Order Number 8814208

The voice of emotion: Acoustic properties of six emotional expressions

Baldwin, Carol May, Ph.D.
The University of Arizona, 1988

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

THE VOICE OF EMOTION: ACOUSTIC PROPERTIES OF SIX EMOTIONAL EXPRESSIONS

by

Carol May Baldwin

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 8

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Carol May Baldwin

entitled THE VOICE OF EMOTION: ACOUSTIC PROPERTIES OF SIX EMOTIONAL EXPRESSIONS

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Mary C. Wetzel	Date 4/11/88
Judith L. Lauter	Date 4-11-88
Robert W. Lansing	Date 4-11-88
Roger J. Daldrup	Date 4-11-88
Oscar C. Christensen	Date 4/11/88

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director — Date 4/11/88

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Carol May Baldwin

ACKNOWLEDGEMENTS

> The shortest and surest way to arriving at real knowledge is to unlearn the lessons we have been taught, to remount first principles, and to take nobody's word about them.
>
> Henry Bolingbroke

For the guidance, encouragement, and support I received during my academic training, and for their friendship and regard, I sincerely thank my graduate committee: Mary Wetzel, Judith Lauter, Robert Lansing, Roger Daldrup, and Oscar Christensen.

For their time, talent, and willingness to express emotions, I offer genuine appreciation to the actors (Cynthia Meier, Tamra Moore, Susan Rush, Andrew Dasher, Daniel Mello, David Williams) and nonactors (Ann Kelley, Nancy Finch, Angela Sorrell, Kelly Aune, Donald Finch, Mark Lowder), who made this study possible.

For the use of the spectrograph equipment, I thank Richard Demers and the Department of Linguistics. For their comments on the data analysis, I thank James King and Peter Facciola.

For their concern during times of self-doubt, humor in times of struggle, compassion when nothing seems to go right, and love without restraint, I extend gratitude and love to my family and many dear friends.

To the hospice clients (with special recognition to Dad, Shoobie, Joan, and Catherine) who, in their dying, taught me to live with dignity, and to my daughter, Jennifer, who so generously taught me a full range of emotions, I dedicate this work.

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS.................................. 8

LIST OF TABLES... 9

ABSTRACT... 10

INTRODUCTION... 12

Need for This study.................................... 13
Statement of the Problem............................... 14
Research Questions..................................... 15
Operational Definitions................................ 17
Assumptions.. 18

REVIEW OF THE LITERATURE............................... 19

Early Studies of Emotional Expressions................. 20
Darwin... 20
Studies of Facial Expressions of Emotion............... 22
Summary.. 24
Linguistic and Speech Science Studies.................. 25
Suprasegmentals.. 25
Prosody.. 28
Intonation... 29
Linguistic Analysis of Emotion......................... 31
Summary.. 31
Acoustic Correlate Studies............................. 32
Research Tools... 32
Physiological Stress Studies........................... 35
Summary.. 38
Emotion Related Acoustical Studies--English............ 38
Emotion Related Acoustical Studies--Foreign............ 42
French... 42
Dutch.. 43
Russian.. 44
Summary for Acoustical Studies......................... 46
American English....................................... 46
Foreign Language....................................... 47

METHODS.. 49

Gender... 49
Actors and Nonactors................................... 49

TABLE OF CONTENTS-Continued

Page

Design and Procedures.............................. 50
Emotion Types................................ 50
Linguistic Carrier of Emotion Types........... 50
Recording Instrumentation.......................... 50
Testing Procedure.................................. 51
Preliminary Instructions...................... 51
Introduction to Experimental Procedures....... 52
Experimental Procedures....................... 53
Evocation and Production of Emotion........... 53
Validation of Emotion Types................... 54
Acoustic Instrumentation........................... 55
Acoustic Measures.................................. 55
Duration Values............................... 55
Intensity Values.............................. 57

RESULTS... 58

Sentence Duration.................................. 58
Effects for Duration.......................... 60
Interaction of Conditions X Role X Sex for Duration...................................... 60
Interaction of Conditions X Emotions X Sex for Duration...................................... 63
Post Hoc Analyses Within Groups--Males.... 65
Post Hoc Analyses Within Groups--Females.. 67
Post Hoc Analyses Between Groups--Males and Females.............................. 67
Intensity..................................... 68
Effects for Intensity..................... 70
Summary....................................... 73

DISCUSSION AND RECOMMENDATIONS......................... 75

Vocal Taxonomy..................................... 76
Intensity Taxonomy for Six Emotions........... 77
Preliminary Production/Perception Comparisons. 81
Recommendations for Future Vocal Taxonomic Studies....................................... 83
Gender Differences................................. 84
Gender Specific Taxonomy...................... 90
Recommendations for Future Gender Studies..... 92
Actor and Nonactor Differences..................... 93
Recommendations for Future Research........... 96

TABLE OF CONTENTS-Continued

Page

APPENDIX A: ACOUSTIC CORRELATES DEMOGRAPHIC FORM..... 97

APPENDIX B: DATA SETS FOR SUBJECTS.................... 99

REFERENCES.. 104

LIST OF ILLUSTRATIONS

Figure | Page

1. Spectrogram, Amplitude Contour, and Waveform Samples for the Sentence "Of Course I Love You" Expressed in a "Happy" Tone of Voice by a 23 Year Old Male Subject........................... 56

2A. Mean and Standard Error Results for Main Effects for Conditions (6 Neutral/6 Emotions) on Sentence Duration (n = 12)...................... 61

2B. Mean and Standard Error Results for Main Effects for Sex on Sentence Duration (n = 12)........... 61

3. Conditions X Role X Sex Interaction on Duration. Shows Main Effects Also: Conditions and Sex (n = 6 Males/6 Females)......................... 62

4. Conditions X Emotions X Sex Interaction. Shows Main Effects Also: Conditions and Sex. HA = Happiness, SU = Surprise, SA = Sadness, FE = Fear, AN = Anger, DI = Disgust (n = 6 Males/ 6 Females)....................................... 64

5. Mean Durations for 6 Male and 6 Female Subjects for Seven Vocal Expressions. Ha = Happiness, SU = Surprise, SA = Sadness, FE = Fear, AN = Anger, DI = Disgust, NU = Neutral Condition..... 66

6. Conditions X Emotions Interaction on Mean Intensity. Shows Main Effects Also: Emotions. N = Neutral, HA = Happiness, SU = Surprise, SA = Sadness, FE = Fear, AN = Anger, DI = Disgust (n = 12)....................................... 71

LIST OF TABLES

Table		Page
1.	Analysis of Variance for Overall Sentence Duration. R = Role (Actor/Nonactor); S = Sex (Male/Female); C = Conditions (6 Neutral Tones/ 6 Emotions); E = Emotions (Happiness, Surprise, Sadness, Fear, Anger, Disgust) (n = 12)	59
2.	Analysis of Variance for Mean Intensity. R = Role (Actor/Nonactor); S = Sex (Male/Female); C = Conditions (6 Neutral Tones/6 Emotions); E = Emotions (Happiness, Surprise, Sadness, Fear, Anger, Disgust) (n = 12)	69
3.	Newman-Keuls Paired Comparisons Results for Mean Intensity Measures for Six Emotional Expressions. HA = Happiness, AN = Anger, SU = Surprise, FE = Fear, DI = Disgust, SA = Sadness	72

ABSTRACT

Studies in the perceptual identification of emotional states suggested that listeners seemed to depend on a limited set of vocal cues to distinguish among emotions. Linguistics and speech science literatures have indicated that this small set of cues included intensity, fundamental frequency, and temporal properties such as speech rate and duration. Little research has been done, however, to validate these cues in the production of emotional speech, or to determine if specific dimensions of each cue are associated with the production of a particular emotion for a variety of speakers.

This study addressed deficiencies in understanding of the acoustical properties of duration and intensity as components of emotional speech by means of speech science instrumentation. Acoustic data were conveyed in a brief sentence spoken by twelve English speaking adult male and female subjects, half with dramatic training, and half without such training. Simulated expressions included: happiness, surprise, sadness, fear, anger, and disgust.

The study demonstrated that the acoustic property of mean intensity served as an important cue for

a vocal taxonomy. Overall duration was rejected as an element for a general taxonomy due to interactions involving gender and role. Findings suggested a gender-related taxonomy, however, based on differences in the ways in which men and women use the duration cue in their emotional expressions. Results also indicated that speaker training may influence greater use of the duration cue in expressions of emotion, particularly for male actors.

Discussion of these results provided linkages to (1) practical management of emotional interactions in clinical and interpersonal environments, (2) implications for differences in the ways in which males and females may be socialized to express emotions, and (3) guidelines for future perceptual studies of emotional sensitivity.

CHAPTER 1

INTRODUCTION

Vocal expressive cues are common to most human relationships and can strongly influence the context of these interactions. Starkweather (1961, p. 63) wrote:

> The tone of voice and the manner of speaking affect the listener's perception of the speaker's feeling state. These vocal guideposts suggest some of the personality characteristics of individuals, often enable a person to recognize a friend without seeing him, and indicate the speaker's emotional state of the moment. During infancy, prior to the learning of language, parents and children communicate largely through nonverbal vocal cues.

Unlike recognition of emotion in natural situations, however, scientific definitions remain ambiguous. While a range of approaches has been taken to identify the effects of emotional states on vocal characteristics, and their concomitant effects on listener's perceptions, few studies have provided a database and conceptual organization for emotional speech. No research has identified a taxonomy of emotional speech for a variety of speakers producing a variety of emotions.

Need for this Study

Despite almost universal endorsement of Starkweather's (1961) position quoted above, definitive research on the acoustical properties of emotional speech is lacking. The rationale for the present work came from Siegman (1985), who suggested that quantitative studies of expressive behavior could lead to a taxonomy of emotions. Support for this approach comes from Ekman's (1973) work demonstrating associations between members of a set of emotions and specific simulated facial patterns.

According to Brown et al. (1985), vocal correlate research has been restricted largely to studies of personality traits and states based on respondent perceptions of vocal characteristics, such as breathiness, or pitch modulation. Some possible reasons for the neglect of the study of emotional speech have been a research emphasis on non-emotional speech patterns and/or methodological difficulties.

Pickett (1980) and Scherer (1981) suggested that the preoccupation with language shown by most social and behavioral scientists has left nonlinguistic vocalizations largely overlooked or disregarded. In addition, methodological problems, such as capturing and recording the fleeting acoustic signals, have been cited as contributing to the disregard for vocal emotional phenomena such as intensity and durational properties of emotional speech. This last reason is no longer acceptable due to the development and availability of contemporary spectrograph equipment and computerized software.

The current study examined several acoustic properties of productions of one short sentence by twelve English speaking adult male and female subjects. Half of the subjects had training in dramatic expression, and half had no such training. The expressions included happiness, surprise, sadness, fear, anger, and disgust (for selection of these six, see Ekman, 1973; Ekman, Friesen & Tomkins, 1971; Ekman, Levenson and Friesen, 1983). This study complemented Ekman's facial taxonomy by providing comparable comparisons of vocal correlates. From this data base, a taxonomy of emotional speech was developed from acoustic properties for the simulated vocal expressions.

This research is significant because a taxonomy of vocal expressions of emotion can help teach people about emotional speech, including pathological speech. In addition, these data can provide linkages to (1) practical management of emotional interactions in natural environments, as well as (2) future perceptual studies of emotional sensitivity.

Statement of the Problem

Current social and behavioral science research have lacked acoustic parameters for emotional speech. Although there was a potential to develop a taxonomy for emotional speech, little standardized research had been done to achieve this goal. The few acoustical studies available in the literature revealed scissures in terms of (1) range of emotions, (2) stimulus characteristics, (3) subject selection (number, gender, training), (4) linguistic carriers of emotion (vowels, or single words, or sentences), and (5) language spoken.

Due to these gaps in the acoustic correlates research, a number of questions remained unanswered. For example, is a vocal taxonomy for a range of emotions even a possibility? Do men and women vocalize differently in their expressions of emotion? Do actors and nonactors differ in their productions of vocal expressions of emotion? What are the characteristics, or patterns for acoustic properties of emotion conveyed through a brief sentence in the English language in comparison to vowels, words, or sentences spoken in foreign languages? If a taxonomy of vocal expressions of emotions is to be realized, then an holistic approach to studying acoustic properties of emotion must be taken.

The aim of this investigation was to (1) develop a taxonomy of vocal expressions of emotion; (2) discern any differences in males' and females' productions of vocal expressions of emotion; (3) detect any differences in the production of vocal emotional expressions for actors and nonactors; and (4) contribute to the body of knowledge regarding acoustic properties of emotional expressions in general.

Research Questions

1. Are the six categories of emotions acceptable as descriptive of the discrete emotions in comparison to the neutral tone of voice? Do the simulated emotional expressions of happiness, surprise, sadness, fear, anger, and disgust conveyed via a short sentence differ among each other in duration and intensity? Finally, do the acoustic properties of sentence duration and mean intensity differ between a short sentence spoken in a neutral tone of voice and the vocal emotional expressions of: happiness, surprise,

sadness, fear, anger, and disgust?

2. What differences exist between males and females in their productions of vocal expressions of emotion for the variables of sentence duration and mean intensity?
3. What differences exist between actors and nonactors in their productions of vocal emotional expressions for the variables of sentence duration and mean intensity?

Operational Definitions

The following terms were defined for the purposes of this research:

Acoustic properties: Those elements of the speech signal, including temporal (measured in seconds) and intensity (measured in decibels) features, that are perceived by the listener as duration and loudness respectively. The acoustic properties of emotional speech have been variously known as "suprasegmental" aspects of speech, "vocalics," "paralanguage," speech "prosody," and "nonlinguistic vocalizations."

Simulated emotional expression: A production of one of the tones of voice selected for study in this research, including happiness, surprise, fear, sadness, anger, disgust, and neutrality. Each subject was asked to vocalize each expression based on the recall of a prior experience that evoked the particular emotion.

Assumptions

The following assumptions were made in this study:

1. Male and female, actor and nonactor subjects, who were asked to simulate six expressions of emotion in addition to a neutral tone of voice, produced expressions uniquely their own.
2. These expressions, evoked from the recollections of previous emotional experiences, would parallel their productions of vocal emotional expressions in natural situations.
3. Each of the vocal expressions of emotion produced would exhibit acoustic properties unique to its emotion type in comparison to a neutral tone of voice.
4. Males and females should produce vocal expressions of emotion differentially based on socialization processes and/or differences in the vocal mechanisms.
5. Actors and nonactors should produce vocal expressions of emotion differentially based on training and experience.

CHAPTER 2

REVIEW OF THE LITERATURE

"Nonverbal vocalizations," "vocalics," "prosody," nonlinguistic vocalizations," "suprasegmentals," "paralanguage," "vocal emotional nuances," "emotional speech," "emotional intonation," "vocal affect displays," "vocal expressions of emotion"--all of the preceding words and phrases have been used to signify the voice of emotion--those aspects of speech that indicate the emotional content carried in the parlance of everyday interaction.

With the clutter of terminology found in the literature on the topic of vocal expressions of emotion, it is important that this chapter be prefaced with a definition for "the voice of emotion." Soskin and Kauffman (1961, p. 73) provided a good start when they wrote:

> Essential to experimentation is the fact that normal human speech consists of two simultaneous sets of cues--the articulated sound patterns forming words, phrases, and sentences and the discriminable qualitative features of voice itself. The former set of cues constitute a rapidly changing succession of stimuli which present semantically meaningful material. The latter, an amalgam of physical properties forming a relatively smoothly flowing, continuous signal,

> is the carrier upon which articulated sounds are imposed. And it is in this "carrier" that major cues to emotional disposition may reside.

It is this "amalgam of physical properties" that carry the vocal cues of emotion on which this literature review will focus. For this study, the review of related literature will emphasize the following topic areas:

1. early studies of emotional expressions
2. linguistic and speech science studies
3. acoustical correlate studies

Early Studies of Emotional Expressions

Darwin

Early contributions to the study of the vocal expressions of emotion can be attributed to Charles Darwin. In his paragraph titled "The emission of sounds," Darwin (1872/1965, p. 83) wrote:

> With many kinds of animals, man included, the vocal organs are efficient in the highest degree as a means of expression. We have seen...that when the sensorium is strongly excited, the muscles of the body are generally thrown into violent action; and as a consequence, loud sounds are uttered, however silent the animal may generally be, and although the sounds may be of no use.

Darwin (1872/1965) provided further description into the production and function of sounds, ranging from observations of his young son's whine of "obstinate determination" (p. 86) to the phylogenetic basis for the "musical character" of the voice when used under any strong emotion (p. 87) to descriptions of sounds of several animals in states of pain, anger, fright, pleasure, as well as for courting rituals, cries for attention, and threats in self-defense (pp. 88-94).

Darwin was also aware of the mutual influence the vocal and facial mechanisms had on each other during the production of expressive behaviors. Darwin (1872/1965, p. 92) provided examples of this interdependence when he wrote:

> If, together with surprise, pain be felt, there is a tendency to contract all the muscles of the body, including those of the face, and the lips will then be drawn back; and this will perhaps account for the sound becoming higher and assuming the character of Ah! or Ach! As fear causes all the muscles of the body to tremble, the voice naturally becomes tremulous, and at the same time husky from the dryness of the mouth, owing to the salivary glands failing to act.

Following Charles Darwin's lead, and from a more recent phylogenetic perspective, Van Hooff provided another example of the relationship between facial expressive movements and vocal mechanisms. Van Hooff (1972, p. 212) hypothesized "that laughter and smiling could be conceived as displays with a different phylogenetic origin, that have converged to a considerable extent in Homo." In his comparative review of the phylogeny of smiling and laughter, especially with respect to data on chimpanzees and humans, Van Hooff (1972, p. 235) reported similarities in productions for the "silent bared-teeth," or smile, and the "relaxed open mouth," or laughter, displays. Relevant to the facial and vocal interaction, Van Hooff posited a two dimensional model to account for variations in lip and mouth posture and the presence of vocalization in which the ordinate portrayed the baring of teeth, and the abscissa portrayed the opening of the mouth and the resultant laughter vocalization (p. 234).

Studies of Facial Expressions of Emotion

Although Darwin placed equal importance on both the vocal and visual channels in his investigations of human and animal behavior, the visual channel in general, and facial expressions of emotion in particular, have received most of the research emphasis. Darwin (1872/1965, p. 93) was aware of the difficulties in studying vocal expressions when he wrote, "the whole subject of the differences of the sounds produced under different states of the mind is so obscure, that I have succeeded in throwing hardly any light on it." The neglect of the vocal channel has continued to the present. Scherer (1982, 1986) has suggested that the difficulty in obtaining, storing, and measuring the vocal signals that convey emotional content encouraged the study of facial expressions over vocal expressions of emotion. Whatever the reasons for this imbalance, facial expressions have been important for all studies of emotion. It is the study of facial expressions that has established the data base for a taxonomy of vocal expressions of emotion.

A major approach in facial expressive research has been that of correlating particular facial muscle patterns with discrete expressions of emotion. The early beginnings for this approach can be attributed to Bell and Henle (in Darwin (1872/1965, pp. 1 - 26). Darwin credited Bell and Henle for their comprehensive descriptions of human dermal facial muscles for various emotions, and provided illustrations of their work.

These early investigations were supported by two contemporary studies of blind and sighted children (Fulcher 1942; Thompson, 1941). Thompson found similarities in spontaneous facial patterns of subjects ranging in age from 7 weeks to 13 years for smiling, laughing, crying, and anger. Fulcher also found parallels in posed expressions of happiness, sadness, fear, and anger for blind and sighted subjects, who ranged from 4 to 21 years of age. Both authors reported maturational effects and, although results indicated differences in muscular movements between the two groups, the differences were in degree of movement, rather than kind.

More recently, evidence for an association between specific muscle configurations and discrete emotions resulted from the development of a measurement tool for facial behavior (Ekman, Friesen, and Tomkins, 1971). The Facial Affect Scoring Technique (FAST) (see Ekman, 1977 for a review of this tool) utilized pictures of each of three areas of the face: 1) brows and forehead area; 2) eyes/lids; and 3) lower face to define movements within each of the three categories that theoretically distinguished among six emotions, including happiness, sadness, surprise, fear, anger, and disgust. The development of the FAST and its diligent application by Ekman and colleagues have provided quantitative studies leading to a taxonomy of facial expressions of emotion.

Summary

Darwin and other nineteenth century investigators, such as Bell and Henle, provided the underpinnings for a taxonomy of emotions. Contemporary comparative phylogenetic studies, which stemmed from Darwin's work, suggested a relationship between facial and vocal mechanisms in the production of emotions. Although Darwin emphasized the importance of both the vocal and visual channels in the expression of emotions, the vocal channel has not received as much research attention as have facial studies. Nevertheless, investigations of associations between specific facial muscle patterns and discrete emotions, such as those of the blind and sighted children, in addition to those of Ekman and his colleagues, have provided (1) an heuristic approach to the study of vocal expressions of emotion, and (2) a foundation for a taxonomy of emotions.

Linguistic and Speech Science Studies

Suprasegmentals

Most linguistic and speech science investigations of acoustical characteristics of speech have been concerned with "segmental" aspects of speech; i.e., cues important for phoneme identification (Borden and Harris, 1984; Pickett, 1980). There is also a literature that describes "suprasegmental" characteristics of speech (for a comprehensive review, see Lehiste, 1970), a category that is particularly important for the field of emotional expression.

Suprasegmentals include quantitative, tonal, and stress features of the speech signal (see Lehiste, 1970, p. 4 for framework). Quantitative features include the time parameters of the acoustic signal, are perceived as duration of the speech signal, and result in tempo of the signal at the sentence level. Tonal features include fundamental frequency (Fo), are perceived as pitch of the voice, and function as intonation at the sentence level of the speech signal. Stress features include dimensions of intensity and amplitude, are perceived as loudness of and emphasis on the speech signal, and provide for syntactic and semantic stress at the word/sentence level.

Broad (1973) indicated that, physiologically, the quantity features, or phonetic segment duration, are determined, for the most part, by supraglottal articulator rates of movement; tonal features, or fundamental frequency of the voice, by the rate of vocal fold vibration; and vocal intensity, which is dependent in part on the intensity of the laryngeal voice source. Broad (1973, pp. 147 - 148) reported of stress features that "Stress differences are in part made in the larynx, though other variables such as vowel duration and vowel quality significantly contribute to syllable stress. Stressed vowels tend to have higher fundamental frequencies, greater durations, and higher acoustic intensities than their unstressed counterparts."

Probably one of the best descriptions of the structure and function of the suprasegmental aspects of speech production was provided by Minifie (1973, p. 281):

> Changes in the intonational patterns of the voice (melody of the voice), changes in linguistic stress (relative emphasis given to syllable within an utterance), and changes in the durational characteristics of utterances (including pausal patterns, tempo, and rate of syllable utterance) all assist in providing vocal variety and contribute to the meaningfulness of the message generated. These changes occur at the suprasegmental level, that is, they occur across a number of phonemes. The regulation of the rate of utterance is primarily controlled by the number and extent of the pauses distributed throughout the discourse...When changing the emotionality of the message, changes in all of the suprasegmental parameters interplay to provide the proper emotional "tone" for the message.

This passage from Minifie (1973) is representative of writings by most linguists and speech scientists. Suprasegmentals are recognized as playing a major role in emotional speech, yet little attention is given to the suprasegmental aspects of speech involved in emotional expression.

Prosody

Suprasegmentals have been of interest to linguists and speech scientists for their lexical and syntactic, rather than their emotional, functions. Some of the linguistic studies, however, have provided information about the ways in which meaning is conveyed via the speech signal--information that is relevant to the study of vocal expressions of emotion. These studies have generally been investigated for their "prosodic," and "intonational" contributions to meaningful speech.

Pickett (1980 p. 80) described prosody as "the general name for the rhythmic and tonal features of speech." Pickett added that since prosodic features generally extended over more than one phoneme segment, they were said to be "suprasegmental" (p. 80). Prosodic studies are a subgroup of suprasegmental aspects of speech, that focus on variations of fundamental frequency, intensity, and duration of the speech signal.

Lieberman (1974) provided a comprehensive review of the study of prosodic features. Although most of the work is devoted to the role of prosody in linguistic studies, Lieberman pointed out that the prosodic aspects of the speech signal, such as vocal intonation, can convey the emotional state of the speaker. Lieberman (1974,

p. 2421) cautioned, however, that these "paralinguistic" cues carried in the prosodic features were, to a degree, arbitrary due to other influencing factors, such as context, culture, and social convention. Lyons (1972, p. 53) pointed out that paralinguistic features differ from prosodic features in that paralinguistics are not as closely integrated with the grammatical structure of an utterance.

Intonation

A major research approach for studying the suprasegmental aspects of speech has been that of intonation. Intonations are measured as fluctuations in fundamental frequency alone, or in combination with variations in amplitude, across the speech signal. Studies of intonation most relevant to emotion included those by Denes and Milton-Williams (1962), Dittmann and Wynne (1961), and Lieberman (1965).

In their investigations of intonation contours for monosyllabic utterance types, such as doubt, emphatic expression, confirmation, and question, Denes and Milton-Williams (1962, p. 1) reported that "Comparisons of the acoustic characteristics of utterances and of the correct recognition by listeners of the intonation classes showed that fundamental frequency, intensity and duration formed a complex pattern of cues: the fundamental frequency often played the dominant part, but in numerous cases recognition was strongly influenced by other characteristics." Some of these other characteristics included sentence structure and context. These authors also found marked similarities between the fundamental frequency and intensity variations with time for many intonation categories--findings that could speak to a taxonomy for statement types.

In a perceptual study using electronically manipulated and linguistically preserved non-emotional and emotional utterances, Lieberman (1965, p. 54) showed that sentence intonation "can be predicted if one considers three sets of factors: (1) the physiological constraints imposed by the human respiratory system, (2) the emotional state of the speaker, and (3) the ultimate recoverability of the Deep Phrase Marker that underlies the final phonological shape of the sentence." Lieberman concludes with the idea that intonation is perceived as an interaction matrix of fundamental frequency and amplitude variations as functions of time.

Linguistic Analysis of Emotion

A final offering from the speech science literature was that of an analysis of emotion in interviews that used linguistic coding techniques. Dittmann and Wynne (1961) coded linguistic phenomena for "junctures," or clause separations, "stress," or accents on syllables of multisyllabic words, and "pitch," or rise and fall of the voice. The paralinguistic phenomena were coded for "vocal characterizers," such as laughing or crying, "vocal segregates," such as "um," "hmm," or "huh?," and "vocal qualifiers," such as extra increase or decrease in loudness, pitch, and duration. Dittmann and Wynne (1961, p. 203) indicated that:

> the Linguistic patterns (juncture, stress, and pitch) can be described reliably with presently available coding techniques, but that these aspects of speech probably have little psychological relevance. By contrast, the Paralinguistic phenomena (vocalizations, voice quality, and voice set) presumably have higher psychological relevance, but cannot be coded reliably. Our explanation for these findings is that the methods developed in traditional linguistic analysis may not be applicable to the analysis of emotional expression, not because of deficiencies in the field of linguistics, but because of fundamental differences in the nature of language and emotional expression.

Summary

Although most linguists and speech scientists emphasize syntactical and lexical investigations in the study of suprasegmental parameters of speech, some studies have addressed the means by which emotional meaning and attitudes are conveyed. This literature indicates that the suprasegmental features used in the lexical and syntactic research carry most, if not all, of the emotional properties conveyed in everyday speech. These properties include fundamental frequency, amplitude/intensity, rhythm, spectral, and other temporal characteristics. Therefore, linguists and speech scientists have provided the baseline for emotional speech studies in terms of the vocal properties that need to be measured in the voicing of emotion.

Acoustic Correlate Studies

Research Tools

Global studies of the acoustic dimensions of emotion have described characteristics related to speech spectrum, fundamental frequency (Fo), amplitude, and temporal aspects of the speech signal. Research tools in speech science that have been used to test these dimensions have included (1) the oscilloscope, (2) the speech sound spectrograph, (3) spectral analysis, and (4) the laryngograph (see Borden and Harris, 1984 for a thorough review of these instruments).

The oscilloscope is, essentially, a cathode ray tube that displays the magnitude of an electrical signal as a function of time, and provides for high amplification of weak signals. The amplitude of the signal is measured on the vertical, or Y axis, and time is measured on the horizontal, or X axis. Hard copies of the data can be produced with polaroid snapshots, or through the use of a graphic level recorder. Oscilloscopes can be used to measure signal amplitude, duration, and to establish the fundamental frequency of complex periodic waveforms such as vowels.

The development of the speech sound spectrograph in the early 1940s revolutionized speech science studies. This instrument is used to produce a hard copy of a signal with frequency on the Y axis, duration on the X axis, and intensity on the Z axis, or grey scale, as relative darkness. Most speech spectrographs provide for the selection of two bandwidth settings. A narrow band setting (for example, a 45 Hertz (Hz) bandwidth for a frequency range of 8 kHz) is of use for tracking fundamental frequency due to better frequency resolution. A wide band setting (for example, a 300 Hz bandwidth for an 8 kHz frequency range) is of use for obtaining details of formants (vocal tract resonances that make up the vowel sounds) due to the enhanced time resolution.

Most speech spectrographs have optional functions, which include productions of amplitude contours and waveforms. The spectrographic productions are similar to those visualized on the oscilloscope, but with the added advantage of producing hard copies of the signals. Amplitude contours are of use in studies of intensity and/or the placement of stress on running speech. Waveforms are of value in voice onset time (VOT) studies, and for measuring total duration of a speech signal.

Spectral analysis provides the researcher with information about the distribution of energy at various frequencies by separating the speech signal into components through the use of a bank of filters. The changing spectra of complex signals, such as running speech, can be displayed through the use of a real time spectral analyzer. This instrument is useful for studying speech sounds, such as vowels and consonants, at different frequencies.

The laryngograph is used to measure impedance across the vocal folds. Two small electrodes are placed on either side of the larynx. Vocal fold movement provides measures of the relative conductance or impedance between the two electrodes, which indicate vocal fold contact for each vibratory cycle. This instrument is used to record fundamental frequency over time.

Physiological Stress Studies

One research approach to uncovering the acoustic parameters of emotion, which utilized various of the speech science instruments described above, is that of effects of physiological stress on the voice. Authors for this type of research include (Friedhoff et al., 1964; Hecker et al., 1968; Simonov and Frolov, 1973).

Friedhoff et al. (1964) recorded changes in the human voice via spectral analysis in combination with measures of blood pressure and skin resistance. The authors devised a number of situations that were stress provoking, such as requests to lie. Friedhoff and colleagues found that the voice appeared to contain information in intensity variations, changes in emphasis and in register that served as cues for reflecting changes in emotional states. The authors indicated that the voice revealed changes in emotions more directly than that of blood pressure or skin resistance.

In another task-induced stress study in which subjects were required to add numbers under time constraints, Hecker et al. (1968) obtained verbal data from ten subjects while they were either under stress or relaxed. Responses were analyzed for amplitude, fundamental frequency, and with comparisons of spectrograms. Results indicated that task-induced stress produced changes predominantly in the amplitude, frequency, and waveform of the glottal pulses. Hecker et al. also reported that although manifestations of stress showed considerable individual differences, test responses of most subjects showed some consistent effects.

A study of Russian cosmonauts presented graphed results of voice frequencies related to emotional stress and states of attention, which were recorded during aviation and space flights (Simonov and Frolov, 1973, p. 257). Vowel formant structures of single words were studied with a one-third-octave spectral analyzer, which showed an augmentation in the first formant range with an increase in emotional stress. In the attention state, results indicated that speech signal parameters may be characterized by a decrease in standard deviation, i.e. stabilization, of spectral components, and a drop in the probability of formant shift in comparison to the resting state.

Summary

The analyses of stress/attention-related effects on the voice indicate discernable changes in fundamental frequency, intensity, temporal patterns, and/or changes in vowel formant structure. Although these speech related studies of general autonomic nervous system arousal have held promise as indicators of emotional states, the larger question remains to be answered. That question is--are there vocal correlates for discrete emotions?

Emotion Related Acoustical Studies--American English

Discrete emotions investigated in the American English acoustical studies included joy, terror, grief, and contempt (Coleman and Williams, 1979), anger, fear, contempt, grief, and indifference (Fairbanks and Hoaglin, 1941; Fairbanks and Pronvost, 1938), happiness, sadness, and ordinary tone of voice (Skinner, 1935), and anger, fear, sorrow, and neutral tone of voice (Williams and Stevens, 1972). No measures have been reported for the expressions of surprise and disgust. Additionally, happiness (vs. joy/elation) and anger (vs. irritation or rage) were not clearly defined (see Scherer, 1986 for comments on these distinctions).

Linguistic carrier and subject selection

differed for these studies. Coleman and Williams (1979) studied 3 females and 10 males reading a portion of a nonsense passage. Fairbanks and Hoaglin (1941) and Fairbanks and Pronvost (1938), in companion research, studied 6 males reading a standard passage. Skinner (1935), studied 1 male and 1 female recording a vowel in response to mood induction. Williams and Stevens (1972), studied 3 males reading dialogue from a short scenario, together with a real-life situation (the Hindenburg crash).

Studies of stimulus characteristics in the production of vocal expressions of emotion focused on temporal (Coleman and Williams, 1979; Fairbanks and Hoaglin, 1941; Williams and Stevens, 1972), intensity (Coleman and Williams, 1979; Skinner, 1935) and fundamental frequency aspects of the speech signal (Coleman and Williams, 1979; Fairbanks and Pronvost, 1938; Skinner, 1935; Williams and Stevens, 1972). As to temporal aspects, Coleman and Williams (1979, p. 9) reported mean durations for emotions by means of a Honeywell Visicorder Oscillograph. Grief showed the longest duration in seconds, followed by contempt, joy, and terror. For speech rate, the fastest average word per minute rate was terror, followed by joy, then contempt, and grief (p. 79). Coleman and Williams (1979, p. 77) indicated that differences in total and phonation times were due to pauses between words and phrases; not to changes in the word lengths themselves.

Using sound-wave photography, Fairbanks and Hoaglin (1941, p. 86) showed contempt to be the longest in duration, followed by grief, then anger, fear, and indifference. Speech rate showed indifference to have the fastest word per minute rate; then fear, anger, grief, and contempt. These authors also pointed out that pauses between words and phrases contributed to differences in total phonation time, rather than changes in word lengths. Williams and Stevens (1972) reported that sadness showed the longest duration with a marked decrease in speaking rate; then fear, then anger. Results were inconsistent in the syllable rates for fear and anger. Duration for the neutral tone of voice was usually shorter compared to the emotion conditions.

In the intensity domain, Coleman and Williams (1979, p. 80) used a graphic level recorder to obtain "average peak SPL [Sound Pressure Level] values." The terror condition showed the greatest amplitude, followed by joy, contempt, and grief. Skinner (1935, p. 92) recorded vocal intensity in response to an evocation of mood by means of an oscillograph. Skinner reported that force of the voice in response to happiness is greater, while in response to sadness is lesser than in an ordinary tone of voice. No measures of intensity variability or intensity range have been reported in the literature.

Studies have also reported fundamental frequency changes for American English productions of several emotions. Coleman and Williams (1979, p. 78) provided rough estimates of overall fundamental frequency by counting consecutive waves at intervals throughout samples of oscillograph traces. Terror had the highest average fundamental frequency, followed by joy, contempt, and grief in that order. Using phono-photographic techniques from phonograph recordings, Fairbanks and Pronvost (1938, p. 382), provided median fundamental frequencies for five simulated emotional conditions. Fear showed the highest median fundamental frequency, then anger, grief, contempt, disgust, and indifference.

Williams and Stevens (1972) utilized narrow-band spectrograms to determine median fundamental frequency, and frequency range. The most consistent acoustic manifestation for anger was a high fundamental frequency that persisted throughout a breath group. This frequency tended to be at least half an octave above the fundamental for the neutral tone of voice, and the range for anger was greater than for neutral. Fear showed an elevated fundamental frequency and range in comparison to neutrality, but did not reach those seen in the anger condition. Sorrow generally showed a reduced fundamental frequency and range, assuming that the speaker's normal fundamental frequency and frequency range were known.

Skinner (1935), using oscillographs of 9 males and 10 females, reported that happiness was characterized by a fundamental frequency considerably higher than that of an ordinary tone of voice. However, the average fundamental frequency produced in response to stimuli for sadness approximated that of the ordinary tone of voice, whether male or female. Skinner (1935, p. 105) reports a corollary:

> if the subject has an ordinary tone of low frequency, his sad state is expressed with one definitely higher; if he has an ordinary tone of high frequency, his sad state is expressed with one decidedly lower; while if he has an ordinary tone of medium or average frequency, his sad state is expressed with one approximately the same. Female subjects exhibit a similar tendency.

Gender and Speaker Training

Of the studies cited, only Skinner (1935) compared differences between productions by male and female speakers. Skinner reported that that males used three times as much force to match women in vocal intensity, which was attributed to the lower fundamental frequency of the male voice. Additionally, all studies used trained actors as subjects. No acoustical production studies have been reported that have compared "trained" actors with speakers who have had no training in acting or speaking performance.

Emotion Related Acoustical Studies--Foreign

Cross-cultural studies of acoustic correlates of emotion are pertinent to this review of literature. Although languages may show differences in syntax and lexicon, the acoustical properties that carry emotional information remain the same across cultures. These properties include fundamental frequency, intensity, and temporal dimensions of the acoustic speech signal. The cross-cultural studies include expressions of emotion in French (Fonagy, 1978), Dutch (Kaiser, 1962), and Russian (Kotlyar and Morozov, 1976).

French

Fonagy (1978) used a laryngograph to study the fundamental frequency changes during emotive passages in French produced by a professional actress. Results indicated that joy was characterized as having a high fundamental frequency and large melodic interval, sorrow as having low average fundamental frequency and narrow interval, and fear as having a mid-high frequency and reduced interval. Fonagy (1978, p.36) also reported that the functions of some contrasting emotive attitudes, such as anger and joy, overlapped. For example, the fundamental frequency for repressed anger (hatred) came closer to tenderness than anger, and approximated sorrow. Further, some emotive attitudes displayed typical melodic configurations. For example, the regularity of the sudden rise of fundamental frequency in stressed syllables differentiated anger from an erratically varying frequency pattern seen in joy.

Dutch

In the Dutch study, Kaiser (1962) used spectrographic analyses of three vowels spoken in different emotional attitudes by student speakers. Durational aspects of the vowels showed sadness to be the longest, then cheerfulness, enthusiasm, and disgust, which were the same in durational value; lastly, kindness and grimness. The durations for men were slightly higher values than for women (p. 305). Intensity (Kaiser, 1962,

p. 309) was greatest for enthusiasm, followed by cheerfulness, disgust, grimness, kindness, then sadness. Both males and females showed similar values in intensity measures.

Kaiser (1962, p. 306) also reported fundamental frequency characteristics for male and female speakers. Three positive affects, or emotions--cheerfulness, enthusiams and kindness--first showed a rise and then a drop in fundamental frequency. This biphasic change was negligible in sadness and disgust. Grimness showed a moderate rise. Females tended to show a rise in frequency toward the end of the kindness condition, which was sometimes interpreted as a question. Kaiser indicated, however, that despite individual differences, characteristic fundamental frequency patterns were indicative of each of the six emotional attitudes.

Russian

A Russian study (Kotlyar and Morozov, 1976) provided acoustic analysis of vocal phrases sung by eleven classically trained singers. Their emotional shadings included, joy, anger, sorrow, fear, and neutrality. Included in the results were reports of temporal properties, including total phrase duration, syllable duration, and coefficient of variation of syllable durations. Intensity properties that contributed to the emotional shadings included average sound pressure level of a syllable within a phrase, the coefficient of variation of the intensity of syllables within a phrase, and the rise and decay time of the sound pressure level in a syllable.

According to Kotlyar and Morozov (1976, p. 209), sorrow was of longest total duration, followed by joy, neutrality, anger, and fear. Average syllable durations revealed sorrow to be the longest, followed by neutrality, joy, anger, then fear. Coefficient of variations for syllable durations showed values of 60.8% for fear, 59.8% for joy, 58.0% for anger, 54.6% for sorrow, and 44.5% for neutrality (p. 210). The minimum value was characteristic of phrases in the neutral state, while various emotional shadings had a much greater coefficient of variation for duration.

The average vocal intensity for emotions showed anger to be most intense, then joy, and sorrow. Fear and neutrality followed sorrow and were equivalent in intensity (Kotlyar and Morozov, 1976, p. 209). The coefficient of variation was highest for sorrow (67.5%) and fear (70.1%), decreased for anger (46.7%) and neutrality (46.1%), and showed an intermediate value (56.4%) for joy (p. 210). Kotlyar and Morozov (1976, p. 210) also reported that the rise and decay times of the sound pressure levels were well correlated with each other except in the neutral condition. The maximum rise and decay time was seen for the expression of sorrow, and the minimum was seen in anger and fear. Neutrality revealed a large rise and small decay time.

Summary for Acoustical Studies

American English. The research on acoustic correlates of emotions shows a number of gaps. There have been a limited range of emotions studied, and few numbers and types of subjects. Further, the linguistic carriers used, such as words and phrases, and the stimulus characteristics measured as indicators of discrete emotions, such as intensity, fundamental frequency and timing of the speech signal, have received peripatetic attention in the literature. Despite these methodological differences, studies appear to show similar qualitative findings for some emotions. For example, most studies have shown a prolonged duration and decreased intensity for sadness. Findings such as the ones for sadness, with empirical validation, hold promise for a taxonomy of vocal expressions of emotion. Of the American English studies reviewed, one (Skinner, 1935) provided qualitative differences in the intensity and fundamental frequency of male and female speakers. No studies validated these findings, nor are there studies in the literature that indicate quantitative differences for gender. Actor and nonactor differences have not been reported. The acoustical studies cited used trained actors, yet findings were generalized to the population-at-large.

Foreign Language. Variances in design and methodology hold true for cross-cultural acoustical correlate studies. For example, of the three studies reviewed, one used an actress, one trained singers, and one used students. Again, despite differences in subject selection, linguistic carriers and, most important, language spoken, studies showed findings similar to the American English descriptions. For example, both the Dutch and Russian studies showed sadness to have the longest duration and lowest intensity. One study (Kaiser, 1972) suggested a difference for duration, but similarities for intensity in productions of emotion by males and females. Actor and nonactor differences have not been reported in the cross-cultural literature. The conformity of reports, such as for sadness, warrant a study investigating a full range of emotions as a bridge toward the development of a general taxonomy of vocal expressive behavior.

CHAPTER 3

METHODS

Subjects

Gender

The subjects recruited for this study consisted of twelve Caucasian adults, ranging in age from 20 to 47 years. Six subjects were males and six were females. All subjects were native speakers of American English, all were intact neurologically and, by self report, had no history of speech and/or auditory deficits. Prior to participating in the experiment, all subjects completed a demographics form (Appendix A) designed for this study.

Actors and Nonactors

Six adults (3 males and 3 females) with training in dramatic expression, who served as "actors" for this study, were recruited from Old Tucson Movie Studio, Tucson, Arizona, and from the Departments of Media Arts and Communication, University of Arizona. Six adults (3 males and 3 females) with no dramatic training were recruited from the University of Arizona population, and served as "nonactor" subjects.

Design and Procedures

Emotion Types

The simulated expressions of emotion investigated in this study included; happiness, surprise, sadness, fear, anger, and disgust. A neutral tone of voice was also included so that each of the emotion types could be compared to this baseline measure. Simulated expressions were selected because they provided an initial advantage, for a preliminary study, in permitting control of such variables as phonetic content and syntactic form.

Linguistic Carrier of Emotion Types

Subjects were asked to produce each of the emotion types in one "semantically relevant" emotional sentence ("Of course I love you"), and one semantically irrelevant "neutral" sentence ("The horse tries one food"). Data on this latter sentence were obtained for a companion study concerning the influence of semantic content on vocal expressive cues, and will not be described here.

Recording Instrumentation

All recording was done in an anechoic chamber located in the Department of Psychology, University of

Arizona. Recordings were made using an AKG Model C451E high fidelity microphone, covered with a foam pop filter that reduced noise from plosive sound stimuli. The subject's face was positioned approximately two feet from the microphone throughout the recording procedure.

The microphone was connected to a General Radio U.S.A. Model 1565-B Sound Level Meter, providing the testor with an approximate sound range in decibels (dB) for calibration purposes. Communication between the testor and the subject was achieved by means of an intercom connection from the chamber to the recording lab.

The emotional expressions were recorded and digitized by means of a Nakamichi Model BMP100 Pulse Code Modulator (PCM) and stored on the video portion of videotape via a Fisher Model 205A Video Cassette Recorder (VCR). This methodology provided high quality recordings of acoustical stimuli with high signal-to-noise ratio.

Testing Procedure

Preliminary Instructions

Approximately one week prior to testing, each subject was given a list of the expressions and sentence conditions. Subjects were instructed to practice both sentences in the six expressions. Subjects were further instructed to vocalize each expression based on the recall of a prior experience that had evoked the particular emotion. In addition to the six emotional expressions, subjects were also requested to practice speaking both sentences in a neutral tone of voice--a tone devoid of any emotional expression.

Additional instructions were provided to the subjects as to further distinctions in their productions of anger and happiness. Subjects were asked to recall a situation that evoked "hot" anger, bordering on rage, rather than a feeling of irritation, or "cold," controlled anger. Similarly, subjects were asked to recall a personal experience that evoked the feeling of happiness, rather than "joy" or "enthusiasm" (see Scherer, 1986 for these distinctions in the literature).

Introduction to Experimental Procedures

Each subject attended one experimental session that lasted approximately one hour. After being comfortably seated, the subject was oriented to the chamber and experimental procedures. Initially, subjects were asked to record a brief statement of history that included; name, age, native language, and places lived until age 10 years. Following this statement, subjects recorded a series of speech syllables, including "ba," "da," "ga," "pa," "ta," and "ka." Both these procedures allowed the subject to adapt to the surroundings and recording equipment.

Experimental Procedures

Recording of expressions consisted of two blocks; (1) a practice block, and (2) the experimental block. Within each block were two trials for the sentence, "Of course I love you." The first trial consisted of the sentence spoken in a neutral tone of voice. The second trial consisted of the sentence spoken with an expression of emotion. Order of emotions was fixed for both blocks; happiness, surprise, sadness, fear, anger, and disgust.

Evocation and Production of Emotion

Subjects were given sets of twelve 4 x 6 inch note cards with the trial and sentence listed, which were used as prompting devices during the experiment. For example, the first card would show--

(Neutral)

Of course I love you

the second--

(Happiness)

Of course I love you

the third--

(Neutral)

Of course I love you

the fourth--

(Surprise)

Of course I love you

and so on. The subject was encouraged to pause for 2 to 3 minutes between each trial for both blocks in order to recall and evoke the particular emotion that was produced vocally.

Validation of Emotion Types

Following recording, all trials for both the practice and experimental blocks produced by the subject were replayed for the subject to confirm, on hearing each trial, the emotion typed expressed for each block. All subjects confirmed their emotional expressions. There were no problems with, or during the procedures. Following the validation procedure by the subject, the experimental session was concluded.

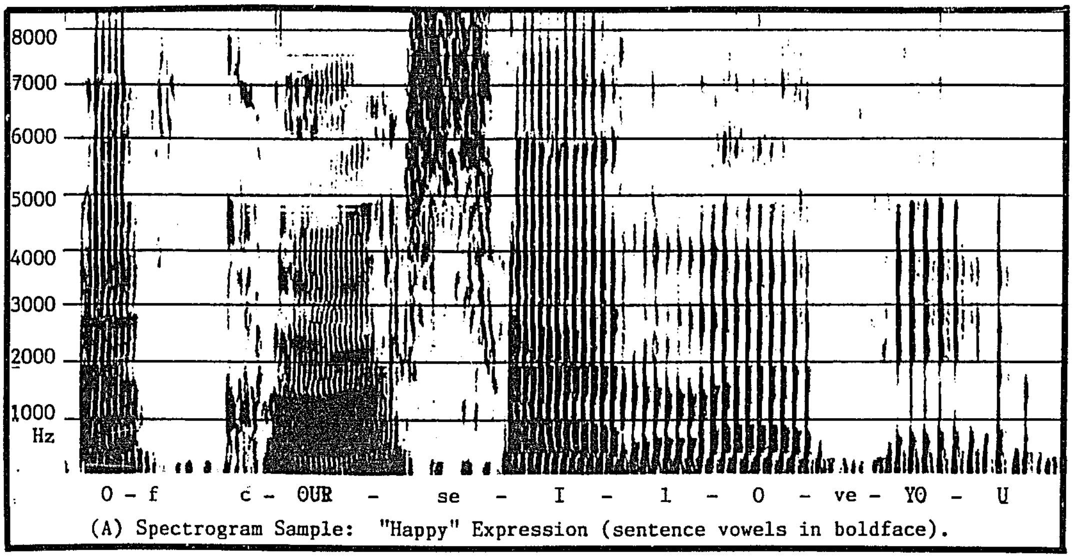
Acoustic Instrumentation

A Kay 7800 Spectrograph, which was connected with the PCM/VCR equipment by means of a phono plug, was used to produce wide-band spectrograms, waveforms, and amplitude contours for all expressions for the "Of course I love you" sentence condition (Figure 1 provides samples of each of these productions). For recordings of all voice prints, the frequency range on the spectrograph was set at 8kHz, which provided for storage of a speech signal with a maximum duration of 2.56 seconds. The filter bandwidth was set at 300Hz for a sampling rate of 25.6kHz. This standard wide-band setting provided for accurate timing resolution of the spectrographic productions.

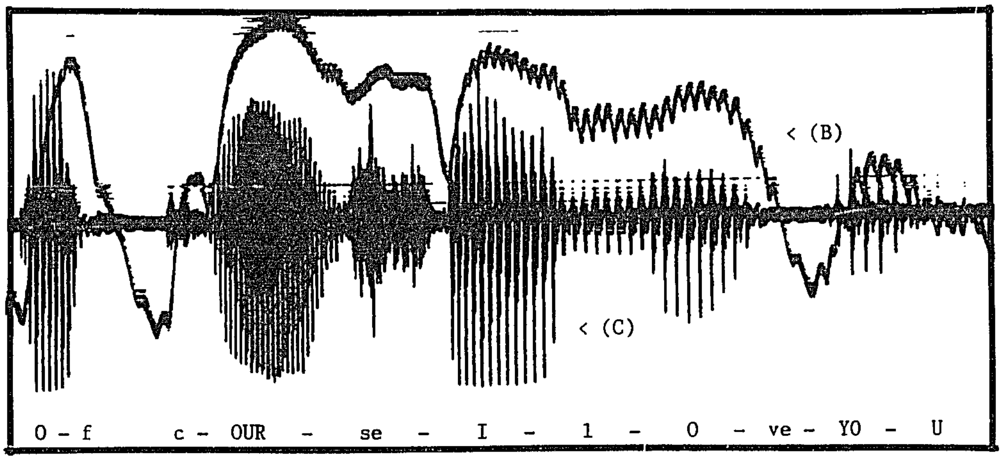
Acoustic Measures

Duration Values

The temporal dimension of the spectrograms, waveforms, and amplitude contours was measured in seconds across the X axis (1 inch = 100 milliseconds). Overall sentence duration was the temporal parameter investigated in this study. Sentence duration for the six emotion types and the six neutral conditions were measured from onset to end of sentence.



(A) Spectrogram Sample: Frequency on the Y-axis, time on the X-axis, intensity as grey scale. (Sentence vowels in boldface).



(B) Relative Amplitude Contour Sample: Amplitude on the Y-axis, time on the X-axis.

(C) Waveform Sample: Frequency on the Y-axis, time on the X-axis.

Figure 1. (A) Spectrogram, (B) Amplitude Contour, and (C) Waveform samples for the Sentence "Of Course I Love You" Expressed in a "Happy" Tone of Voice by a 23 Year Old Male Subject (Xerox Reduction to 50% Actual Size).

Intensity Values

The intensity values were based on calculations of the amplitude contours for each emotional expression and each neutral tone of voice. Amplitude was calibrated by means of a pure tone generator. A pure tone was fed into the Kay Spectrograph, and the height of the recorded deflection was used as the calibration standard (1 centimeter = 10 decibels (dB)) for measurements of amplitude. Mean intensity for each emotion type and each neutral tone of voice was obtained by calculating the average peak amplitude for the five syllables of each sentence (complete data sets for all subjects are shown in Appendix B).

Data were analyzed using ANOVA for duration and mean intensity in a four factor design with two two-level between subjects variables (gender and role) and two within subjects variables (condition and emotion). In addition, post hoc analyses using the Newman-Keuls test of significance were obtained for comparisons among emotions and between neutral to emotion conditions.

CHAPTER 4

RESULTS

Major results of this investigation addressed a taxonomy of vocal expressions of emotion based on duration and intensity measures. Emotional expression types included: happiness, surprise, sadness, fear, anger, and disgust. Related questions were:

1. What differences exist between males and females in their productions of vocal emotional expressions for duration and intensity?

2. What differences exist between actors and nonactors in their vocal expressions of emotion for the duration and intensity variables?

Sentence Duration

The duration data were analyzed in a four factor design with two two-level between subjects variables (role and gender) and two within subjects variables (conditions and emotions). The ANOVA for sentence duration is summarized in Table 1.

Table 1. Analysis of Variance for Overall Sentence Duration. R = Role (Actor/Non-actor); S = Sex (Male/Female); C = Conditions (6 Neutral Tones/ 6 Emotions); E = Emotions (Happiness, Surprise, Sadness, Fear, Anger, Disgust) (n = 12).

Source	df	Mean Square	F
Role	1	.07854	.41
Sex	1	1.22600	6.38*
RS	1	.18155	.95
Error	8	.19209	
Conditions	1	1.59666	14.37**
CR	1	.02158	.19
CS	1	.05495	.49
CRS	1	.66872	6.02*
Error	8	.11110	
Emotions	5	.02282	1.34
ER	5	.02427	1.43
ES	5	.03833	2.25
ERS	5	.00119	.07
Error	40	.01702	
CE	5	.02617	1.40
CER	5	.01846	.99
CES	5	.05995	3.21*
CERS	5	.00344	.18
Error	40	.01865	

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

Effects for Duration

Data for sentence duration showed a significant three-way interaction for Conditions X Role X Sex (F (1, 8) = 6.02, p <.05), and a significant three-way interaction for Conditions X Emotion X Sex (F (1, 8) = 3.21, p <.05), with significant main effects for Conditions (F (1, 8) = 14.37, p <.01) and Sex (F (1, 8) = 6.38, p <.05). Figure 2A plots ..e main effects for conditions, which indicated significantly longer durations for real emotions (M = 1.29 seconds) compared to neutral productions (M = 1.08 seconds). Figure 2B plots the main effects for sex, which showed significantly shorter durations for males (M = 1.22 seconds) compared to females (M = 1.37 seconds). There were no significant main effects for role or emotions.

Interaction of Condition X Role X Sex for Duration

Interpretation of the duration data was complicated by the significant three-way interaction of conditions (C) by role (R) by sex (S). To assist in interpretation of these data, mean scores for each level of the variables involved in this interaction were plotted in Figure 3. Results indicated that the relationship between conditions (neutral and emotion trials) and role

Figure 2A. Mean and Standard Error Results for Main Effects for Conditions (6 Neutral/6 Emotions) on Sentence Duration (n = 12).

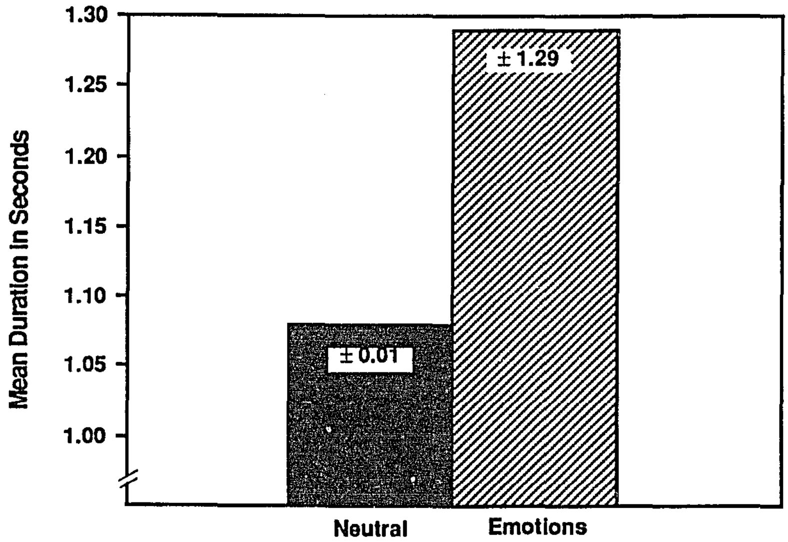


Figure 2B. Mean and Standard Error Results for Main Effects for Sex on Sentence Duration (n = 12).

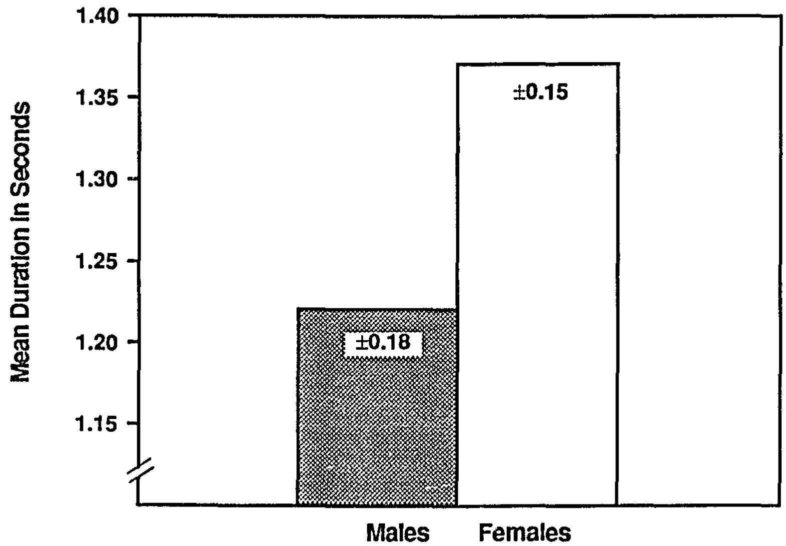
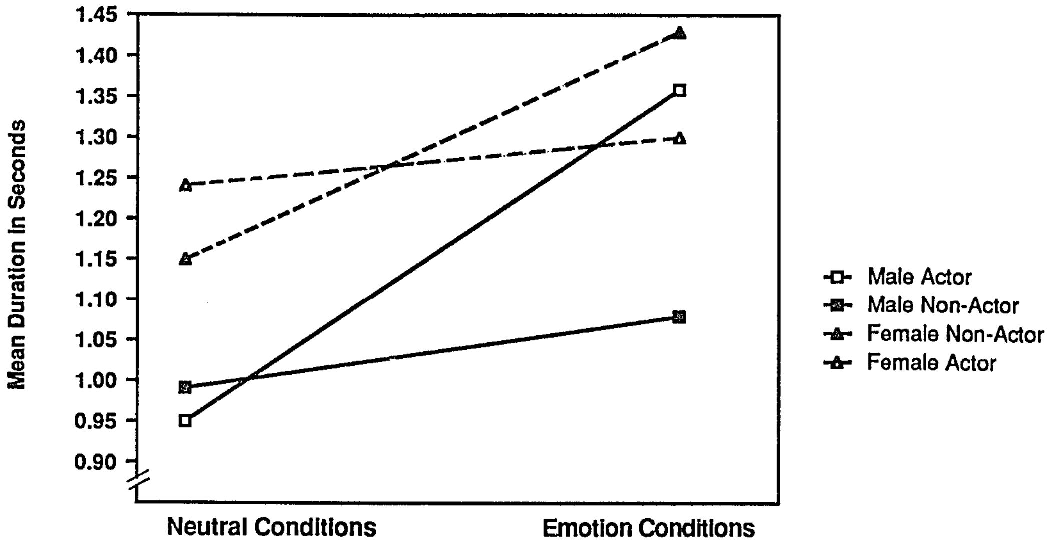


Figure 3. Conditions X Role X Sex Interaction on Duration. Shows Main Effects Also: Conditions and Sex (n = 6 Males/6 Females).

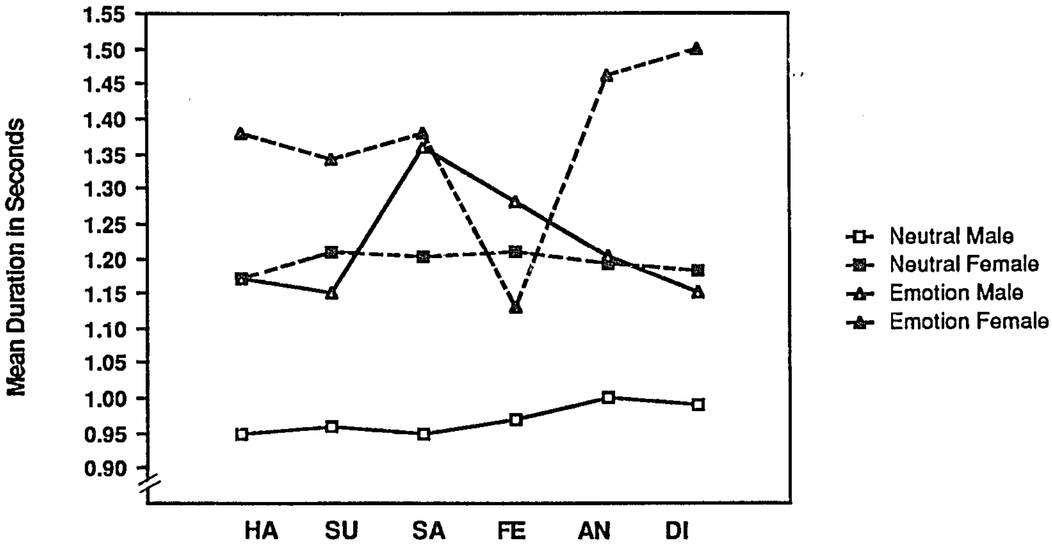


(actors and nonactors) differed according to sex of the speaker. Under the neutral conditions, male actors and nonactors showed similar durations to each other, but were shorter in duration compared to female actors and nonactors. However, under the emotion conditions, male actors' durations exceeded those of male nonactors and female actors, while the durations of female actors were less than those of female nonactors. Paired comparisons using the Newman-Keuls test supported these results, which showed (1) male actors, and male and female nonactors produced significantly shorter durations compared to female actors in the neutral conditions, (2) female nonactors, and male and female actors produced significantly longer durations compared to male nonactors in the emotion conditions, and (3) durations were longer with expressions of emotion than with neutral expressions ($p < .05$ in each case).

Interaction of Conditions X Emotions X Sex for Duration

Interpretation of the duration data was complicated further by a significant conditions (C) by emotions (E) by sex (S) interaction. To assist with interpretation of these data, mean scores for each level of the variables involved in this three-way interaction were plotted in Figure 4.

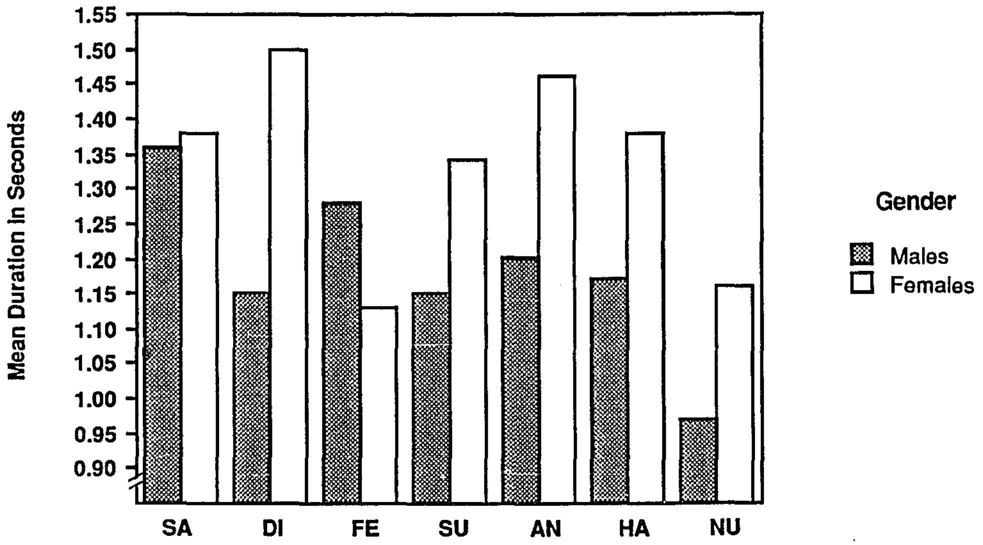
Figure 4. Conditions X Emotions X Sex Interaction. Shows Main Effects Also: Conditions and Sex. HA = Happiness, SU = Surprise, SA = Sadness, FE = Fear, AN = Anger, DI = Disgust (n = 6 Males/6 Females).



The relationship between conditions (differences in duration between the neutral to emotion trials) and among the six emotions were influenced by the sex of the speaker. In the neutral to emotion conditions, female speakers showed longer durations for all neutral trials compared with males' neutral trials. Four of the six durations produced by males in their emotion trials, however, came close to matching the corresponding neutral trials produced by female speakers, whereas females showed moderate to prolonged durations from their neutral counterparts for five of six emotions. In their productions of emotions, males and females showed comparable durations for sadness. However, durations for happiness, surprise, anger, and disgust showed longer durations when produced by female speakers. Duration for fear increased when produced by male speakers, but decreased when produced by female speakers. To analyze these differences further, post hoc analyses using the Newman-Keuls test were performed (1) within groups, and (2) between groups. In addition, mean scores for seven vocal expressions produced by males and females are graphed in Figure 5.

Post Hoc Analyses Within Groups--Males. Paired comparisons confirmed the results that males produced

Figure 5. Mean Durations for 6 Male and 6 Female Subjects for Seven Vocal Expressions. HA = Happiness, SU = Surprise, SA = Sadness, FE = Fear, AN = Anger, DI = Disgust, NU = Neutral Condition.*



* Sums shown for the neutral conditions were averaged across the six trials for male and female speakers.

significantly longer durations for their productions of sadness (*M* = 1.36 seconds) and fear (*M* = 1.28 seconds) compared to their neutral/sadness (*M* = 0.95 seconds) and neutral/fear (*M* = 0.97 seconds) trials (p <.05 in each case). There were no significant differences among the males' productions of discrete emotions, or their neutral productions.

Post Hoc Analyses Within Groups--Females. Paired comparisons supported the results that females produced longer durations for disgust (*M* = 1.50 seconds), anger (*M* = 1.46 seconds), and happiness (*M* = 1.38 seconds) compared to their neutral/disgust (*M* = 1.18 seconds), neutral anger (*M* = 1.19 seconds), and neutral happiness (*M* = 1.17 seconds) trials. Results also confirmed that females produced a duration for fear (*M* = 1.13 seconds) that was significantly shorter than those in their other emotion trials, including disgust (*M* = 1.50 seconds) anger (*M* = 1.46 seconds), sadness (*M* = 1.38 seconds), happiness (*M* = 1.38 seconds), and surprise (*M* = 1.34 seconds) (p <.05 in each case). There were no significant differences among the neutral trials produced by females.

Post Hoc Analyses Between Groups--Males and Females. Post hoc results using the Newman-Keuls test (p <.05 for all comparisons) supported the results that females' durations were significantly longer than males' durations for the neutral/surprise (female *M* = 1.21/male *M* = 0.96 seconds), neutral/sadness (female *M* = 1.20/male *M* = 0.95 seconds), and neutral/happiness (female *M* = 1.17/male *M* = 0.95 seconds). Post hoc analyses also supported the finding that while females produced significantly longer durations for five of six emotions compared with males' neutral durations, none of the males' durations for the emotion trials differed significantly from the females' neutral durations. The females' duration for fear was not significantly different from the males' neutral counterpart. Results also confirmed the finding that females' durations were significantly longer than males' durations for the emotional expressions of disgust (female *M* = 1.50/male *M* = 1.15 seconds) and anger (female *M* = 1.46/male *M* = 1.20 seconds).

Intensity

The intensity data were analyzed in a four factor design with two two-level between subjects variables (role and gender) and two within subjects variables (conditions and emotions). The ANOVA for intensity is summarized in Table 2.

Table 2. Analysis of Variance for Mean Intensity. R = Role (Actor/Non-actor); S = Sex (Male/Female); C = Conditions (6 Neutral Tones/6 Emotions); E = Emotions (Happiness, Surprise, Sadness, Fear, Anger, Disgust) (n = 12).

Source	df	Mean Square	F
Role	1	242.84028	1.38
Sex	1	91.84028	.52
RS	1	.34028	.00
Error	8	175.58333	
Conditions	1	122.84028	2.55
CR	1	207.84028	4.32
CS	1	126.56250	2.63
CRS	1	33.06250	.69
Error	8	48.09722	
Emotions	5	152.34583	17.09***
ER	5	18.92361	2.12
ES	5	9.15694	1.03
ERS	5	16.59028	1.86
Error	40	8.91667	
CE	5	103.49028	8.22***
CER	5	19.02361	1.51
CES	5	14.77917	1.17
CERS	5	23.61250	1.87
Error	40	12.59722	

*** $p < .001$

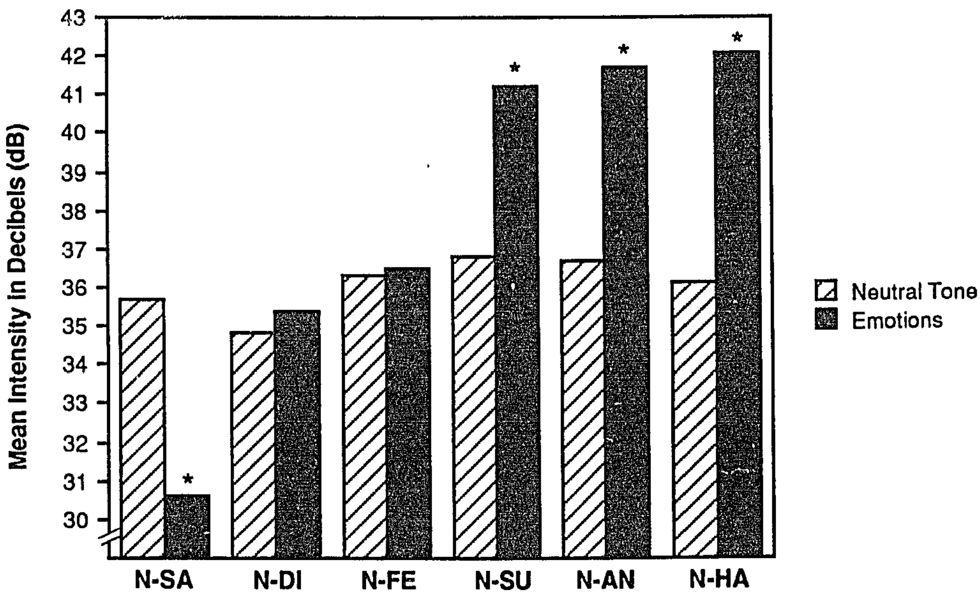
Effects for Intensity

Data analysis for mean intensity showed a significant two-way interaction for Conditions X Emotions (F (5, 40) = 8.22, p <.001), with a significant main effect for Emotions (F (5, 40) = 17.09, p <.001). There were no other interaction effects. There were no main effects for role, sex, or conditions.

To assist with the interpretation of these data, mean scores for the neutral and emotion conditions were plotted in Figure 5. Results indicated that intensity varied as a function of emotion type, and in relation to the neutral trials. Within emotions, mean intensity for sadness was lowest in comparison to all other emotions. Happiness, anger, and surprise were similar to each other in intensity, and of high intensity compared to fear, disgust, and sadness. Disgust and fear were similar to each other in intensity, moderately high compared to sadness, and moderately low compared to suprise, anger, and happiness. Paired comparisons using the Newman-Keuls test supported significant differences in mean intensity among most of the emotions. These results are provided in Table 3.

Intensity differed also between the emotion and neutral trials. Sadness showed a low intensity compared

Figure 6. Conditions X Emotions Interaction on Mean Intensity. Shows Main Effects Also: Emotions. N = Neutral, HA = Happiness, SU = Surprise, SA = Sadness, FE = Fear, AN = Anger, DI = Disgust (n = 12). (Table 3 Provides Significant Results Among Emotions).



* p <.05 for Emotions to Neutral Conditions

Table 3. Newman-Keuls Paired Comparisons Results for Mean Intensity Measures for Six Emotional Expressions. HA = Happiness, AN = Anger, SU = Surprise, FE = Fear, DI = Disgust, SA = Sadness, NU = Neutrality** (n = 12).

	HA (42.1)	AN (41.7)	SU (41.2)	FE (36.5)	NU (36.1)	DI (35.4)	SA (30.6)
HA (42.1)	--	0.4	0.1	5.6*	6.0*	6.7*	11.5*
AN (41.7)	--	--	0.5	5.2*	5.6*	6.3*	11.1*
SU (41.2)	--	--	--	4.7*	5.1*	5.8*	10.6*
FE (36.5)	--	--	--	--	0.4	1.1	5.9*
NU (36.1)	--	--	--	--	--	0.7	5.5*
DI (35.4)	--	--	--	--	--	--	4.8*
SA (30.6)							

* $p < .05$

** Sums shown for neutrality were averaged across subjects for intensity.

condition. Mean intensities for surprise, anger, and happiness were higher than those shown for the neutral trials. Paired comparisons supported significantly higher mean intensities for happiness (M = 42.1 dB), anger (M = 41.7 dB), and surprise (M = 41.2 dB), and a significantly lower mean intensity for sadness (M = 30.6 dB) compared to the neutral happiness (M = 36.1), neutral anger (M = 36.7 dB), neutral surprise (M = 36.8 dB), and neutral sadness (M = 35.7 dB) conditions (p <.05 in each case).

Chapter Summary

The results of the ANOVA for sentence duration did not support the acoustic property of duration as a component for a taxonomy of vocal expressions of emotion. However, the significant three-way interaction of condition x emotion x sex, and the significant main effects for conditions and sex, indicated that duration among emotions and between neutral and emotion conditions was affected by the sex of the speakers. The three-way interaction of condition x role x sex indicated that durations produced by actors and nonactors in the neutral and emotion conditions were influenced by sex of the speaker.

The results of analysis of variance for mean intensity did support the acoustic property of intensity as a component toward a taxonomy of six vocal emotional expressions. The conditions x emotions effect, and the main effect of emotions were highly significant at the $p < .001$ level. There were no main effects for role or sex. There were no three-way interactions. Results indicated that, collapsed across males and females, actors and nonactors, mean intensity showed consistent patterns among emotions, and between emotion and neutral conditions. Post hoc analyses using Newman-Keuls tests for paired comparisons among emotions and conditions supported the results for a vocal taxonomy of emotional speech based on mean intensity.

CHAPTER 5

DISCUSSION AND RECOMMENDATIONS

In his introduction to "Speech and Emotional States," Scherer (1981, p. 189) wrote:

> Most laymen agree that there are discrete emotions, such as anger, fear, or joy, and they seem to agree on what it feels like to be angry, fearful, or joyous. Psychologists also recognize these discrete emotions, but find it difficult to define the theoretical construct and to reach a consensus about its relevant aspects and about the relationships of this construct to other psychological constructs.

To respond to this lack of definition and agreement, the present study provided a data base and conceptual organization for emotional speech. This chapter will interpret findings that can most usefully form a taxonomy of emotion, with due regard for role and gender. A variety of speakers produced a variety of simulated emotions. Results of the analysis suggested a vocal taxonomy of emotions based on one principle variable, intensity. Another potential criterion variable, sentence duration, was rejected as a component for a vocal taxonomy because there were interactions involving role and sex. The aspect of duration in emotional expressions did have strong implications, however, for the ways in which males and females are differentially socialized in their expressions of emotion.

Vocal Taxonomy

A predominant approach in the study of emotional speech by most social and behavioral scientists has been that of forced-choice aural identification of selected emotions (usually sadness and anger, then happiness/joy and contempt) and/or emotional states, such as boredom, anxiety, and confidence. Linguistic carriers of emotion used in these studies, usually vowels or brief sentences, were manipulated in a variety of ways to test the listener's capacity to detect the correct emotion. These manipulations included original versions; synthetic speech in which intensity, duration, or fundamental frequency were systematically varied; and content-masking techniques in which speech frequencies above 500 Hertz were removed, thereby eliminating most of the syntactic information. Despite all manipulations, subjects consistently showed correct rates of identification at greater than chance (see Scherer et al., 1972 for a review of these data).

The rates of accuracy for emotion identification in these perceptual studies led researchers to believe that there could be a limited set of vocal cues on which the listener depends to distinguish among emotions. The linguistics, speech science, and acoustic correlates literature have suggested that this set of cues included intensity, perceived as loudness; fundamental frequency, perceived as pitch; and temporal variables, such as rate and duration of speech.

Intensity Taxonomy for Six Emotions

Despite provisional identification of the intensity, fundamental frequency, and temporal cues as the elements of emotional speech, little empirical research has been done to validate these cues as components in the production of emotion and, more importantly, to determine dimensions of each cue in the production of discrete emotions. However, results of analyses of variance and post hoc paired comparisons between conditions and emotions in this study have extended qualitative results of previous studies to provide a strong rationale for a taxonomy of vocal expressions based on the criterion variable of mean intensity. Just as Ekman's (1982) work has contributed to our knowledge of the expressions of discrete emotions through the patterning of specific facial muscles, so can findings in the current study serve as the first step toward a vocal taxonomy complementary to Ekman's facial categories.

Findings in this study confirmed intensity as a key element in the vocal production of emotion. Of greater importance, results provided gradations of the intensity cue across speakers for the six categories. These data can be used to formulate a baseline taxonomy to differentiate between emotions. Based on the results, descriptions in both acoustic and perceptual terms can be provided for six emotions, in addition to a neutral tone of voice. Sadness can be described as being of low intensity; fear and disgust, moderate intensity; and happiness, "hot" anger, and surprise being of high intensity. Neutrality is also seen to be of moderate intensity. Perceptually, sadness can be described as soft; fear, disgust, and neutrality as moderately loud; and happiness, "hot" anger, and surprise as loud.

Results from this taxonomic study conform to perceptual descriptions for some emotions. For example, sadness (Davitz, 1964; Eldred and Price, 1958; Huttar, 1968) was judged as being of soft volume, and happiness (Davitz, 1964; Huttar, 1968) and anger (Costanzo et al., 1969; Davitz, 1964; Eldred and Price, 1958; Huttar, 1968) were both judged as being loud. No perceptual parameters have been outlined for fear, disgust, or surprise in the literature.

This taxonomy of acoustic properties can serve as a viable tool for teaching people about emotional speech, including pathological speech. Although there are few studies using speech science paradigms in psychotic disorders, there are some data regarding voice character and quality (Moses, 1954; Moskowitz, 1951; Spoerri, 1966), and vocal regulation (Ostwald and Skolnikoff, 1966). None of these studies provided indices of emotional speech. There have also been a small number of investigations into vocal indicators of depression (see Scherer, 1979 for a good review of these data), which have shown a significant, albeit expected, reduction in vocal intensity. No acoustical correlate studies for bipolar (manic-depressive) disorder have been identified.

A vocal taxonomy of emotions can play a critical role in the clinical setting in both diagnosis and treatment of affective and thought disorders. At present, clinicians must base their diagnoses and therapeutic interventions on information perceived from the client's affective state--perceptions derived from clinical hunches. Acoustical studies of subjects with psychotic disorders could be obtained (perhaps from tape recorded transcripts) using the methods outlined in this research. Comparisons can be made of characteristics between and within psychotic types to establish a baseline of acoustic properties characteristic of the different vocal behaviors to serve as more reliable diagnostic indicators. These data can be compared to the taxonomy identified in this study to assess in more detail means by which the different psychotic behaviors deviate from the norms identified in this work, and determine changes in the client's affect concomitant with therapy or training. Basic behavioral therapies can be implemented based on the taxonomy. For example, a client with a thought disorder could "practice" a soft tone of voice, which can be equated with a sad facial expression in order to learn the appropriate vocal affect with the appropriate facial expression.

Following validity and reliability checks, a tape recording of selected emotional expressions developed for the current study could become a valuable tool in neuropsychological assessment. Currently, there are no valid and reliable aural tests in the clinical setting that can be used to assess a brain-injured client's capacity to identify vocal expressions of emotion. Testing such as this, in combination with the client completing a forced-choice analysis for perceived acoustical properties, would provide the advantages of (1) more accurate assessment of remaining functions, (2) more detailed information regarding the processing of prosody, and (3) more effective diagnosis and rehabilitation for the client, and education for the primary care giver for maximum recovery of function.

Preliminary Production/Perception Comparisons

Intensity patterns for the emotions examined in this study revealed some interesting implications in light of a preliminary perceptual study (Baldwin and Lauter, unpublished). Fourteen of the expressions studied here and six additional ones recorded from speakers of other languages were re-recorded in 2 versions: original, and low-pass-filtered at 135 Hertz (males) and 150 Hertz (females), to produce a total of 40 trials. Seven subjects (5 male and 2 female), ranging in age from 22 to 47 years, were tested for forced-choice identification of the seven expression types. Performance ranged from 40% correct (22 year old male) to 70% correct (47 year old female) [with chance at 14% correct].

Correct responses and errors were analyzed using a confusion matrix, which indicated that subjects most often confused surprise with happiness (both of which showed high intensity), fear (moderate) with sadness (low intensity), anger (high intensity) with disgust and neutrality (both showed moderate intensity), disgust with sadness (both showed reduced intensities) and anger (high intensity), and neutrality with disgust (both of moderate intensity) and anger (high intensity). These results encourage further use of these recorded stimuli, which have been empirically tested on their emotion dimensions, in perceptual studies such as the one outlined here. Taxonomic data can be compared with these perceptual results to provide a means by which we can understand better the listener strategies and cue utilizations in the detection and interpretation of emotional expressions.

In summary, a taxonomy for emotional speech eventually can be used as a tool to teach clients ways to produce expressions of emotion without ambiguity. Audiotapes of the data from this study can be used in neuropsychological testing to determine better the nature of the central nervous system damage, and to implement therapies that utilize the tapes as teaching devices for producing and perceiving affect. Methods similar to those used here could also be employed in developmental studies, since little is known about the changes in the vocal expressions of emotion from birth to senescence. Finally, this research would be pertinent to cross-cultural studies, to examine whether there is a "universal" taxonomy of vocal emotional expressions.

Recommendations for Future Vocal Taxonomic Studies

Based on the results of the analysis of variance for the intensity variable, which supported a vocal taxonomy of emotions, and in light of the results for the preliminary perceptual study, the following recommendations are suggested for future research:

1. The subject pool should be expanded to include cultural groups (Black and Native American speakers, for example) to provide results of greater generalizability.

2. Taxonomic studies should be expanded to include variations of some of the emotions investigated in this work. The subjects in this study were asked to produce "hot" anger. As Scherer (1986) indicated, distinctions for anger ("hot" vs. "cold") have not been clearly defined in previous studies. Future production studies should include "cold" anger to determine if this production approximates a neutral tone of voice. This could account, in part, for the anger/disgust/neutrality confusions.

3. Other intensity measures should be investigated

to determine any additional contributions this variable could provide toward the vocal taxonomy. These measures include intensity variability and intensity range.

4. Another variable not considered in this study that is undoubtedly important in emotional speech is fundamental frequency (Fo), or voice pitch. A data base should be analyzed for mean Fo, Fo variability, and Fo range. Analyses of these data could contribute to the acoustic patterns for the discrete emotions, and may assist further in differentiating between emotions, such as happiness and surprise, and disgust, anger, and neutrality.

Gender Differences

Her voice was ever soft,
Gentle, and low, an excellent thing in woman.

King Lear, V,iii

The criterion variable of sentence duration was qualified as an element for a vocal taxonomy due to two significant second-order interactions involving gender and role. Findings suggested (1) that women show greater manipulation of the duration cue in their expressions of emotion, (2) that males show little use of the duration cue in their expressions of emotion, (3) that durations for mens' expressions of emotion approximate womens' neutral tones of voice. In addition, an interesting difference was found in the way males and females use the duration cue to express the emotion of fear.

Results from the current investigation, which reflected womens' greater manipulation of duration in their expressions of emotion, appear to support a literature that addressed differential socialization processes for males and females from infancy through adulthood. Research indicated that parents tended to vocalize with daughters more than sons. For example, Lewis and Freedle (1973) studied 3-month-old infants and their mothers in natural play situations. Results showed that mothers of girls vocalized more with their infants than did mothers of boys. Others (Hall, 1979; Henley, 1977; Hickson and Stacks, 1985) have documented an advantage women have shown in their production and perception of expressive behaviors compared with men.

These results also seem to support differences in nurturing practices for males and females with respect to spontaneous expressive behavior. Affect displays are generally subject to shaping by means of social reinforcement from childhood on through modeling and imitation--usually by parental example. These practices hold an implication for males in our culture. This implication was summarized best by Buck (1984, p. 143) when he wrote, "Thus a young boy in our culture is likely to find relatively few male models for the open expression of many emotions, and is likely to experience punishment when openly expressing them; as girls learn they must not hit, boys learn that they must not cry."

The duration differences for males and females are also of interest with respect to a literature regarding personality and attribution characteristics inferred from temporal characteristics of speech. In two field experiments, Miller et al. (1976) found that speech rate in persuasive discourse functioned as a general cue to augment credibility, and that rapid speech enhanced persuasion. The work by Miller and colleagues supported earlier findings (Brown et al., 1973; Smith et al., 1975). In a study that included duration characteristics of oral reading of 14 males on a masculinity-femininity dimension, Terango (1966, p. 593) reported a slower mean reading rate ($\underline{M}$ = 185 words per minute) for effeminate males compared to a slightly faster mean rate ($\underline{M}$ = 194 words per minute) for masculine males. Apple et al. (1979) also reported that slow-talking speakers were judged less persuasive and more "passive," whereas fast-talking speakers were judged as more persuasive and more "active."

These descriptions parallel those found in the literature on gender communication. According to Pearson (1985, p. 202), "Men are viewed as instrumental, task-oriented, aggressive, assertive, ambitious, and achievement oriented. Women, on the other hand, are seen as relational, socio-emotional, caring, nurturing, affiliative, and expressive." The results for durations between male and female speakers appeared to support these sex-role related perceptions and behaviors, with males showing a more "instrumental" style in their shorter vocalizations, and females a more "expressive" style in their longer vocalizations. This finding could account for some males being perceived by their partners in interpersonal encounters as "callous," "indifferent," or "neutral." Figure 5 (Chapter 4) provides evidence for this observation in view of the fact that a number of durations produced by males in their expressions of emotion are comparable to females' neutral tones of voice.

Results of these personality and attribution studies combined with the magnitude of difference in males' and females' durations for some of the emotions might also suggest an application of these cues in dichotic listening studies. A series of experiments have demonstrated a consistent right ear advantage relative to a decrease in the duration of individual sound within a sequence (see Lauter, 1982, for details of these absolute and relative ear advantages). Dichotic studies that incorporate the different temporal dimensions produced by males and females in their productions of emotion might suggest distinctions in respective ear advantages for the detection of "instrumental" (relative right ear advantage), and "expressive" (relative left ear advantage) components of speech transmitted via the durational component of the sound signal.

An interesting difference was also noted between males' and females' expressions of fear. While males showed an increase in duration from their neutral trial for the fear expression, females showed a decrease in duration that fell below all productions except for the males' neutral tones of voice. This reversal could indicate an interesting gender-dependent vocal response to a threatening situation. Electrodermal studies (Craig and Lowrey, 1969; McCracken, 1969; Prokasy and Raskin, 1973) showed greater skin conductance among males in response to a number of emotional-inducing situations. This increase in conductance was associated with the male's need to mask and inhibit overt expressions of emotion in our culture.

Results suggested that males showed an internalizing mode, and females showed an externalizing mode in response to emotion inducing events.

These dermal responses were consistent except for situations involving aggression. In these situations, females showed less of a tendency to be as physically aggressive as males. In addition, females showed an increase in skin conductance similar to those of males in emotionally evocative events. Buck (1976) has suggested that in aggressive situations, males seem to use an externalizing mode and females an internalizing mode.

More generally, Scherer (1976, p. 507) has suggested that "Differences in vocalization mechanisms may be based on differential excitation or inhibition of the peripheral neuromuscular systems involved in the regulation and control of various structures responsible for respiration, phonation, or articulation." Based on Scherer's observations, in combination with results from the internalizer/externalizer studies, the finding for a reversal in the direction of the duration cue produced by males' and females' in their expressions of fear warrants further investigation (1) to determine whether males' and females' vocalizations might reflect differences in coping mechanisms, such as fight vs. flight, and (2) to provide patterns of changes in neuromotor responses under different emotion conditions, with concomitant changes in vocalizations of emotional expressions.

Gender Specific Taxonomy

The duration variable was rejected as a cue for a general taxonomy of emotions due to significant main and interaction effects. However, based on the nature of differences within groups, and the pattern of differences between groups, results for duration suggest a potential gender specific taxonomy. A taxonomy of this sort would be of value (1) as a basis for further research, and (2) as an emotional speech training tool. Figure 5 (Chapter 4) shows durations for males and females, and Figure 6 (Chapter 4) shows the intensity taxonomy.

As a basis for additional research in gender related productions, the gender-specific taxonomy can provide descriptors of males and females durations with the general intensity taxonomy. For example, since the duration for sadness is similar for both males and females, the acoustical description for the expression of sadness can be described as being of long duration and low intensity. The consistency of these two variables for both groups could account, in part, for the high rate of accuracy in perceptual studies. The remaining five emotions would be described in gender specific terms: for males, happiness and surprise are of moderate duration and high intensity, and for females, both emotions are of long duration and high intensity; for males, fear is of moderately long duration and moderate intensity, while for females, fear is of short duration and moderate intensity; for males, anger is of moderate duration and high intensity, while for females, anger is of prolonged duration and high intensity; for males, disgust is of moderate duration and intensity, while for females, disgust is of very prolonged duration and moderate intensity.

These descriptions could be used as a data base for additional gender-related vocal expressive studies. Results can be compared with these descriptions for consistency of effects. Additionally, these descriptions can be used for comparison with perceptual results. For example, confusion matrices can be cross-tabulated with the gender-specific taxonomy to determine if the acoustic variables produced by the sex of the speaker influence, or interfere with the listener's ability to detect the emotion being transmitted.

A second advantage to a gender-related taxonomy of emotions is that of speaker and listener training. If males are utilizing an instrumental style of speech in the transmission of their affective cues, a great deal of misunderstanding can occur, with devastating consequences to interpersonal relationships. A vocal taxonomy can be used as a tool to teach men and women about emotional speech. Interpersonal relations can be improved when a speaker learns to utilize the appropriate cues in emotional contexts and, just as importantly, the listener is able to detect without ambiguity the emotional context to which he or she will respond.

Recommendations for Future Gender Studies

Based on the quantitative and qualitative results found for the duration variable between males' and females' productions of emotions, suggestions for future research are as follows:

1. A larger subject pool incorporating equal numbers of males and females should be utilized to confirm consistency of results noted in this study.

2. Males and females from different cultural groups (Blacks, Hispanics, and Native Americans, for example) should be included to determine if these effects are specific only to Caucasian males and females, or

generalizable to the population.

3. Use of both males and females in future acoustic productions is strongly encouraged due to the magnitude of differences found in the current study. The limited number of acoustical studies extant have employed male speakers. However, descriptions of duration for various emotions have been described in general terms.

4. Other temporal properties of the emotional speech signal should also be investigated to determine if overall duration is the only temporal variable that is gender-specific, or if other temporal properties, such as pause and consonant durations, are indicative of gender differences in the expression of emotions.

Actor and Non-Actor Differences

> Speak the speech, I pray you, as I pronounced it to you, trippingly on the tongue. But if you mouth it, as many of your players do, I had as lief the town crier spoke my lines.
>
> Hamlet, III,ii

Results did not confirm differences between actors and nonactors in their emotional expressions. However, results did indicate, albeit indirectly, that training in dramatic expression could influence greater use of the duration variable in emotional conditions.

This finding is supported by results which showed that durations in the emotion conditions, particularly for male actors, surpassed those of male nonactors and female actors compared with durations in the neutral conditions. Support for manipulation of duration in actor training as a learned behavior was provided by Stern (1983, p. 199) when he wrote, "In working with actors, I frequently am surprised by their resistance to adopting a rate of speech sufficiently slow to allow the audience time not only to hear but also to process what is happening to the characters."

Of particular interest in this interaction is that of sex of the speaker type in the neutral and emotional conditions. All speakers showed an increase in duration from the neutral to the emotional conditions, and all the female subjects showed a longer duration compared to all the male speakers for the neutral condition. However, in the emotion condition, female non-actors showed the longest duration, then male actors, then female actors, then male non-actors. This finding could indicate that in training, male actors learn to "emote" with durations akin to females in general, and female actors learn to "emote" with durations in the emotion condition that, although significantly longer, show a pattern similar to the male non-actor style (see Figure 3, Chapter 4 for an illustration of these effects).

These results suggest that sex and speaker training must be considered in tandem. Although no empirical studies have appeared in the literature to confirm the effects of training in the production of vocal expressions of emotion, it has been reported that an individual's training or occupation could affect the ability to more sensitively detect the cues of others. Rosenthal et al. (1974) reported that men who trained for, or worked in, occupations requiring expressiveness, nurturing, or artistic skill, performed as well as women in decoding the feelings of others.

Results of the conditions x role x sex interaction appear to support the gender-related differences in duration between the neutral and emotion conditions discussed in the previous section, and indicate that male nonactors, in particular, may not be fully utilizing their durational cues in the production of vocal expressions of emotion. This difference could be attributed to the ways in which men and women are differentially shaped to express emotions in our culture, and suggests that dramatic training can help teach people in general, and untrained male speakers in particular, to use their vocal mechanisms more skillfully in their emotional speech.

Recommendations for Future Research

Based on the results for the second-order interaction discussed in this section, recommendations for future research using actors and non-actors include:

1. A pre and post test design in which male nonactors produce vocal expressions of emotion prior to and after a standard course in dramatic expression to determine changes in duration with training.

2. A longitudinal study in which boys and girls produce vocal expressions of emotion at ages 5, 10 and 15 years to determine gender differences during maturation, education, and socialization.

APPENDIX A

ACOUSTIC CORRELATES DEMOGRAPHIC FORM

ACOUSTIC CORRELATES DEMOGRAPHIC FORM

Name ______________________________

Birth Date ______________ Female ________ Male __________

Place of Birth ______________________________

Places Lived to Age 10 Years ______________________________

____

Native Language ______________________________

Other Languages (Fluent) ______________________________

Occupation (s) ______________________________

Have You Had Any of the Following? (Indicate Time):

Speech Therapy (List Type) ______________________________
Acting Lessons ______________ Voice Lessons ____________
Singing Lessons ______________ Other ______________
Professional Acting Experience ______________________________

Courses in Drama/Theatre/Performance (List Type and Time):

____

Major/Minor in Drama/Theatre/Performance (If So, Which?):

____

Membership in professional acting/performance groups, guilds, unions, etc.: ______________________________

____

If you need additional space, please use the back of this form. All information will be kept confidential. Thank you for participating. This research could not be done without you.

Signature ______________________________

APPENDIX B

DATA SETS FOR SUBJECTS

SUBJECT DATA:

Subject #	Description
1	Male.....Actor......47 Years Old
2	Male.....Actor......21 Years Old
3	Male.....Actor......27 Years Old
4	Male.....Nonactor...23 Years Old
5	Male.....Nonactor...31 Years Old
6	Male.....Nonactor...31 Years Old
7	Female...Nonactor...29 Years Old
8	Female...Nonactor...27 Years Old
9	Female...Nonactor...26 Years Old
10	Female...Actor......29 Years Old
11	Female...Actor......43 Years Old
12	Female...Actor......21 Years Old

CODES:

HA = HAPPINESS
SU = SURPRISE
SA = SADNESS
FE = FEAR
AN = ANGER
DI = DISGUST
N = NEUTRAL

DURATION DATA (in seconds)

	NHA	NSU	NSA	NFE	NAN	NDI
1.	0.850	0.800	0.850	0.825	0.925	0.950
2.	0.975	1.000	0.975	1.025	1.075	1.025
3.	0.950	1.000	0.950	0.975	0.975	0.950
4.	0.950	1.000	0.950	0.975	0.950	0.975
5.	0.950	0.950	0.900	0.950	0.950	0.950
6.	1.025	1.025	1.075	1.075	1.100	1.100
7.	1.050	1.100	1.050	1.050	1.000	1.025
8.	1.025	1.050	1.025	1.100	1.050	0.975
9.	1.200	1.475	1.375	1.325	1.350	1.475
10.	1.400	1.350	1.400	1.425	1.450	1.425
11.	1.200	1.150	1.150	1.175	1.125	1.050
12.	1.175	1.125	1.200	1.200	1.150	1.125

DURATION DATA-Continued

	HA	SU	SA	FE	AN	DI
1.	1.725	1.563	2.250	2.050	1.175	1.300
2.	1.225	1.125	1.050	1.050	1.300	1.100
3.	1.175	1.200	1.325	1.250	1.425	1.175
4.	0.900	0.925	1.350	1.125	1.225	1.125
5.	0.950	1.000	1.050	1.100	1.000	1.050
6.	1.025	1.075	1.150	1.100	1.100	1.200
7.	1.200	1.450	1.400	1.150	1.125	1.450
8.	1.375	1.250	1.350	1.125	2.000	1.675
9.	1.550	1.450	1.625	1.250	1.550	1.825
10.	1.675	1.675	1.475	1.225	1.675	1.400
11.	1.275	1.100	1.250	1.025	1.225	1.300
12.	1.200	1.125	1.200	1.000	1.175	1.350

MEAN INTENSITY DATA (in decibels)

	NHA	NSU	NSA	NFE	NAN	NDI	HA	SU	SA	FE	AN	DI
1.	32	33	35	35	36	34	43	44	19	19	38	36
2.	40	42	41	42	42	41	45	38	15	24	44	25
3.	42	42	41	42	42	44	47	43	40	42	45	38
4.	28	28	27	30	29	28	34	36	28	34	40	38
5.	31	32	32	30	30	27	37	34	31	37	36	34
6.	39	40	41	43	43	39	43	39	38	40	39	39
7.	39	41	34	40	41	39	38	41	40	37	42	42
8.	34	35	35	33	33	30	46	44	33	40	45	37
9.	30	30	30	31	30	26	43	45	23	41	38	29
10.	39	40	42	39	40	39	42	45	31	41	45	32
11.	36	38	32	34	33	32	41	39	32	39	42	38
12.	43	40	38	37	41	39	46	46	37	44	46	37

AMPLITUDE MEASURES ACROSS 5 SYLLABLES
(in decibels)

	NHA	NSU	NSA	NFE	NAN
1.	3241353221	1747393625	2544413727	2745403625	3244413726
	NDI	HA	SU	SA	FE
	2942403523	3042495044	3549464840	0933083214	0923093023
	AN	DI			
	3944304532	2745304038			

AMPLITUDE MEASURES-Continued

	NHA	NSU	NSA	NFE	NAN
2.	3545404139	3449474238	3347444139	4148424136	4048484033
	NDI	HA	SU	SA	FE
	4047434037	4248494539	1444494540	1024201804	1328222927
	AN	DI			
	3549464544	2034282223			
	NHA	NSU	NSA	NFE	NAN
3	4545464528	4549464326	4649443827	4548484029	4548484328
	NDI	HA	SU	SA	FE
	4347494039	4948484943	3546494540	2247424642	3547434342
	AN	DI			
	3650474747	2242414441			
	NHA	NSU	NSA	NFE	NAN
4.	3141332710	3239272812	2844292610	3039393012	3043322912
	NDI	HA	SU	SA	FE
	3142302513	3847393015	3147443620	3539272513	3645342927
	AN	DI			
	4147404033	4049443821			
	NHA	NSU	NSA	NFE	NAN
5.	0743403926	0446424028	0346424229	0443403826	0443394022
	NDI	HA	SU	SA	FE
	0343403714	0649474635	1322495036	0342404228	0650454739
	AN	DI			
	0850484529	1345414032			
	NHA	NSU	NSA	NFE	NAN
6.	3143434732	3346444534	3445454732	3946464934	4045455033
	NDI	HA	SU	SA	FE
	2144464737	4249434524	3849424024	3545404029	2950454926
	AN	DI			
	4050444419	4050483719			
	NHA	NSU	NSA	NFE	NAN
7.	3943404332	3745394242	3445373323	3645404534	3945434631
	NDI	HA	SU	SA	FE
	3645384233	2943494621	2547434646	4043434231	2539454529
	AN	DI			
	3350504929	3650474830			

AMPLITUDE MEASURES-Continued

	NHA	NSU	NSA	NFE	NAN
8.	2643353730	3546393222	3945333426	3641333123	3540363122
	NDI	HA	SU	SA	FE
	2737382822	4847504838	4148484238	3940362920	3747474130
	AN	DI			
	4747484339	4045413523			
	NHA	NSU	NSA	NFE	NAN
9.	3035343217	2535323525	3038343217	2345363118	1941353024
	NDI	HA	SU	SA	FE
	2434292817	3248494641	4146474744	2227252219	3742434638
	AN	DI			
	3648313737	3325323122			
	NHA	NSU	NSA	NFE	NAN
10.	3947413832	4044433833	3745464539	3442424235	4043414234
	NDI	HA	SU	SA	FE
	3643414136	3948433939	3949504741	2839352827	3045454935
	AN	DI			
	4349474343	2347373418			
	NHA	NSU	NSA	NFE	NAN
11.	4147443315	4048453520	4447372605	3947363413	4442363212
	NDI	HA	SU	SA	FE
	3641343120	4548364035	3950403929	3745282226	3445464820
	AN	DI			
	3448494632	3548464023			
	NHA	NSU	NSA	NFE	NAN
12.	4049454635	3648424032	3748393631	4049343034	4046424135
	NDI	HA	SU	SA	FE
	4348413530	3950504942	3748504945	2442414040	3547494843
	AN	DI			
	3650505042	2248453437			

REFERENCES

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. Journal of Personality and Social Psychology, 37, 715-727.

Baldwin, C. M., & Lauter, J. L. (1987). [Perceptual identification of vocal expressions of emotion: A preliminary study]. Unpublished raw data.

Borden, G. J., & Harris, K. S. (1984). Speech science primer. Baltimore, MD: Williams and Wilkins.

Broad, D. J. (1973). Phonation. In F. D. Minifie, T. J. Hixon, & F. Williams (Eds.), Normal aspects of speech, hearing, and language. NJ: Prentice-Hall.

Brown, B. L., Strong, W. J., & Rencher, A. C. (1973). Perceptions of personality from speech: Effects of manipulations of acoustical parameters. Journal of the Acoustical Society of America, 54, 29-35.

Brown, B. L., Warner, C. T., & Williams, R. N. (1985). Vocal paralanguage without unconscious processes. In A. W. Siegman, & S. Feldstein (Eds.), Multichannel integrations of nonverbal behavior. Hillsdale, NJ: Lawrence Erlbaum Associates.

Buck, R. (1984). The communication of emotion. NY: Guilford.

Buck, R. (1976). Human motivation and emotion. NY: Wiley.

Coleman, R. F., & Williams, R. (1979). Identification of emotional states using perceptual and acoustic analysis. In V. Lawrence & B. Weinberg (Eds.), Transcript of the eighth symposium: Care of the professional voice (Part I). NY: The Voice Foundation.

Costanzo, F.S., Markel, N.N., & Costanzo, P.R. (1969). Voice quality profile and perceived emotion. Journal of Consulting Psychology, 16, 267-270.

Craig, K., & Lowrey, H. J. (1969). Heart rate components of conditioned vicarious autonomic responses. Journal of Personality and Social Psychology, 11, 381-387.

Darwin, C. (1872/1965). The expression of the emotions in man and animals. London: John Murray. (Reprinted in Chicago, IL: University of Chicago Press).

Davitz, J.R. (1964). The communication of emotional meaning. NY: McGraw-Hill.

Denes, P., & Milton-Williams, J. (1962). Further studies in intonation. Language and Speech, 5, 1-14.

Dittman, A. T., & Wynne, L. C. (1961). Linguistic techniques and the analysis of emotionality in interviews. Journal of Abnormal and Social Psychology, 1, 201-204.

Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), Darwin and facial expression: A century of research in review. New York: Academic Press.

Ekman, P. (Ed.). (1982). Emotion in the human face (2nd ed.). NY: Cambridge University Press.

Ekman, P., Friesen, W. V., and Tomkins, S. S. (1971). Facial affect scoring technique (FAST): A first validity study. Semiotica, 3, 37-58.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. Science, 221, 1208-1210.

Eldred, S. H., & Price, D. B. (1958). A linguistic evaluation of feeling states in psychotherapy. Psychiatry, 21, 115-121.

Fairbanks, G., & Hoaglin, L.W. (1941). An experimental study of the durational characteristics of the voice during the expression of emotion. Speech Monographs, 8, 85-90.

Fairbanks, G., & Pronvost, W. (1938). Vocal pitch during simulated emotion. Science, 88, 382-383.

Fonagy, I. (1978). A new method of investigating the perception of prosodic features. Language and Speech, 21, 34-49.

Friedhoff, A. J., Alpert, M., & Kurtzberg, R. L. (1964). An electro-acoustic analysis of the effects of stress on the voice. Journal of Neuropsychiatry, 5, 266-272.

Fulcher, J. S. (1942). "Voluntary" facial expression in blind and seeing children. Archives of Psychology, 38, No. 272.

Hall, J. A. (1979). Gender, gender roles, and nonverbal communication skills. In R. Rosenthal (Ed.), Skill in nonverbal communication. Cambridge, MA: Oelgeschlager, Gunn and Hain.

Hecker, M. H. L., Stevens, K. N., Bismarck, G. Von., & Williams, C. E. (1968). Manifestations of task-induced stress in the acoustic speech signal. Journal of the Acoustical Society of America, 44, 993-1001.

Henley, N. M. (1977). Body politics. Englewood Cliffs, NJ: Prentice Hall.

Hickson, M. L., & Stacks, D. W. (1985). Nonverbal communication. Dubuque, IA: W. C. Brown.

Hooff, Van, J. A. R. A. M. (1972). The phylogeny of laughter and smiling. In R. A. Hinde (Ed.), Non-verbal communication. NY: Cambridge University Press.

Huttar, G.L. (1968). Relations between prosodic variables and emotions in normal American English utterances. Journal of Speech and Hearing Research, 11, 481-487.

Kaiser, L. (1962). Communication of affects by single vowels. Synthese, 14, 300-319.

Kotlyar, G. M., & Morozov, V. P. (1976). Acoustical correlates of the emotional content of vocalized speech. Journal of Acoustics of the Academy of Sciences of the USSR, 22, 208-211.

Lauter, J.L. (1982). Dichotic identification of complex sounds: Absolute and relative ear advantages. Journal of the Acoustical Society of America, 71, 701-707.

Lehiste, I. (1970). Suprasegmentals. Cambridge, MA: M.I.T. Press.

Lewis, M., & Freedle, R. (1973). Mother-infant dyad: The cradle of meaning. In P. Pliner, L. Krames, & T. Alloway (Eds.), Communication and affect: Language and thought, pp. 127-155. NY: Academic Press.

Lieberman, P. (1965). On the acoustic basis of the perception of intonation by linguists. Word, 21, 40-54.

Lieberman, P. (1974). A study of prosodic features. In T. A. Sebeok (Ed.), Current trends in linguistics (Vol. 12: Linguistics and adjacent arts and sciences), pp. 2419-2449. The Hague: Mouton.

Lyons, J. (1972). Human language. In R. A. Hinde (Ed.), Non-verbal communication. NY: Cambridge University Press.

McCracken, S. R. (1971). Comprehension for immediate recall of time-compressed speech as a function of sex and level of activation of the listener. In E. Foulke (Ed.), Proceedings of the second Louisville conference on rate and/or frequency-controlled speech. Louisville, KY: University of Louisville.

Miller, N., Maruyama, G., Beaber, R. J., & Valine, K. (1976). Speed of speech and persuasion. Journal of Personality and Social Psychology, 34, 615-624.

Minifie, F. D. (1973). Speech Acoustics. In F. D. Minifie, T. J. Hixon, & F. Williams (Eds.), Normal aspects of speech, hearing, and language. NJ: Prentice-Hall.

Moses, P. J. (1954). The voice of neurosis. NY: Grune & Stratton.

Moskowitz, E. (1952). Voice quality in the schizophrenic type. Abstracted by D. Mulgrave in Speech Monographs, 19, 118-119.

Ostwald, P. F., & Skolnikoff, A. (1966). Speech disturbances in a schizophrenic adolescent. Postgraduate Medicine, 40-49.

Pearson, J. C. (1985). Gender and communication. Dubuque, IA: W. C. Brown.

Pickett, J.M. (1980). The sounds of speech communication. Baltimore, MD: University Park Press.

Prokasy, W., & Raskin, D. (1973). Electrodermal activity in psychological research. NY: Academic Press.

Rosenthal, R., Archer, D., DiMatteo, M., Koivumaki, R., Hall, J., & Rogers, P. L. (1974). Body talk and tone of voice: The language without words. Psychology Today, 8, 64-68.

Scherer, K. R. (1979). Nonlinguistic vocal indicators of emotion and psychopathology. In C. Izard (Ed.), Emotions in personality and psychopathology. New York: Plenum.

Scherer, K. R. (1981). Speech and emotional states. In J. K. Darby (Ed.), Speech evaluation in psychiatry. New York: Grune and Stratton.

Scherer, K. R. (1982). Methods of research on vocal communication: Paradigms and parameters. In K. R. Scherer & P. Ekman (Eds.), Handbook of methods in nonverbal behavior research. Cambridge, UK: Cambridge University Press.

Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. Psychological Bulletin, 99, 143-165.

Scherer, K.R., Koivumaki, J., & Rosenthal, R. (1972). Minimal cues in the vocal communication of affect: Judging emotions from content-masked speech. Journal of Psycholinguistic Research, 1, 269-285.

Siegman, A. W. (1985). Expressive correlates of affective traits and states. In A. W. Siegman & S. Feldstein (Eds.), Multichannel integrations of nonverbal behavior. Hillsdale, NJ: Lawrence Earlbaum Associates.

Simonov, P. V., & Frolov, M. V. (1973). Utilization of human voice for estimation of man's emotional stress and state of attention. Aerospace Medicine, 44, 256-258.

Skinner, E.R. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness. Speech Monographs, 2, 81-137.

Smith, B. L., Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Effects of speech rate on personality perception. Language and Speech, 18, 145-152.

Soskin, W. F., & Kauffman, P. E. (1961). Judgment of emotion in word-free voice samples. Journal of Communication, 11, 73-80.

Spoerri, T. H. (1966). Speaking voice of the schizophrenic patient. Archives of General Psychiatry, 14, 581-585.

Starkweather, J. A. (1961). Vocal communication of personality and human feelings. Journal of Communication, 11, 63-72.

Stern, D. A. (1983). Teaching and Acting: A vocal analogy. In A. M. Katz & V. T. Katz (Eds.), Foundations of nonverbal communication. Carbondale, IL: Southern Illinois University Press.

Terango, L. (1966). Pitch and duration characteristics of the oral reading of males on a masculinity-femininity dimension. Journal of Speech and Hearing Research, 9, 590-595.

Thompson, J. R. (1943). Development of facial expression of emotion in blind and seeing children. Archives of Psychology, 37, No. 264.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. Journal of the Acoustical Society of America, 52, 1238-1250.