechreading + Fo Quantized at 8 Levels
- V: Speechreading + Fo Quantized at 12 Levels
- VI: Speechreading + Unprocessed Fo

Note: Letters (A-F) indicate sentence set groups shown in Table 3.1.

Results

The subjects' responses were analyzed as a function of condition and replication, and also as a function of condition and sentence length. The results of the former analysis will be presented first.

The percentage of words correctly perceived by each subject in each of the two replications under a condition was calculated. The scores (each based on 3 sentence sets) obtained for each subject, as well as the group means and standard deviations, will be found in Table 3.3. The group means are also illustrated in Figure 3.6.

An arcsine transformation was applied to the raw data to increase homogeneity of variance (Brownlee, 1965). Table 3.4 shows the results of a repeated measures analysis of variance in the arcsine transformed data. The main effect of condition was highly significant ($F(5,55)=48.25, p<.01$). However, the main effect of replication was not significant ($F(1,55) = 1.62, p>.05$), nor was the condition by replication interaction ($F(5,55) = 1.94, p>.05$). The lack of significant replication and interaction effects will also be seen in Figure 3.7, which illustrates the mean scores obtained for each sentence set presentation.

Table 3.3
Word Recognition Scores as a Function of Test Condition
and Replication for 12 Subjects Tested in the Main
Experiment

SUB- JECT	TRIAL	CONDITION					
		SR	SR+1	SR+4	SR+8	SR+1	SR+UNP
1	1	18.0	23.2	47.9	38.9	57.5	81.7
	2	16.3	29.1	39.6	44.8	57.8	80.4
2	1	48.3	64.7	42.8	56.5	66.0	83.0
	2	54.6	71.5	61.4	62.4	56.2	86.6
3	1	37.9	43.1	52.6	66.3	53.4	76.2
	2	39.2	54.2	61.4	59.1	57.5	88.7
4	1	21.6	11.4	23.8	38.9	19.0	40.5
	2	10.1	11.4	11.8	44.8	15.7	44.5
5	1	19.9	35.3	34.7	36.4	59.5	69.9
	2	25.5	42.8	40.8	34.3	46.7	67.6
6	1	52.3	35.9	38.2	57.8	48.0	63.7
	2	40.5	42.1	44.4	67.0	49.7	67.3
7	1	15.4	42.2	36.2	55.8	58.5	75.5
	2	14.0	35.3	42.9	38.9	68.3	76.8
8	1	51.0	65.0	59.5	64.1	84.0	85.2
	2	61.1	66.4	59.2	67.3	86.9	91.2
9	1		46.4	39.9	62.4	44.4	88.2
	2	36.9	62.8	51.3	50.0	46.8	80.7
10	1	16.7	6.5	25.8	37.3	42.8	59.8
	2	8.5	10.1	21.2	21.6	41.5	69.3
11	1	38.6	51.3	54.6	45.9	67.1	78.1
	2	36.9	56.2	56.7	54.0	64.5	82.3
12	1	34.3	38.8	42.8	51.0	54.5	69.0
	2	36.7	56.5	46.1	47.1	50.3	72.2
MEAN	1	32.2	38.1	41.3	49.7	54.6	72.9
	2	32.2	43.1	44.8	47.1	54.5	75.7
STAND. DEV.	1	14.4	18.4	10.9	13.3	15.8	13.6
	2	16.9	19.9	14.2	16.4	17.5	13.8

Key to conditions: SR: Speechreading alone; SR+1:
Speechreading + Fo Quantized at 1 Level; SR + 4:
Speechreading + Fo quantized at 4 levels; SR + 8:
Speechreading + Fo Quantized at 8 Levels; SR + 12:
Speechreading + Fo Quantized at 12 Levels; SR + UNPR:
Speechreading + Unprocessed Fo (EGG Signal)

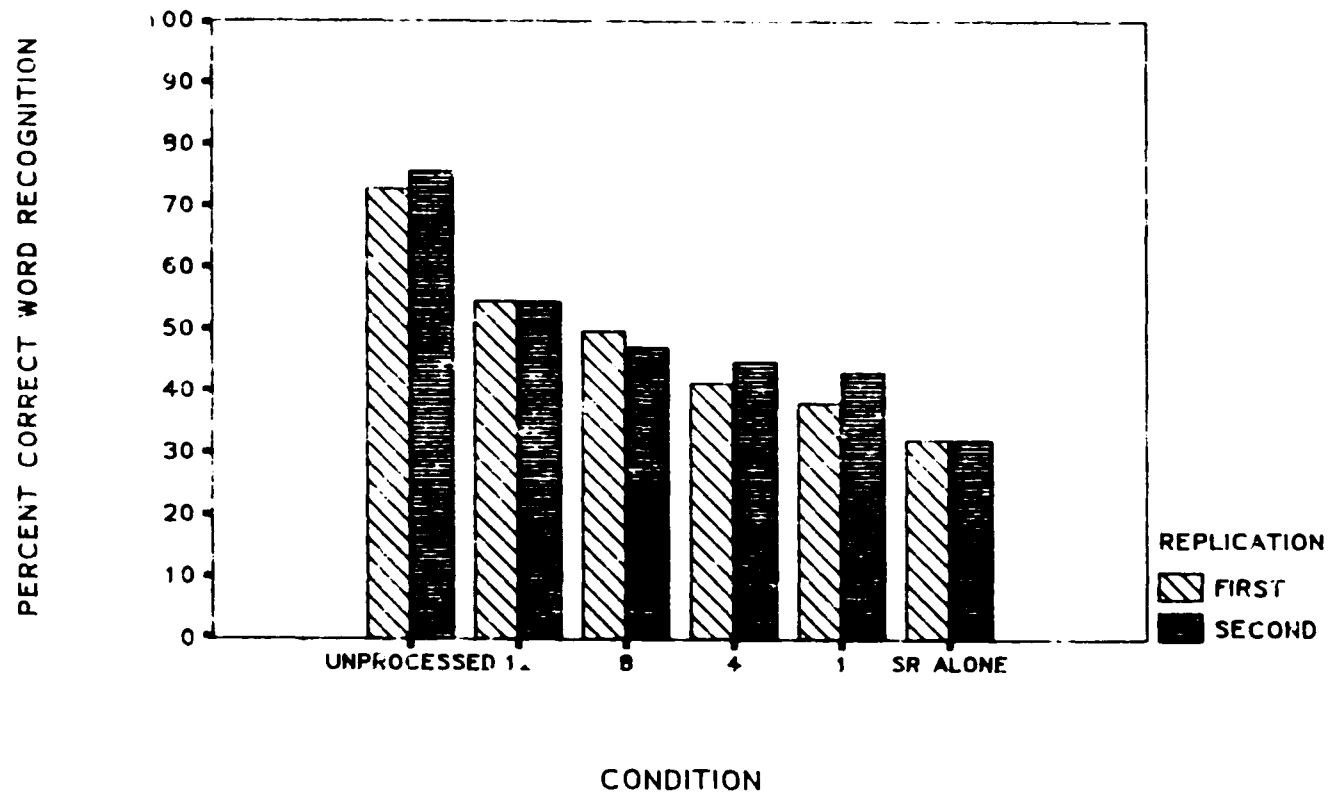


FIGURE 3.6: MEAN PERFORMANCE AS A FUNCTION OF CONDITION AND REPLICATION IN THE MAIN EXPERIMENT

Table 3.4
ANOVA Summary for the Main Effects of Condition and
Replication and the Condition by Replication
Interaction in the main experiment

SOURCE	dF	MEAN SQUARE	F
CONDITION (C)	5	.6031	48.25*
error	55	.0125	
REPLICATION (R)	1	.0080	1.62
error	11	.0049	
C x R	5	.0060	1.94
error	55	.0029	

* Significant at $p < .01$

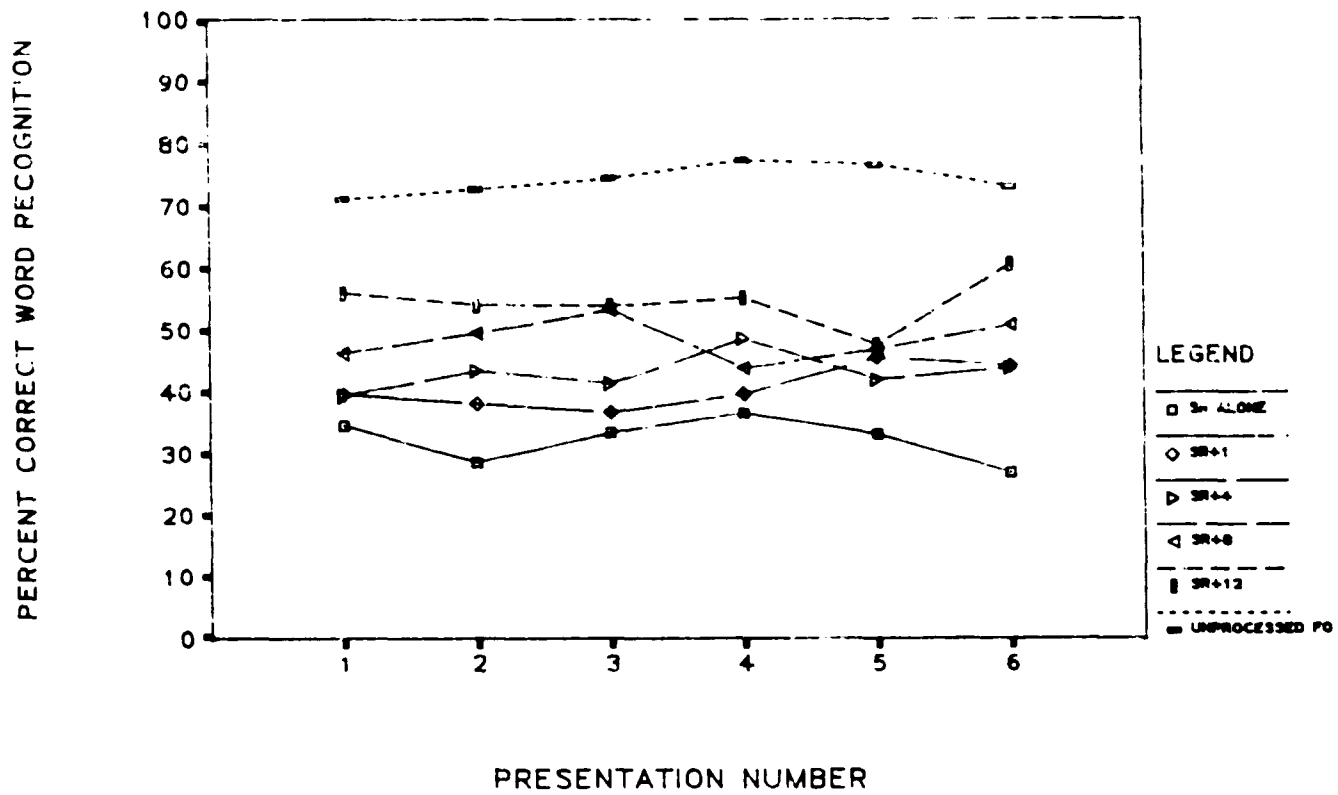


FIGURE 3.7: MEAN PERFORMANCE AS A FUNCTION OF CONDITION AND SENTENCE SET PRESENTATION IN THE MAIN EXPERIMENT

In order to determine the significance of differences among the means for the conditions, the Student Newman-Keuls procedure was applied. The results are summarized in Table 3.5. Performance in all auditorily supplemented speechreading conditions was significantly better than that obtained by speechreading alone. Significant differences were found among all quantized conditions except between the 1 level and 4 level conditions and the 4 level and 8 level conditions. However, 8 levels was significantly better than 1 level. Performance in the speechreading + unprocessed Fo condition was significantly better than performance in any of the quantized Fo conditions.

Table 3.5
Results of Student-Newman Keuls Procedure Applied to
Differences Between the Condition Means of the Arcsine
Transformations of the Raw Data

CONDITION:							
	SR	SR+1	SR+4	SR+8	SR+12	SR+UNP	CRITICAL VALUE
CONDI- TION							
SR	-	.09*	.12*	.18*	.24*	.46*	.09
SR+1	-	-	.03	.12*	.15*	.37*	.09
SR+4	-	-	-	.05	.12*	.34*	.08
SR+8	-	-	-	-	.07*	.28*	.07
SR+12	-	-	-	-	-	.21*	.06

* Significant at $p < .05$.

Key to Conditions:

SR: Speechreading alone

SR + 1: Speechreading + Fo Quantized at 1 Level

SR + 4: Speechreading + Fo Quantized at 4 Levels

SR + 8: Speechreading + Fo Quantized at 8 Levels

SR + 12: Speechreading + Fo Quantized at 12 Levels

SR + UNP: Speechreading + Unprocessed Fo (EGG Signal)

Each of the sentences sets used in this investigation contained sentences varying in length from 3-12 words and each set contained one sentence of each length. Thus, an equal number of sentences of each length were presented in all conditions in the main experiment. The responses of the subjects were rescored to determine the percentage of words correctly recognized in short (3-5 words), medium (6-9 words), and long (10-12 words) sentences.

The raw data for this evaluation are presented in Appendix F. The means for each sentence length in each condition are presented in Table 3.6 and are also illustrated in Figure 3.8. The results of a repeated measures analysis of the variance in the arcsine transforms of the raw data are shown in Table 3.7. The most important finding of this analysis is that there was a highly significant interaction between sentence length and condition ($F(10,110) = 3.21, p < .01$). It will be seen, from an examination of Figure 3.7, that sentence length affected performance in all conditions except speechreading + unprocessed Fo. The mean percentage correct scores are also shown in Figure 3.9. This figure highlights the effect of condition on performance for each of the sentence lengths. It will be seen that condition appears to have less of an effect on the performance measured for short sentences

than it did for that measured for medium and longer length sentences.

Table 3.6
Mean Percentage Correct Word Recognition Scores as a
Function of Sentence Length and Condition

CONDITION	SHORT	MEDIUM	LONG
SR + UNP Fo	74.1	73.3	72.4
SR + 12	63.8	54.5	49.9
SR + 8	54.5	49.7	44.2
SR + 4	56.5	45.7	36.9
SR + 1	53.3	41.0	35.3
SR ALONE	46.6	32.6	27.2

Key to Conditions:

SR: Speechreading alone

SR + 1: Speechreading + Fo Quantized at 1 Level

SR + 4: Speechreading + Fo Quantized at 4 Levels

SR + 8: Speechreading + Fo Quantized at 8 Levels

SR + 12: Speechreading + Fo Quantized at 12 Levels

SR + UNP: Speechreading + Unprocessed Fo (EGG Signal)

Table 3.7
ANOVA Summary for the Effect of Sentence Length and Condition

SOURCE	df	MEAN SQUARE	F
CONDITION (C)	5	.7223	32.59*
error	55	.0221	
LENGTH (L)	2	.3739	33.84*
error	22	.0111	
C x L	10	.0261	3.21*
error	110	.0081	

* Significant at $p < .01$

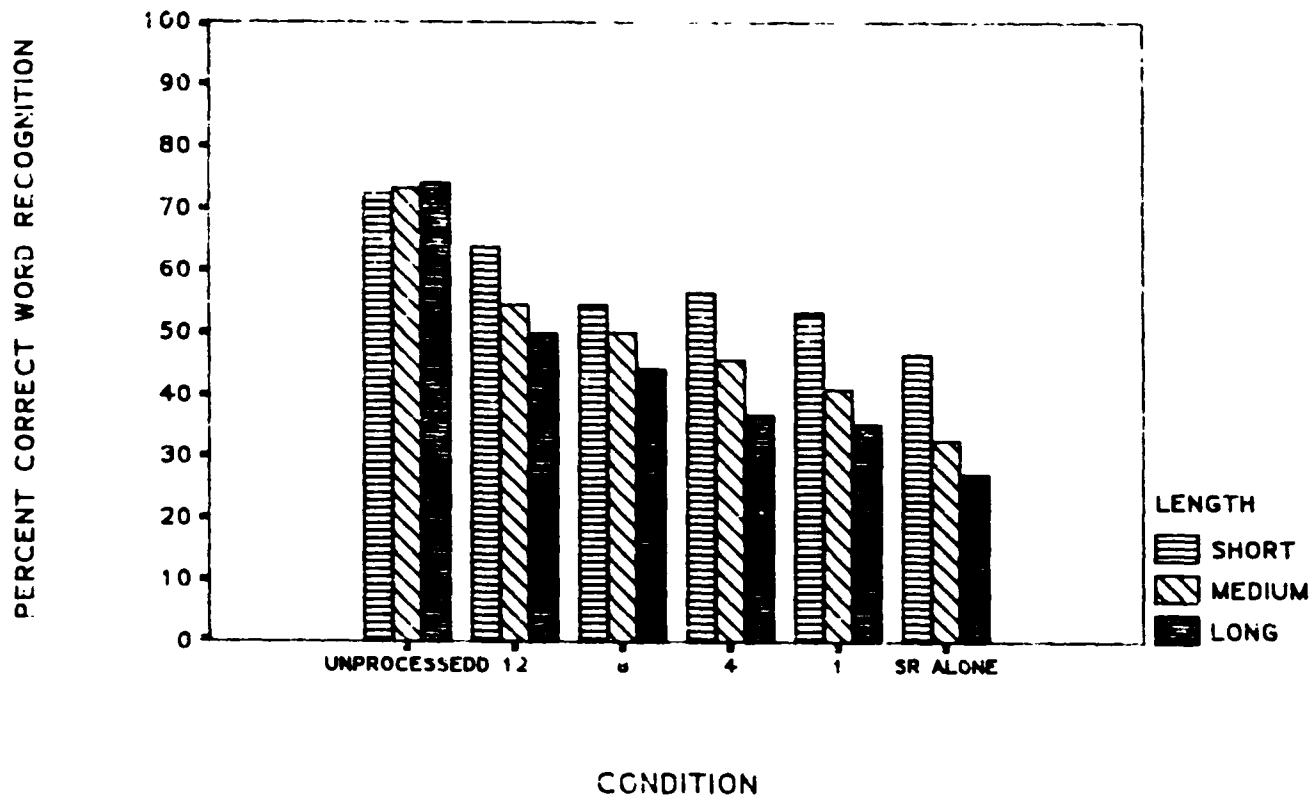


FIGURE 3.8: MEAN PERFORMANCE AS A FUNCTION OF SENTENCE LENGTH WITHIN CONDITIONS IN THE MAIN EXPERIMENT

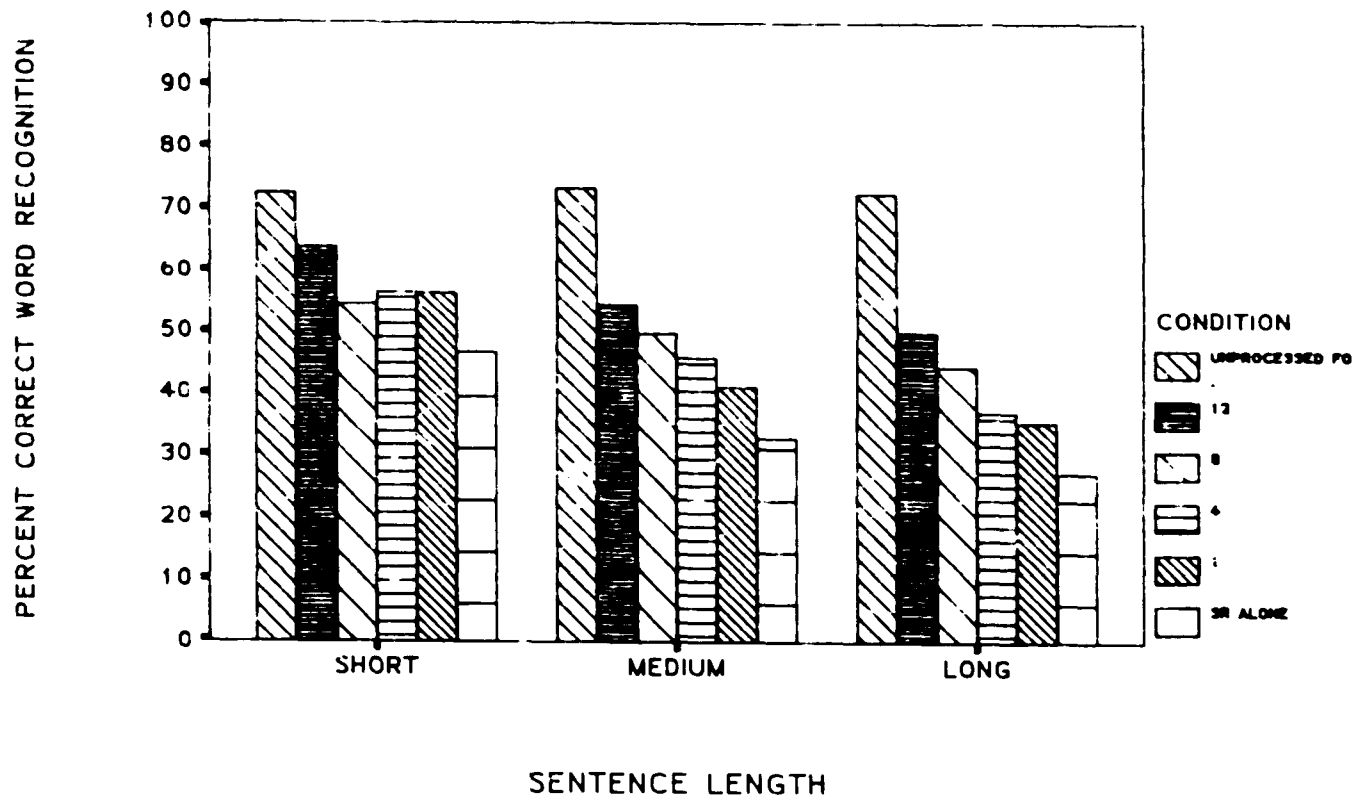


FIGURE 3.9: MEAN PERFORMANCE AS A FUNCTION OF CONDITION WITHIN SENTENCE LENGTH IN THE MAIN EXPERIMENT

Summary of results

1. As hypothesized, the speechreading performance of normal subjects significantly increased with the addition of auditorily presented fundamental frequency information. The best supplemented speechreading performance was obtained when fundamental frequency was presented using the EGG signal, that is, without quantization.
2. The results failed to support the hypothesis that a critical level of quantization would be found at which auditorily supplemented speechreading performance is equivalent to that obtained with unprocessed F_0 contours.
3. As hypothesized, as the number of frequency quantization levels diminished supplemented speechreading performance decreased in a monotonic fashion.
4. As hypothesized, utterance length was an important factor in measuring the effects of supplementing speechreading with auditorily presented F_0 .

Preliminary Discussion of Results

The results of this experiment supported all of the hypotheses of this investigation with one exception. A critical level of quantization was not found. One possible interpretation for this is that even the finest degree of quantization used in this

study (i.e. 12 levels) represents a serious form of degradation. This interpretation implies that if a critical level of quantization of F_0 contours exists, it must be greater than 12. However, it will be recalled that forms of degradation other than quantization were introduced through the processing technique used in this study. Although it was assumed that these other forms of degradation would have minimal influence on performance, the validity of this assumption was not tested. It is possible that this assumption was erroneous and that degradation introduced through the processing technique alone may have rendered the quantized F_0 contours less useful than the unprocessed contours. Thus, drawing any conclusions about the effects of quantization at 12 levels would be premature without first empirically evaluating the efficacy of the signal processing technique. This was the purpose of the supplementary experiment to be described in Chapter 4.

CHAPTER 4
SUPPLEMENTARY EXPERIMENT

Purpose

In the main experiment the signal quantized at 12 levels proved to be a significantly poorer speechreading supplement than the unprocessed EGG signal. It was not clear, however, what caused the difference. Was quantization at 12 levels sufficient to account for the effect, or were there unexpected degradations introduced by processing itself, even before quantization was introduced? The purpose of the supplementary experiment was to answer the second question.

Methodology

Preparation of test material

The test material consisted of the same topic-related sentence sets used in the main experiment and described in detail in Chapter 3. A test tape was made which contained a Fo signal which was processed but not quantized into a discrete number of levels. The instrumentation used to prepare this test tape was the same as was used to prepare the test tapes for the

main experiment (see Figure 3.3). To obtain this processed F_0 signal the digital representation of the Visipitch (Kay 6397) output was translated via the IBM PC-XT and the D/A converter (ISSAC 91A) for input to the voltage controlled oscillator (Coulbourn S24-05). The frequency of the constant amplitude sine wave output of the VCO was the same as the sampled F_0 values. Sample spectrograms showing the filtered EGG signal which provided input into Visipitch (Kay 6397) and the processed F_0 signal are shown in Figure 4.1.

Experiment

1. Instrumentation

The instrumentation used for the experiment remain unchanged from that used in the main experiment. It is described in detail in Chapter 3 and is illustrated in Figure 3.5

2. Subjects

Subject selection criteria were the same as that used in the main experiment. Individual subject history for the twelve native English speaking adults who served as subjects will be found in Appendix C. None of the subjects had participated in the main experiment. The subjects were between the ages of 18 and 38 years. Two were male and ten were female. All were from middle class backgrounds. All but one was a

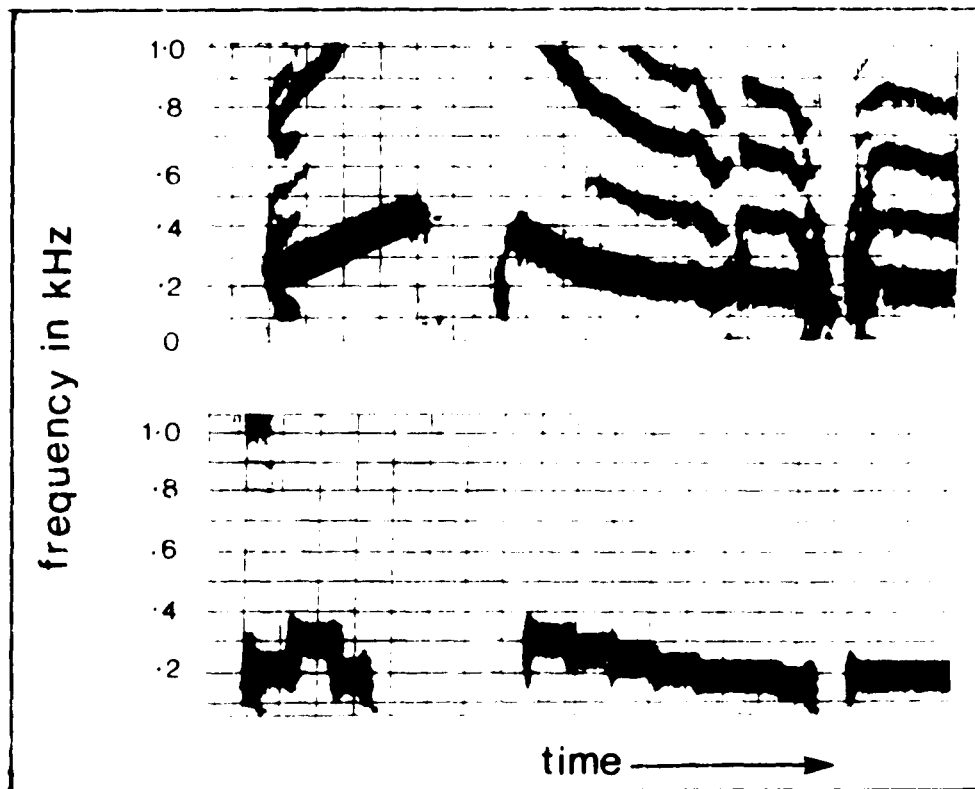


FIGURE 4.1: SAMPLE SPECTROGRAMS OF UNPROCESSED F_0 SIGNAL (UPPER) AND PROCESSED F_0 SIGNAL (LOWER).

college graduate. None had any formal training in speechreading; however ten had backgrounds in speech and hearing sciences. None had a known history of hearing loss and all had reportedly normal or corrected normal vision. To determine language, hearing, and vision history each subject was required to complete a questionnaire prior to participation in the experiment (Appendix D).

3. Procedure

To evaluate the effects of processing without quantization into a discrete number of levels subjects were tested in the following conditions:

1. speechreading alone
2. speechreading + processed Fo
3. speechreading + unprocessed Fo

The test tape prepared for use in the main experiment which contained the EGG signal was used for the unprocessed Fo condition in this experiment. The speechreading alone condition was obtained by simply turning off the audiosignal of one of the test tapes.

The procedures for the supplementary experiment were the same as for the main experiment except that subjects were seen only one time and only five sentence sets were presented in each of the three conditions. As in the first experiment, the first two sentence sets in each group of five were presented to provide the

subjects with practice in a condition. Only one score was obtained for each condition. This score was based on the average of the last three sentence sets (306 words). The order of presentation of sentence sets was randomized across conditions for each subject. The order of presentation of conditions was counterbalanced across subjects as shown in Table 4.1.

Table 4.1
Order of Presentation of Conditions for Subjects in the
Supplementary Experiment

SUBJECT	CONDITION		
	FIRST	SECOND	THIRD
1	SR	SR+PROC	SR+Fo
2	SR+PROC	SR+Fo	SR
3	SR+Fo	SR	SR+PROC
4	SR	SR+PROC	SR+Fo
5	SR+PROC	SR	SR+Fo
6	SR+Fo	SR+PROC	SR
7	SR	SR+PROC	SR+Fo
8	SR+PROC	SR+Fo	SR
9	SR+Fo	SR	SR+PROC
10	SR	SR+PROC	SR+Fo
11	SR+PROC	SR	SR+Fo
12	SR+Fo	SR+PROC	SR

Key to conditions:

SR: Speechreading alone

SR+PROC: Speechreading + Processed Fo Contours

SR+Fo: Speechreading + Unprocessed Fo Contours

Results

The percentage of words correctly perceived by each subject in the three conditions is presented in Table 4.2. The mean scores and standard deviations for each condition are also presented in this table. The mean scores are also illustrated in Figure 4.2. It will be seen that although performance in both auditorily supplemented conditions was better than that by speechreading alone, performance with the unprocessed F₀ signal was better than that with the processed F₀ signal.

The results of a repeated measures analysis of the variance in the arcsine transforms of the raw data are summarized in Table 4.3. The main effect of condition was found to be highly significant ($F_{2,22} = 54.06, p < .01$). Further statistical analysis, using the Student Newman-Keuls procedure, revealed that the differences among performance in the three conditions were significant ($p < .05$). These results are shown in Table 4.4. Thus, the results failed to support the hypothesis that processing F₀ contours without quantization into a discrete number of levels would have little or no effect on performance.

Table 4.2
Word Recognition Scores as a Function of Condition
Obtained by 12 Subjects in the Supplementary Experiment

SUBJECT	CONDITION		
	SR ALONE	SR + PROC	SR + UNPROC
1	70.6	97.7	97.1
2	58.2	69.3	89.9
3	30.4	33.7	57.9
4	58.8	79.7	87.2
5	72.9	89.5	89.9
6	2.9	4.2	23.2
7	31.7	57.7	85.3
8	31.4	35.0	74.2
9	60.8	81.4	93.8
10	32.0	81.7	80.4
11	60.8	81.4	93.8
12	8.7	36.9	61.4
MEANS	43.3	62.4	77.8
S.D.	23.5	28.7	21.3

Key to conditions:

SR: Speechreading alone

SR+PROC: Speechreading + Processed Fo Contours

SR+Fo: Speechreading + Unprocessed Fo Contours

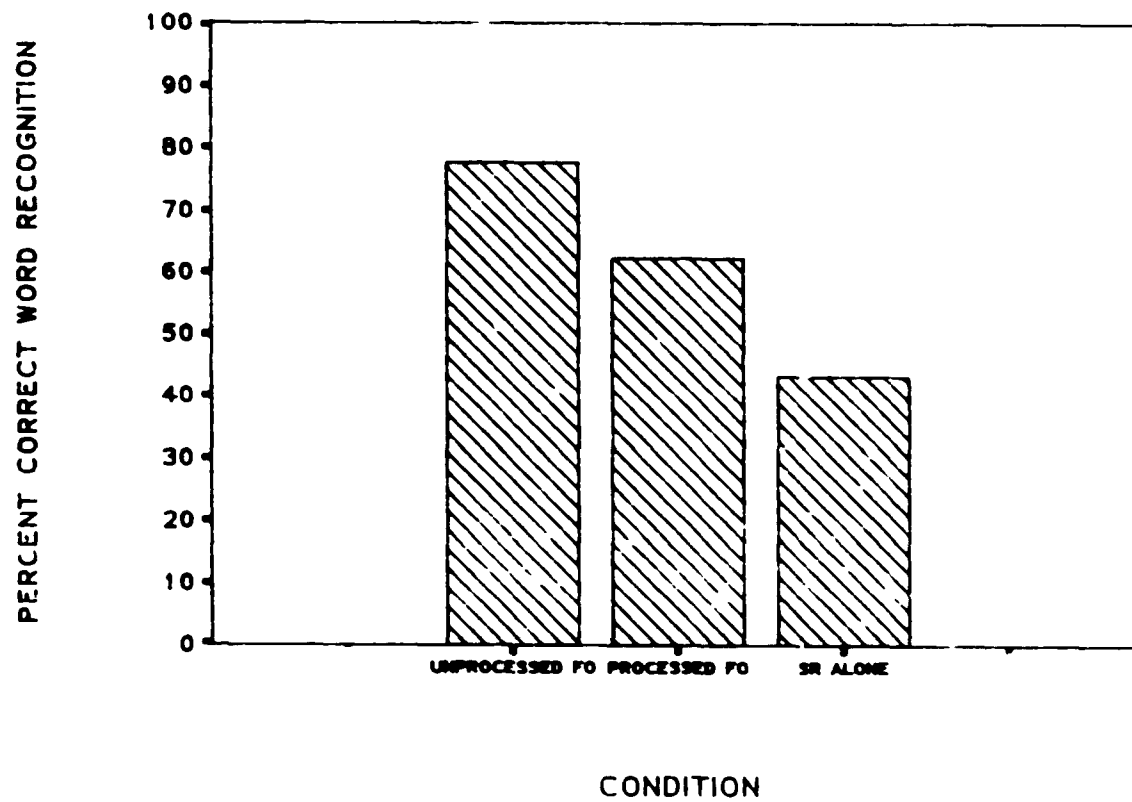


FIGURE 4.2: MEAN PERFORMANCE AS A FUNCTION OF CONDITION IN THE SUPPLEMENTARY EXPERIMENT

Table 4.3
ANOVA Summary for the Main Effect of Condition in the
Supplementary Experiment

SOURCE	dF	MEAN SQUARE	F
CONDITION	2	.5085	54.06*
error	22	.0094	

* Significant at $p < .01$

Table 4.4
Results of Student-Newman Keuls Procedure Applied to
the Means of the Arcsine Transforms of the Raw Data in
the Supplementary Experiment

CONDI- TION	CONDITION			CRITICAL VALUE
	SR ALONE	SR + PROC Fo	SR + UNPROC Fo	
SR	-	.22*	.41*	.10
SR+PROC	-	-	.18*	.08

* Significant at $p < .01$

Key to conditions:

SR: Speechreadingg alone

SR+PROC: Speechreading + Processed Fo Contours

SR+Fo: Speechreading + Unprocessed Fo Contours

Preliminary Discussion

The results of this experiment demonstrated that the processing technique used in this study introduced sufficient degradation to reduce markedly the effectiveness of F_0 contours as an acoustic speechreading supplement. This finding aids in the interpretation of the results of the main experiment. It will be recalled that the results of the main experiment left unresolved the question of why a critical level of frequency quantization was not found. Since processing alone significantly reduced the effectiveness of the F_0 contours, it would not have been possible to find a level of frequency quantization at which performance was equivalent to that obtained with unprocessed F_0 contours (i.e. EGG signal). It may, however, be possible to find a level of frequency quantization at which performance is equivalent to that obtained with processed F_0 contours. A comparison of the results of the main and supplementary experiments would be necessary to determine if this is the case. It was to make such a comparison that the post-hoc analysis described in Chapter 5 was conducted.

CHAPTER 5

COMPARISON OF RESULTS OF THE MAIN AND SUPPLEMENTARY
EXPERIMENTS

Purpose

The results of the supplementary experiment demonstrated that it would not have been possible, in the main experiment, to find a quantization level giving equivalent results to the EGG signal. It might, however, have been possible to find a quantization level equivalent to the processed, but unquantized signal. The purpose of this comparison is to address this issue.

Methodology

The difference between the mean speechreading alone scores obtained in the main (32.2%) and supplementary (43.3%) experiments made a direct comparison of the results difficult. There are many possible methods of comparing the changes in speechreading performance resulting from supplementation. The one chosen here was to calculate the percentage reduction of speechreading error that resulted from the introduction of the acoustic supplement. This method has theoretical merit and has

been previously used. In a 1975 report, Boothroyd, investigated the effects of combining speechreading with band-pass filtered speech. Using probability theory he showed that:

$$(1-P_{hs}) = (1-P_h) (1-P_s) (1-P_i) \quad (1)$$

where P_h = the probability of word recognition by hearing alone

P_s = the probability of word recognition by speechreading alone

P_i = the probability of word recognition on the basis of cues that only become available when hearing and speechreading interact.

P_{hs} = the probability of word recognition by hearing and speechreading in combination.

Note that equation (1) assumes that P_h , P_s , and P_i are statistically independent.

In the present study it may be assumed that $P_h = 0$, i.e. that subjects presented with F_0 contours alone will recognize no words. Thus the supplementary effect of F_0 becomes attributable solely to an interaction between vision and hearing and

$$(1-P_{hs}) = (1-P_s) (1-P_i) \quad (2)$$

The probability of word recognition on the basis of cues that only become available through auditorily presented F_0 interacting with speechreading P_i can be calculated as follows:

$$P_i = 1 - \frac{1 - P_{hs}}{1 - P_s} \quad (3)$$

$$= \frac{1 - P_s}{1 - P_s} - \frac{1 - P_{hs}}{1 - P_s} \quad (4)$$

$$= \frac{(1 - P_s) - (1 - P_{hs})}{1 - P_s} \quad (5)$$

$$= \frac{1 - P_s - 1 + P_{hs}}{1 - P_s} \quad (6)$$

$$= \frac{P_{hs} - P_s}{1 - P_s} \quad (7)$$

The quantity on the right of Equation 7 is the reduction of error divided by the maximum possible reduction of error. Multiplying by 100 gives the percentage reduction of error.

Results

The mean percentage reduction of error calculated for each of the supplemented conditions in both experiments is presented in Table 5.1 and illustrated in Figure 5.1. It will be seen that the scores for the two conditions tested in both experiments (i.e. speechreading alone and speechreading + unprocessed Fo) were equated using the comparison method chosen here. Thus, the method has empirical as well as theoretical merit. The most important finding of this comparison is that the mean percentage reduction of errors was equivalent for the processed Fo condition and the Fo

Table 5.1
Percent Reduction of Error in Supplemented
Speechreading Conditions in the Main and Supplementary
Experiments

CONDITION	MAIN EXPERIMENT	SUPPLEMENTARY EXPERIMENT
SPEECHREADING + F ₀ QUANTIZED AT 1 LEVEL	12.4	NOT TESTED
SPEECHREADING + F ₀ QUANTIZED AT 4 LEVELS	16.1	NOT TESTED
SPEECHREADING + F ₀ QUANTIZED AT 8 LEVELS	23.9	NOT TESTED
SPEECHREADING + F ₀ QUANTIZED AT 12 LEVELS	32.9	NOT TESTED
SPEECHREADING + PROCESSED	NOT TESTED	33.5
SPEECHREADING + UNPROCESSED F ₀	62.1	60.8

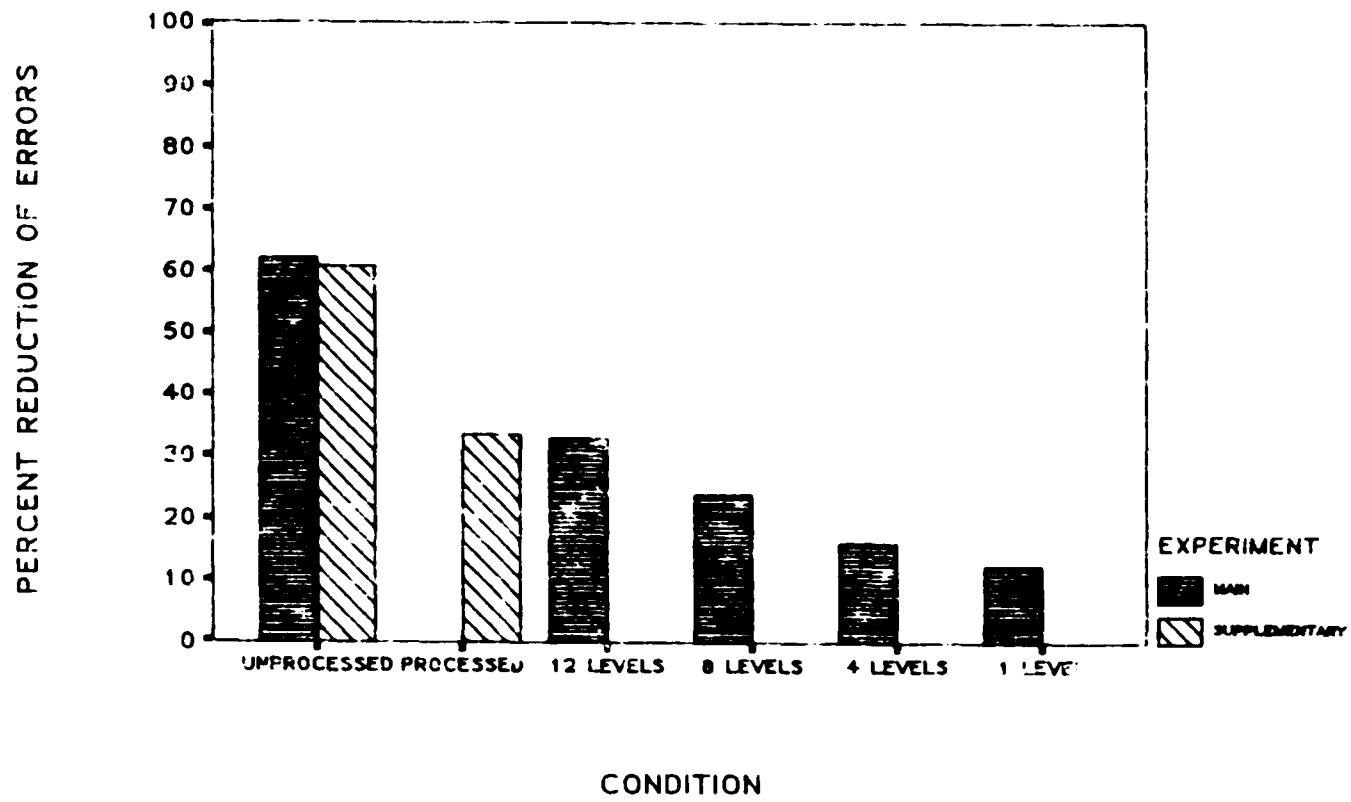


FIGURE 5.1: PERCENT REDUCTION OF ERROR DUE TO SUPPLEMENTATION IN THE MAIN AND SUPPLEMENTARY EXPERIMENTS

quantized at 12 levels condition. Thus, the addition of quantization at 12 levels does not appear to have further reduced the effectiveness of previously processed Fo contours.

Preliminary Discussion

The finding that the mean percentage reduction of errors was equivalent for the processed Fo and 12 levels of quantization conditions strongly suggests that the failure to find a critical level of quantization was due to degradation introduced through the processing technique. It is also possible, however, that the effects of processing were masking the effects of quantization at 12 levels. That is, quantization at 12 levels does have an effect, but the effect was weaker than the effects of processing. It is clear, however, that as the number of quantization levels decreased beyond 12 the utility of the Fo contours diminishes in a monotonic fashion. That is, we can conclude that 8 levels of quantization is worse than 12.

CHAPTER 6

EFFECTS OF PITCH EXTRACTION ERRORS ON PERFORMANCE

Purpose

As previously noted the processing technique used in this study was not ideal as it introduced several forms of degradation besides frequency quantization. Thus, the decrements in performance as a result of processing the F_0 contours may have been due to any one or some combination of the following factors: (1) pitch extraction errors; (2) loss of amplitude information; (3) loss of harmonic structure; (4) sampling time effects. Although it was not possible to independently examine the contribution of each of these factors, it was believed that the relative effects of pitch extraction errors might be evaluated using the available data.

The purpose of this evaluation was to examine the relative effect of pitch extraction errors by comparing performance for two types of sentences: (1) those which contained pitch extraction errors after processing (to be called "error sentences"); (2) those which did not contain pitch extraction errors after processing (to be called "non-error sentences").

It was hypothesized that if pitch extraction errors were not important than there would be no significant difference in the scores obtained on the two sentence types under any of the conditions. If, however, pitch extraction errors were important there would be a significant difference in performance for the two types of sentences under the speechreading + processed Fo condition.

Methodology

Spectrographic analysis of a sample of sentences (see Appendix G) allowed for the selection of six error sentences and six non-error sentences. Each of the twelve subjects who participated in the supplementary experiment had received each of the twelve sentences under only one of the three conditions tested in the supplementary experiment. These were: (1) speechreading alone; (2) speechreading + processed Fo; and, (3) speechreading + unprocessed Fo (i.e. EGG).

The percentage of words correctly perceived in each of these sentences by each subject was calculated. The results are shown in Tables 6.1 and 6.2. It will be seen that for any particular sentence the number of subjects receiving it in one of the three conditions was not equal. For example, sentence #8 from Set 6 was received under the speechreading alone condition by two

Table 6.1
Percentage Words Correctly Recognized for each of the
Non Error Sentences by Each Subject in Each Condition

SENT COND	SUBJECT											
	1	2	3	4	5	6	7	8	9	10	11	12
A												
SR		100				0	33	100	0	17		
SR+PR	100		100									100
SR+UN				100	100						83	
B												
SR				80	100							0
SR+PR			0					0	100	80		
SR+UN	100	100				0	100		100			
C												
SR		100		0							75	
SR+PR					100	0	100	75	88			
SR+UN	100		100							100		88
D												
SR	78		0								44	
SR+PR		100				11	67		78			
SR+UN				100	89			100		100		89
E												
SR					77							
SR+PR	100		0							69		0
SR+UN		77		100		8	100	92	85		77	
F												
SR	100				100		75	25				
SR+PR				67						100		
SR+UN		100	75			8					100	100
												8

Key to Conditions: SR: Speechreading alone; SR+PR: Speechreading + Processed Fo; SR + UN: Speechreading + Unprocessed Fo. Key to Sentences: A:SET 8,#4; B:SET 10,#10; C:SET 9,#10; D:SET 13,#8; E:SET 14,#6; F:SET 22,#1

Table 6.2
Percentage Words Correctly Recognized for each of the
Error Sentences by Each Subject in Each Condition

SENT COND	SUBJECT											
	1	2	3	4	5	6	7	8	9	10	11	12
G												
SR				100		0			25		100	0
SR+PR			100		100			75		100		
SR+UN	100	100					100					
H												
SR						0	100					
SR+PR		0						50			100	100
SR+UN	100		100	100	100				100	100		
I												
SR	100		43	0								0
SR+PR		0			100				29		100	
SR+UN						0	43	86		100		
J												
SR	75				100		50	63				
SR+PR				75					100			
SR+UN		100	88			0				100	100	50
K												
SR	0		0									0
SR+PR		20				0	0		0			
SR+UN				80	40			40		10		0
L												
SR	67	25			92	8			58			
SR+PR				100			0			92	83	83
SR+UN			58					67				

Key to Conditions: SR: Speechreading alone; SR+PR: Speechreading + Processed Fo; SR + UN: Speechreading + Unprocessed Fo. **Key to Sentences:** G:SET 5, #9; H:SET 6, #8; I:SET 2, #3; J:SET 22, #4; K:SET 13, #5; L:SET 19, #4.

subjects, under the speechreading + processed F₀ condition by four subjects, and under the speechreading + unprocessed F₀ condition by six subjects. This resulted in an unequal number of data points for each of the two types (i.e. error and non-error) of sentences under each of the three conditions. There are several possible ways to examine the effect of pitch extraction errors from the available data. Two of these methods were used here. These were: (1) two-way analysis of variance for a randomized design; and, (2) T-test on differences between matched-pairs of supplemented speechreading scores.

1. Two-way analysis of variance for a randomized design.

In order to perform the analysis of variance the data sets were truncated to provide an equal number of observations in each cell. (Decisions about which observations to eliminate were made at random.) The raw data for this analysis are presented in Table 6.3. The group means are illustrated in Figure 6.1

The results of a two-way analysis of variance on the raw data are shown in Tables 6.4. It will be seen that the main effect of condition was significant ($F(2, 119) = 7.62, p < .05$). As would be expected performance for both sentence types was best in the

Table 6.3
Raw Data for ANOVA for the Effect of Pitch Extraction
Errors

	NON-ERROR SENTENCES			ERROR SENTENCES		
	SR	SR+PROC	SR+UNPROC	SR	SR+PROC	SR+UNPROC
	100	100	100	75	100	100
	33	100	100	0	100	100
O	0	100	83	25	75	100
B	0	0	100	100	100	100
S	100	0	100	0	0	100
E	17	100	0	0	50	100
R	80	80	100	100	100	100
V	100	100	100	100	100	100
A	0	0	100	43	0	100
T	100	100	100	0	100	0
I	0	75	100	0	29	43
O	75	88	100	75	100	86
N	78	100	88	100	75	100
S	0	11	100	50	100	100
	44	67	89	0	20	88
	77	78	100	0	0	0
	100	100	100	0	0	100
	100	0	89	67	0	100
	75	69	77	25	100	50
	25	0	100	92	0	80

Key to conditons: SR: Speechreading alone SR+ PROC:
 Speechreading + Processed Fo; SR + UNPROC:
 Speechreading + Unprocessed Fo

Table 6.4
ANOVA SUMMARY for the Main Effect of Condition and Sentence Type and the Interaction Between Condition and Sentence Type. (Condition: Speechreading alone, Speechreading + Processed Fo, & Speechreading + Unprocessed Fo; Sentence Type: ERROR, NON-ERROR)

SOURCE	df	MEAN SQUARE	F
CONDITION (C)	2	11660.6	7.62*
error	119	1526.2	
SENTENCE TYPE (ST)	1	2457.1	1.61
error	119	1526.2	
C x ST	2	224.1	0.07
error	119	1526.2	

* Significant at $p < .05$

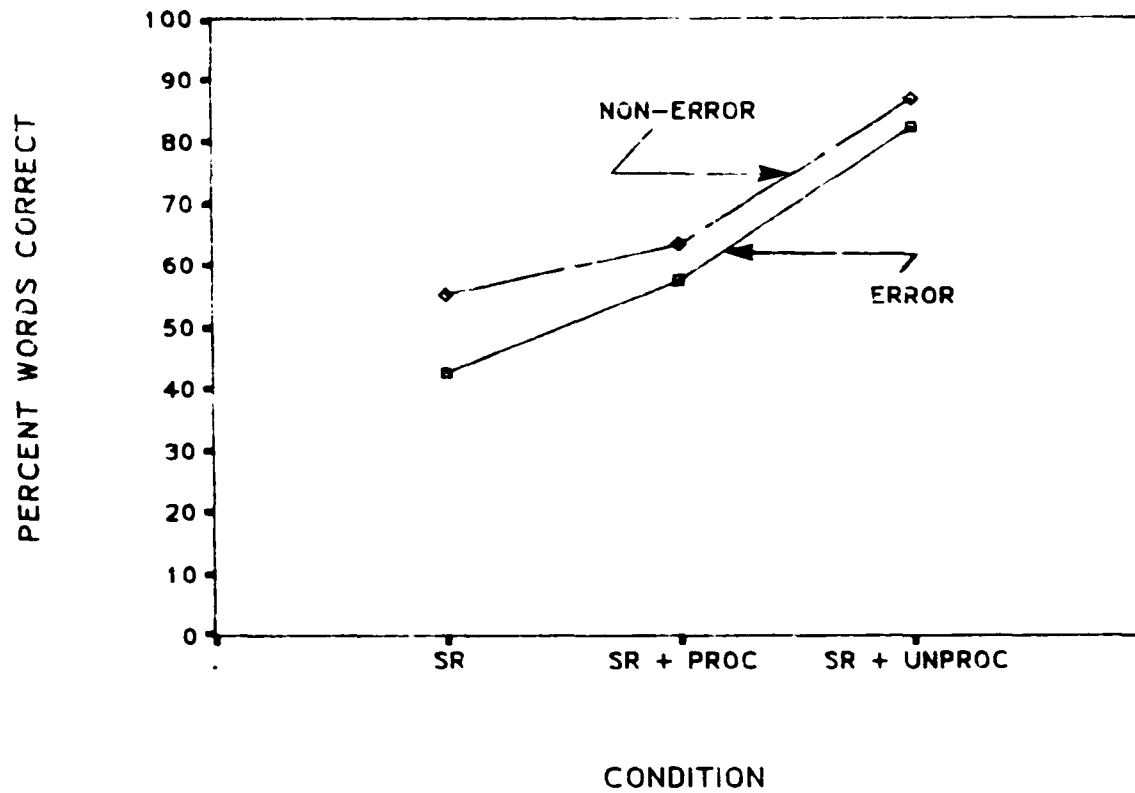


FIGURE 6.1: PERCENT WORDS CORRECTLY RECOGNIZED FOR EACH SENTENCE TYPE IN EACH OF THREE CONDITIONS

speechreading + unprocessed Fo condition and poorest in the speechreading alone condition, with the mean scores for the speechreading + processed Fo condition falling in-between. The main effect of sentence type was not significant ($F(1,119) = 1.61, p > .05$). The most important finding of this analysis was that the interaction between sentence type and condition was not significant ($F(1,119) = 0.07, p > .05$). Thus there was no evidence from this analysis that the pitch extraction errors had a significant effect in the speechreading + processed Fo condition.

2. t-test on differences between matched-pairs of supplemented speechreading scores for error and non-error sentences.

A problem with the previous analysis is that the randomized groups were not truly randomized. All subjects appeared in more than one condition/sentence type category and some subjects provided more than one observation within each of these categories. Thus subject differences may have been confounded with the sentence type x condition interaction. An additional problem is the inherent insensitivity of the randomized design in which inter-subject differences provide the error term for the ANOVA. In an attempt to deal with these two problems the following analysis was performed.

First it was necessary to locate pairs of subjects who had received a given error sentence in both the unprocessed and processed conditions and also a non-error sentence in both of these conditions. This combination is illustrated in Table 6.5. The difference between scores W and X provides a measure of the effect of processing in an error sentence but also includes the effect of the difference between subjects 1 and 2. The difference between scores Y and Z provides a measure of the effect of processing in a

Table 6.5
Illustrating the Method used for Arriving at an Estimate of the Effects of Pitch Extraction Errors from which Subject and Sentence Effects have been Cancelled

SENTENCE		SUBJECT	
		1	2
A - ERROR	PROCESSED	SCORE W	
	UNPROCESSED		SCORE X
B - NON-ERROR	PROCESSED	SCORE Y	
	UNPROCESSED		SCORE Z

Note: The difference scores was obtained by subtracting score X from score W and by subtracting score Z from score Y.

non-error sentence, but also includes the effect of the difference between subjects 1 and 2. When the (W - X) score and the (Y - Z) score are subtracted the effects of processing and subject (which, assuming no interactions, should be identical) are cancelled, thus leaving the effect of pitch extraction error alone. Eleven such combinations of scores were found, and are presented in Table 6.6. The difference scores are shown in Table 6.7 and illustrated in Figure 6.2. It will be seen that the mean effect of pitch extraction errors is 1.1 percentage points. This difference is not statistically significant ($t(10) = 0.07, p > .05$). This result therefore supports that of the previous analysis in that it provides no evidence that pitch extraction errors were a primary cause for the decrements in performance in the processed Fo condition.

Table 6.6
Blocks of Individual Subject (sub) Scores of Error and
Non-Error Sentences (sent) Under the Processed and
Unprocessed Fo Conditions (cond)

		BLOCK I			BLOCK II			BLOCK III		
TYPE		SENT	COND	SUB	SENT	COND	SUB	SENT	COND	SUB
				1 3			2 8			2 4
E	G	P		100	G	P	75	A	P	0
		UP	100			UP	100		UP	100
NE	B	P		0	B	P	0	D	P	100
		UP	100			UP	100		UP	100
		BLOCK IV			BLOCK V			BLOCK VI		
TYPE		SENT	COND	SUB	SENT	COND	SUB	SENT	COND	SUB
				2 4			3 8			3 9
E	J	P		75	H	P	50	J	P	100
		UP	100			UP	100		UP	88
NE	F	P		67	C	P	100	F	P	100
		UP	100			UP	75		UP	75
		BLOCK VII			BLOCK VIII			BLOCK IX		
TYPE		SENT	COND	SUB	SENT	COND	SUB	SENT	COND	SUB
				5 12			5 7			5 10
E	H	P		100	K	P	0	I	P	100
		UP	100			UP	40		UP	100
NE	A	P		100	D	P	67	C	P	100
		UP	100			UP	89		UP	100
		BLOCK X			BLOCK XI					
TYPE		SENT	COND	SUB	SENT	COND	SUB			
				7 10			8 9			
E	G	P		100	I	P	29			
		UP	100			UP	86			
NE	B	P		100	D	P	78			
		UP	100			UP	100			

Key to Sentence Types: E: Error Sentences; NE: Non-Error Sentences

Key to Conditions: P:Speechreading + Processed Fo
 UP:Speechreading + Unprocessed Fo

Key to Sentencees: A: SET 8, #4; B: SET 10, #10; C: SET 9, #10; D: SET 13, #8; E: SET 14, #6; F: SET 22, #1; G: SET 5, #9; H: SET 6, #8 I: SET 2, #3; J: SET 22, #4; K: SET 13, #5; L: SET 19, #4.

Table 6.7
Difference Between Processed and Unprocessed Scores for
Matched Pairs of Error and Non-Error Sentences

BLOCK											MEAN DIFF.
I	II	III	IV	V	VI	VII	VIII	IX	X	XI	
<u>ERROR SENTENCES</u>											
0	25	100	25	50	-22	0	40	0	0	57	25.0
<u>NON-ERROR SENTENCES</u>											
100	100	0	3	33	-25	0	22	0	0	22	26.1
<u>DIFFERENCE</u>											
-100	-75	100	-10	17	3	0	18	0	0	35	-1.1

STANDARD ERROR 15.8
 $t(10) = 0.07, p > .05$

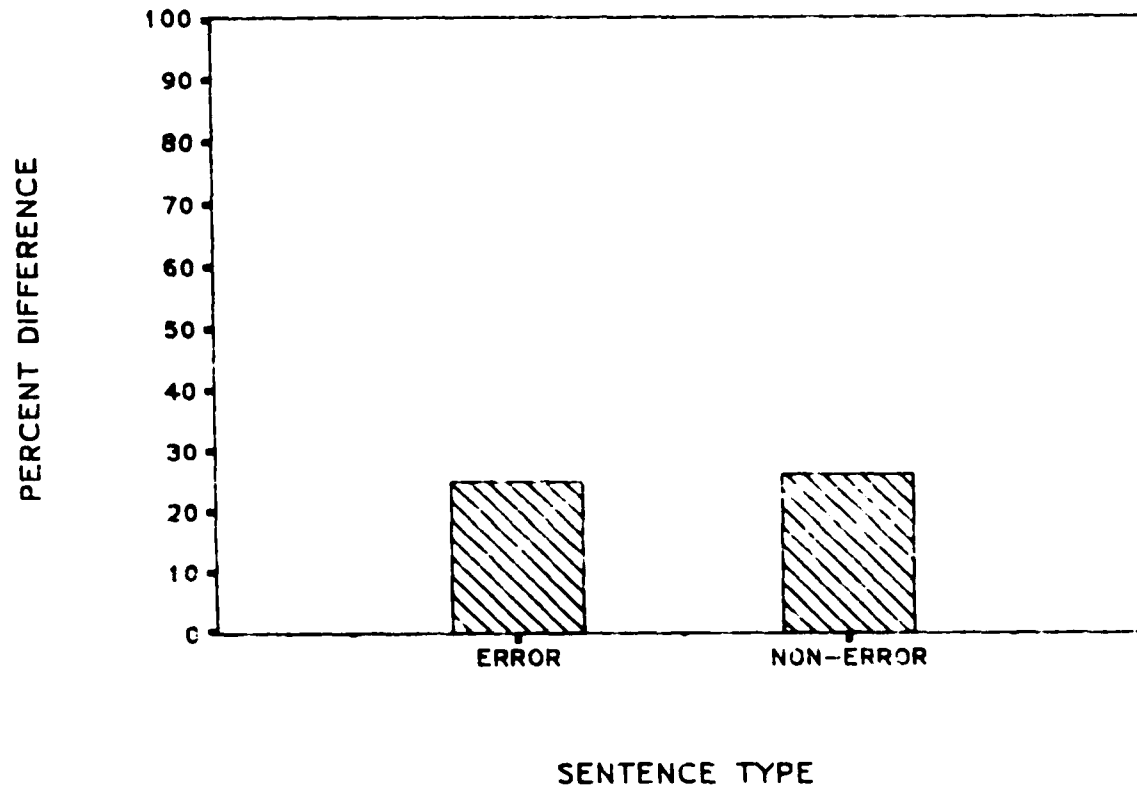


FIGURE 6.2: MEAN DIFFERENCE BETWEEN MATCHED-PAIRS OF UNPROCESSED & PROCESSED SCORES FOR TWO SENTENCE TYPES

Preliminary Discussion

Both the results of the analysis of variance and the t-test for differences between matched pairs of error/non-error sentences suggest that the presence of pitch extraction errors did not have an influence on performance.

CHAPTER 7
DISCUSSION

This investigation was designed and conducted on two premises: (1) that voice fundamental frequency (F_0) serves as a significant acoustic speechreading supplement; and, (2) that frequency quantization would seriously limit the effectiveness of F_0 as an acoustic speechreading supplement. The results of the main experiment support these premises. Performance in all auditorily supplemented speechreading conditions was better than performance by speechreading alone, with the unprocessed F_0 signal (i.e. EGG signal) providing the strongest supplement. Performance under all quantized conditions was poorer than under the unprocessed F_0 condition. In addition, performance decreased as the number of quantization levels diminished.

The finding that the unprocessed F_0 signal serves as a significant supplement to speechreading of sentence length stimuli is not a new one. Similar findings have been reported by Rosen, et al (1981), Grant, et al (1985), and Breeuwer and Plomp (1986). It is difficult to compare the results of the present investigation to those of Rosen, et al and Grant, et al since different methods of measuring performance were

used. Breeuwer and Plomp (1986) conducted the only investigation utilizing a measurement method similar to the one used in the present study. They found a mean increase of 34 percentage points for syllable recognition in Dutch sentences with the addition of a constant amplitude F_0 signal. This is very similar to the mean increase of 32 percentage points found in the present study for word recognition in English sentences. (See Table 7.1.)

TABLE 7.1
Comparison of Methods and Results of Studies which
Investigated the Perception of Sentence Length Stimuli
by Speechreading Alone and Speechreading Supplemented
with Acoustic Fo

STUDY	EVALUATION METHOD	CONDITION	
		SR ALONE	SR + Fo
Rosen, et al (1981)	Connected	14.0, 10.0,	27.1, 16.6,
	Discourse	6.9, 3.2,	10.7, 5.0,
	Tracking	12.9	31.7
	(words per minute for 5 subjects)*		
Grant, et al (1985)	Connected		
	Discourse Tracking (scores reported re: normal tracking rate)	37.9%	67.9%
Beeuwer & Plomp (1986)	Percent correct		
	syllable recognition in sentences **	16.7%	51.4%
Present Investi- gation	Percent correct word		
	recognition in sentences	32.2%	71.9%

* Data was also reported re: speechreading alone as increases of 55% to 110%.

** Scores are for second trial in each condition.

There is little basis for a comparison of the results obtained under the quantized conditions with those of earlier research. It has been shown that quantizing the F_0 range of a single speaker into 16 steps is just detectable by normal-hearing subjects (Rosenberg, et al., 1971). Rosenberg, et al's study did not, however, address the issue of the maximum amount of F_0 quantization that could be introduced without measurably affecting intelligibility under a supplemented speechreading condition. The results of the main experiment suggested initially that quantization at 12 levels (the least degraded condition) negatively affected intelligibility. However, before this interpretation could be accepted it was necessary to conduct a supplementary experiment examining the efficacy of the signal processing technique used in this investigation.

The results of the supplementary experiment demonstrated that the processing technique introduced serious degradation even before systematic frequency quantization. (The effects of processing will be discussed in detail later.) Thus, it would not have been possible to find a level of frequency quantization which did not result in a decrement in supplemented speechreading performance in the main experiment.

The next question of interest was whether

quantization at 12 levels gave results equivalent to the processed Fo signal. A post-hoc analysis using the percentage reduction of error to measure the effects of supplementation was conducted to examine this question. The results demonstrated that quantization at 12 levels did not introduce sufficient extra degradation, compared with the processing used here, to have a measureable effect on performance. In order to find out whether quantization at higher levels is superior to quantization at 12 levels it will be necessary to improve the processing technique. What can be concluded from these data, however, is that the utility of Fo contours as an acoustic speechreading supplement diminishes in a monotonic fashion as the number of quantization levels is decreased beyond 12. The smallest increase in performance was found with the fundamental frequency signal quantized at one level (i.e. a constant frequency signal).

The Fo signal quantized at one level provides information on the onsets and duration of voicing but not on the patterns of change in Fo. The finding that a constant frequency signal supplements speechreading of sentence length stimuli has been previously reported (Grant, et al., 1985; Breeuwer and Plomp, 1986). A major difference, however, between Breeuwer and Plomp's and the present results and those of Grant, et al lies in the relative effectiveness of the unprocessed Fo

signal and the constant frequency signal. The results for these conditions from the present and from the two previous investigations are presented in Table 7.2. It will be seen that in both the present study and Breeuwer and Plomp's, greater increases in supplemented speechreading performance were found with the fundamental frequency signal than with constant frequency signal. However, the results of Grant, et al. study suggest that the effects of supplementing speechreading with a constant frequency signal is similar in magnitude to those of supplementing speechreading with a fundamental frequency signal.

The apparent equivalence of the F_0 and constant frequency signals shown in Grant, et al.'s data may be attributed to their experimental design. The data for all conditions was the average of performance obtained during 5-hours of tracking/testing. While the investigators measured speechreading alone performance after testing with the F_0 signal, they did not do so after testing with the constant frequency signal. That is, it is unclear as to whether the score reported for the constant frequency condition reflects changes in performance due to supplementation or changes in performance due to improved speechreading ability.

Thus, there is very little evidence to suggest that a constant frequency signal and a F_0 signal are

equivalent acoustic speechreading supplements. Rather, the evidence supports a conclusion that a F_0 signal is a much more effective speechreading supplement than a constant frequency signal for sentence length stimuli.

Table 7.2
Methods and Results of Studies which Investigated the Perception of Sentence Length Stimuli by Speechreading Supplemented with an Acoustic Fo Signal and by Speechreading Supplemented with a Constant Frequency Signal

STUDY	EVALUATION METHOD	CONDITION	
		SR+ CF	SR+ Fo
Grant, et al (1985)	Connected Discourse Tracking (scores reported re: normal tracking rate)*	62.8%	67.9%
Beeuwer & Plomp (1986)	Percent correct syllable recognition in sentences**	29.5%	51.4%
Present Investigation	Percent correct word recognition	40.6%	71.9%

Key to conditions:

SR + CF: Speechreading + Constant Frequency
 SR + Fo: Speechreading + Fo Frequency

* Data reported for SR + CONSTANT FREQUENCY condition is only for one subject while for SR + Fo condition is the mean of five subjects.

** Scores are for second trial in each condition.

One of the most interesting findings of the main experiment concerned the role of sentence length. In the speechreading + unprocessed Fo condition sentence length had no effect on performance. In the degraded Fo conditions, however, length was important, and became more so as the degree of degradation was increased. In the speechreading alone condition sentence length had the largest effect on performance. The effect of sentence length on speechreading alone performance has been noted previously (Taaffe and Wong, 1957). This effect may reflect the difficulties involved in segmenting a long continuous stream of visual speech information into word and phrase sub-units. The fact that length is unimportant when speechreading is supplemented by unprocessed Fo suggests that Fo aids in the process of segmentation. It will be recalled that the most important role of intonation is believed by Lieberman (1967) to be the provision of acoustic cues that can be used in the parsing of utterances for higher order syntactic processing. The present data support that view.

The findings related to sentence length indicate that this factor needs to be considered in future studies. Had only short (3-6 words) sentences been used in the present study, different conclusions may have been drawn as to the magnitude of the effects of quantization. That is, quantization would not have

appeared to have had as serious a negative effect on the utility of F_0 . It should be noted that in previous investigations the issue of utterance length has been ignored. Researchers have used nonsense syllable stimuli (Rosen, et al., 1979; Breeuwer and Plomp, 1986); sentences and/or phrases of unspecified length (Risberg, 1974; Risberg and Lubker, 1978; Rosen, et al., 1981; Grant, et al., 1985); or, only short sentences (Breeuwer and Plomp, 1986).

One important issue that remains to be addressed is the cause of the decrements in performance as a result of processing the F_0 contours. The decrements in performance may have been due to any one or some combination of the following factors: (1) loss of amplitude information; (2) loss of harmonic information; (3) sampling time effects; and/or, (4) pitch extraction errors.

Information related to amplitude variations resulting from differences in the area of vocal fold contact was recorded via the EGG. The quantized signals were generated with constant amplitude and thus did not contain any information about amplitude variations. The importance of this information during supplemented speechreading is not known. However, the possibility exists that by using a constant amplitude signal in the regeneration of the F_0 contours critical

information was removed.

Since the processed Fo contours were replicated using sinusoidal signals they did not have a harmonic structure. Such a structure was present in the natural voice Fo contours recorded via the EGG. This resulted in a qualitative difference between the two signals. It is possible that this qualitative difference reduced the effectiveness of the processed Fo contours.

The sampling time of the processing technique used in this investigation was rather long (40 msec). The results of using such a long sampling time will be seen in a comparison of the spectrograms shown in Figure 7.1. The processed signal has a stepped contour rather than a smooth one. Thus, as a result of sampling time constraints the contour was quantized already in the frequency domain. Another result of constraints imposed by the sampling time was that small, abrupt frequency changes in the EGG waveform (i.e. the microstructure) were not reproduced in the processed signal. An example of the failure to produce the microstructure will be seen by comparing point A in the EGG spectrogram and point B in the processed Fo spectrogram. The relative importance of the differences in the EGG signal and the processed Fo signal which resulted from time sampling constraints on supplemented speechreading performance are as yet

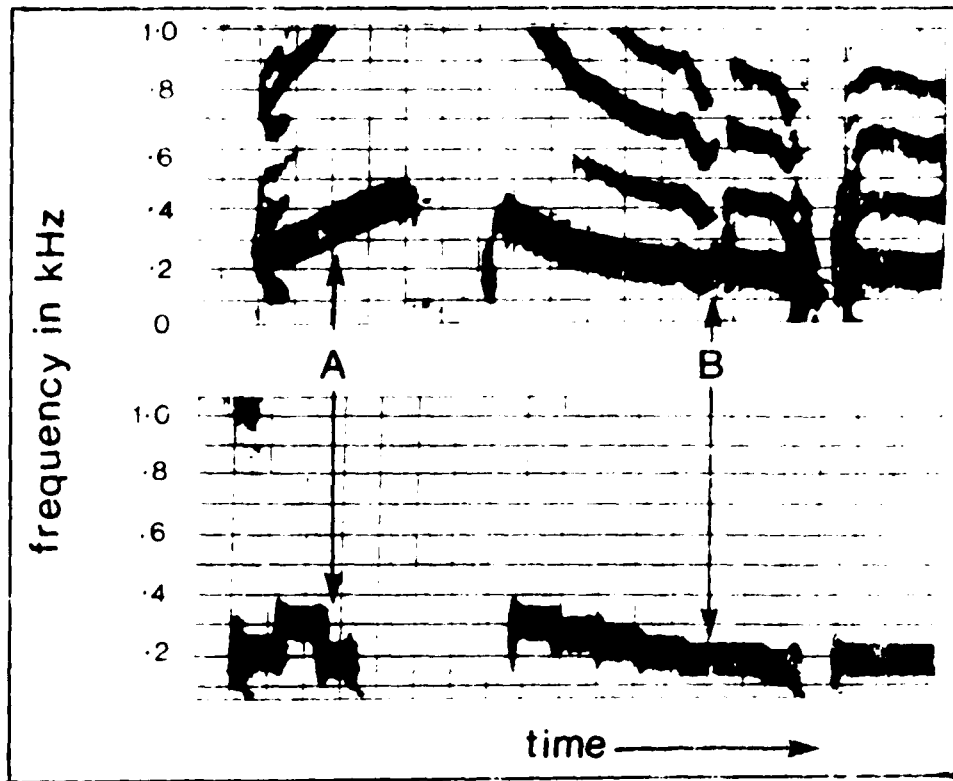


FIGURE 7.1: SPECTROGRAMS ILLUSTRATING THE EFFECTS OF PITCH EXTRACTION ERROR (A) AND THE EFFECTS OF TIME SAMPLING CONSTRAINTS (B).

unknown.

Errors in pitch extraction were introduced when the EGG signal was passed through Visipitch. An example of a pitch extraction error will be seen by comparing point C on the EGG spectrogram and point D on the processed Fo spectrogram illustrated in Figure 7.1. An estimate of the percentage of pitch extraction errors was made by comparing spectrograms of the EGG signal and the processed signal from a random 5% sample of all sentences. A pitch extraction error was believed to have occurred when (1) the direction of the pitch contour in the processed Fo spectrogram did not follow the direction of the pitch contour in the EGG spectrogram; or, (2) the spectrogram of the processed Fo contour indicated energy at a time when the spectrogram of the EGG contour showed no energy. The data for this examination are included in Appendix G. There did not appear to be any systematic pattern as to the time of occurrence of pitch extraction errors (i.e. they did not appear to occur more at the beginning than at the end of an utterance (or vice versa)). However, all of the pitch extraction errors related to a difference in the direction of the pitch contour were in a downward direction (i.e. a falling in frequency). The examination revealed that pitch extraction errors occurred in approximately 67% of the sentences (14% of the words). These percentages appear to be quite high.

Although the results of this examination suggest the possibility that pitch extraction errors may have been a major cause of the decrements in performance as a result of processing, the post-hoc analyses on the effects of pitch extraction errors on performance failed to support this suggestion. Caution must, however, be taken in interpreting the results of these post-hoc analyses.

Several aspects of design could not be accounted for in the post-hoc analyses on the effects of pitch extraction errors. The analysis of variance used was based on a randomized experimental design. Such a design normally assumes that subjects will be randomly assigned to only one condition. In order to fit the available data into such an experimental design each score for a sentence was treated as if it had been obtained from a totally randomly selected subject. That is, each point was treated as if it were independent of the other points. In reality, however, these data points were not independent. Twelve individuals each provided twelve data points and these data points were not evenly distributed amongst the conditions. The number of data points provided by each subject in each condition included in the analysis is shown in Table 7.3. The data points provided by any one individual were intrinsically related to that individual's speechreading ability. It is well-known

Table 7.3
Number of Data Points Contributed by each of the 12
Subjects Tested in the Supplementary Experiment for
Each of the Conditions Evaluated in the Post-Hoc Data
Analysis on the Effects of Pitch Extraction Errors on
Performance

SUBJECT	NON-ERROR SENTENCES			ERROR SENTENCES		
	SR	SR+PROC	SR+Fo	SR	SR+PROC	SR+Fo
1	2	2	2	4	0	2
2	2	1	2	1	3	2
3	1	3	1	2	1	2
4	2	0	3	2	2	2
5	3	1	2	2	2	1
6	1	2	2	2	1	2
7	2	2	1	2	2	2
8	2	1	1	0	2	1
9	1	3	1	1	3	1
10	1	2	3	0	1	3
11	2	1	1	2	1	2
12	1	2	2	2	1	1

Key to conditions: SR: Speechreading alone
SR+PROC: Speechreading + Processed Fo
SR+UNPROC: Speechreading + Unprocessed Fo

that inter-subject speechreading ability is very high. Since many of the individuals contributed more data points to one condition than to another, this may have skewed results by reducing the within-group variance and also by confounding subject effects with error and condition effects.

The data for the sentences used in the analysis of variance also violated the assumption of normal distribution. That is, there was a proportionally large number of scores equal to 100% or equal to 0% in the data pool. This was because each score was based on only one sentence and there was a tendency throughout this investigation for subjects to perceive either all of a sentence or none of a sentence. Normally, when percentage scores tend to be either very high or very low an arcsine transformation can be used to counteract the reduced error variance for scores close to 100% or 0%. However, an arcsine transformation does not help with scores that equal 100% or 0%. Thus, "within group" variance is artificially low because of the large number of 0% and 100% scores, and the chance of finding significance is artificially increased.

In analyzing the data based on a t-test on differences between matched-pairs of scores an assumption was made that only the presence or absence

of a pitch extraction error contributed to the difference between the unprocessed and processed scores. That is, it was assumed that there were no interactions between sentences and subjects, sentences and condition, or condition and subjects and the difference score. In addition there were so few blocks of scores (11) that any difference between them would have to be very large in order to show significance. Note that the standard error of the difference in Table 6.6 was 15.8 percentage points. Thus to be significant at the 5% level a difference of roughly 30 percentage points would be required. It would be wrong to conclude, however, that pitch extraction errors were not contributing to the decrements in performance found in the processed F_0 condition.

The question of the effects of pitch extraction errors on the utility of F_0 as an acoustic speechreading supplement is an important one, but it could not be addressed adequately from the available data.

The results of the present investigation show that both quantization and processing (as done in this study) fundamental frequency seriously limits its utility as a speechreading supplement. Certain cautions, however, must be taken in interpreting the results of this investigation. There is for example, the generalizability of the results to populations

other than the one from which subjects were selected for this study. This investigation used randomly selected normal hearing and normal sighted adults (aged 18-50 years) as subjects, thus the results should be generalizeable to the normal adult population. There is no a priori reason, however, to believe that groups such as the elderly or children have the same ability to integrate auditorily presented F_0 frequency information and visual speech information as do adults. Moreover, the reduced frequency discrimination exhibited by many persons with sensorineural hearing loss precludes generalizing results obtained with normal hearing subjects to those that might be obtained if fundamental frequency information was made available to the hearing-impaired through amplification. In addition, since the results obtained for the effect of sentence length suggest that the benefits of supplementing speechreading with F_0 may be due in part to an enhancement in the ability to parse utterances for higher level processing, the typically poorer linguistic knowledge of the pre-lingually deaf may reduce their ability to use the cues provided by F_0 .

It should also be noted that only one speaker was used to present stimuli. The effectiveness of unprocessed, processed and quantized F_0 as acoustic speechreading supplements might be different for other speakers. That is some, as yet unknown, relationships

may exist between a speaker's precision of articulation and the range of that speaker's F_0 , that could influence the supplementary nature of F_0 . For example, it may be the case that auditorily presented F_0 does not assist in the speechreading of a speaker with very precise articulation, while it does in the speechreading of a speaker who has very poor articulation. It may also be the case that providing access to the F_0 produced by a speaker who uses a limited range of F_0 results in little difference in the ability to speechread that person regardless of precision of articulation. The issue of speaker differences remains to be investigated.

The methods of quantization and regeneration used in this study were not the only possible ways to accomplish the signal manipulations desired. Had other methods been used results may have been different. For example, in this experiment F_0 was quantized on a logarithmic scale. It is possible to quantize on a linear or some non-linear scale other than logarithmic. The choice of scale may have influenced results. In addition, although we reduced transient spectral spread at the frequency transitions, we did not try to smooth the voltage output of the processing. It is possible that such a smoothing would have affected results by making the contours sound more natural. Thus, the results obtained in the present study may be specific

to the quantization and regeneration techniques used.

The possible need for experience in listening to the quantized Fo signals was only minimally acknowledged (by the use of "practice" presentations of sentence sets) in the current study. Many of the subjects commented on the non-speechlike quality of the quantized Fo signals. The perception of the quantized signals as non-speechlike may have influenced performance. It is possible that with increased experience listening to the quantized Fo signals the subjects would have learned to perceive them as speech and that performance would have improved.

Even with these cautions in mind, the following conclusions may be drawn from the results of the present study:

(1) Auditorily presented fundamental frequency serves as a significant supplement to speechreading.

(2) The pitch extraction and regeneration (i.e. processing) techniques used in this investigation seriously reduced the effectiveness of fundamental frequency as an acoustic speechreading supplement.

(3) The effectiveness of processed fundamental frequency is further reduced by quantization at 8 levels.

(4) Although processing and quantization reduces the effectiveness of auditorily presented fundamental

frequency cues, even performance with the most extreme of the quantized signals (i.e 1 level) is superior to that of speechreading alone.

(5) Since the F_0 signal quantized at one level provides only information about binary voicing, the above finding demonstrates that acoustic voicing information can serve to supplement speechreading.

(6) Sentence length is an important factor in evaluating the effects of supplementing speechreading with fundamental frequency cues.

(7) Fundamental frequency is a more efficient acoustic speechreading supplement for longer sentences than for shorter ones.

CHAPTER 8

CLINICAL IMPLICATIONS, RESEARCH IMPLICATIONS,
AND FUTURE RESEARCH

The primary impetus for the present investigation was the limited benefits obtained in previous research with existing tactile spatial displays of voice fundamental frequency as supplements to speechreading (Boothroyd and Hnath, 1986; Grant, et al., 1986). It was believed that the frequency quantization which occurred in the 8-channel device designed by Boothroyd (1985) and the 10-channel device designed by Grant (1980) was limiting their effectiveness. The results of the present study support this belief. That is, the results demonstrate that at least 12 levels of quantization is needed if quantization is to have no effect on the utility of voice F_0 as a supplement to speechreading. For the speaker used in this study, 12 levels was equivalent to 6 levels of frequency quantization per octave. Thus, in order to encode the frequency range of both male and female voices (approximately two and one-half octaves) at least 15 levels would probably be required. Based on the present findings, it appears that the limited performance of existing spatial tactile displays of F_0 is due not only to possible limitations in the tactile

system, but also to the deleterious effects of frequency quantization.

The present findings also suggest that the relatively poor frequency resolution of the skin might limit the effectiveness of a temporal tactile display of F_0 as a speechreading supplement. It will be recalled that Rothenberg, et al (1977) found approximately seven discriminable steps in frequency for stimuli applied to the arm and ten for stimuli applied to the finger. If, as the current results suggest, frequency discrimination is an important factor in the effectiveness of F_0 information as a supplement to speechreading sentence length, then limitations in the frequency resolution of the tactile system would argue for spatial rather than temporal encoding.

Although voicing information is not as significant an acoustic supplement to speechreading as is F_0 information, the finding that speechreading performance improves when access to voicing information is provided is also important. This finding suggests that any sensory aid which presents voicing information may serve, albeit minimally, to provide assistance during speechreading.

The results of this study suggest that it is

important to separate the effects of processing and coding from modality effects in future sensory aid research. Although limitations in the processing capacity of the tactile system may in part account for minimal supplementation that tactually presented Fo provides the speechreader, the results of the present investigation clearly demonstrate that processing and coding severely limits the effectiveness of auditorily presented Fo. It is highly probable that the processing and coding of Fo for presentation to the tactile system also limits its effectiveness. Researchers may choose to accept such limitations because they cannot be overcome easily or because they believe that it might be possible to reduce the negative effects through extensive training. On the other hand, they may choose to invest a considerable amount of time and effort to develop processing and coding techniques which result in minimal degradation, while providing salient sensory stimuli. Exploring the effects of processing and coding, separately from modality effects, will assist researchers in making such decisions.

An important finding of this study was that small in all but the SR + Unprocessed Fo condition performance was affected by sentence length. Specifically, longer sentences were harder to speechread than shorter sentences. In these test

materials, however, length and syntactic complexity covaried. That is, the longer sentences tended to be more complex. It is not clear, therefore, whether it was length per se that affected performance, or the increased complexity of the longer sentences. One potentially useful avenue for further research would be an investigation of the independent effects of length and complexity in suitably designed test material.

The failure to find a significant effect of sentence length in the speechreading + unprocessed Fo condition also has important implications for the design of future research into the effectiveness of sensory displays of Fo during speechreading. The ultimate test of a sensory speechreading aid is its effectiveness during real communication situations. In real communication situations, sentences are of different lengths. Thus, evaluation of the effectiveness of a sensory display of Fo needs to be considered at for sentences of different lengths, and this has not previously been done.

The results obtained in the present investigation suggest that further research should be conducted to evaluate the effects of other forms of degradation that are introduced during the processing and coding of Fo, as well as other speech parameters, for alternative sensory transmission. In the case of Fo a first step

would be to explore the effects of pitch extraction errors and other forms of degradation introduced through the processing technique used in the current study. Such experiments would not only provide an understanding of the effects of these forms of degradation, but would also provide data on the minimum degree of degradation that can be tolerated if there is to be no effect on the utility of sensory displays of F_0 .

The results of the present investigation also suggest that an evaluation of the relationship between frequency discrimination performance of persons with profound hearing losses and auditorily supplemented speechreading performance is warranted. There are many profoundly deaf persons whose audiograms show residual hearing below 500 Hz, and who thus may be predicted to have access to voice F_0 frequency from amplification during speechreading. However, these individuals often have reduced frequency discrimination capacity (see Chapter 2, Section 2 for review). The present finding, that auditorily supplemented speechreading performance diminished with a decrease in the number of discriminable steps in frequency (i.e. quantization levels) suggests that auditory frequency discrimination capacity might be an important determinant of the magnitude of improvements which can be expected. If frequency discrimination capacity is an important

factor in the utility of F₀ as an acoustic speechreading supplement, then this may account for why many of these individuals gain limited or no benefit from the use of amplification during speechreading. Such an evaluation should examine performance in two groups of subjects: pre-lingually and post-lingually deaf. The inclusion of both pre- and post-lingually deaf subjects could provide data on the relationship between linguistic competence and the ability to use of F₀ information as a speechreading supplement. This knowledge would assist in the determination of the appropriate information to be presented through sensory aids to these two populations, as well as having implications for auditory training for the pre-lingually deaf.

APPENDIX A

TOPIC RELATED SENTENCE SETS

Set 1

- | | |
|---------------------|---|
| food | 1. Have you eaten yet? |
| family | 2. How many of your brothers still live at home? |
| work | 3. How many years of school did it take to become a nurse? |
| clothes | 4. Where can I get my suit cleaned? |
| animals | 5. Cats are easy to take care of because you don't have to walk them. |
| homes | 6. She just moved into a three room apartment. |
| sports
hobbies | 7. I like to play tennis. |
| weather | 8. We couldn't fly home yesterday because of the big snowstorm. |
| health | 9. Remember to get plenty of rest and drink lots of fluids. |
| seasons
holidays | 10. Carve the turkey. |
| money | 11. Make sure you deposit that check. |
| music | 12. Please make sure that you practice a lot before your next piano lesson. |

Set 2

- | | |
|--------|---|
| food | 1. Do you want to have a barbecue this evening? |
| family | 2. When was the last time that you went to |

- visit your parents?
- work 3. When will you be taking your vacation?
- clothes 4. Remember to take enough hangers with you when you go to do your laundry.
- animals 5. You can see deer near my country house.
- homes 6. We're looking for an apartment.
- sports
hobbies 7. The football field is right next to the baseball field.
- weather 8. Can you remember the last time we had so much snow?
- health 9. See your doctor.
- seasons
holidays 10. Put these lights on the tree.
- money 11. Don't use your credit card if you can't pay the bill on time.
- music 12. I like that song.

Set 3

- food 1. Do you want fried chicken or do you want pizza for dinner.
- family 2. How long have your parents been married?
- work 3. Take the job only if you feel that the work will be more challenging.
- clothes 4. Buy that bathing suit because it fits you.
- animals 5. That dog likes to run.
- homes 6. We're going to paint the guest bedroom next week.
- sports
hobbies 7. How many miles a day do you run before a marathon.
- weather 8. Did it rain?
- health 9. Remember to take all your vitamins.

seasons

holidays 10. You buy the food and she will buy the drinks for the barbecue.

money 11. That place is expensive.

music 12. The rock concerts on the pier have been very successful.

Set 4

food 1. Do you want gravy on your potatoes?

family 2. You have to invite all of your aunts and uncles to the wedding reception.

work 3. Make sure you get to work on time.

clothes 4. Throw away your old shoes.

animals 5. My cat was chasing the bird all around the yard.

homes 6. Did it take you a long time to find an apartment?

sports

hobbies 7. Do you jog?

weather 8. Did you bring an umbrella today?

health 9. Don't drink alcohol while you are pregnant because it's harmful to your baby.

seasons

holidays 10. Summer is finally here.

money 11. I lost my checkbook so I closed the account.

music 12. He did not learn how to play the piano until very recently.

Set 5

food 1. Please get the groceries out of the car and put them away for me.

family 2. Make sure you call your brother this week.

- work 3. Finish that report this afternoon.
- clothes 4. Take the dress back and buy something that you like.
- animals 5. Did your landlord give you any trouble about having a dog?
- homes 6. Where's the bathroom?
- sports
hobbies 7. Do you like to go hiking?
- weather 8. Do you know if it rains a lot this time of year?
- health 9. She has a fever.
- seasons
holidays 10. I like going to the mountains in the fall.
- money 11. Many of my friends had to take out school loans this year.
- music 12. Their last song was a big hit.

Set 6

- food 1. Make sure that you don't overcook the shrimp.
- family 2. Visit your grandmother on Sunday.
- work 3. If you want to quit, please give two weeks notice.
- clothes 4. The store is having a sale on all their bathing suits.
- animals 5. Where's your dog?
- homes 6. How many bedrooms do you have?
- sports
hobbies 7. Did you ever think of trying to sell some of your better paintings.
- weather 8. Put on your raincoat.
- health 9. All my friends seem to be catching summer

colds.

- seasons
holidays 10. It seems like we haven't had a decent spring in several years.
- money 11. She wants to start buying some stock.
- music 12. Did you know that they both get season tickets for the opera every year?

Set 7

- food 1. Please turn on the coffee.
- family 2. You have to let your children make their own mistakes.
- work 3. If you work for an airline you get great travel benefits.
- clothes 4. My zipper broke.
- animals 5. How many cats do you have?
- homes 6. Would you ever consider building an extra room or two onto the house?
- sports
hobbies 7. Throw me the ball.
- weather 8. Be careful driving on the bridge when it's windy.
- health 9. Children get sick more often when they start to go to school.
- seasons
holidays 10. I always eat alot on Thanksgiving.
- money 11. What type of investment plan did they think would be the best for you?
- music 12. Have you seen the new musical on Broadway?

Set 8

- food 1. Make sure you wash the fruit before you eat it.

- family 2. My brother and his two children are coming to visit me.
- work 3. Here's my office.
- clothes 4. I need to buy a suit.
- animals 5. Did you see any interesting birds when you went bird watching last week?
- homes 6. Clean the kitchen first.
- sports
hobbies 7. Don't try to run unless you have good shoes.
- weather 8. Always take your sunglasses with you when you drive on sunny days.
- health 9. You have to start losing some weight.
- seasons
holidays 10. Do you want to take your vacation in the winter or summer this year?
- money 11. Do you always balance your checkbook every month?
- music 12. Do you play the guitar?

Set 9

- food 1. The desserts at that restaurant are very good but very expensive.
- family 2. She's my daughter.
- work 3. He doesn't like his new boss.
- clothes 4. I bought three pairs of shoes and a pair of boots last week.
- animals 5. Let the dog out.
- homes 6. Move the couch to the wall facing the window.
- sports
hobbies 7. If you want to take good pictures, take a few photography classes.

- weather 8. Wear a heavy sweater since it's cold.
- health 9. Do you go to a gym to exercise, or do you exercise at home?
- seasons
holidays 10. Have you finished all your Christmas shopping yet?
- money 11. Can I pay by check?
- music 12. Do you think he'll play the piano at the party?

Set 10

- food 1. Cake is sweet.
- family 2. My sister has a new boyfriend.
- work 3. He got a promotion at work and now he doesn't want to leave.
- clothes 4. Is that coat new?
- animals 5. Please don't go that close to the lion's cage.
- homes 6. Remember to lock all the doors and windows before you go away.
- sports
holidays 7. Take your baseball glove to the game.
- weather 8. All the roads into the city were closed yesterday due to the heavy rains.
- health 9. I know a doctor who makes house calls.
- seasons
holidays 10. Where are you spending Thanksgiving?
- money 11. How much interest are they charging you on your loan?
- music 12. Don't stand so close to the microphone when you are singing.

Set 11

- food 1. She baked a big apple pie.
- family 2. She shares a two bedroom apartment in the city with her two sisters.
- work 3. How's your new job?
- clothes 4. How much did they charge you for the alterations?
- animals 5. Take the dog to the vet every year for it's rabies shot.
- homes 6. Clean the guest bedroom before next weekend.
- sports hobbies 7. Professional football players usually have to train in the summer before the season begins.
- weather 8. When the humidity is high, it's uncomfortable outside.
- health 9. What medicine are you taking?
- seasons holidays 10. Do you want to go Christmas shopping with me today?
- money 11. You really should be able to save more of your salary.
- music 12. Play that song.

Set 12

- food 1. We had roast beef, baked potatoes with sour cream and broccoli for dinner.
- family 2. How's your brother doing?
- work 3. How long have you been working for this company?
- clothes 4. Did you know that she made her own dress for the wedding?
- animals 5. Don't let the dog off his leash.
- homes 6. We decided to paint the living room light blue and the dining room yellow.

- sports
hobbies 7. He plays softball with his friends on Sundays.
- weather 8. The sun is finally shining.
- health 9. Did he get sick from eating some bad food yesterday?
- seasons
holidays 10. Be careful if you use fireworks on the Fourth of July.
- money 11. Take your change.
- music 12. Put your albums where they belong.

Set 13

- food 1. When did you learn how to bake bread?
- family 2. When is your aunt getting married?
- work 3. Did your boss give you a raise?
- clothes 4. Where's my jacket?
- animals 5. We saw alot of deer while driving last night.
- homes 6. They bought a new house.
- sports
hobbies 7. She was on the swimming team and the track team while in college.
- weather 8. I heard that there will be rain this Sunday.
- health 9. Get the flu shot.
- seasons
holidays 10. Make sure you remember to buy your boyfriend a Valentine's Day gift.
- money 11. You have to keep one hundred dollars in this bank account for six months.
- music 12. Don't forget to buy the tickets for the concert next week.

Set 14

- food 1. What do you want for lunch?
- family 2. Did you remember to call your grandmother?
- work 3. Where's your office?
- clothes 4. Remember to use bleach when you wash that white underwear.
- animals 5. Goldfish are easy to keep.
- homes 6. We are trying, but it's difficult to find an apartment we can afford.
- sports
hobbies 7. I really like going to major league baseball games.
- weather 8. Has it started raining?
- health 9. Give the insurance forms to the dentist when you go next week.
- seasons
holidays 10. If you want to fly to Florida during spring break, make your reservations early.
- money 11. Have twenty dollars taken from your paycheck for savings each week.
- music 12. The free concerts in the park are wonderful.

Set 15

- food 1. Do you cook big dinners every night?
- family 2. Where's your sister?
- work 3. When you want to quit give them two week's notice.
- clothes 4. Iron all of your shirts.
- animals 5. It was really very easy to teach the parrot how to say hello.

- homes 6. We are really excited about looking for a house.
- sports
hobbies 7. Do you like camping?
- weather 8. Does it look as if the sun is trying to come out?
- health 9. Remember to take your medicine every four hours and rest in bed all day.
- seasons
holidays 10. Don't wait until the last minute to do your Christmas shopping.
- money 11. I always pay my credit card bills promptly.
- music 12. I really like that new album.

Set 16

- food 1. Are you hungry?
- family 2. Try to visit your mother at least once a week.
- work 3. Finish that assignment by Friday.
- clothes 4. Put all your winter suits in storage and take out your spring suits.
- animals 5. The cats keep each other company during the day.
- homes 6. Where do you live?
- sports
hobbies 7. Did you get to see any of the gymnastics competition last week?
- weather 8. Will it rain this weekend or is it supposed to be clear and sunny?
- health 9. You have to exercise as well as diet to lose weight.
- seasons
holidays 10. There is always a parade for Memorial Day.

- money 11. I think I have some change.
 music 12. She's the one who gives singing lessons.

Set 17

- food 1. Defrost the freezer first, then go do your grocery shopping.
 family 2. Don't forget your father's birthday.
 work 3. Remember to keep all your receipts when you go on a business trip.
 clothes 4. Buy yourself a new coat for the cold months.
 animals 5. Where's the new puppy?
 homes 6. Will you be putting a swimming pool in the backyard this year?
 sports hobbies 7. How long did it take you to learn how to crochet these beautiful blankets?
 weather 8. Will you go to the beach even if it's still cloudy?
 health 9. My allergies are really acting up right now.
 seasons holidays 10. I like New Year's Eve parties.
 money 11. She's taken out a lot of loans.
 music 12. It's my guitar.

Set 18

- food 1. Pass the bread and butter.
 family 2. Call your grandmother and ask her if she needs anything from the store.
 work 3. You should have your resume printed instead of xeroxed.
 clothes 4. That's my new dress.

- animals 5. Did you see many lions and tigers when you went on safari?
- homes 6. Are you going to hire someone to paint the apartment before you move in?
- sports
hobbies 7. Do you want to go to the baseball game with us?
- weather 8. Close the windows in case it starts raining.
- health 9. Smoking is bad for your lungs.
- seasons
holidays 10. We're going to the beach on Sunday.
- money 11. Here's a quarter.
- music 12. Do you know of a good band for the party?

Set 19

- food 1. Take the steaks out of the freezer and put them on the counter.
- family 2. Remember to take your sister to the airport tomorrow.
- work 3. I quit my job.
- clothes 4. I have a sweater that would look great with this plaid skirt.
- animals 5. Did you go to see the bird exhibit when you went to the zoo?
- homes 6. Where did you buy all that new furniture for the house?
- sports
hobbies 7. Put all the golf clubs in the cart.
- weather 8. Don't be afraid of the thunder.
- health 9. I really don't like to get injections.
- seasons
holidays 10. The flowers bloomed.

money 11. Do you think we should buy him a savings bond?

music 12. Have you heard her sing?

Set 20

food 1. Let the grill get really hot before you barbecue.

family 2. My mother is working.

work 3. My commute each morning is just a little bit over an hour.

clothes 4. That suit is very expensive but I think the style is perfect for you.

animals 5. Is it very difficult to take care of all those fish?

homes 6. Clean up the bedroom before you watch television.

sports hobbies 7. Throw him a few good pitches.

weather 8. Take a sweater in case it's chilly.

health 9. I am sick.

seasons holidays 10. Are you going to go apple picking in the Fall?

money 11. How much is the bill?

music 12. Do you think you would enjoy going to see an opera with me?

Set 21

food 1. That pizza smells great.

family 2. It's hard to believe that she has four daughters and three sons.

work 3. It's difficult for me to leave the office on time during our busy season.

- clothes 4. It's okay to dress casually for the party this Saturday night.
- animals 5. Give the ducks some bread from your sandwich.
- homes 6. Hide the key under the mat.
- sports
hobbies 7. Don't run for a couple of days.
- weather 8. Take an umbrella.
- health 9. How often do you go for a complete physical examination?
- seasons
holidays 10. Can you build a snowman?
- money 11. How long did it take you to save enough money for a car?
- music 12. Would you like to go to a rock concert?

Set 22

- food 1. I was so hungry this morning that I ate a huge breakfast.
- family 2. Every spring we have a big picnic with all my aunts, uncles, and cousins.
- work 3. The job is great except for the business trips every month.
- clothes 4. Did you buy those shoes in that store?
- animals 5. Take the dog for a walk.
- homes 6. Polish the furniture in the living room.
- sports
hobbies 7. Finish the race.
- weather 8. The days have been cool for this time of year.
- health 9. She has the chicken pox.
- seasons
holidays 10. Do you think you'll go to the mountains for your vacation this year?

money 11. Is there a minimum balance for a checking account?

music 12. Turn the radio off.

Set 23

food 1. Each family will bring a special dish to the pot luck dinner on Saturday.

family 2. My grandfather's house is only a few blocks away from here.

work 3. Do you have to take many business trips?

clothes 4. Can I borrow your red sweater?

animals 5. You should really have your cat declawed.

homes 6. Lock the door.

sports hobbies 7. We're going to buy season tickets for the football games.

weather 8. It's very warm out tonight.

health 9. Do you think that it's risky to smoke cigarettes while you are pregnant?

seasons holidays 10. Will you be going on vacation in the spring?

money 11. Keep all the change.

music 12. Please turn down the stereo before the neighbors start complaining about it.

Set 24

food 1. We go out to eat dinner at least once a week.

family 2. Are you going to visit your parents tomorrow?

work 3. Have you heard about your promotion?

- clothes 4. Where did you buy your new hat?
- animals 5. Feed the dog.
- homes 6. We had all the floors scraped and polished this weekend.
- sports
hobbies 7. They like to go fishing.
- weather 8. There's a water shortage because it hasn't rained in the last eight weeks.
- health 9. Did your leg take a long time to heal?
- seasons
holidays 10. Mail your Christmas cards.
- money 11. Remember to get travelers checks before you go on vacation next week.
- music 12. Buy those records you want now while the store is having a big sale.

Set 25

- food 1. Do you eat bacon and eggs for breakfast every Sunday morning?
- family 2. Does he still live with his parents?
- work 3. Are you learning a lot on your new job?
- clothes 4. Do you need a new tie to go with that shirt you bought?
- animals 5. My landlord won't let me have a dog but I can have a cat.
- homes 6. Her apartment is big so she shares it with roommates.
- sports
hobbies 7. I play tennis every Tuesday.
- weather 8. I'm not going to work tomorrow if it snows a lot tonight.
- health 9. Cover your mouth when you cough.
- seasons
holidays 10. Buy champagne for the New Year's Eve

party.

money 11. Don't cash that check.

music 12. Practice the guitar.

Set 26

food 1. Can you bake the chocolate cake tonight?

family 2. Are his sisters still single or are they married?

work 3. Do you think that our secretary should be given a raise this year?

clothes 4. Get your new pants shortened or you won't be able to wear them Saturday.

animals 5. We went to the lake and fed the baby ducks.

homes 6. The kitchen needs new wallpaper.

sports
hobbies 7. My friends are having a party after the football game tomorrow night.

weather 8. Did you shovel the snow yet?

health 9. Go to the doctor and get a physical.

seasons
holidays 10. Vote on Election Day.

money 11. Lend him money.

music 12. He always buys subscriptions to both the symphony and the ballet.

Set 27

food 1. Are you having soup and a sandwich for lunch?

family 2. Who takes care of your children when they get home from school early?

work 3. Get a key made for both of the offices that you will be using.

- clothes 4. Buy yourself a new jacket now while they're on sale.
- animals 5. I never saw a bear.
- homes 6. My friends bought an old house and are starting to renovate it.
- sports
hobbies 7. Who taught you to play chess?
- weather 8. Did you see the full moon last night?
- health 9. Go see the nurse.
- seasons
holidays 10. Mow the lawn.
- money 11. It's better to pay with cash than to use credit cards.
- music 12. He plays chamber music on Wednesday nights.

Set 28

- food 1. What should we make for dinner when our friends come to visit us?
- family 2. You must spend more of your free time with your younger sisters and brothers!
- work 3. Finish all the filing by the end of the day.
- clothes 4. Hang up your good suit.
- animals 5. It's not a good idea to go that close to the pigeons.
- homes 6. Where is your new beach house?
- sports
hobbies 7. Did your father teach you to play chess?
- weather 8. How much snow fell?
- health 9. Brush your teeth.
- seasons
holidays 10. Families take vacations during the

summer when there is no school.

- money 11. He put his savings into the bank.
 music 12. I like to listen to the stereo at home.

Set 29

- food 1. Don't eat that cake because I baked it to
 take to the birthday party.
 family 2. Please take your little brother to the
 store with you.
 work 3. Don't take a long lunch.
 clothes 4. Do your laundry so that you have a clean
 shirt for tomorrow.
 animals 5. When did your dog get sick?
 homes 6. Where did you get all that new furniture?
 sports
 hobbies 7. Can she ice skate?
 weather 8. What's the temperature?
 health 9. She told me to take this medicine for my
 upset stomach.
 seasons
 holidays 10. We're going trick or treating on
 Halloween.
 money 11. I have to call the bank about my
 statement.
 music 12. He asked me to go to the concert with him
 next Saturday night.

Set 30

- food 1. Eat a good breakfast every morning before
 you leave home.
 family 2. Give your uncle the money.
 work 3. Don't make so many personal telephone
 calls while you are at work.

- clothes 4. I need to buy new boots.
- animals 5. Do you want to buy that cute puppy?
- homes 6. Where's your new house?
- sports/hobbies 7. Where's the football?
- weather 8. Put some salt on the sidewalk before someone slips and falls.
- health 9. My doctor said I'll be better soon.
- seasons
holidays 10. A lot of candy is sold on Valentine's Day.
- money 11. I'm getting a big tax return this year so I'm buying a car.
- music 12. Do you still like to play the violin or have you given it up?

Set 31

- food 1. Make my steak well done.
- family 2. Take these cookies with you when you go to visit your grandfather.
- work 3. Computers make typing reports much easier.
- clothes 4. The store is having a sale on nightgowns.
- animals 5. Who fed the goldfish?
- homes 6. What's your address?
- sports
hobbies 7. Remember to stretch before you try to run a long distance.
- weather 8. Don't go outside if it's too cold.
- health 9. Crash diets can really make a person very sick.
- seasons
holidays 10. Passover and Easter always seem to occur at the same time of the year.

- money 11. Will you lend me ten dollars until I can go to the bank?
- music 12. Did you tape the concert they broadcast on the radio?

Set 32

- food 1. Please slice the meat but be careful not to cut your fingers.
- family 2. My nephew is having a party.
- work 3. My friend was just fired from his job.
- clothes 4. My new shoes hurt.
- animals 5. Do parrots fly?
- homes 6. Vacuum the rugs and polish the furniture before the party tonight.
- sports hobbies 7. Bring your sneakers to the exercise class.
- weather 8. Don't ever stand under a tree during a thunderstorm.
- health 9. More people seem to catch colds in the winter than in the summer.
- seasons holidays 10. Will you be eating Thanksgiving dinner at a restaurant or at home this year?
- money 11. Did you put your savings in a high interest account?
- music 12. Where do you store albums?

Set 33

- food 1. I ate a big lunch today.
- family 2. My aunt and uncle live two blocks away.
- work 3. My boss is quitting.
- clothes 4. I need pants.

- animals 5. Remember to let the dog out before you go to school.
- homes 6. Paint the outside of the house first.
- sports
hobbies 7. Take off your skis when you leave the slopes.
- weather 8. Make sure you dress warmly on days when the temperature drops below freezing.
- health 9. How long will you have to keep that big cast on your broken leg?
- seasons
holidays 10. Why do people drink so much on New Year's Eve?
- money 11. Does he get an allowance?
- music 12. How many people are going to try out for the marching band?

Set 34

- food 1. The seafood restaurant had a special on lobster.
- family 2. My sister is moving.
- work 3. Filing is boring.
- clothes 4. Are you going to buy a new down coat this year?
- animals 5. Don't feed the squirrels in the park.
- homes 6. Clean out your closet before the painters get here.
- sports
hobbies 7. Buy yourself a better racket now that you have lost your old one.
- weather 8. The wind was so strong yesterday that it blew the branches off the tree.
- health 9. I have a sore throat and a very bad cough.
- seasons
holidays 10. When is April Fool's Day?

- money 11. Do you still think you will get a big bonus this year?
- music 12. Don't play your radio that loud.

Set 35

- food 1. I want french fries.
- family 2. My grandfather died.
- work 3. What did they give you for a present when you retired?
- clothes 4. Did you bring a sweater with you?
- animals 5. Close the windows before you open the bird cage.
- homes 6. Move the furniture into the middle of the room before you paint it.
- sports
hobbies 7. I went camping with my friends last week and we forgot the tent poles.
- weather 8. She thinks it's warm enough to go to the beach.
- health 9. How are you feeling today?
- seasons
holidays 10. Why do leaves always change color before they fall off the trees?
- money 11. Pay the bill at the store.
- music 12. Be careful when you use the new stereo.

Set 36

- food 1. I want coffee.
- family 2. Do your parents live near you or have they moved away?
- work 3. When did you arrive at the office?
- clothes 4. Will you take my suit to the cleaners

tomorrow?

- animals 5. When you go to the store buy a scratching post for the cat.
- homes 6. My new summer house at the beach needs to be painted before it's used.
- sports
hobbies 7. My friend went up in a hot air balloon yesterday.
- weather 8. It's very windy outside today.
- health 9. Do you still need to get a vaccination before you go overseas?
- seasons
holidays 10. Give her candy on Valentine's Day.
- money 11. Have all the checks mailed to your house.
- music 12. Don't scratch that record.

Set 37

- food 1. Are you going to bake a cake for the birthday party tomorrow?
- family 2. Are your cousins invited to the wedding?
- work 3. Is there a new person answering phones in your office?
- clothes 4. Where did you buy your jacket?
- animals 5. When I go to the zoo I like watching the chimps.
- homes 6. Here's my house.
- sports
hobbies 7. A lot of kids play soccer in school.
- weather 8. It's been very cold in the South this winter.
- health 9. If you still have fever tomorrow make sure you go to the doctor.
- seasons
holidays 10. Give her flowers.

- money 11. Make sure that you don't forget to send your tax returns in on time.
- music 12. Please change the channel.

Set 38

- food 1. Do you want a toasted English muffin?
- family 2. Are you going to take your brother to the movies?
- work 3. Do you walk to the office?
- clothes 4. Don't wear your new boots today because the streets are slushy.
- animals 5. Cats have whiskers.
- homes 6. We have new carpeting in the living room.
- sports
hobbies 7. I don't like to swim in an indoor pool.
- weather 8. Was it snowing when you left the house or had it already stopped?
- health 9. Don't take so much aspirin.
- seasons
holidays 10. Remember to plant your garden early so that the vegetables are ripe by summer.
- money 11. Don't spend so much.
- music 12. Many people like to listen to the radio before going to bed.

Set 39

- food 1. Where do you want to go to have dinner tonight?
- family 2. Did your sister have the baby?
- work 3. Tell your boss that you need to take three days off.
- clothes 4. Tie your shoelaces.

- animals 5. I walk my dog three times a day.
- homes 6. The elevator in my apartment building broke last night.
- sports
hobbies 7. Do you want to go horseback riding with us at the dude ranch?
- weather 8. Will it be sunny tomorrow?
- health 9. Make sure you go to the dentist and have him look at that tooth.
- seasons
holidays 10. Buy a Halloween costume.
- money 11. My friend's wedding reception will cost five thousand dollars, including the band.
- music 12. She plays the clarinet in a band.

Set 40

- food 1. What did you eat for breakfast?
- family 2. Your father has to go with you to buy the car.
- work 3. Call your boss.
- clothes 4. Iron that blouse before you go to school.
- animals 5. My friend has a horse that I can ride.
- homes 6. Which do you like better, living in a house or in an apartment?
- sports
hobbies 7. Was your tennis racket stolen?
- weather 8. Do you think that we can go away even if it snows a lot?
- health 9. Give blood this year.
- seasons
holidays 10. We can start going to our summer house on Memorial Day weekend.

- money 11. I always forget to balance my checkbook.
- music 12. The concert we went to last night was too loud.

Set 41

- food 1. Put the roast in the oven while I peel the potatoes.
- family 2. Ask your mother.
- work 3. Ask them to give you a new contract.
- clothes 4. Bring your shoes to the shoemaker for new heels.
- animals 5. Isn't it very hard to keep a big dog in a small apartment?
- homes 6. What color is your bathroom?
- sports
hobbies 7. How much does it cost to rent a sailboat for a half a day?
- weather 8. How much rain fell?
- health 9. If you try to exercise every day you won't be as tired.
- seasons
holidays 10. I'm having a costume party on Halloween.
- money 11. I'm going to the bank today to cash my check.
- music 12. He started to take singing lessons.

Set 42

- food 1. Frost the cake.
- family 2. Call your father if you'll be home late.
- work 3. Next time, don't come back from lunch so late.
- clothes 4. They bought that sweater for my roommate but I like to wear it.

- animals 5. Did you clean the fishtank?
- homes 6. Did it cost a lot of money to have the master bedroom made larger?
- sports
hobbies 7. Do you play golf?
- weather 8. I have to put the storm windows in before it becomes cold.
- health 9. I had the mumps two years ago.
- seasons
holidays 10. We are having a big party for New Year's Eve.
- money 11. I need to borrow five dollars.
- music 12. Did you buy new speakers when you bought the new turntable?

Set 43

- food 1. Cut some tomatoes and cucumbers for the salad.
- family 2. Don't fight so much with your sisters and brothers.
- work 3. I'm so busy at my new job that I'm always tired at night.
- clothes 4. The sleeves are too long.
- animals 5. Did you ever take the children to ride on the camels in the zoo?
- homes 6. Is the basement clean?
- sports
hobbies 7. Don't throw the baseball bat when you are getting ready to run.
- weather 8. Take the dog out before it snows.
- health 9. After your tonsils come out you can eat ice cream.
- seasons
holidays 10. People eat too much on Thanksgiving.

money 11. What day of the week does the bank stay open late?

music 12. Can you sing?

Set 44

food 1. Give me a hot corned beef sandwich on rye.

family 2. Let's take the children to the movies after we get home from work.

work 3. I'm getting a new office.

clothes 4. During the summer I put all my winter coats in storage at the cleaners.

animals 5. How can birds talk?

homes 6. Please clean all the storm windows before you hang up the curtains.

sports hobbies 7. Put away your chess set right now.

weather 8. Take in the laundry before it starts to rain again.

health 9. I stay home when I'm sick.

seasons holidays 10. Are most schools closed between Christmas Eve and New Year's Day?

money 11. Where's the check?

music 12. Do you have time to practice the piano?

Set 45

food 1. I try not to eat red meat more than three times a week.

family 2. My father has two jobs.

work 3. It's important to make a good impression when you go on a job interview.

clothes 4. That's a beautiful dress.

- animals 5. Remember not to go too close to stray dogs in the street.
- homes 6. Please don't eat in the living room.
- sports
hobbies 7. If you're skiing for the first time, take a lesson.
- weather 8. Wax the car if it's sunny.
- health 9. Did you go to the emergency room when you fell down?
- seasons
holidays 10. When is Thanksgiving?
- money 11. How much does your car insurance cost you?
- music 12. Where did your friend buy that new tape recorder?

Set 46

- food 1. She never eats red meat.
- family 2. My brother and his wife try to visit me at least once a month.
- work 3. The office is busy.
- clothes 4. Who bought you that beautiful green blouse that you were wearing yesterday?
- animals 5. Please don't give the dog that bone.
- homes 6. Clean the basement if you want to have that party.
- sports
hobbies 7. Make sure your running shoes fit.
- weather 8. This past summer was the hottest one that I can remember.
- health 9. My throat hurts.
- seasons
holidays 10. How will you celebrate the Fourth of July?

- money 11. Do you get paid once or twice a month?
- music 12. Please lower the radio when you listen to it this late at night.

Set 47

- food 1. We decided to make hamburgers and hot dogs for the barbecue this Sunday afternoon.
- family 2. I have ten cousins.
- work 3. Is your new job closer to your house than the old one?
- clothes 4. Did you buy a bathing suit yet?
- animals 5. Give the cat some food before you go out tonight.
- homes 6. Meet me at my house later.
- sports
hobbies 7. I bought new cross country skis to take on my trip.
- weather 8. It's hot today.
- health 9. When are you going to the eye doctor?
- seasons
holidays 10. Do you go to the beach during the summer?
- money 11. Don't spend so much or you'll never be able to buy that car.
- music 12. Buy that record for him.

Set 48

- food 1. I like ice cream.
- family 2. Did your brothers and sisters come to visit you on your birthday?
- work 3. How long have you been employed there?
- clothes 4. What should I wear to the party next Saturday night?

- animals 5. Don't pet that big black dog.
- homes 6. I think we should replace the carpeting in the living room.
- sports
hobbies 7. I like baseball.
- weather 8. It is going to be very windy today.
- health 9. Did you break your wrist or was it sprained?
- seasons
holidays 10. You better be very careful when you go trick or treating on Halloween.
- money 11. Give me five dollars tomorrow.
- music 12. Stop playing the radio so loud at night or I will take it away.

APPENDIX B

PITCH QUANTIZATION SOFTWARE

This program allows for the frequency quantization of voice fundamental frequency. It requires the use of an ISAAC interface to an IBM PC-XT computer. The program is currently saved under the name PITQ. It is menu driven. The main menu options are:

1. (at line 2000) Display and select quantization level.
2. (at line 3000) Convert digital representation of input dc voltage to an equivalent output value for the VCO.
3. (at line 4000) Convert input to appropriate digital representation to produce quantized fundamental frequency contours.

On the following pages will be found:

- A. Table of representative input/output voltages and frequencies generated by VCO.
- B. Program listing.

Table B.1
Representative Input Voltages to IBM PC-XT and
Equivalent Voltage Input to VCO Needed to Replicate
Pitch Contour

INPUT FREQUENCY	INPUT VOLTAGE TO IBM	VOLTAGE INPUT TO VCO	OUTPUT FREQUENCY
90	75*	2105	90
100	192	2181	100
150	817	2571	150
200	1402	2961	200
250	2028	3350	250
300	2612	3740	300
350	3156	4129	350

*Note: Formula for Conversion of Input Voltage to IBM
to Appropriate Input Voltage for VCO:

$$\text{INPUT TO VCO} = \text{INPUT TO IBM} * 657 + 2055.725$$

```
5 REM Pitch Quantization Program
6 Version N*2: Written by Edmundo A. Tereso, Manila
9 REM Set Constant
10 CL=16
15 FLAG=0
24 CLS
25 LOCATE 2,35
28 REM Main Menu
30 PRINT "MAIN MENU"
40 LOCATE 5,11
70 PRINT SPC(10)"2: PARAMETER CHANGE";PRINT
70 PRINT SPC(10)"3: REAL TIME CHECK";PRINT
80 PRINT SPC(10)"4: QUANTIFICATION";PRINT
90 PRINT SPC(10)"5: EXIT TO SYSTEM";PRINT
100 LOCATE 20,10
110 PRINT "WHICH NUMBER?"
120 A$=INPUT$(1)
130 X=VAL(A$)
140 ON X GOSUB 1000, 2000, 3000, 4000, 5000
150 GOTO 24
```

```

8000 CLE
8005 NEW Parameter Change
8010 LOCATE 8, 35
8020 PRINT "PARAMETERS"
8030 LOCATE 9, 10
8040 PRINT "11 QUANTIZATION LEVEL (0-9)"; DL
8050 LOCATE 20, 10; PRINT "WHICH NUMBERS? (M FOR MENU)
8100 AS=INPUT(1)
8110 IF AS = "M" OR AS = "" THEN RETURN
8120 X=VAL(AS)
8130 ON X GOTO 2500
8200 LOCATE 20, 10; INPUT "CHANGE LEVEL TO (0-9)"; DL
8205 NEW CHECK IF VALID
8210 IF DL<0 OR DL>9 THEN PRINT "TRY AGAIN"; GOTO 2500
8220 GOTO 2000

```

```
3000 CLS
3005 REM Real Time Check
3010 LOCATE 2,20
3020 PRINT "REAL TIME CHECK"
3030 LOCATE 20,20:PRINT "START TAPE"
3035 REM Turn off interrupts not associated with Isiad
3040 OPTS="INTSOFF"
3045 REM Set input channel 0 as input of A/D
3050 INCHANX=0
3055 REM Set output channel 0 as output of D/A
3060 OUTCHANX=0
3065 REM Get an input value
3070 CALL AINS(INCHANX,BLUEX,OPTS)
3075 REM Check if it is valid
3080 IF BLUEX<75 THEN 3100
3085 REM Convert to an output value for the VLD
3090 BLUEX=BLUEX*.657 + 2055.725
3095 REM Output value to D/A
3100 CALL AOUS(OUTCHANX,BLUEX,OPTS)
3105 REM Check for keypress; if keypress then quit
3110 AS=INKEYS;IF LEN(AS)=1 THEN RETURN
3115 REM Get another sample
3120 GOTO 3070
```

```

4000 CLR
4005 REM Quantization
4010 LOCATE 2,20
4020 PRINT "QUANTIZATION LEVEL:DL"
4025 REM Check if it is the right quantization level
4030 LOCATE 3,20:PRINT"IS THIS CORRECT? Y/N"
4040 A$=INPUT$(1)
4045 REM IF NOT CORRECT GET CORRECT VALUE
4050 IF A$="N" OR A$="n" THEN GOSUB 3000: GOTO 4000
4055 REM IF CORRECT CONTINUE
4060 IF A$="Y" OR A$="y" THEN GOTO 4070
4070 GOTO 4030
4080 LOCATE 3,20:PRINT"START TAP"
4085 REM TURN OFF INTERRUPTS
4090 OPTS="INTBOFF"
4095 REM SET INPUT CHANNEL
4100 INCHANX=0
4101 IF DL=0 THEN GOSUB 4000
4102 REM CALCULATE CONSTANT FOR QUANTIZE LEVEL DESIRED
4104 IF FLAG=1 THEN FLAG=0: RETURN
4105 Y=.693/DL
4106 REM CALCULATE CONSTANT FOR QUANTIZE LEVEL DESIRED
4107 Z=.693/Y
4108 REM SET OUTPUT CHANNEL
4110 OUTCHANX=0
4115 REM GET INPUT
4120 CALL AINS (INCHANX,INX,OPTS)
4125 REM CHECK IF ITS VALID
4130 IF INX=5 THEN OTX=INY:GOTO 4155
4135 REM CONVERT INPUT VALUE TO APPROPRIATE QUANTIZE LEVEL
4140 FX=LOG(INX/1210-.8413)+Z+.5
4145 REM CONVERT QUANTIZE LEVEL TO APPROPRIATE OUTPUT VALUE FOR VCO
4146 OTX=780*EXP(Y*(FX-.5))+1401
4150 OTX=780*EXP(Y*(FX-.5))+1401
4155 REM OUTPUT VALUE
4160 CALL AOUTS (OUTCHANX,OTX,OPTS)
4165 REM LOOK FOR KEYPRES: IF KEYPRES THEN TURN OFF VCO AND RETURN TO MAIN
U
4170 A$=INKEY$:IF LEN(A$)=1 THEN OTX=25:CALL AOUTS(OUTCHANX,OTX,OPTS):RETURN
4175 REM GET ANOTHER VALUE
4180 GOTO 4120

```

```

4500 OUTCHN%0: REM SET OUTPUT CHANNEL
4510 REM GET INPUT
4520 CALL GAINS (INCHAN, INVAL, IPT%)
4530 REM CHECK IF VALID
4540 IF INK<75 THEN DIM=2735
4550 REM OUTPUT VALUE
4560 CALL AOUS (OUTCHN%, 0, OPT%)
4565 REM LOOK FOR REQUEST IF KEYS ARE IN TURN, WE GO AND GET IN TURN
4570 REM LOOK FOR KEYS IF LEN(A%)=1 THEN DIM=2735:GOTO 4565:GOTO 4565:GOTO 4565
4580 GOTO 4510
5000 END

```

APPENDIX C

INDIVIDUAL SUBJECT HISTORY

Table C.1
Individual Subject History for Subjects Who
Participated in the Main Experiment

SUBJECT	SEX	AGE	COLLEGE GRADUATE	SPEECH & HEARING BACKGROUND
1	F	24	YES	NO
2	F	31	YES	YES
3	F	31	YES	YES
4	F	40	YES	NO
5	M	32	YES	YES
6	F	33	YES	YES
7	F	33	YES	NO
8	M	32	YES	YES
9	M	32	YES	NO
10	M	31	YES	NO
11	F	31	YES	YES
12	F	25	YES	NO

Table C.2
Individual Subject History for Subjects Who
Participated in the Supplementary Experiment

SUBJECT	SEX	AGE	COLLEGE GRADUATE	SPEECH & HEARING BACKGROUND
1	M	38	YES	YES
2	F	38	YES	YES
3	F	34	YES	NO
4	F	38	YES	YES
5	F	35	YES	YES
6	M	34	YES	YES
7	F	28	YES	YES
8	F	33	YES	YES
9	F	36	NO	NO
10	F	36	YES	YES
11	F	26	YES	YES
12	F	18	NO	NO

APPENDIX D

FORM A: INFORMAL HISTORY QUESTIONNAIRE

NAME: _____
DOB _____

1. Have you ever been evaluated for a speech and/or language problem?
2. Have you ever been treated for a speech, language or hearing problem?
3. Is English your native language?
4. Do you now or have you ever worn glasses?
5. When was the last time you had your eyes examined?
6. Do you feel you have any trouble seeing?
7. Do you have a driver's license?

APPENDIX E
INSTRUCTIONS

The following are the written instructions given to subjects prior to testing in each condition. The first set of instructions was for the speechreading alone condition. The second set was used for all auditorily supplemented conditions.

YOU WILL BE WATCHING A SPEAKER SAY SENTENCES. SHE BEGINS EACH SENTENCE BY SAYING "NUMBER ____". WHEN SHE SAYS "NUMBER 1" THE SENTENCE WILL BE ABOUT FOOD, "NUMBER 2" IS ABOUT FAMILY, ETC. PLEASE LOOK AT THE LIST OF TOPICS BEFORE YOU WATCH EACH SENTENCE. AFTER THE SPEAKER FINISHES SAYING THE SENTENCE, YOU WILL REPEAT AS MUCH OF THE SENTENCE AS POSSIBLE. DON'T BE AFRAID TO GUESS IF YOU ARE NOT SURE.

YOU WILL BE WATCHING AND LISTENING TO A SPEAKER SAY SENTENCES. SHE BEGINS EACH SENTENCE BY SAYING "NUMBER ____". WHEN SHE SAYS "NUMBER 1" THE SENTENCE WILL BE ABOUT FOOD, "NUMBER 2" IS ABOUT FAMILY, ETC. PLEASE LOOK AT THE LIST OF TOPICS BEFORE YOU WATCH EACH SENTENCE. AFTER THE SPEAKER FINISHES SAYING THE SENTENCE, YOU WILL REPEAT AS MUCH OF THE SENTENCE AS POSSIBLE. DON'T BE AFRAID TO GUESS IF YOU ARE NOT SURE.

APPENDIX F

RAW DATA FOR SENTENCE LENGTH ANALYSIS

Table F.1
Percent Correct Word Recognition as a Function of
Sentence Length by Speechreading alone

<u>SUBJECT</u>	<u>AVERAGE SENTENCE LENGTH</u>		
	<u>SHORT</u>	<u>MEDIUM</u>	<u>LONG</u>
1	28.9	17.6	10.7
2	65.5	50.8	43.1
3	66.2	37.1	33.9
4	21.2	11.5	12.8
5	33.6	24.5	24.4
6	60.9	56.9	41.6
7	21.7	19.7	10.2
8	68.3	48.3	59.8
9	48.5	33.0	35.0
10	19.0	17.2	8.2
11	61.5	38.5	21.8
12	63.5	37.6	24.8

Table F.2
Percent Correct Word Recognition as a Function of
Sentence Length by Speechreading + Fo Quantized at one
level

<u>SUBJECT</u>	<u>AVERAGE SENTENCE LENGTH</u>		
	<u>SHORT</u>	<u>MEDIUM</u>	<u>LONG</u>
1	33.7	19.8	21.0
2	77.1	69.9	65.7
3	64.6	54.2	44.1
4	25.6	9.0	7.3
5	54.3	49.1	27.9
6	65.5	47.7	28.1
7	44.8	28.7	31.5
8	80.6	54.7	67.6
9	70.5	46.3	34.1
10	22.6	7.4	5.3
11	61.5	55.3	51.0
12	39.0	49.4	40.3

Table F.3
Percent Correct Word Recognition as a Function of
Sentence Length by Speechreading + Fo Quantized at four
Levels

<u>SUBJECT</u>	<u>AVERAGE SENTENCE LENGTH</u>		
	<u>SHORT</u>	<u>MEDIUM</u>	<u>LONG</u>
1	72.3	42.6	31.8
2	54.0	42.4	51.0
3	67.1	54.7	51.2
4	31.9	21.6	15.0
5	53.8	49.3	26.0
6	42.9	47.3	37.9
7	40.7	44.9	39.7
8	76.7	65.2	51.5
9	57.4	53.1	40.0
10	44.0	27.3	14.7
11	73.2	56.0	47.4
12	64.0	43.7	38.8

Table F.4
Percent Correct Word Recognition as a Function of
Sentence Length by Speechreading + Fo Quantized at
Eight Levels

<u>SUBJECT</u>	<u>AVERAGE SENTENCE LENGTH</u>		
	<u>SHORT</u>	<u>MEDIUM</u>	<u>LONG</u>
1	50.2	38.1	45.7
2	15.3	63.2	56.7
3	74.5	63.3	57.9
4	36.1	16.2	17.8
5	52.1	42.3	26.2
6	68.3	75.7	53.2
7	49.8	45.3	50.6
8	75.4	70.5	59.9
9	69.7	53.8	41.8
10	36.0	24.5	28.5
11	66.7	55.0	43.6
12	60.0	48.1	48.1

Table F.5
Percent Correct Word Recognition as a Function of
Sentence Length by Speechreading + Fo Quantized at
Twelve Levels

<u>SUBJECT</u>	<u>AVERAGE SENTENCE LENGTH</u>		
	<u>SHORT</u>	<u>MEDIUM</u>	<u>LONG</u>
1	68.9	47.8	58.3
2	85.9	69.6	61.5
3	67.7	52.6	53.1
4	28.5	19.9	11.5
5	57.9	55.5	40.3
6	55.4	57.6	42.0
7	81.4	63.4	58.7
8	85.0	71.8	89.8
9	38.4	54.9	42.2
10	51.9	42.2	37.6
11	79.3	67.6	58.9
12	65.1	50.8	44.9

Table F.6
Percent Correct Word Recognition as a Function of
Sentence Length by Speechreading + Raw Fo

<u>SUBJECT</u>	<u>AVERAGE SENTENCE LENGTH</u>		
	<u>SHORT</u>	<u>MEDIUM</u>	<u>LONG</u>
1	82.2	77.8	83.1
2	77.0	88.9	84.4
3	84.6	82.6	79.9
4	53.1	39.9	46.2
5	76.4	69.1	69.9
6	73.7	59.2	68.0
7	78.0	68.0	80.0
8	86.4	80.4	93.3
9	88.3	82.4	81.4
10	8.7	66.0	57.0
11	83.7	85.8	72.4
12	76.7	79.9	73.2

APPENDIX G

EVALUATION OF PERCENTAGE OF PITCH EXTRACTION ERRORS

In order to estimate the percentage of pitch extraction errors that occurred during the preparation of the test tapes spectrograms were made from the raw and the processed fundamental frequency recordings for a random 5% sample of the sentences. Analysis was conducted for an equal number of questions, statements and imperative. Listed on the following page are the sentences analyzed and the results.

Table G.1
Sentences Analyzed to Determine the Percentage of Pitch
 Extraction Errors

SET/SENTENCE	# OF WORDS	# OF ERRORS
1/1	4	1
2/2	12	1
2/3	7	2
2/6	5	2
5/9	4	2
6/8	4	0
7/5	6	2
8/4	4	0
9/5	4	0
9/10	7	0
10/10	5	0
11/12	11	3
12/12	6	1
13/4	8	1
13/5	9	2
13/8	9	0
14/6	13	0
18/1	5	2
19/2	12	1
20/2	4	1
22/1	12	1
22/4	8	1
23/7	10	2
25/2	7	0
29/11	9	1
33/6	7	2
34/10	5	1
36/10	6	1
40/1	6	1
42/2	8	0
43/6	4	1
45/8	6	0
47/2	4	0
TOTAL:	231	33
PERCENTAGE ERROR WORDS:		13.9%
PERCENTAGE ERROR SENTENCES:		66.7%

REFERENCES

- Bekesy, G. von (1959). Similarities between hearing and skin sensations. Psychological Review, 66, 1-22.
- Bekesy, G. von (1967). Sensory inhibition. Princeton, N.J.: Princeton University Press.
- Borden, G., & Harris, K. (1980). Speech Science Primer. Baltimore: Williams and Wilkens.
- Boothroyd, A. (1975). SARP Report # . The Clarke School for the Deaf. Northampton, Mass.
- Boothroyd, A. (1981). Development of a sentence test of speech perception. SARP Report #39.3, The Clarke School for the Deaf. Northampton, Mass.
- Boothroyd, A. (1984). Experiments with a wearable tactile pitch display. Proceedings of the International Conference on Rehabilitation Engineering. Ottawa, Canada.
- Boothroyd, A. (1985). A wearable tactile intonation display for the deaf. IEEE Trans. Acoust. Speech & Signal Processing, 33, 111-117.
- Boothroyd, A. (1986). Personal communication.
- Boothroyd, A., & Cawkwell, S. (1970). Vibrotactile thresholds in pure tone audiometry. Acta Otolaryngology, 69, 381-387.
- Boothroyd, A., & Jiang, Q. (1985). An interpulse interval timer with binary output. Interim Report TSL#1, NIH Grant #PO1N17764, City University of New York.
- Boothroyd, A., & Hnath, T. (1986). Lipreading with tactile supplements. Journal of Rehabilitation Research and Development, 23, 139-146.
- Boothroyd, A., Hnath, T., & Hanin, L. (1986). Tactile Detection of Pitch Falls. Interim Report TSL#7, NIH Grant #PO1N17764, City University of New York.
- Breeuwer, M., & Plomp, R. (1986). Speechreading supplemented with auditorily presented speech parameters. Journal of the Acoustical Society of America, 79, 481-489.

- Brownlee, K. (1965). Statistical Theory and Methods in Science and Engineering. New York: Dunn Wiley & Sons.
- Butler, D., & Albrite, M. (1956). The pitch discrimination function of the pathological ear. Archives of Otolaryngology, 63, 411-418.
- Craig, W. (1964). Effects of pre-school training on the development of reading and lipreading skills of deaf children. American Annals of the Deaf, 109, 48-63.
- deFilippo, C., & Scott, B. (1978). A method for training and evaluating the reception of ongoing speech. Journal of the Acoustical Society of America, 63, 1186-1192.
- DiCarlo, L. (1962). Some relationship between frequency discrimination and speech reception performance. Journal of Auditory Research, 2, 37-49.
- Douek, E., Fourcin, A., Moore, B., & Clarke, G. (1977). A new approach to the cochlear implant. Proceedings of the Royal Society of Medicine, 70, 379-383.
- Erber, N. (1972). Speech envelope cues as an acoustic aid to lipreading for profoundly deaf children. Journal of the Acoustical Society of America, 51, 1224-1227.
- Erber, N. (1979). Speech perception by profoundly hearing-impaired children. Journal of Speech and Hearing Disorders, 44, 255-270.
- Ericson, L., & Risberg, A. (1977). Threshold of hearing, vibration and discomfort in a group of severely hard of hearing and profoundly deaf students. STL Quarterly Progress & Status Report, 4, 22-28.
- Flannigan, J. (1957). Estimates of the maximum precision necessary in quantizing certain "dimensions" of vowel sounds. Journal of the Acoustical Society of America, 29, 553-554.
- Fourcin, A., Rosen, S., Moore, B., Douek, E., Clarke, G., Dodson, H., & Bannister, L. (1979). External stimulation of the cochlea: Clinical, physiological, speech perceptual and histological findings. Brittish Journal of Audiology, 13, 85-107.
- Fourcin, A., Douek, E., Moore, B., Abberton, E., Rosen, S., & Walliker, J. (1984). Speech perception

- element stimulation in electrical hearing. Archives of Otolaryngology, 110, 145-153.
- Geldard, F. (1972). The human senses. New York: Wiley.
- Geldard, F., & Sherrick, C. (1972). The cutaneous "rabbit": a perceptual illusion. Science, 178, 178-179.
- Geldard, F., Sherrick, C., & Cholewaik, R. (1981). Princeton cutaneous project progress report no. 38. Princeton, N.J.: Princeton University.
- Gengel, R. (1969). Practice effects in frequency discrimination by hearing-impaired listeners. Journal of the Acoustical Society of America, 54, 11-15.
- Gilson, R. (1969). Vibrotactile masking: effects of multiple maskers. Perception and Psychophysics, 5, 181-182.
- Gold, T. (1978). Speech and hearing skills: A comparison between hard-of-hearing and deaf children. Unpublished doctoral dissertation, City University of New York.
- Grant, K. (1980). Investigating a tactile speechreading aid: the transmission of prosodic information in connected discourse and sentences. Unpublished masters thesis, University of Washington, Seattle.
- Grant, K., Ardell, L., Kuhl, P., & Sparks, D. (1985). The contribution of fundamental frequency, amplitude envelope, and voicing duration cues to speechreading in normal-hearing subjects. Journal of the Acoustical Society of America, 77, 671-677.
- Grant, K., Ardell, L., Kuhl, P., & Sparks, D. (1986). The transmission of prosodic information via an electrotactile speechreading aid. Ear and Hearing, 7, 328-335.
- Hopkins, L. (1953). The relationship between degree of deafness and response to acoustic training. Volta Review, 55, 32-35.
- Hudgins, C. (1954). Auditory training: it's possibilities and limitations. Volta Review, 61, 316-319.
- Johnson, K., & Phillips, J. (1981). Tactile spatial

- resolution I: Two-point discrimination, gap detection, grating resolution, and letter recognition. Journal of Neurophysiology, 45, 1177-1191.
- Kirman, J. (1973). Tactile communication of speech: a review and analysis. Psychological Bulletin, 80, 54-74.
- Kirman, J. (1974). Number of stimulators and tactual apparent movement. Journal of Experimental Psychology, 103, 1175-1180.
- Lieberman, P. (1967). Intonation, perception, and language. Cambridge: MIT Press.
- McGarr, N. (1976). The production and reception of prosodic features. Paper presented A.G. Bell Assoc. Boston, Mass.
- Merznich, M., & Harrington, T. (1969). The sense of flutter-vibration evoked by stimulation of the hairy skin of primates: comparison of human sensory capacity with response of mechanoreceptive efferents innervating the hairy skin of monkeys. Experimental Brain Research, 9, 236-260.
- Most, T. (1985). Assessment of the perception of intonation by severely and profoundly hearing-impaired children. Unpublished doctoral dissertation, City University of New York.
- Nober, E. (1970). Cuticle air and bone conduction thresholds of the deaf. Exceptional Children, 56, 571-579.
- Numbers, M., & Hudgins, C. (1948). Speech perception in present day education for deaf children. Volta Review, 50, 449-456.
- Pickett, J. & McFarland, W. Auditory implants and tactile aids for the profoundly deaf. Journal of Speech and Hearing Research, 28, 134-150.
- Plant, G., & Risberg, A. (1983). The transmission of fundamental frequency variations via a single channel vibrotactile aid. STL Quarterly Progress & Status Report, 2-3, 61-84.
- Reed, C., Durlach, N., & Braida, L. (1982). Research on tactile communication of speech: a review. ASHA monographs, 20.
- Risberg, A., (1974). The importance of prosodic speech

- elements for the lipreader. In H. Birk-Nielsen and E. Klapm (eds.), Visual and audiovisual perception of speech, Sixth Danavox symposium, Scandinavian Audiology, Sup. 4, 153-164.
- Risberg, A., & Agelfors, E. (1978). On the identification of intonation contours by hearing-impaired listeners. STL Quarterly Progress & Status Report, 2-3, 51-61.
- Risberg, A., & Lubker, J. (1978). Prosody and speechreading. STL Quarterly Progress & Status Report, 4, 1-16.
- Rosen, S., Moore, B., & Fourcin, A. (1979). Lipreading with fundamental frequency information, Proceedings of Inst. Acoust., 5-8.
- Rosen, S., Fourcin, A., & Moore, B. (1981). Voice pitch as an aid to lipreading, Nature, 291, 150-152.
- Rosenberg, R., Schafer, R., & Rabiner, L. (1971). Effect of smoothing and quantizing formant coded voiced speech. Journal of the Acoustical Society of America, 50, 1532-1538.
- Rothenberg, M., Verillo, R., Zohorian, S., Brachman, M., & Bolanowski, S. (1977). Vibrotactile frequency for encoding a speech parameter. Journal of the Acoustical Society of America, 62, 1003-1012.
- Rothenberg, M., & Molitor, R. (1979). Encoding voice fundamental frequency into vibrotactile frequency. Journal of the Acoustical Society of America, 66, 1029-1038.
- Sanders, D., & Goodrich, S. (1971). The relative contribution of visual and auditory components of speech to speech intelligibility as a function of three conditions of frequency distortion. Journal of Speech and Hearing Research, 14, 172-178.
- Sherrick, C. (1968). Studies of tactile apparent movement. In D.R. Kenshalo (ed.), The skin senses. Springfield, Ill: Charles C. Thomas.
- Sherrick, C. (1984). Basic and applied research on tactile aids for deaf people: progress and prospects. Journal of the Acoustical Society of America, 75, 1325-2342.
- Sherrick, C., & Rogers, R. (1966) Apparent haptic movement. Perception and Psychophysics, 1, 175-180.

- Sherrick, C., & Craig, J. (1982). The psychophysics of touch. In W. Schiff and E. Foulke (eds.), Tactual perception: a sourcebook. Cambridge: Cambridge University Press.
- Stark, R., & Levitt, H. (1974). Prosodic feature reception and production in deaf children. Paper presented at the 87th meeting of the Acoustical Society of America.
- Taaffe, G. and Wong, W. (1957). Studies in variables in lipreading stimulus materials. Studies in visual communication, Los Angeles: John Tracey Clinic.
- Talbot, W., Darian-Smith, H., Kornhuber, H., & Mountcastle, V. (1968). The sense of flutter-vibration: comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. Journal of Neurophysiology, 31, 301-335.
- Taylor, B. (1978). Dimensional redundancy in the processing of vibrotactile temporal order. Unpublished doctoral dissertation, Princeton University.
- Valbo, A., & Johansson, R. (1978). The tactile sensory innervation of the glabrous skin of the human hand. In G. Gordon (ed.), The mechanism of recognition of objects by manipulation: a multidisciplinary approach. Oxford: Pergamon Press.
- Verillo, R. (1963). Effect of contactor area on the vibrotactile threshold. Journal of the Acoustical Society of America, 35, 1962-1966.
- Verillo, R. (1968). A duplex mechanism of mechanoreception. In D.R. Kenshalo (ed.), The skin senses. Springfield, Ill: Charles C. Thomas.
- Weinstein, S. (1968) Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In D.R. Kenshalo (ed.), The skin senses. Springfield, Ill: Charles C. Thomas.
- Yeung, E., & Boothroyd, A. (1985). PITCHEXT: Analysis and display software for use with interpulse interval timer. Interim Report TSL#2, NIH Grant #PO1N17764, City University of New York.