

University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

1977

Pitch level and pitch range as a function of client-therapist interaction in psychotherapy

Terrel Lee Templeman
The University of Montana

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Templeman, Terrel Lee, "Pitch level and pitch range as a function of client-therapist interaction in psychotherapy" (1977). *Graduate Student Theses, Dissertations, & Professional Papers*. 3228.
<https://scholarworks.umt.edu/etd/3228>

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

PITCH LEVEL AND PITCH RANGE AS A FUNCTION
OF CLIENT-THERAPIST INTERACTION IN
PSYCHOTHERAPY

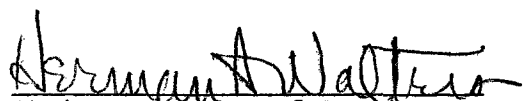
By

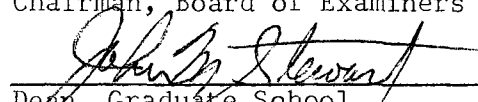
Terrel L. Templeman

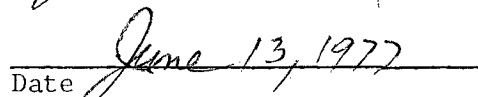
B.A., University of Oregon, 1971

Presented in partial fulfillment of the requirements
for the degree of
Master of Arts
UNIVERSITY OF MONTANA
1977

Approved by:


Chairman, Board of Examiners


Dean, Graduate School


Date June 13, 1977

UMI Number: EP34759

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI EP34759

Copyright 2012 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ABSTRACT

Templeman, Terrel L., M.A., June 12, 1977

Psychology

Pitch Level and Pitch Range as a Function of Client-Therapist Interaction in Psychotherapy (68 pps)

Director: Herman A. Walters *HW*

The purpose of this study was to investigate the functional significance of vocal pitch in psychotherapy. Pitch means and ranges of a female dyad were measured across three sessions of psychotherapy and those of a male dyad were measured within one psychotherapy session by means of a psychophysical technique requiring human judges to match the tones of therapist and client speech to pure tones. Three specific hypotheses were tested: 1) mean pitch and pitch range of therapist and client are not independent of one another in the course of any given psychotherapy session; 2) pitch range of client can be modified by extremes of therapist pitch range; and 3) mean pitch and pitch range of therapist and client vary across sessions as a function of the therapeutic relationship. No significant interactions of pitch were found for the female dyad across sessions, and no significant interaction of pitch was found for either dyad within any psychotherapy session. This result held true both for those therapist statements which followed client statements and those which preceded client statements. Both therapists showed great difficulty in modifying their pitch range upon request in therapy, but one instance of extreme broadening of pitch by the male therapist resulted in no modification of male client pitch range. All pitch fluctuations exhibited temporal characteristics which could best be described through time series analysis by an integrated moving averages model. The results and their implications are discussed in terms of process variables and their importance for understanding therapeutic relationship. The usefulness of the psychophysical technique for measuring pitch in psychotherapeutic research is also discussed.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
LIST OF ILLUSTRATIONS	iv
LIST OF TABLES	v
ACKNOWLEDGMENTS	vi
CHAPTER	
I. INTRODUCTION	1
II. ACOUSTICAL AND TEMPORAL PROPERTIES OF SPEECH IN PSYCHOTHERAPY	3
III. METHOD	19
IV. RESULTS	27
V. DISCUSSION AND CONCLUSIONS	48
VI. SUMMARY	59
LIST OF REFERENCES	62

LIST OF ILLUSTRATIONS

	Page
1. Experimental Design	21
2. A Flow Chart of the Pitch Measurement Procedure Used in This Design	25
3. Average Pitch Means and Ranges for the Female Therapist and Client across 3 Sessions of Psychotherapy	29
4. Plots of the Autocorrelations and Partial Autocorrel- ations after First Order Differencing for Female Therapist and Client Mean Pitch and Pitch Range Series in Session 1	32
5. Plots of the Autocorrelations and Partial Auto- correlations after First Order Differencing for Female Therapist and Client Mean Pitch and Pitch Range Series in Session 5	33
6. Plots of the Autocorrelations and Partial Auto- correlations after First Order Differencing for Female Therapist and Client Mean Pitch and Pitch Range Series in Session 8	34
7. Mean Pitch and Pitch Range Fluctuations of Female Therapist and Client across Session 1	35
8. Mean Pitch and Pitch Range Fluctuations of Female Therapist and Client across Session 5	36
9. Mean Pitch and Pitch Range Fluctuations of Female Therapist and Client across Session 8	37
10. Correlation Coefficients Computed at Lags 0 and 1 for Female Therapist and Client Mean Pitch and Pitch Range Levels, Plotted for 3 Sessions of Psychotherapy.	39
11. Mean Pitch and Pitch Range Fluctuations of Male Therapist and Client across Session 4.	43
12. Plots of the Autocorrelations and Partial Auto- correlations after First Order Differencing for Male Therapist and Client Mean Pitch and Pitch Range Series in Session 4	44
13. Standard Scores Derived from Therapist and Client Personal Reaction Questionnaire Responses across 3 Sessions of Psychotherapy for the Female and Male Dyads	47

LIST OF TABLES

1. Percent Agreement among Three Judges Computed
across Each Pair of Judges for Each Subject
and Each Dependent Variable 28
2. A 3X3 Contingency Table and a 3X3 Estimated
Frequency Table of Directional Changes in the
Pitch Inflections of the Female Therapist as
a Function of Such Changes by the Female
Client in Session 1 41
3. Chi Square Tests of Independence of Directional
Changes between Pitch Inflections of Female
Therapist and Client for 3 Sessions of
Psychotherapy 41
4. Chi Square Tests of Independence of Directional
Changes between Pitch Inflections of Male Therapist
and Client in Session 4 45

ACKNOWLEDGEMENTS

I wish to thank the members of my Thesis Committee, Dr. Kellogg O. Lyndes, Department of Communication Sciences and Disorders, Dr. John R. Means, Department of Psychology, Dr. James A. Walsh, Department of Psychology, and especially its director, Dr. Herman A. Walters, Department of Psychology, for their steadfast support and encouragement in this endeavor. I am also grateful to Dr. Richard M. Boehmler, Chairman of the Department of Communication Sciences and Disorders, for providing me with the space and equipment necessary for conducting the research; to Mr. Elmer Doll for assembling the apparatus and getting it to work properly; and to Dr. Donald W. Simmons, Chairman of the Department of Music, for actively helping me recruit vocal music students as judges in the experiment. Finally, I wish to thank Ms. Sue Carlson, Ms. Joan Darchuk, and Ms. Stephanie Smith for their participation as judges in a rather tedious psychophysical task.

CHAPTER I

INTRODUCTION

Psychotherapy generally involves two or more persons in communication. The study of this communication has been called "process research" (Kiesler, 1973) because 1) it focuses on the interactions of therapist and client behaviors, 2) it focuses on within-therapy changes in behavior, and 3) it measures "variables having to do with movement or flow" (Shlien & Zimring, 1970) in therapy. A variety of process variables, ranging from body kinesics to paralinguistic to verbal content of speech, have been studied (Duncan, 1969). The experiment presented in this paper represents one attempt to quantify changes in a paralinguistic variable that occurs in ongoing psychotherapy.

Pitch levels and pitch ranges of therapist and client speech were monitored across three psychotherapy sessions, and three hypotheses were tested:

1. Pitch level and pitch range of therapist and client are not independent of one another in the course of any given psychotherapy session.
2. Pitch range of client can be modified by extreme pitch ranges (that is, very narrow or very broad ranges) of the therapist.
3. Mean pitch levels and ranges of therapist and client change across time as the number of therapy sessions increases. These changes may represent a convergence or divergence, depending upon the nature of the therapeutic relationship.

All pitch measurements were made utilizing a novel psycho-physical technique which required judges to match vocal pitch from speech segments to pure tones. The reliability, precision, empirical and psychological validity, and general usefulness of this technique for process research is discussed and compared to other measures of pitch used previously. In addition to the pitch measurements, transcripts of each psychotherapy session were prepared, and modified versions of the Client and Therapist Personal Reaction Questionnaire (Ashby, Ford, Guerney, Guerney, & Snyder, 1957; Goldstein, 1971) were administered after each session to assess therapeutic relationship.

CHAPTER II

ACOUSTICAL AND TEMPORAL PROPERTIES OF SPEECH IN PSYCHOTHERAPY

Because of the predominantly verbal nature of psychotherapy, properties of speech have long been considered important variables for study. Speech may most simply be dichotomized into 1) what one says, and 2) how one says it (Davitz, 1964). This chapter is concerned with the latter class of variables, specifically the acoustical and temporal properties of speech and how they interact in psychotherapy.

Speech Correlates of Personality

The first experimental studies of speech properties began to proliferate with the advent of the phonograph and radio, the first instruments to make the analysis of speech in the absence of the speaker possible. These early studies attempted to demonstrate that from listening to voice recordings alone one could accurately assess a variety of personality traits of the speaker, such as introversion-extroversion (Fay & Middleton, 1942), social values (Allport & Cantrill, 1934), dominance-nondominance (Eisenberg & Zabowitz, 1938), neuroticism (Taylor, 1934), sociability (Fay & Middleton, 1941a), truth-telling (Fay & Middleton, 1941b), Kretschmerian body types (Fay & Middleton, 1940), leadership (Fay & Middleton, 1943), and even occupation (Fay & Middleton, 1939). These early findings were

disappointing: judges' ratings exhibited a high degree of interrater reliability but did not correlate well with the external measures of the traits, suggesting that they were in fact based upon stereotyped associations of voice and personality (Starkweather, 1961).

Other researchers concentrated on separating and studying the component properties of speech, such as "intonation", "rhythm", "continuity", and "speed" (Sapir, 1927), and correlating these with personality traits. Stagner (1936) found that voice intensity of speakers reading passages to an audience is positively correlated with the speaker's self-confidence as rated on the Bernreuter Scale. Chapple, Chapple, and Repp (1955) published an entire taxonomy of personality traits correlated with rate of speaking alone, as measured by the Interaction Chronometer (Chapple, 1939). Similarly, clinicians have sometimes tried to classify psychiatric disorders on the basis of voice quality. Newman and Mather (1938) classified various affective disorders on the basis of such qualities as pitch level, pitch range, tempo, and length of speech. Moses (1954) differentiated between neurotic and psychotic voices on the basis of what he called "the 5 R's": respiration, range, register, resonance, and rhythm. Markel, Meisels, and Houck (1964) differentiated schizophrenic from nonschizophrenic voices in terms of three dimensions of voice quality which they called the "evaluative" quality, "potency", and "activity". Markel, Bein, and Phillis (1973) differentiated hostility-directed-inward versus hostility-directed-outward in patients on the basis of rate of speaking alone. Fundamental frequency of voice, defined as

periodicity of vocal fold vibrations and denoted by F_0 (Fant, 1968), has also been used to classify psychiatric types. Ostwald (1963, 1965) has defined four vocal stereotypes based on F_0 profiles which he associates with different kinds of patients: "the sharp voice", which is found in schizophrenic and hysterical patients; "the flat voice", found in patients with strong dependency needs and feelings of helplessness; "the hollow voice", found in debilitated or stuporous patients; and "the robust voice", found in aggressive or extroverted patients. Likewise several investigators have attempted to define vocal correlates of depression. Condon and Ogston (1966) described the speech of a depressed patient as having a narrowed pitch range, "a lack of quality or timbre", and "a marked pitch rhythm...characterized by a rise in pitch at the termination of most subsegments". Hargreaves and Starkweather (1964) have even shown a concomitant shift in the F_0 profile of a depressed patient that accompanied his recovery. Finally, Kanfer (1960) reported a negative correlation between rate of speaking and elevated Hs, D, Pa, Pt, and Si scale scores on the Minnesota Multiphasic Personality Inventories (MMPI) of hospitalized patients.

Not all studies in this area have achieved positive results. Luft (1951) asked clinicians to predict how clients would respond on objective and projective personality tests after listening to recordings of or reading transcripts of clinical interviews. He found that readers were as accurate as listeners in predicting objective test responses but less accurate in predicting projective test responses. In a related study Grigg (1958) found that reading

transcripts allowed for better prediction of responses on some tests (for example, the Q-Sort) than did listening to recordings of the interviews. Cohen (1961) has also shown that nonclinicians are better able to assess the pathology of schizophrenic patients after reading transcripts of interviews than after listening to recordings of the interviews. These results all highlight the importance of verbal content in conveying personality information about a speaker, and they suggest that for some kinds of information (for example, what kinds of self-statements a person is likely to endorse), knowing what one says is more important than knowing how he says it.

To summarize this section, research has provided at least tentative evidence that acoustical properties of speech can convey personality or trait information about an individual. However, it has also demonstrated that such correlates of personality are often stereotyped and may not always reflect how the speaker perceives himself. Verbal content of speech continues to be an important variable for understanding human beings and in some situations is probably more enlightening than vocal properties. Methodologically, this line of research also needs refinement. Personality estimates have often been based upon very restricted samples of speech within very structured frameworks (for example, the interview situation). More research is needed on the question of how stable these acoustical profiles are across time and within various situations. Also, most of the evidence presented here has been strictly correlational. More controlled research like that of Hargreaves and Starkweather (1964), in which independent variables are manipulated to change personality

and vocal properties are measured as dependent variables, is needed to define the relationship between voice and personality more precisely.

Speech Correlates of Emotion

Research into the relationship between speech variables and affective states has generally yielded more conclusive results than that into personality traits. In one of the earliest experiments of this kind, E. R. Skinner (1935) measured vocal pitch in terms of sound frequency and reported that sighs of happiness were higher in pitch than sighs of sadness. Fairbanks (1940) and Fairbanks and Provonost (1939) also found characteristic fundamental frequency profiles for the expression of anger, grief, fear, and indifference when they analyzed the speech of actors reading prose passages with these different feeling expressions. Fairbanks furthermore demonstrated that these feeling expressions could be reliably identified by judges listening to recordings of the speech (Fairbanks & Hoaglin, 1941). More recently, however, Lieberman (1961) has shown that the same emotion expressed more than once yields varying pitch profiles, and Williams and Stevens (1972) have evidence that such variability may be a function of the context in which the emotion is expressed. Thus, although there seems to be a positive relationship between fundamental frequency and emotion, the relationship is probably not as fixed as early investigators assumed.

Several studies have examined the relationship between speech and affective state in clinical interviews. Eldred and Price (1958) asked therapists to listen to 15 interviews of the same client-therapist

dyad, transcribe the interviews into linguistic and paralinguistic notation (Trager, 1961), and note any relationship between paralinguistic cues and affect. They found patterns of pitch, rate of speaking, vocal intensity, and disrupted speech that seemed to vary consistently with suppressed versus overt feelings of anger and depression in the client. Other studies using this procedure have produced suggestive although inconclusive results indicating some relationship between speech and affective state in psychotherapy (McQuown, 1957; Pittenger, Hockett, & Danehy, 1960). One study, however, found no such relationship in clinical interviews and even reported difficulties in trying to transcribe the interviews into paralinguistic notation reliably (Dittman & Wynne, 1961). Studies like these, which ask therapists to peruse recorded interviews, code the speech for paralinguistic variables, and look for affective states are generally poorly controlled and do not involve any a priori rules for differentiating emotions.

More sophisticated methods for standardizing interviews, measuring speech variables, and rating affective states have led to more conclusive results. Using the standardized interview (Matarazzo, Saslow, & Matarazzo, 1956), in which the temporal characteristics of interviewer speech are held constant and those of the interviewee are measured as dependent variables, Mahl (1956) found a significant positive relationship between therapist ratings of client anxiety during psychotherapy sessions and the disruptedness of client speech in those sessions. The standardized interview has also been used to demonstrate a positive relationship between rate of speaking and

anxiety (Kanfer, 1959), and between rate of speaking and induced elation (Natale, 1977). Pope, Blass, Siegman, and Rahe (1970) have reported a negative correlation between rate of speaking and ratings of high versus low states of depression in hospitalized patients. Finally, Roessler and Lester (1976) have reported being able to predict ratings of affect (depression, anger, and fear) made of recorded client speech in psychotherapy on the basis of equations based upon several parameters of fundamental frequency of those speech samples. They conclude that the physical parameters of voices are indeed "objective indices of emotion, emotional intensity, and emotional conflict in psychotherapy" (Roessler & Lester, 1976).

Another approach in this area of research involves manipulating the parameters of speech themselves and asking judges to rate the affect of different kinds of voices. In one of the earliest studies of this kind, Knower (1941) demonstrated that emotion can be correctly identified from whispered speech. Speech can also be electronically filtered so that those frequencies necessary to understand what is being said are missing (Davitz, 1964). There is a substantial body of evidence now indicating that emotion can be correctly identified from this "content-free speech" (Davitz, 1964; Kramer, 1964; Soskin & Kauffman, 1961; Starkweather, 1956a, b, 1961). Lieberman and Michaels (1962) have electronically dampened pitch ranges of speech segments to monotonic levels to show that accuracy in identifying emotion drops by 86% after this manipulation. Finally, Brown, Strong, and

Rencker (1973, 1974) electronically increased and decreased both F_0 and rate of speaking in recorded speech samples and then asked judges to rate the different voices on fifteen emotional variables. These ratings were factor analyzed into "benevolence" and "competence" dimensions. Results showed that increasing the F_0 variance increased the benevolence ratings and decreasing it decreased both benevolence and competence ratings, but that manipulating rate of speaking affected both sets of ratings more than the F_0 manipulations.

To summarize, acoustical properties of voice, particularly pitch measures and rate of speaking, have been shown to convey emotion. Pitch functions vary for any given emotion, however, and the relationship between pitch and affect may be a function of the context in which it is expressed. Also, the work of Brown and his coworkers (Brown et al., 1973, 1974) suggests that rate of speaking may be more important for affective expression in some cases than is pitch. More research is needed regarding which vocal properties are important for expressing which emotions and how various properties interact in the expression of any given emotion. Methodologically, studies utilizing the most sophisticated measures of speech properties have tended to examine nonspontaneous speech samples, while those investigating spontaneous, conversational speech have tended to be poorly controlled and less precise in their measurements. In the former case results may be valid only for stereotyped examples of emotional expression, and in the latter case few hard conclusions about which parameters are important for affective

expression can be drawn. More attempts to elicit specific emotions within spontaneous interactions like those of Natale (1977), and more efforts to subject larger samples of spontaneous speech to precise acoustical analysis, such as the study of Roessler and Lester (1976), are needed before a full understanding of the natural expression of emotion can be achieved.

The Interaction of Vocal Properties in Psychotherapy

As early as 1949 Reik proposed that "vocal modulations that do not strike us" nevertheless convey a great deal of information about the client in therapy. He included "the particular pitch and timbre of [the client's] voice, his particular speech rhythm", and "...variations of tone, pauses, and shifted accentuation" (Reik, 1949) as important process variables, and he suggested that therapists learn to listen for them "with the third ear" (Reik, 1949). Since then much research has investigated the role these variables play in psychotherapy. Thompson and Bradway (1950) attempted to develop third ear listening in graduate students in clinical psychology by having them observe role-played therapy sessions in which the client and therapist expressed themselves by means of nonverbal vocalizations only. They reported that students achieved some success in learning to correctly assess the content of the sessions. Other researchers have found that successful psychotherapy outcome can be predicted from vocal qualities of either client (Rice & Wagstaff, 1967) or therapist (Rice, 1965) speech in therapy. Duncan, Rice and Butler (1968) have shown that ratings of peak and

poor therapy hours correlated positively with certain stress patterns, vocal intensities, and pitch levels of therapist speech.

In recent years more and more research has examined how non-verbal variables interact in psychotherapy. This line of enquiry has largely grown out of work with standardized interview procedures (Goldman-Eisler, 1951, 1954, 1968; Matarazzo, et al., 1956). Reviews of this research can be found elsewhere (Goldman-Eisler, 1968; Matarazzo & Wiens, in press; Matarazzo, Wiens, Matarazzo, and Saslow, 1968). The results strongly suggest that given a positive therapeutic relationship client speech and silence behaviors tend to imitate those of the therapist; that is, empathic or positive modes of responding by the therapist tend to reinforce the same modes of responding by their clients during psychotherapy (Matarazzo & Wiens, in press). This extraordinary result is not only the first direct evidence that therapist speech serves as a model for clients, but it also represents one of the first behavioral descriptions of a therapist variable long considered important for successful psychotherapy (Truax & Mitchell, 1971).

The interactions of other speech variables in psychotherapy have only just begun to be studied, but the results reported so far have been equally positive. Natale (1975) has demonstrated that when vocal intensity of therapist voice is manipulated, client vocal intensity levels are modified so as to match those of the therapist. When client and therapist are able to speak to each other in the same room, their vocal intensities begin to converge over time. Most

recently, Wexler and Butler (1976) have reported that increased expressiveness of therapist voice, as measured by the Expressiveness Scale (Wexler, 1975), which includes assessments of pitch variability, resulted in increased expressiveness in the voice of a depressed patient. In addition, after 20 sessions of psychotherapy, client and therapist expressiveness had converged to where each person was receiving almost identical scores by the raters.

These kinds of studies are important because they begin to describe for the first time which speech variables are operating in the process of psychotherapy and how they are related to traditional notions of empathy, rapport, and therapeutic relationship. Presumably if one can determine which therapist variables are effective for successful behavior change in the client, the potential exists for training more effective therapists. The experiment reported in this paper is one of the first efforts to examine the interaction of therapist pitch levels and ranges with those of the client in psychotherapy.

Defining Pitch

Before describing the experiment itself, it is first necessary to discuss ways in which pitch is sometimes defined, for what one can conclude about this particular variable is greatly determined by how it is operationalized. The American Standards Association has defined pitch as

that attribute of auditory sensation in terms of which sounds may be ordered on a scale from low to high. Pitch depends primarily upon the frequency of the sound stimulus, but it also depends upon the sound pressure and waveform of the stimulus (Crystal, 1969, P. 108).

Note that pitch is here defined as a psychological variable, that is, it is an attribute of human perception and not simply a physical entity. Crystal (1969) elaborates on this with the following corollaries: 1) No one-to-one relationship between frequency and pitch exists. 2) Frequency change is not directly proportional to pitch change. 3) Pitch may be perceived in the absence of F_0 . 4) Amplitude differences affect pitch. 5) Individual differences in pitch perception exist. 6) Environmental variables such as the size and shape of a room, the recording characteristics of the sound, and the location of the stimulus with respect to the listener affect pitch. 7) Any given pitch level is actually a vibrato which fluctuates about a mean. Stevens (1975) has emphasized the psychological nature of pitch by demonstrating that psychophysically determined pitch scales are not simple functions of either musical scales or the frequency of sounds.

In attempting to measure a dependent variable like pitch the process researcher should strive for empirical and psychological validity. Empirical validity has to do with the objectiveness of the measurement, how well it assesses the physical parameters of the sound and whether or not it can be calibrated with some external referent, such as a musical scale or frequency in cycles per second (hertz). Empirical validity can be checked by comparing the measurement with those of instruments which have been demonstrated to measure the physical properties of the variable precisely. Psychological validity has to do with the relevance of the measurement to how the

variable is actually perceived in everyday life. The measurement must be meaningful in the context in which the variable is operating. This is particularly important for pitch because, as noted above, the human perception of pitch is not a simple function of the physical parameters of sound frequency. Psychological validity can be checked by comparing one's measurement with human ratings of the same variable.

Operational definitions of pitch currently used in research generally manifest either empirical or psychological validity but rarely both of them. For example, subjective measures of "high" versus "low" pitch levels or "wide" versus "narrow" pitch ranges (Markel, 1965; McQuown, 1957; Pittenger et al., 1960; Wexler, 1975) achieve a high degree of psychological validity in that human ratings of these variables generalize well to everyday distinctions of pitch. These measures lack empirical validity, however, in that they cannot describe the physical characteristics of the variable or determine just how high is high pitch or how wide is wide range in terms of some external referent. Empirical validity is better manifested by instruments such as the sound spectrograph (Fant, 1968), pitch meters (Anderson, 1960; Dempsey, Draegert, Siskind, & Steer, 1950), computer-generated fast fourier transform analysis (Gold, 1962; Harris & Weiss, 1963; Miller, 1970; Noll, 1964; Roessler & Lester, 1976), all of which measure speech sounds in terms of sound wave form and frequency, and techniques for measuring vocal fold movements during the production of speech (Fujimura & Lindqvist, 1971; Holmer & Rundqvist, 1975). As noted above, however, such precise measures of F_0 are only

approximations of human pitch perception and thus may be less psychologically valid than the subjective rating scales.

Physical versus subjective measures of pitch also differ in their capacity to handle large samples of speech. Human judges can rate pitch levels for relatively large samples of speech, including entire therapy sessions. Determinations of F_0 require the averaging of many wave form analyses on very small samples of speech (no more than a few seconds each). Pragmatically then both kinds of measures involve a tradeoff: the more precise the measurement the more restricted the focus of the analysis, and the greater the amount of analysis necessary to measure the same sample size. The less precise the measurement, the less restricted the focus of analysis, and the less total amount of analysis necessary. More precise measurements are hence usually more time-consuming and expensive to make than less precise ones.

In their clinical practice speech pathologists sometimes use a pitch measure that is both psychologically and empirically valid. The procedure involves matching the voiced sounds of their clients to tones on a pitch pipe or piano (Boone, 1971). Converted to semi-tone values, vocal pitch can be empirically measured, and changes in the habitual pitch of the client, that is, the pitch at which the person speaks most of the time (Boone, 1971), can be monitored throughout the course of therapy. This procedure has the added attraction of being inexpensive, requiring no elaborate apparatus, and taking little time to perform. Unfortunately, to date no systematic

assessment of the reliability of this technique has been reported. Thus one purpose of the experiment reported in this paper was to assess the reliability of a modification of this procedure and appraise its potential for process research.

Conclusions and Hypotheses

This chapter has discussed the importance of acoustical properties of speech as process variables in psychotherapy. There is at least tentative evidence that voice qualities such as pitch, intensity, and rate of speaking convey affective information about a speaker, and there is promising evidence that the latter two properties as expressed by client and therapist interact in a mutually reinforcing manner in successful psychotherapy. The experiment reported in this paper was designed to look for a similar interaction between pitch levels and ranges of client and therapist in ongoing psychotherapy.

The specific hypotheses tested were as follows:

1. Mean pitch and pitch range of therapist and client are not independent of one another in the course of any given psychotherapy session. This was tested by examining whether pitch means or ranges of client and therapist were correlated in any way in any given session. High correlations (+0.70 or better) would indicate some interaction was occurring and would argue against functional independence.
2. Pitch range of client can be modified by extreme pitch range of therapist. This was tested by asking therapists to speak with either very broad or very narrow pitch ranges in the course of a single

therapy session and by looking for deviations from normal in the pitch ranges of the clients. Any deviation from normal would indicate some effect of therapist manipulation on client pitch. A deviation in the direction of the therapist manipulations would suggest some following or convergence by the client.

3. Mean pitch levels and ranges of therapist and client change across time as the number of therapy sessions increases. This was tested by examining the correlations between therapist and client pitch for each session as they varied across sessions, as well as their average pitch values across sessions. Any increasing or decreasing correlations or any convergence or divergence of average pitch levels across sessions would indicate some long term interactions of pitch between therapist and client. Any concomitant changes in therapeutic relationship, as measured independently, could be examined to determine if pitch changes are themselves a function of the relationship.

The superordinate purpose of all three hypotheses was to determine whether vocal pitch operates as a process variable in psychotherapy. The experiment then was not only designed to look for patterns in pitch fluctuations of two people engaged in psychotherapy, but also to examine the functional significance of any patterns that could be found. It represents only a first step, but a necessary step, toward understanding the role of vocal pitch in human communication.

CHAPTER III

METHOD

Subjects

Two therapist-client dyads, one consisting of two males and the other of two females, were selected for study. The therapists were two graduate students in a clinical psychology doctoral program. Each had over one year of graduate training and clinical experience. The clients were selected by the therapists from the University of Montana Clinical Psychology Center referral file under the practicum supervision of a faculty member. The female client was a 25 year old woman with the presenting problem of depression. The male client was a 27 year old man with the presenting problem of chronic feelings of social inadequacy. Neither client had been in psychotherapy before, and neither dyad had ever met before their first session. At the first session permission to record the therapy sessions for this research was obtained from each client.

All four subjects were U.S. citizens who spoke common American English. Hence any confounding due to cultural differences in pitch inflections (Fry, 1968; Lieberman, 1967) was avoided. The potential pitch ranges for each therapist were defined as the distance in musical half-tones (or semitones) from the lowest to the highest note they could hum. These tones were matched to those of a pitch pipe

and converted to their musical note values. The female therapist exhibited a potential range of 31 semitones on the musical scale (from E_3 to C_6^\sharp) and the male therapist exhibited a range of 28 semitones (from G_2 to B_4).

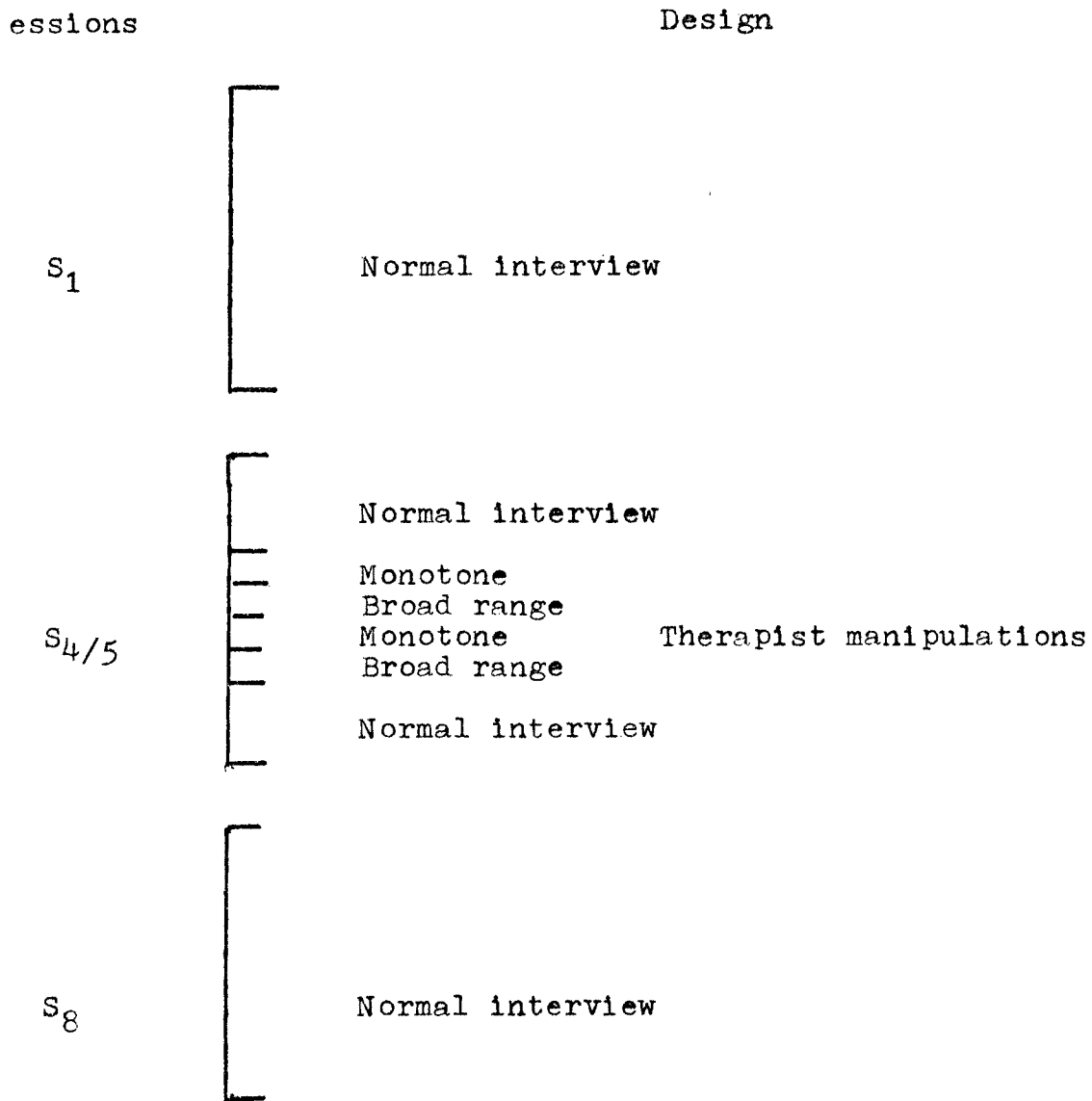
Apparatus

Psychotherapy sessions were recorded unobtrusively by means of hidden microphones feeding into a stereophonic tape recorder (Roberts 400 X) in the next room. During one psychotherapy session a red and green light were positioned in the psychotherapy room behind the client's chair but within viewing range of the therapist. These lights could be switched on and off from the next room and served as cues for therapist pitch inflections.

Procedure

The first, fifth, and eighth sessions of the female dyad and the first, fourth, and eighth sessions of the male dyad were recorded for pitch analysis (see Figure 1). During the middle sessions therapists were asked to restrict their pitch ranges to monotonic levels when signaled by the red light and to broaden their ranges as much as possible when signaled by the green light. They were further asked to make these manipulations without affecting the content of their speech. Each was given approximately one-half hour of practice in narrowing and broadening ranges prior to the therapy session. Under the experimental conditions the lights were presented in an ABCBCA reversal design (Leitenberg, 1973), illustrated in Figure 1. Each therapist was allowed 15 minutes of baseline interview at the beginning

Figure 1. Experimental design. Three sessions of psychotherapy for each dyad were recorded for pitch analysis. The first and eighth sessions consisted of normal psychotherapy interviews. In the middle session each therapist was asked to broaden or narrow his or her pitch range in an ABCBCA design of 15 min. of normal interview at the beginning and end of the session and 5 min. under each experimental condition.



of the session, then 5 minutes of red light and 5 minutes of green light, each repeated once, followed by a lights off or return to baseline condition at the end of the session.

Pitch Analysis

A recording of each psychotherapy session was parsed into a series of tape loops, each representing a segment of either therapist or client speech. Speech segments were selected on the basis of three criteria: 1) the segment must be from declarative statements to avoid confounding due to large pitch inflections characteristic of nondeclarative statements (Lieberman, 1967); 2) each speech sample must just precede or just follow a speech sample from the other communicant; 3) the series of speech segments for each communicant must be equally distributed across the psychotherapy session, preferably with one sample per every 2 minute interval. Tape loops meeting all three criteria were then assigned to either of two groups for pitch analysis: one in which therapist speech preceded client speech (T-C) and one in which therapist speech followed client speech (C-T).

After screening several candidates, three female undergraduates, each with formal training and experience in vocal music, were selected to be judges of pitch. All three were able to hum pitch levels within the speaking ranges of the female subjects (D_3 to F_4) easily. The humming ranges of the judges differed as to lower limits, however, with Judge #1 exhibiting a lower limit at D_3 , Judge #2 at C_3 , and Judge #3 at A_2 . None of the judges were capable of humming the lower limits of the male voice ranges (G_2 to A_3), but they were capable of identifying tones as being exactly one octave higher or

lower than a target tone produced by the electronic tone generator. All three judges were screened for hearing loss at the University of Montana Speech, Language, and Hearing Clinic, and all exhibited normal hearing. Judges were trained to match pitch levels from speech segments to pure tones produced by an electronic tone generator (Audio Oscillator, Hewlett-Packard 2000R) until at least a 75% agreement was reached among all three of them.

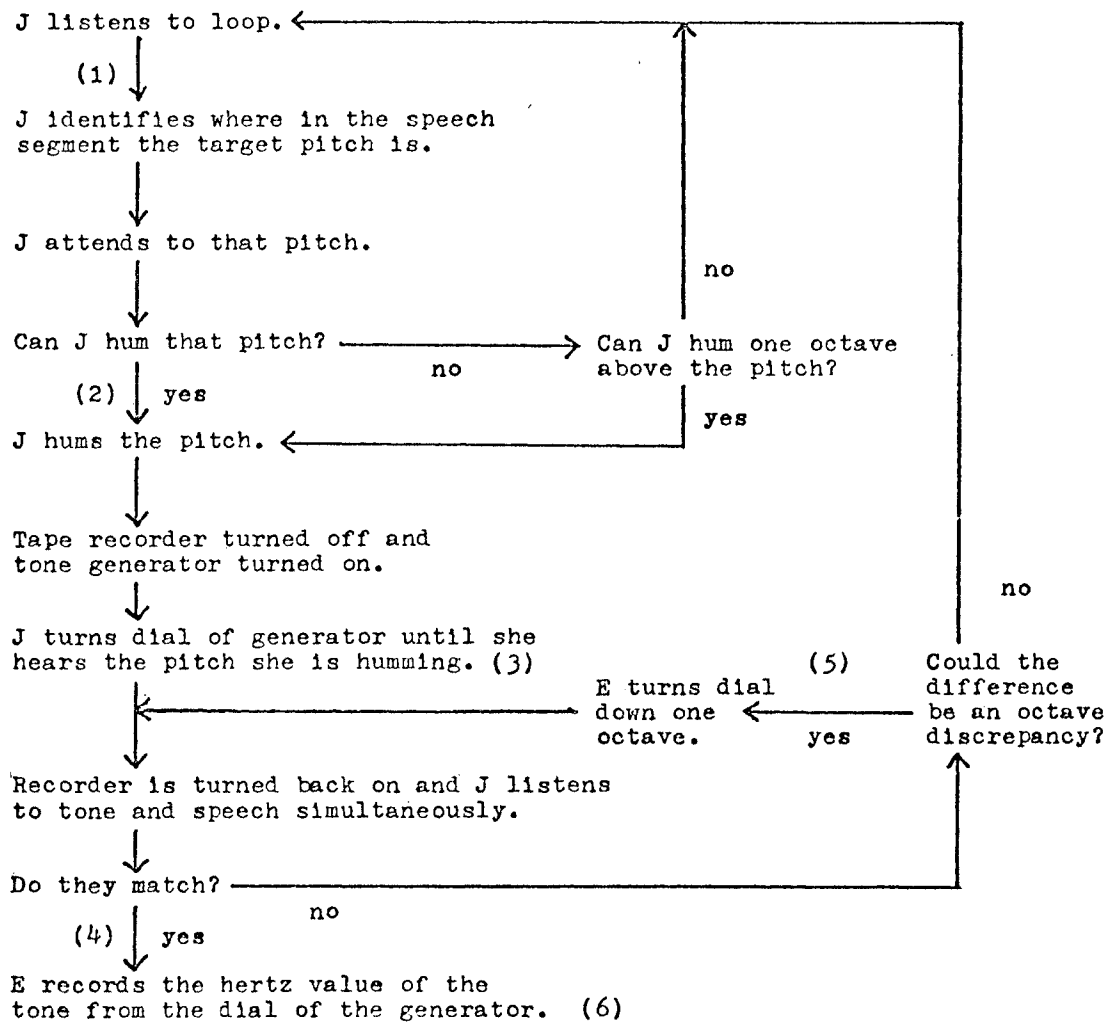
Pitch analysis was conducted as follows. Each tape loop was threaded through a stereophonic tape recorder (AKAI 1722 W) so that the speech segment played repeatedly when the machine was turned on. The continuously repeating speech was fed into one ear of a headphone set. Pure tones ranging in frequency from 60-600 hertz (hz) were conducted from the tone generator to the other ear of the headphones. Each judge was allowed to control both the frequency of tones produced and their intensity by turning the appropriate dials on the generator. For each loop the judge was asked to determine the maximum, minimum, and modal pitch of the segment. When one of these target pitches was identified, the judge was asked to fix the pitch in her own mind by humming it aloud or subvocally and then to turn the frequency dial until she found a tone that matched her internal representation of the pitch. An accuracy check was then made by having her simultaneously listen to the pure tone in one ear and the speech segment in the other to adjust the generator for a positive match. After each positive match the experimenter recorded the frequency value of the tone from the face of the tone generator. A slightly different procedure was necessary for those male speech samples that were out of the ranges

of the judges: instead of trying to hum or fix in her mind the actual pitch of the male voice, each judge was asked to hum the same note an octave higher to bring it within her own range, match that pitch with a pure tone, and then listen to a one octave lower pure tone simultaneously with the male voice for a positive match. The complete matching procedure is illustrated algorithmically in Figure 2.

Pitch levels were converted from their hertz values to semitone values, which transforms the pitch measurements from an exponential frequency scale to a linear musical scale and allows for equal comparisons between upper and lower pitch ranges. Pitch range was calculated by subtracting minima from maxima. Mean pitch was calculated as the arithmetic mean of each range. Thus the raw data yielded three dependent measures of pitch: mean, mode, and range.

Tape loops were divided into male and female groups and randomized for presentation to the judges. Each judge rated approximately one-third of the loops in each group. They were informed that some of the segments would be rated by all three judges but were not told which ones, and each was asked to maintain the same set in judging the experimental loops that she had acquired in training. In fact, every fifth loop presented was rated by all three judges. Inter-judge reliability was calculated as the percent agreement within one full tone (2 semitones) for each pair of judges per subject, averaged across the three pairs of judges. In this way one could compare rates of agreement among judges for each of the subjects.

Figure 2. A flow chart of the pitch measurement procedure used in this experiment. J = Judge; E = Experimenter. Numbers refer to possible sources of error in making the measurements. 1) Deciding where the target pitch is. 2) Humming the pitch accurately. 3) Matching the generated tone with what one is humming accurately. 4) Deciding whether the tone and the vocal pitch match. 5) Deciding whether the difference is simply an octave discrepancy. 6) Reading the generator dial correctly.



Other Dependent Variables

In addition to the pitch analysis, two measures of verbal content and therapeutic relationship were taken. First, transcripts of speech for all 6 psychotherapy sessions were typed for a post hoc content analysis. Second, modified versions of the Client and Therapist Personal Reaction Questionnaire (Ashby et al., 1957; Goldstein, 1971) were administered to clients and therapists immediately after each recorded session. These questionnaires were comprised of 40 personalized statements about the psychotherapy experience which the respondent could endorse or reject on a Likkert type 5 point scale from -2 to +2 (strongly disagree, disagree, no opinion, agree, strongly agree). The statements were afterwards divided into groups of positive and negative reactions (Ashby et al., 1957) and each response was assigned a positive value if it agreed with a positive statement or disagreed with a negative statement, or a negative value if it agreed with a negative statement or disagreed with a positive one. These values were weighted as to the strength of each response, summed for each session, and converted to a standard score based on the percent of maximum positive or maximum negative responding possible.

CHAPTER IV

RESULTS

Reliability

Percent agreement varied across pairs of judges, across subjects, and across dependent variables (see Table 1). All but three of the percent agreement values reached 75% or better. Poorest agreement occurred for the measurements of mode and range of male therapist pitch (58% and 42%, respectively) and for the measurements of mode for female client pitch (69%). In judging segments of male therapist speech, all judges expressed difficulty and frustration in attempting to match their voices to his. Humming the pitch in their own ranges and then matching his pitch with the same tone an octave lower on the generator did not help, as the judges expressed incredulity that the voice was actually as low as the tone on the generator. This situation resulted the judges' dissatisfaction with their own ratings of this voice. For this reason and because of the lowered reliability of ratings for the male voices, only speech segments from the male dyad's fourth session, in which specific therapist manipulations of voice were attempted, were included for analysis.

Agreement among judges for measures of mode was generally poorer than for measures of maxima and minima from which pitch means and ranges were calculated. This was true for male and female

TABLE I

PERCENT AGREEMENT AMONG THREE JUDGES (1,2,3) COMPUTED
ACROSS EACH PAIR OF JUDGES, FOR EACH SUBJECT AND
EACH DEPENDENT VARIABLE. T = THERAPIST;
C = CLIENT; M = MALE SUBJECT; F =
FEMALE SUBJECT.

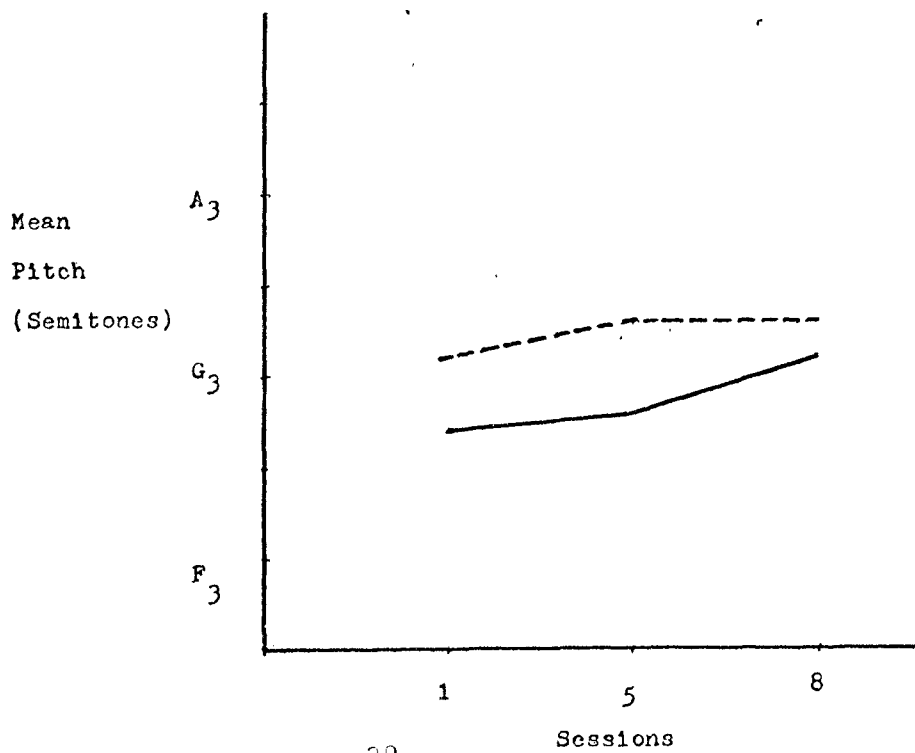
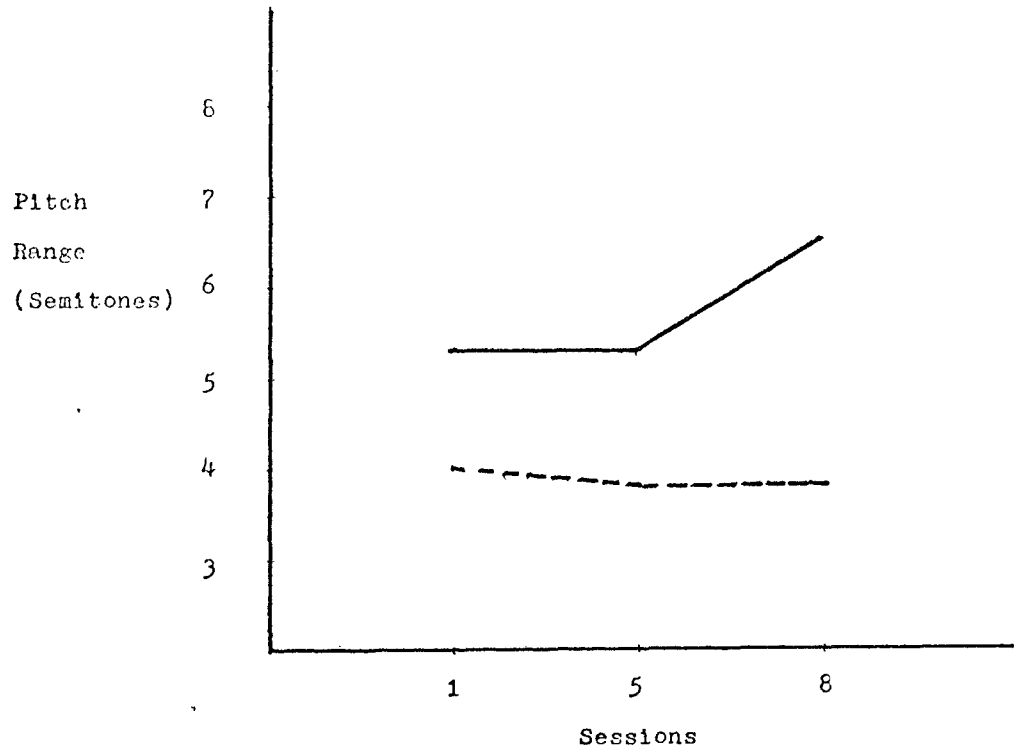
Judges =		MODE			RANGE			MEAN		
		1,2	1,3	2,3	1,2	1,3	2,3	1,2	1,3	2,3
M	T	100	25	50	100	0	25	100	50	75
		$\bar{x} = 58$			$\bar{x} = 42$			$\bar{x} = 75$		
	C	100	80	80	80	100	60	80	80	80
		$\bar{x} = 87$			$\bar{x} = 80$			$\bar{x} = 80$		
F	T	78	89	78	67	100	67	100	100	100
		$\bar{x} = 82$			$\bar{x} = 78$			$\bar{x} = 100$		
	C	100	56	50	78	100	90	100	89	90
		$\bar{x} = 69$			$\bar{x} = 89$			$\bar{x} = 93$		

subjects alike. The judges reported having difficulty determining which pitch in a given speech segment was heard most frequently. Sometimes two different pitch levels were perceived to be equally modal, and the judge had to arbitrarily pick one as the mode. Because of these difficulties in discriminating pitch modes, only pitch means and ranges were included for data analysis.

Pitch Interactions in the Female Dyad

Results of pitch analysis failed to support the hypothesis that mean pitch levels or pitch ranges would converge or diverge across sessions. Average pitch means and pitch ranges for therapist and

Figure 3. Average pitch means and ranges for the female therapist (dashed lines) and client (solid lines) across 3 sessions of psychotherapy.



client are plotted across sessions in Figure 3. Analysis of variance showed no significant sessions effect for either pitch means, $F(2, 148) = 2.07$, $p > 0.10$, or pitch ranges, $F(2, 148) = 1.28$, $p > 0.25$. A significant subjects effect was found for both pitch means, $F(1, 74) = 7.38$, $p < 0.05$, and pitch ranges, $F(1, 74) = 20.67$, $p < 0.05$. The therapist's voice exhibited generally higher pitch levels than did the client's voice across all sessions. The client spoke with generally broader pitch ranges than did the therapist across all sessions.

A mathematical description of the temporal characteristics of pitch inflections for each person across each therapy session was accomplished by means of time series analysis. A complete description of time series analysis and its applications can be found elsewhere (Glass, Willson, & Gottman, 1975), but it essentially enables one to assess the temporal characteristics of a dependent variable as it varies across time. As used in this study, the analysis consisted of computing autocorrelations and partial autocorrelations for each series of pitch inflections for lags 1 through N-1, where N equaled the number of pitch levels in the series. These computations were made for 0 through 4 orders of differencing, which for any given series was performed by subtracting the pitch value at time t from that at time t+1 (first order differencing), subtracting the value at time t from that at t+2 (second order differencing), and so forth. Differencing determines whether or not a series is stationary, that is, whether or not it oscillates regularly about a given pitch level. Nonstationary series yield series of non-zero values after first

order differencing and tend to drift widely from any given level across long periods of time (Glass et al., 1975). All computations were made by means of a DEC-10 digital computer. Plots of the autocorrelations and partial autocorrelations after first order differencing for each subject's mean pitch and pitch range in session 1 are presented in Figure 4, those for session 5 in Figure 5, and those for session 8 in Figure 6. All curves shift from negative to positive, with the autocorrelations generally climbing to within one standard error of zero abruptly after lag 1 and the partial autocorrelations approaching zero more gradually. All the time series depicted in Figures 4, 5, and 6 can best be described by an integrated moving averages or an ARIMA (0,1,1) model, which exhibits no autoregressive components and one moving averages component after first order differencing (Glass et al., 1975).

Results of the correlational analyses of therapist and client pitch fluctuations failed to reject the null hypothesis that the fluctuations are independent of each other. Pitch means and ranges for session 1 are plotted in Figure 7, those for session 5 in Figure 8, and those for session 8 in Figure 9. Visual inspection of the data is sufficient to reveal no salient relationships between therapist and client pitch levels, either when the therapist was responding to the client or when the client was responding to the therapist. Correlation coefficients computed at lags 0 and 1 are presented for each pair of curves. A lag 1 coefficient was computed because of the strong lag 1 component of the autocorrelations

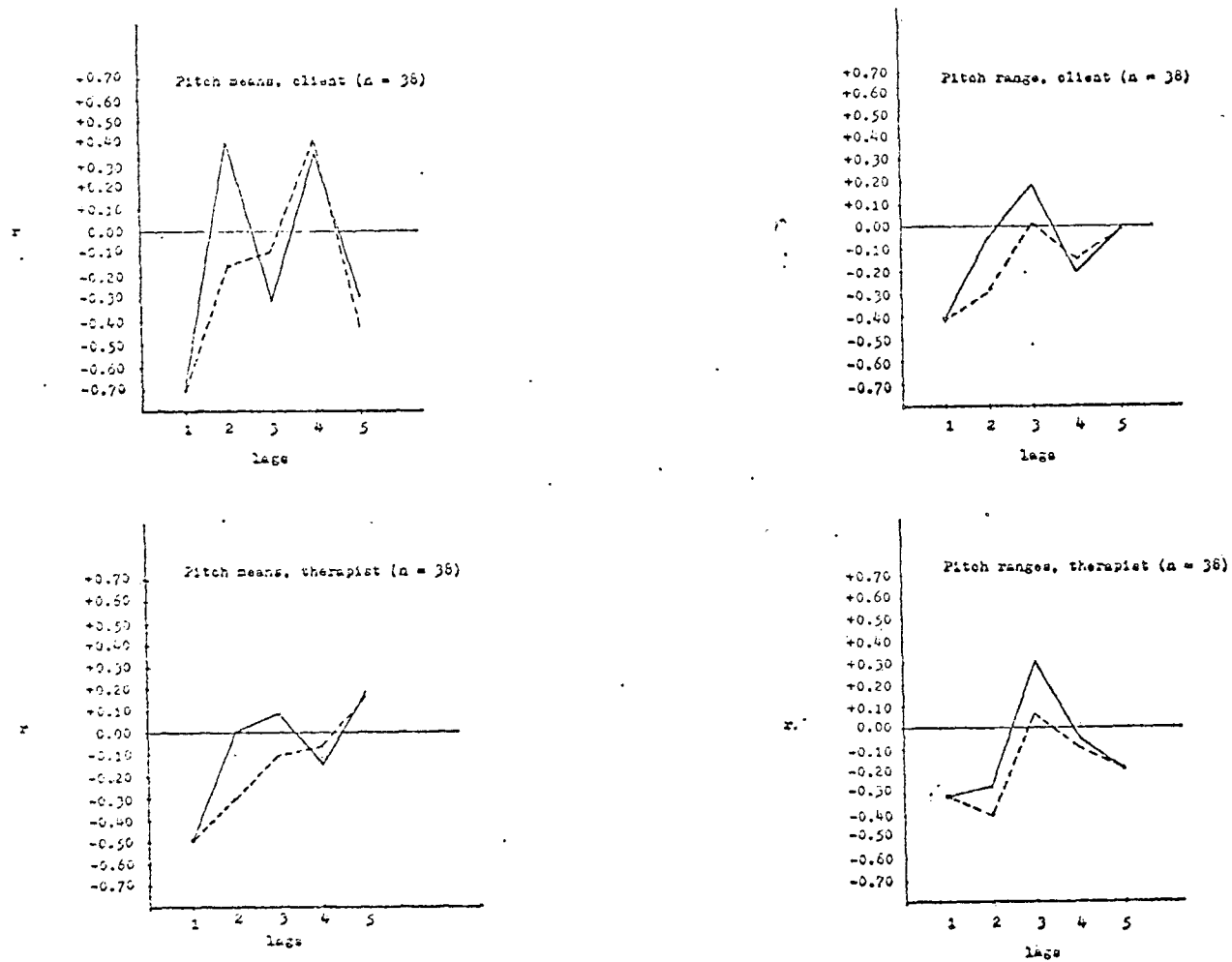


Figure 4. Plots of the first 5 lags of autocorrelations (solid lines) and partial autocorrelations (dashed lines) after first order differencing for female therapist and client mean pitch and pitch range series in session 1.

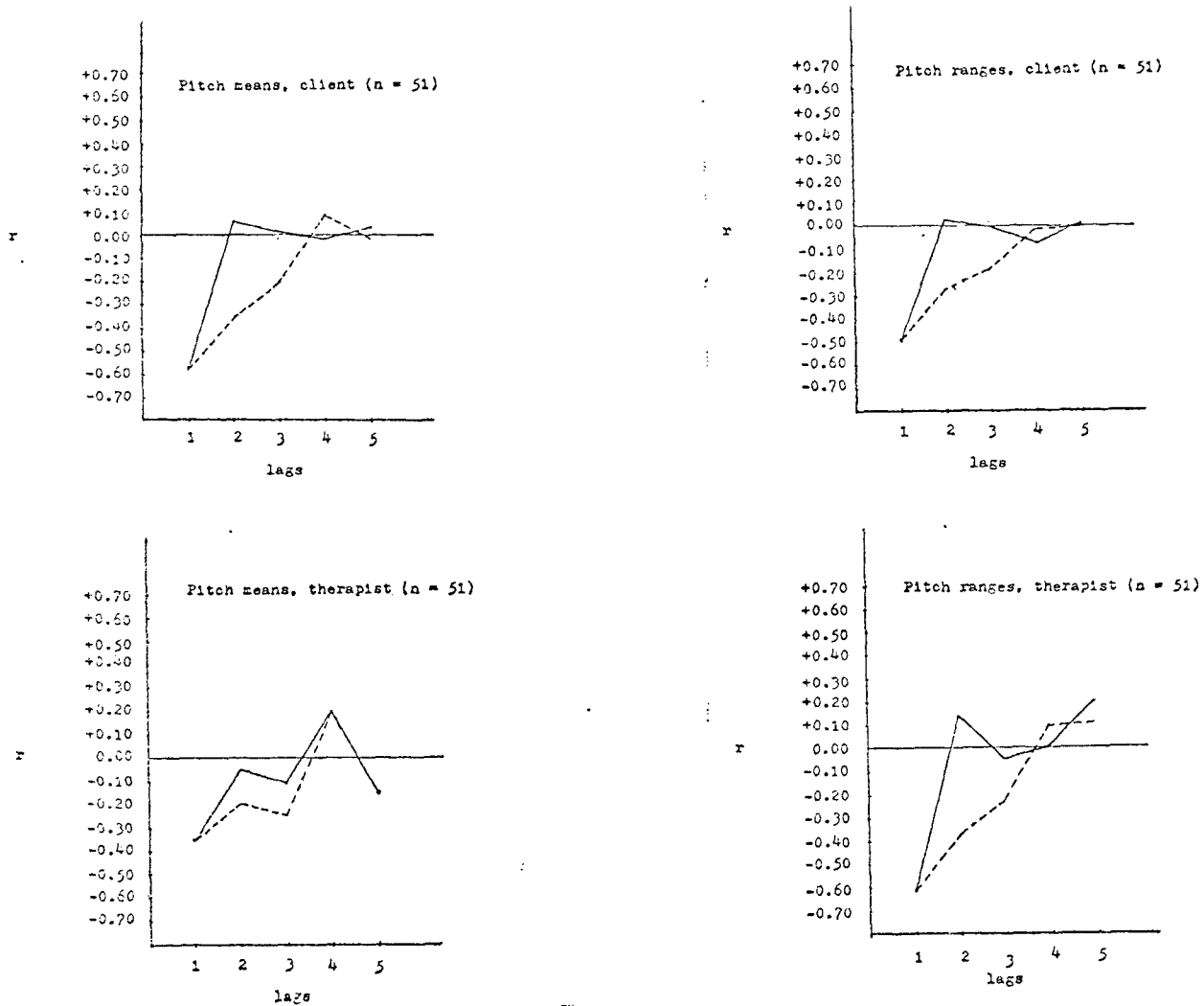


Figure 5. Plots of the autocorrelations (solid lines) and partial autocorrelations (dashed lines) after first order differencing for female therapist and client mean pitch and pitch range series in session 5.

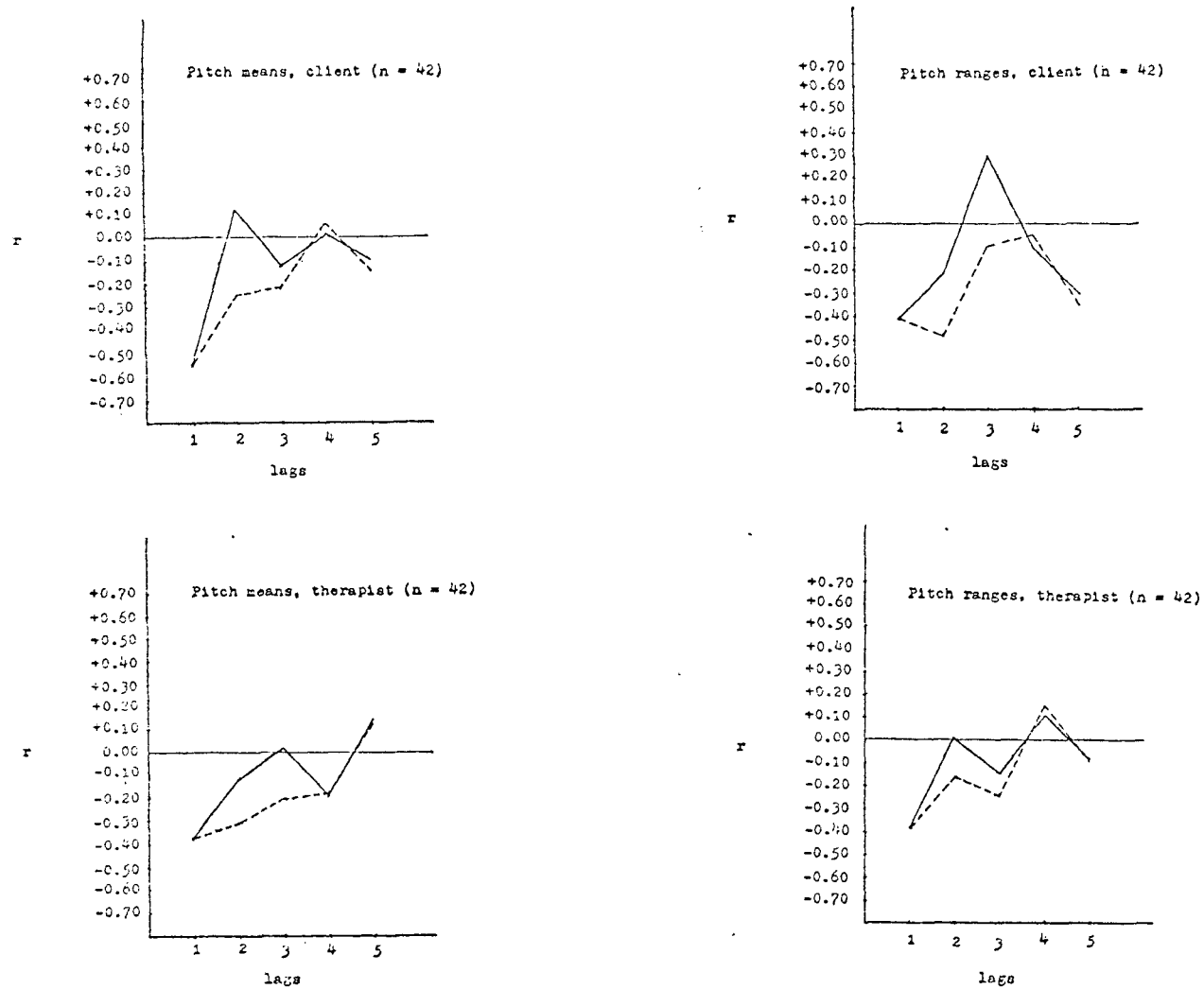


Figure 6. Plots of the autocorrelations (solid lines) and partial autocorrelations (dashed lines) after first order differencing for female therapist and client mean pitch and pitch range series in session 8.

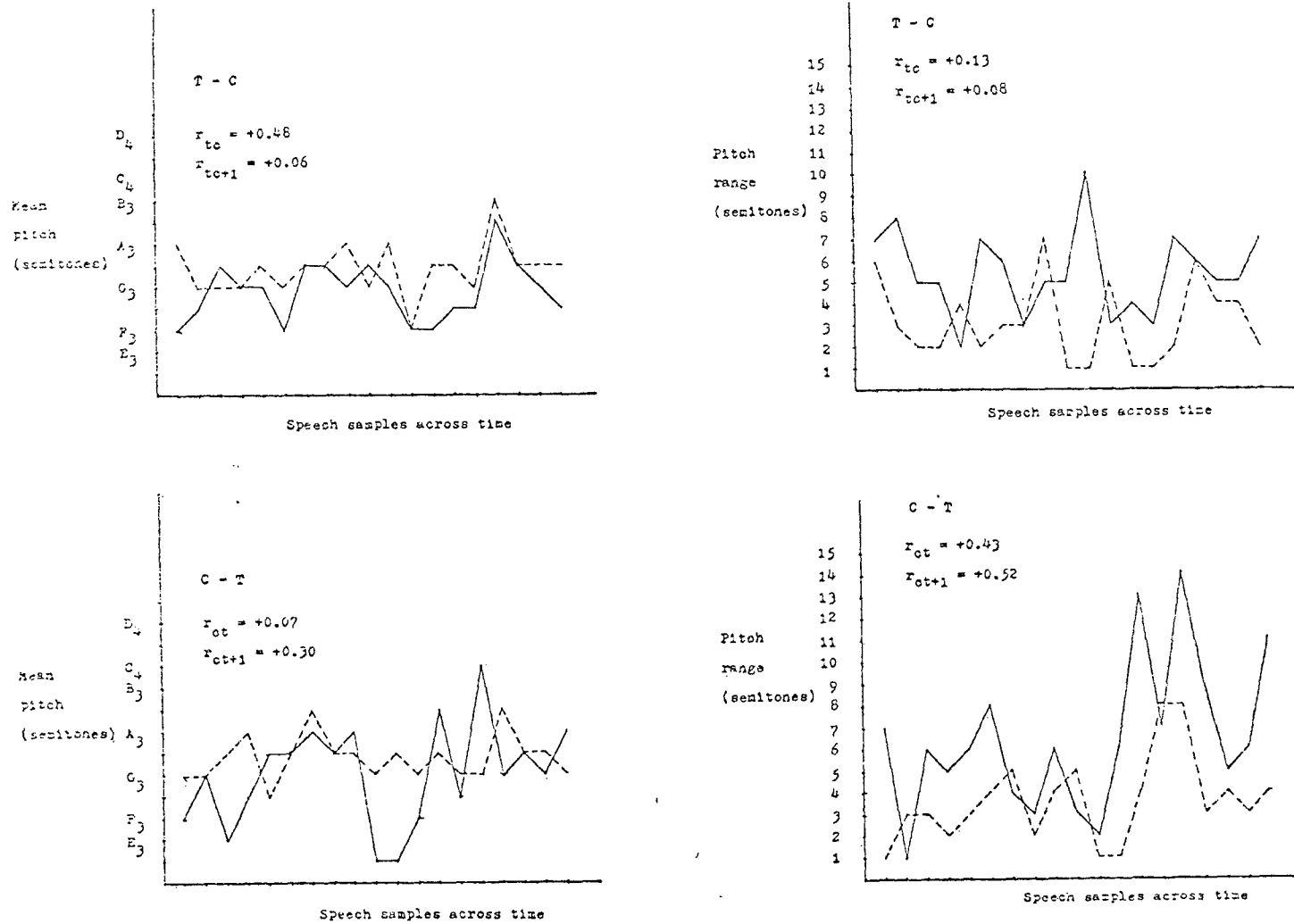


Figure 7. Mean pitch and pitch range fluctuations of female therapist (dashed lines) and client (solid lines) sampled across session 1. T-C: Client statements followed therapist statements; C-T: Therapist statements followed client statements.

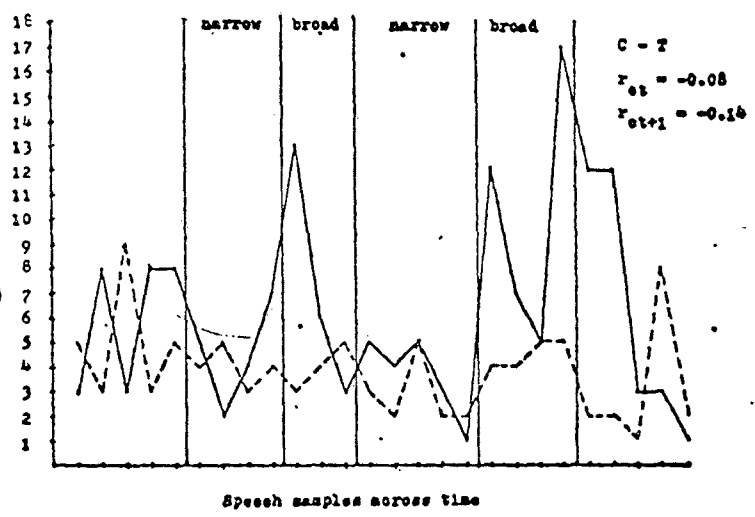
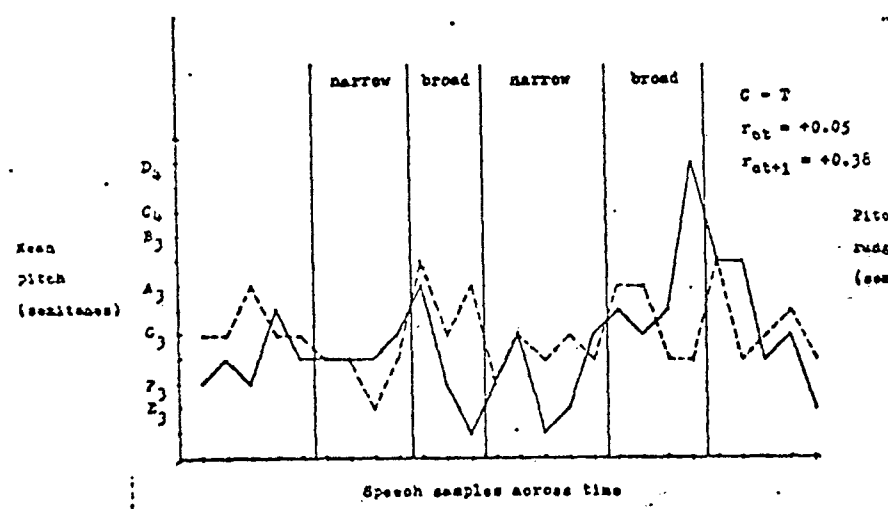
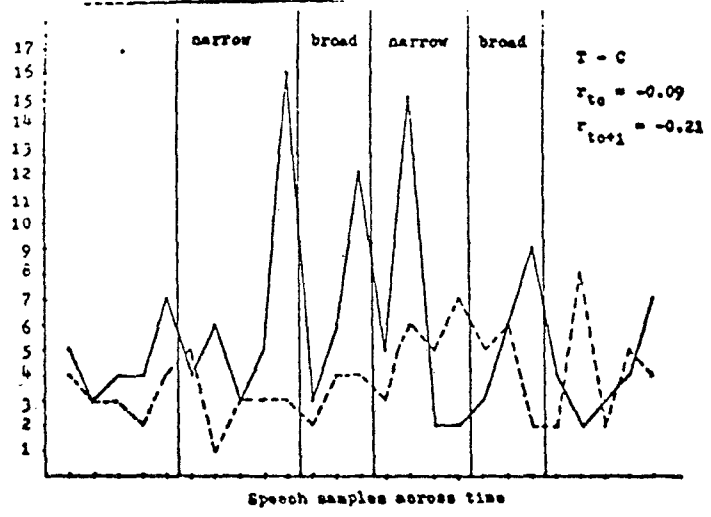
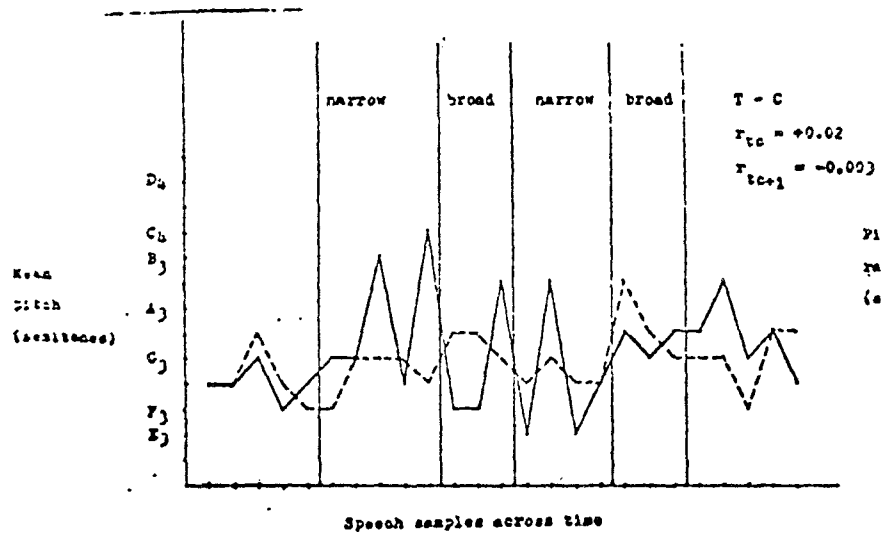


Figure 8. Mean pitch and pitch range fluctuations of female therapist (dashed lines) and client (solid lines) sampled across session 5. T-C: Client statements followed therapist statements; C-T: Therapist statements followed client statements. Dividing lines show the time intervals under each experimental condition governing therapist pitch range.

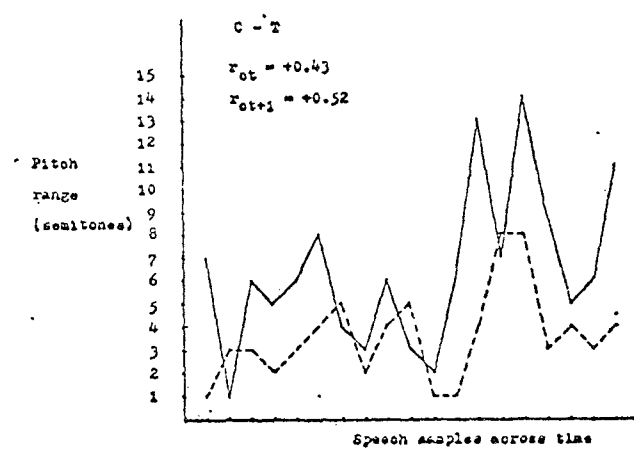
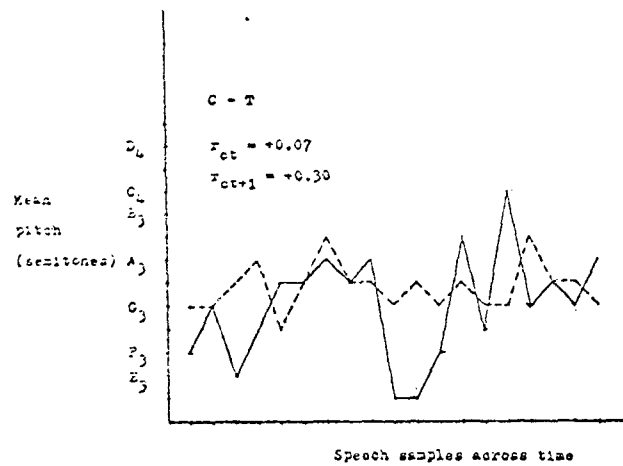
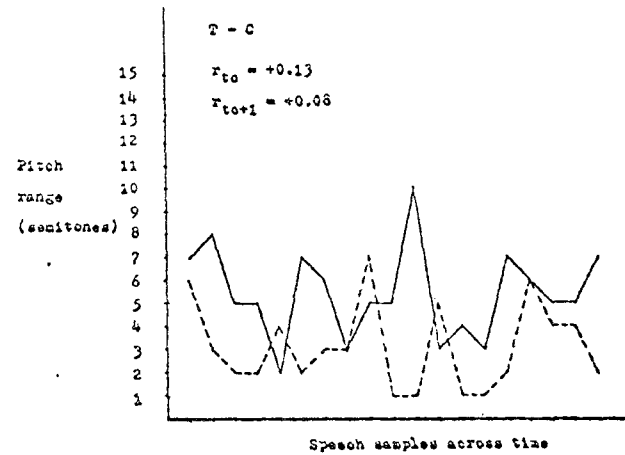
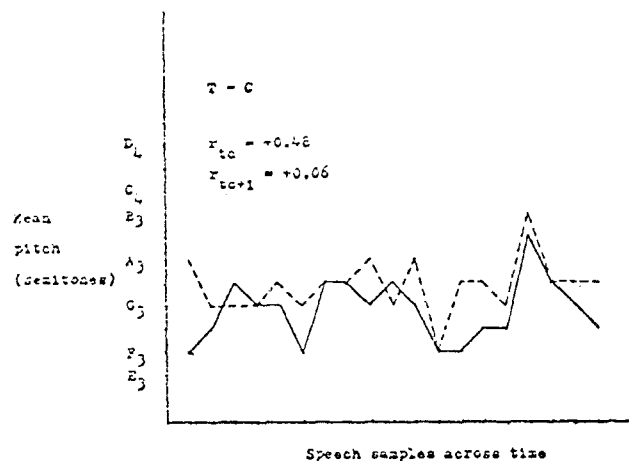
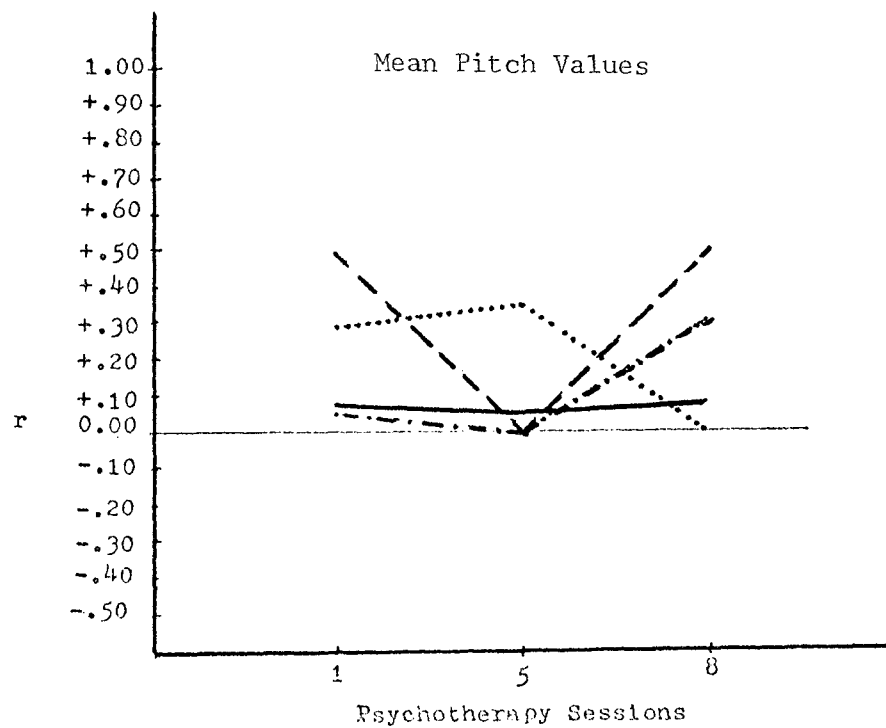
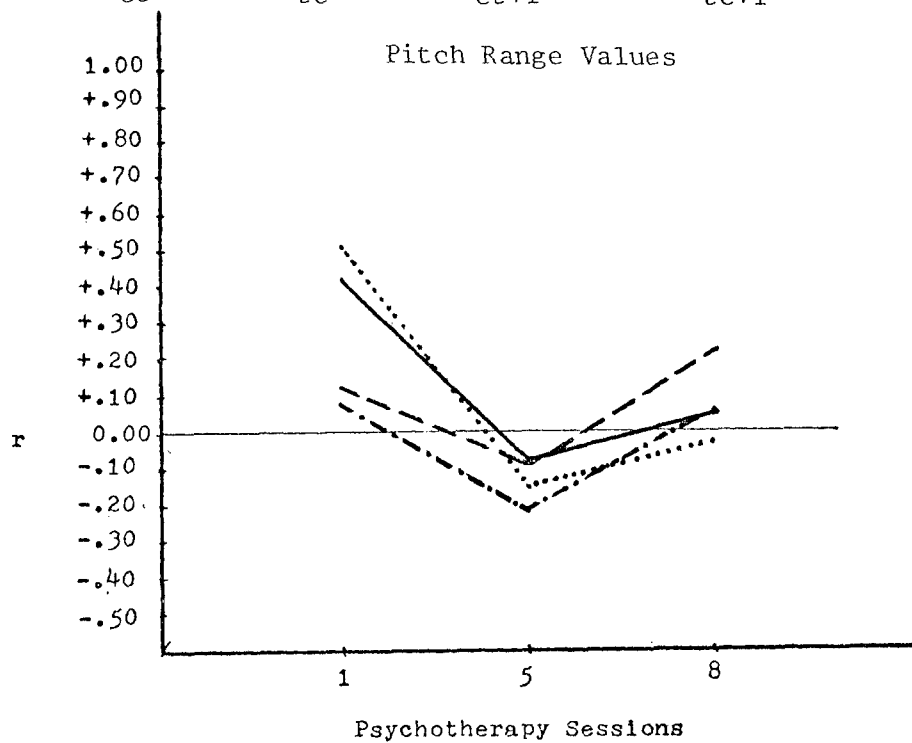


Figure 9. Mean pitch and pitch range fluctuations of female therapist (dashed lines) and client (solid lines) sampled across session 8. T-C: Client statements followed therapist statements; C-T: Therapist statements followed client statements.

found in the time series analysis. No pair of curves is strongly correlated at either lag 0 or lag 1, and no correlation was able to account for more than 27% of the variance in pitch inflections in any session. Moderate correlations were found between therapist and client pitch means when client statements followed therapist statements (T-C) in session 1 ($r_{tc} = +0.48$) and session 8 ($r_{tc} = +0.48$), and between therapist and client pitch ranges when therapist statements followed client statements (C-T) in session 1 ($r_{ct+1} = +0.52$). All other correlations were either weak or nonexistent. A plot of the correlation coefficients for C-T and T-C statements is presented for pitch means and ranges in Figure 10. A chi square test of the null hypothesis that $r_1 = r_5 = r_8$ for each correlation coefficient across sessions (Marascuilo, 1971) was performed on those coefficients which appeared to vary the most across sessions. No trend in the correlations of pitch means was significant, and although all correlations between pitch ranges exhibited the same V pattern across sessions, only one trend approached statistical significance: pitch ranges of therapist statements which followed client statements in session 1 were more highly correlated at lag 1 with the ranges of client statements than they were in later therapy sessions, $\chi^2 (2) = 5.15, p < 0.10$.

A chi square test of independence for a 3X3 contingency table (Marascuilo, 1971) also failed to reject the null hypothesis that pitch inflections of therapist and client are independent of one another. This test was conducted by first defining the direction

Figure 10. Correlation coefficients computed at lags 0 and 1 for female therapist and client mean pitch and pitch range for each of 3 sessions of psychotherapy. Coefficients represent correlations where therapist statements followed client statements (r_{ct}) and where client statements followed therapist statements (r_{tc}).
 $r_{ct} = \text{---}$, $r_{tc} = \text{---}$, $r_{ct+1} = \text{.....}$, $r_{tc+1} = \text{-.-.-}$.



of a given pitch change from time t to $t+1$ as either an increase, a decrease, or no change, and then counting the number of such direction changes for each pair of curves in Figures 7, 8, and 9. These frequency counts were next summarized in 3X3 contingency tables from which the estimated frequencies of each direction change could be computed. An example of the contingency and estimated frequency tables is presented in Table 2. Note that only directionality of pitch changes was utilized in this analysis; amplitude of pitch changes was not considered. Table 3 presents the results of the chi square test for all data in the female psychotherapy sessions. In no case could changes in directionality by one communicant be predicted from those of the other.

The hypothesis that extreme pitch ranges in therapist speech can influence pitch ranges of client speech could not be tested because the therapist was unable to broaden or narrow her pitch range beyond her normal variability. Visual inspection of both the T-C and C-T graphs of pitch range in Figure 8 reveals that therapist pitch ranges did not vary more than 8 semitones (or 26% of her potential range) during session 5 and moreover that this variation was not noticeably affected by the experimental conditions. There was some tendency for mean pitch to vary consistently under the experimental conditions: pitch means rose when the therapist was signaled to broaden her range and fell when she was signaled to narrow her range. The therapist herself reported after the session that she believed she had followed instructions correctly and had in

TABLE 2

A 3X3 contingency table and a 3X3 estimated frequency table of directional changes in the pitch inflections of the female therapist as a function of such changes by the female client in session 1. The first matrix summarizes the number of increases (↑), decreases (↓), and no changes (-) in the range fluctuations of therapist voice (T) that followed each change of direction by the client (C). The second matrix presents the estimated expected frequencies of directional changes for the same data. A chi square test of independence failed to reject the null hypothesis that directionality of therapist inflections is independent of that of client inflections, $\chi^2 (4) = 5.96, p > 0.05$.

	T					T			
		↑	↓	-			↑	↓	-
C	↑	5	5	0	10	↑	5.0	5.0	0
	↓	1	4	0	5	↓	2.5	2.5	0
	-	3	0	0	3	-	1.5	1.5	0
		9	9	0	18				

TABLE 3

Chi square tests of independence of directional changes between pitch inflections of therapist and client for 3 sessions of psychotherapy. C-T: tests of independence in those cases where therapist changes followed those of the client. T-C: tests of independence for those cases in which client changes of direction followed those of the therapist.

Sessions	C - T		T - C	
	Mean	Range	Mean	Range
1	$\chi^2 = 3.96$ $p > 0.05$	$\chi^2 = 5.96$ $p > 0.05$	$\chi^2 = 0.75$ $p > 0.05$	$\chi^2 = 5.79$ $p > 0.05$
5	$\chi^2 = 3.54$ $p > 0.05$	$\chi^2 = 2.95$ $p > 0.05$	$\chi^2 = 5.61$ $p > 0.05$	$\chi^2 = 1.47$ $p > 0.05$
8	$\chi^2 = 3.64$ $p > 0.05$	$\chi^2 = 0.92$ $p > 0.05$	$\chi^2 = 6.68$ $p > 0.05$	$\chi^2 = 1.94$ $p > 0.05$

fact been apprehensive that the client might consciously notice the change in her voice.

Pitch Interactions in the Male Dyad

In general, an analysis of the male pitch interactions supports those of the female pitch interactions in failing to reject the null hypotheses. The data for the fourth psychotherapy session are presented in Figure 11. A time series analysis was performed on therapist and client mean pitch and pitch range series in this session, and plots of the autocorrelations and partial autocorrelations after first order differencing are presented in Figure 12. These curves are quite similar in form to those for the female data, and again an integrated moving averages model was deemed most appropriate for these pitch series.

Correlation coefficients between therapist and client pitch means and ranges for lags 0 and 1 are also presented in Figure 11. One strong relationship between therapist and client ranges when therapist statements followed client statements was found ($r_{ct} = +0.72$). Other correlations were either weak or nonexistent.

A chi square test of independence for a 3X3 contingency table was conducted for directionality changes in male pitch inflections and the results are presented in Table 4. In no case could changes in directionality by one communicant be predicted from those of the other.

Inspection of the graphs in Figure 11 reveals that the male therapist also had difficulty broadening and dampening his pitch

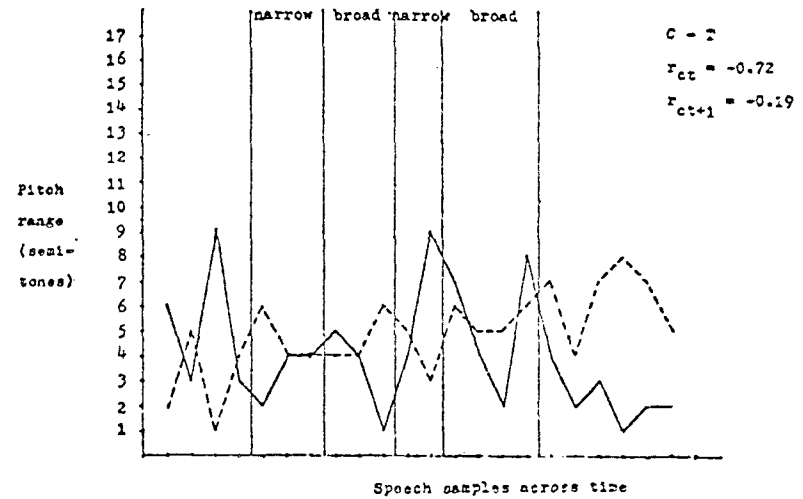
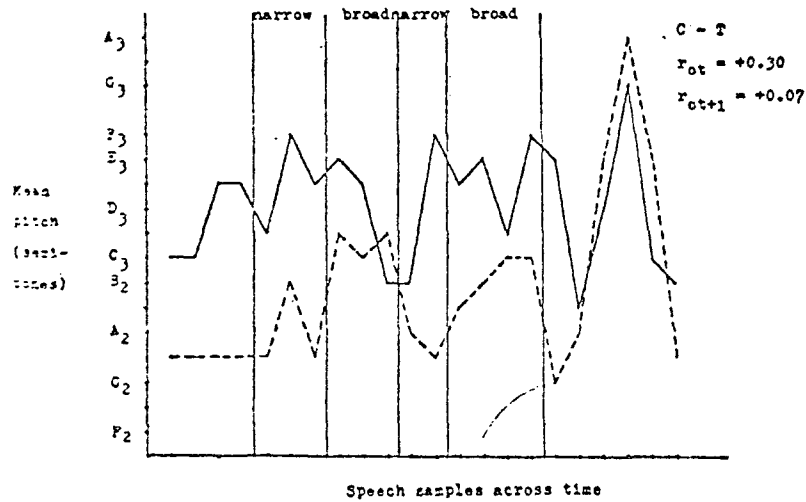
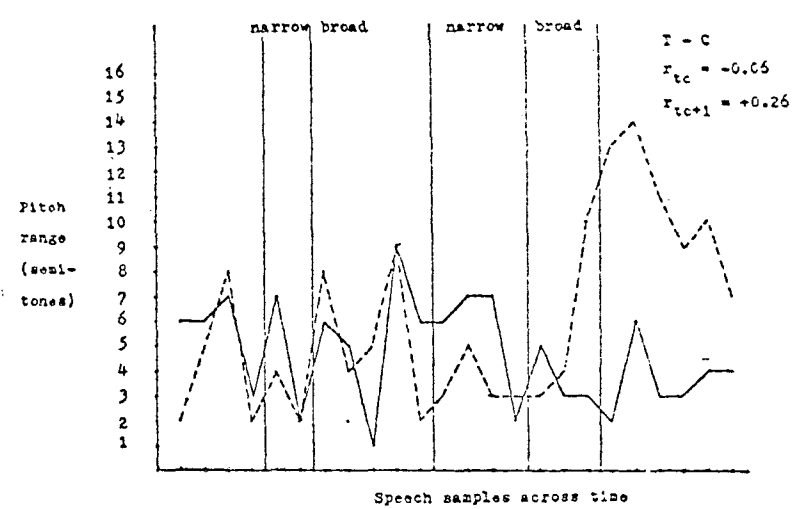
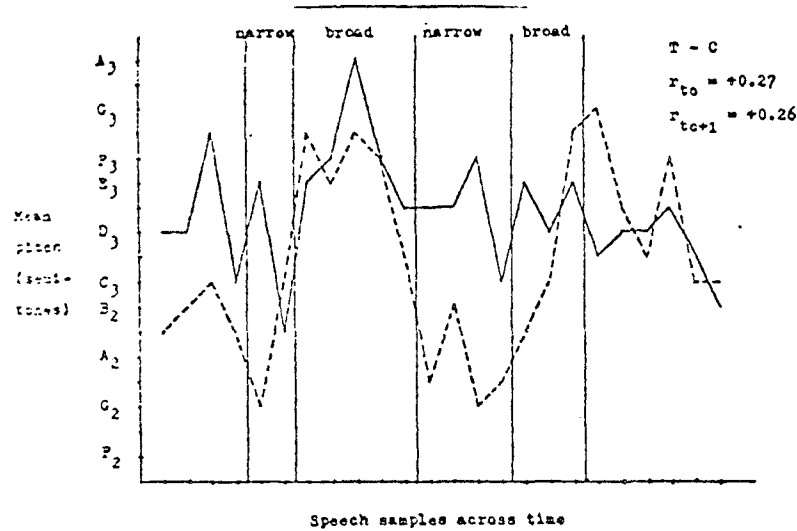


Figure 11. Mean pitch and pitch range fluctuations of male therapist (dashed lines) and client (solid lines) sampled across session 4. T-C: Client statements followed therapist statements; C-T: Therapist statements followed client statements. Dividing lines show time intervals under each experimental condition governing therapist pitch range.

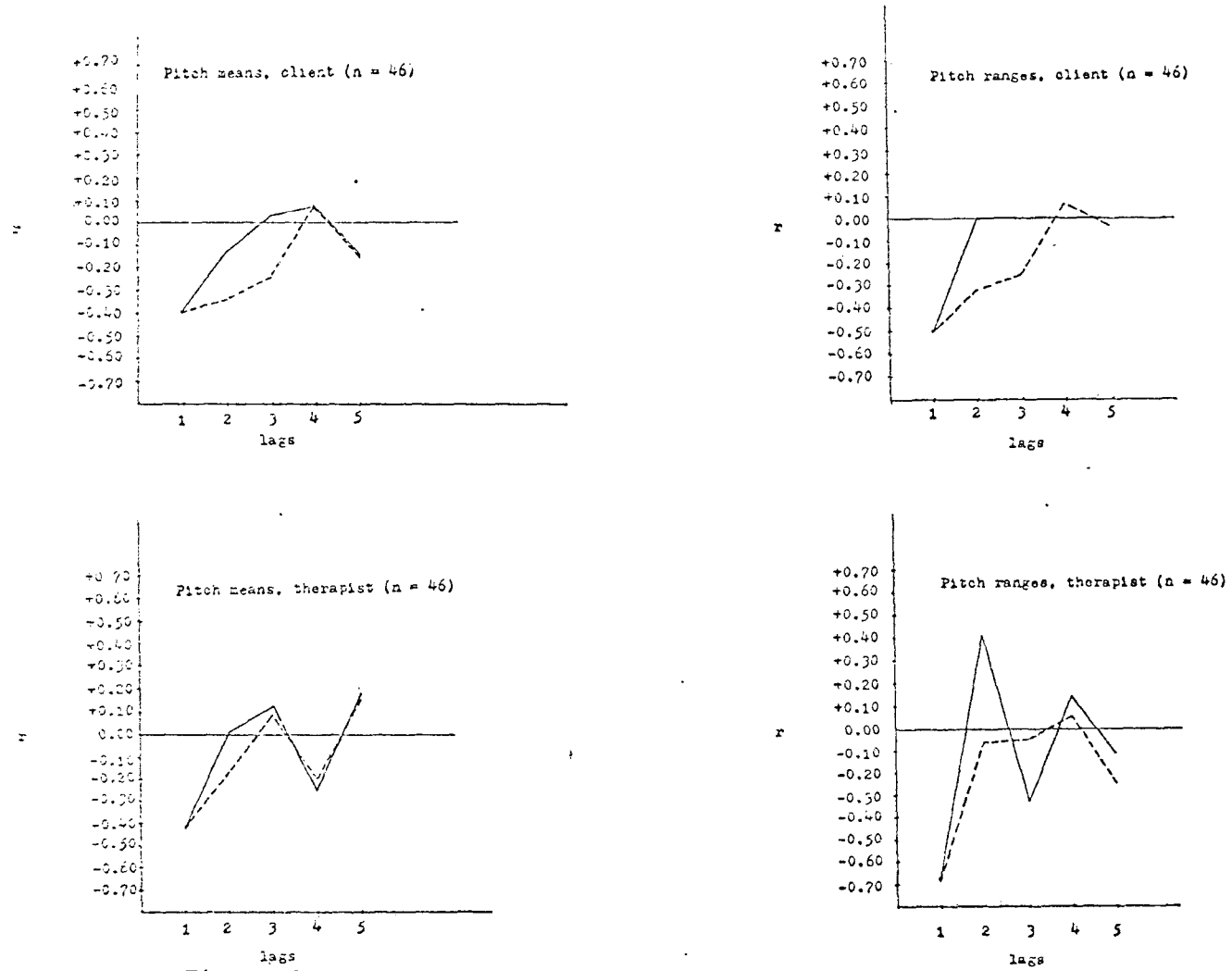


Figure 12. Plots of the first 5 lags of autocorrelations (solid lines) and partial autocorrelations (dashed lines) after first order differencing for male therapist and client mean pitch and pitch range series in session 4.

TABLE 4

Chi square tests of independence for 3X3 contingency tables of directionality of pitch changes for male psychotherapy dyad. The tests were conducted for directionality of mean pitch and pitch range changes for those segments in which therapist statements preceded client statements (T-C) and those in which therapist statements followed client statements (C-T).

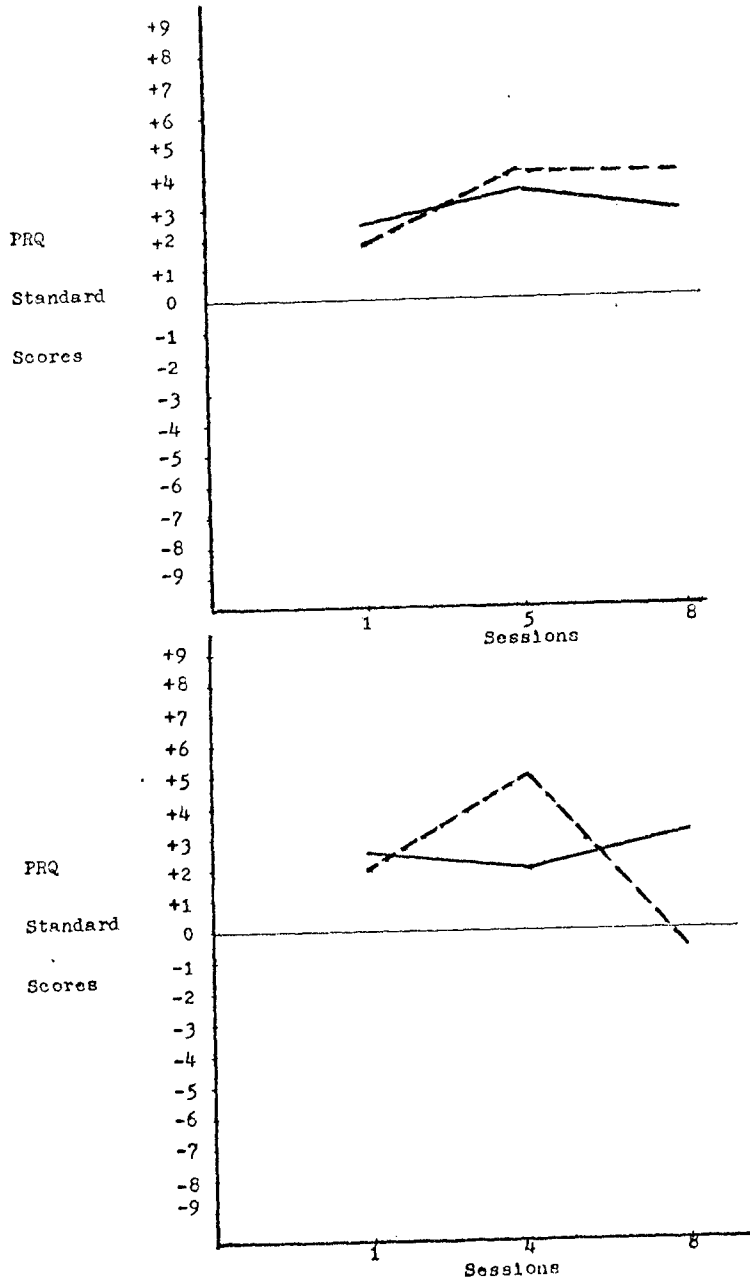
C - T		T - C	
Mean	Range	Mean	Range
$\underline{\chi}^2 = 7.06$	$\underline{\chi}^2 = 6.35$	$\underline{\chi}^2 = 1.84$	$\underline{\chi}^2 = 6.06$
$\underline{p} > 0.05$	$\underline{p} > 0.05$	$\underline{p} > 0.05$	$\underline{p} > 0.05$

range when cued to do so. Like the female therapist, he instead tended to raise the pitch of his voice when signaled to broaden his range and lowered it when signaled to narrow his range. However, unlike the female therapist, the male therapist did broaden his range considerably beyond normal after the experimental conditions had ended, especially for his statements that preceded those of the client (T-C). The therapist reported after the session that he believed he was following the instructions correctly while in the interview. When later informed of the results for this session, he mentioned that he had also been trying to be more emphatic with the client towards the end of the hour. Whatever the cause of the broadened pitch range, it is apparent from the data in Figure 11 that it did not noticeable influence the pitch range of the client.

Other Dependent Variables

Responses to the Therapist and Client Personal Reaction Questionnaires are summarized as a standard score per person per session and are plotted in Figure 13. In general, both dyads began psychotherapy with moderately positive feelings about the experience. By mid-course, however, the male dyad exhibited some divergence of feelings, with the therapist appearing more optimistic about treatment than the client. The transcript of this fourth session also reveals this discrepancy in levels of optimism. Much of the latter half of the session is occupied with the therapist offering the client positive self-statements in an attempt to increase the client's own positive regard and with the client resisting this strategy. Typical therapist statements included, "Now you see, that's what you're saying to yourself"; "Once you start doing it, and realizing that people are responding to you positively, then you'll start feeling happy about it too"; "You know you can do it"; and "You can change your thoughts". Typical client statements included, "I know I should be positive but..."; "I'm not sure we're getting anywhere"; and "I feel I can't do it". By session 8 the TPRQ and CPRQ responding in the male dyad had reversed, with the therapist exhibiting more ambivalent feelings about treatment. This psychotherapy dyad terminated at the client's request after the eleventh session. TPRQ and CPRQ responding in the female dyad remained quite congruent and relatively unchanged across 8 sessions of psychotherapy. Both persons maintained moderately positive feelings about the relationship. This dyad remained in treatment for 25 sessions before a mutual decision was reached to terminate.

Figure 13. Standard scores derived from Therapist (dashed lines) and Client (solid lines) Personal Reaction Questionnaire responses for 3 sessions of psychotherapy for the female (top graph) and male (bottom graph) dyads. Standard scores along the ordinates represent the range of reactions from strong negative (-9) to strong positive (+9).



CHAPTER V

DISCUSSION AND CONCLUSIONS

The results generally indicate the lack of a direct relationship between pitch means and ranges of therapist and client in the psychotherapy sessions examined. No strong or consistent patterns of leading or following by either client or therapist were observed either within or across sessions. Perhaps the most remarkable finding of the experiment was the inflexibility of pitch patterning in both dyads. The speaking ranges of both therapists averaged only around 5 semitones, so that they were utilizing only about 16% of their potential speaking ranges. In addition, neither therapist seemed able to consciously broaden or dampen pitch range at will. The male therapist did broaden his range more than usual at the end of session 4, which demonstrates he was capable of some vocal flexibility, but this flexibility was not a direct function of client pitch fluctuations.

There are two approaches to explaining such negative results, one strictly methodological and the other in terms of the dynamics of communication. The first has to do with how the experiment was designed. It is possible that the speech segments sampled were biased in such a way that they did not represent the pitch interactions that were actually occurring. This is unlikely, however, as the

samples taken were equally distributed across each session and represented nearly every 2 minute interval in each session. It is also possible that those speech segments which were not grammatically complete (that is, which lacked a subject or predicate) may have exhibited artificially narrower pitch ranges and masked any real pitch interactions. However, a post hoc examination of the three female sessions showed that only 5%, 15%, and 10% of the segments, respectively, were grammatically incomplete. In addition, those incomplete statements did not exhibit consistently narrower ranges than the complete ones. Thus it does not appear likely that the lack of relationship was due to artificially narrowed pitch ranges. Another possible explanation is that pitch measurements were not precise enough to detect subtle pitch interactions. This argument is of course always difficult to disprove. The judges achieved an overall 90% agreement within 2 semitones of each other for the means and ranges of the female subjects. The average range of the female voices across all sessions was only 5 semitones (see Figure 3), so that on the average only pitch variability greater than 40% of the total variability (that is, greater than 2 semitones out of 5) was detected with 90% agreement. Any interactions involving pitch fluctuations of less than 2 semitones would have been masked by the error associated with the poor reliability at that precision of measurement. Such interactions would not have been reliably detected in the present experiment. Limits to the precision of one's measurements can always be invoked to explain negative results, and

at the present time there is no way of deciding beforehand how precise a measurement should be to detect some meaningful interaction.

However, it should be noted that the measures in this study were sufficient to detect a deviation from baseline in the pitch range of the male therapist at the end of session 4 and could have detected comparable deviations from baseline by the client had they occurred.

Subject variables operating uniquely in this experiment offer another possible explanation for the negative results. The sample size was extremely small, and the results may not generalize well to other kinds of clients or to other therapists with different orientations, levels of experience, and styles of conducting therapy. Indeed they may not even generalize to the same subjects assigned to different dyads (for example, a male-female dyad). It can also be argued that the convergence of pitch levels within and across sessions could not be adequately assessed in this experiment, at least for the female dyad, because of the closeness of their speaking pitch levels to begin with. Female therapist and client exhibited only a 2-4 semitone disparity in mean pitch at the beginning of session 1 (see Figure 7), which was very close to the limits of precision of the measurements. All that could have been realistically assessed across sessions was a divergence or lack of divergence from these incipient values. The male participants were better selected in this regard because of the more distinct difference in their pitch levels at the beginning of the session. This session produced a divergence of pitch ranges and some convergence of mean pitch by the time it ended. The argument that the disparity of female

therapist and client pitch should have been greater at the beginning is weakened somewhat by the negative results of the correlational analysis across sessions for this dyad, which was capable of detecting some convergence or divergence in the relationship by comparing correlation coefficients across sessions.

Assuming that the experiment was capable of detecting a relationship in pitch between therapist and client if it existed, it may seem paradoxical that no such relationship was found when pitch has been considered a primary component of vocal expressiveness (Wexler, 1975), which has been found to interact in psychotherapy (Wexler & Butler, 1976). First, it may be that pitch is not the active or even necessary component of vocal expressiveness, and that judges who rate such nonverbal variables are not attending to pitch but to other qualities of voice such as intensity or rate of speaking, which have been demonstrated to interact in psychotherapy (Matarazzo & Wiens, in press; Natale, 1975). Research into just what nonverbal variables judges may be attending to when they rate vocal expressiveness is obviously needed. Second, it is possible that pitch may be an integral part of vocal expressiveness without exhibiting any interactions itself. In other words, the whole that is vocal expressiveness may be greater than the sum of its parts, one of which is pitch range, so that the configuration of a host of vocal qualities, including pitch range, stress patterns, intensity, rhythm, timbre, and so forth, all comprise a signal that is capable of interacting with other signals like it during communication. In

this case judges rating expressiveness (or any other process variable) may be attending to some gestalt rather than to the additive effects of interacting components of speech. Third, it may be that pitch range and mean pitch are more a function of the content or intent of speech than they are a function of pitch variations in another speaker. It may be just too simplified a view of human communication to think that a change in the voice of one communicant will lead to an immediate and reciprocal change in the voice of the other. Thus the male therapist probably broadened his range not because of any conscious effort to do so nor as an unconscious reaction to the pitch variability of the client, but rather because a broad range was an integral part of driving home a point emphatically. The male client may well have perceived the deviation from baseline in the therapist's voice, and it may indeed have communicated something meaningful to him about the therapist, but it did not necessitate that the client vary his own pitch range greatly. Hence the results do not necessarily argue against the functional significance of pitch in psychotherapy or in any human communication, but they do argue against any simple notions about the mutual reinforcement of pitch changes in two communicants.

There is one final explanation why neither therapist was consciously able to control pitch variability, and it involves a possible confusion between mean pitch and pitch range. Both persons stated afterwards that they believed they had manipulated their voices appropriately under the experimental conditions. Results

indicated, however, that they were in fact raising mean pitch when they thought they were broadening pitch range and lowering it when they thought they were narrowing pitch range. One might speculate whether this confusion is perhaps inherent in the perception of pitch. Fairbanks (1940) reported that subjects who were asked to read prose passages at high pitch levels also exhibited greater variability in their pitch ranges, which suggests that the two variables are potentially confounding. The nature of this confounding should be explored further because of its implications not only for speakers attempting to manipulate pitch but also for judges attempting to rate pitch changes.

This experiment has described for the first time with time series analysis the temporal properties of pitch fluctuations in spontaneous speech in psychotherapy. All pitch series were identified as integrated moving averages or ARIMA (0,1,1) models. Such uniformity from speaker to speaker and from session to session lends support to the assumption that all of the series were sampled in the same way and also provides some basis for generalizing about other pitch series in psychotherapy. First, these series were nonstationary in level; that is, they oscillated about a mean level for a time and then dropped or rose to a new temporary level (Glass et al., 1975). Second, they contained no autoregressive and one moving averages component. The presence of a moving averages component indicates that some random shocks entered the series with full strength at time t , and some portion of those shocks remained in the system after their

initial occurrence to influence the course of the series. Any subsequent pitch level is to some extent dependent upon and thus predictable from these random shocks. Another important property of the ARIMA (0,1,1) model is that its lag 1 autocorrelation is non-zero and the autocorrelation for greater lags is within standard error of zero. This suggests that any correlation coefficients computed between two such series should be performed at lag 1 as well as the usual lag 0 to check for some hysteresis effect in the relationship. Finally, for future experiments knowing which time series model is appropriate for pitch series will allow one greater power in testing for intervention effects on pitch than relying on pre-post t tests or analysis of variance alone (Glass et al., 1975).

The psychophysical technique of measurement employed in this study has the potential of being quite useful in process research. First, the matching procedure offers a precision of measurement that is reliable. In this experiment judges achieved an overall 90% agreement for female mean pitch and pitch range within 2 semitones of each other. This compares favorably with the Expressiveness Scale, whose reliability has been reported to be 0.75-0.80 (Wexler, 1975) and with Markel's paralinguistic scale, with its reliability in the 0.86-0.97 range (Markel, 1965). In addition, the precision of pitch discriminations with these subjective scales is much less than with the psychophysical technique due to their broad divisions of range and scale. The matching procedure also has the potential for improving precision through increased training of judges

so that they agree within smaller and smaller limits. Indeed a 94% agreement within 1 semitone was attained in this experiment by two of the judges for the female voices. Second, the psychophysical technique appears to be an empirically valid means of quantifying human pitch ratings. Mean pitch levels and pitch ranges as measured in this experiment fall easily within the norms for female and male voices established previously using electronic analysis of fundamental frequency (Boone, 1971; Duffy, 1970; Provonost, 1942; Winckel, 1968). Third, although not tested directly, this procedure should prove psychologically valid in its reliance on the human perception of pitch as the basis for its measurements. In this regard there is no difference between these ratings and those with subjective rating scales. The difference is in the objective of the discriminations: in the Expressiveness Scale judges are asked to discriminate wide from narrow ranges and in the psychophysical task judges are asked to pick out the highest and lowest pitches uttered in a given segment. Assuming that the subjective rating scales more closely approximate how people discriminate pitch in every day life, the psychological validity of the matching procedure could be formally assessed by grouping its estimates of pitch range into wide and narrow categories and comparing these with subjective ratings of the same speech samples. To date this has not been done.

Certain modifications of this procedure will be necessary before its potential can be practically realized. First, it will be helpful if the judges and the subjects have comparable pitch ranges.

The results here indicate that when a judge was unable to hum the pitch of the subject she had great difficulty matching it to a pure tone. This finding supports the information processing position that sensory patterns are only recognized and identified if they can be compared to some internal representation or template (Lindsay & Norman, 1972), and it further suggests that without some internal representation matching two sets of stimuli cannot be accomplished. The practical significance of this is that female judges are probably best able to discriminate pitch levels in female voices and that male judges are probably best able to discriminate pitch levels in male voices. Second, to avoid extensive training time, it will be helpful if the judges have had training in vocal music. Such persons are already practiced in calibrating human pitch to the musical scale. Third, the highest reliability will probably be achieved for measurements of pitch maxima and minima per segment, from which mean pitch and pitch range can be derived. In this experiment judges were much more capable of identifying pitch extremes (that is, the highest and lowest pitches uttered) than they were measures of central tendency. Fourth, the matching procedure can probably be accomplished with other, less costly means of tone production than an electronic tone generator. Matching tones to a pitch pipe, for example, also allows one to calibrate pitch levels in semitone values. It is portable enough for use in the clinic or office. Unlike the tone generator, though, it produces discrete instead of continuous tones, restricting precision to the nearest semitone

value. This may or may not be a disadvantage, depending on how precise one wants the measurements to be. As noted already, even with the tone generator the limits of precision in this experiment was only about 2 semitones. Finally, considerable time and effort could be saved by eliminating the task of producing tape loops. Presumably after some training a judge could listen to marked portions of a tape recording and make the discriminations without the experimenter having to chop the tape physically into segments. In the present experiment judges required progressively less time to make the measurements until toward the end of the study they were listening to each loop only 2-3 times for each measurement. All of the foregoing modifications should of course be tested for reliability and validity in actual practice, but the results of this study nevertheless demonstrate the potential of the psychophysical matching procedure as a quick, convenient, reliable, and empirically quantifiable measure of pitch in human interactions.

Conclusions

1. Pitch means and ranges of therapist and client do not directly influence each other in psychotherapy. This was true for one dyad studied across 8 weeks of psychotherapy and two dyads studied within 4 individual psychotherapy sessions. If pitch interactions are present in psychotherapy, they do not exhibit the large, mutually reinforcing effects reported for other nonverbal variables.
2. Pitch changes in psychotherapy are more likely a function of the content or intent of communication than they are a function of the pitch fluctuations in the other communicant.

3. Pitch fluctuations of both therapist and client comprise nonstationary time series over long periods of time. These series are best described by an integrated moving averages or ARIMA (0,1,1) model.

4. The psychophysical technique of pitch measurement utilized in this experiment has the potential of being a useful nonverbal measure in future process research. Its reliability of measurement compares favorably with more subjective rating scales of pitch, and its precision compares adequately with electronic measures of fundamental frequency. With the modifications suggested, it represents a convenient and inexpensive procedure for monitoring pitch fluctuations in spontaneous speech.

CHAPTER VI

SUMMARY

This study was designed to accomplish two things. First, it sought to detect interactions of vocal pitch between therapist and client in ongoing psychotherapy. Previous research has shown that the pitch of a person's voice can communicate aspects of his personality, his affective state, or the context in which he is speaking. Other studies have shown that certain vocal properties, such as vocal intensity (Natale, 1975), rate of speaking (Matarazzo & Wiens, in press) and vocal expressiveness (Wexler & Butler, 1976), exhibit a strong convergence between therapist and client in psychotherapy. The present study was one of the first to look for a similar convergence phenomenon in the pitch fluctuations of therapist and client. The second purpose of the study was to assess the reliability and psychological and empirical validity of a psychophysical technique for measuring vocal pitch.

Pitch means and ranges of a female dyad were measured across three sessions of psychotherapy, and those of a male dyad were measured within a single psychotherapy session. Three specific hypotheses were tested: 1) mean pitch and pitch range of therapist and client are not independent of one another in the course of any given psychotherapy session; 2) pitch range of client can be modified by extremes of therapist pitch range; and 3) mean pitch and pitch range of therapist

and client vary across sessions as a function of therapeutic relationship. Correlation coefficients computed for therapist and client pitch fluctuations and a chi square test of independence for pitch change directionality were employed to test for interaction effects within sessions. An analysis of variance and a chi square test that $r_1 = r_5 = r_8$ were conducted to assess for interaction effects across sessions. Raw data were plotted and examined to assess the degree of following by the client during conditions of extreme pitch ranges by the therapist.

No significant interactions were found either within or across sessions, and neither therapist was able to modify his or her pitch range beyond their habitual levels under the experimental conditions. These results suggest that no simple convergence of pitch levels by therapist and client occurs in ongoing psychotherapy, at least within the unit of analysis utilized in this study. A time series analysis performed for each pitch series per subject per session indicated that pitch fluctuations were not strictly random, however, and could be described by an integrated moving averages model: ARIMA (0,1,1). Thus, the temporal patterning of pitch fluctuations was predictable for both therapist and client.

The psychophysical technique used to measure pitch in this experiment was found to be at least 90% reliable across three judges with a lower limit of precision at 2 semitones. The procedure was also shown to be both psychologically and empirically valid. It should prove useful for further process research into pitch fluctuations given the following modifications: 1) judges should probably be of the same sex as the subjects; 2) judges who have had training in vocal music

will likely produce the greatest reliability; 3) discriminations of pitch should be restricted to maxima and minima to achieve greatest reliability; 4) less expensive sources of pure tones (for example, pitch pipes) should result in greater convenience without sacrificing precision of measurement; and 5) tape recordings need not be physically parsed into tape loops for analysis, as judges can learn to pick out maxima and minima after listening to each speech segment as little as 2-3 times.

LIST OF REFERENCES

- Allport, G. W., & Cantril, H. Judging personality from voice. Journal of Social Psychology, 1934, 5, 37-55.
- Anderson, F. An experimental pitch indicator for training deaf scholars. Journal of the Acoustical Society of America, 1960, 32, 1065-1074.
- Ashby, J. D., Ford, D. H., Guernsey, Jr., B. G., Guernsey, L. F., & Snyder, W. U. Effects on clients of a reflective and a leading type of psychotherapy. Psychological Monographs, 1957, 71 (24), 1-32.
- Boone, D. R. The voice and voice therapy. Englewood, N.J.: Prentice-Hall, 1971.
- Brown, B. L., Strong, W. J., & Rencker, A. C. Perceptions of personality from speech: effects of manipulations of acoustical parameters. Journal of the Acoustical Society of America, 1973, 54, 29-35.
- Brown, B. L., Strong, W. J., & Rencker, A. C. Fifty-four voices from two: the effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech. Journal of the Acoustical Society of America, 1974, 55, 313-318.
- Chapple, E. D. Quantitative analysis of the interaction of individuals. Proceedings of the National Academy of Science, 1939, 25, 58-67.
- Chapple, E. D., Chapple, M. F., & Repp, J. A. Behavioral definitions of personality and temperament characteristics. Human Organization, 1955, 13 (4), 34-39.
- Cohen, A. J. Estimating the degree of schizophrenic pathology from recorded interview samples. Journal of Clinical Psychology, 1961, 17, 403-406.
- Condon, W. S., & Ogston, W. D. Sound film analysis of normal and pathological behavior patterns. Journal of Nervous and Mental Disease, 1966, 143, 338-347.
- Crystal, D. Prosodic systems and interactions in English. Cambridge, U.K.: Cambridge University Press, 1969.

- Davitz, J. R. The communication of emotional meaning. New York: McGraw Hill, 1964.
- Dempsey, M. E., Draegert, G. L., Siskind, R. P., & Steer, M. D. The Purdue Pitch Meter--a direct reading fundamental frequency analyzer. Journal of Speech and Hearing Disorders, 1950, 15, 135-141.
- Dittman, A. T., & Wynne, L. C. Linguistic techniques and the analysis of emotionality in interviews. Journal of Abnormal and Social Psychology, 1961, 63, 201-204.
- Duffy, R. J. Fundamental frequency characteristics of adolescent females. Language and Speech, 1970, 13, 14-24.
- Duncan Jr., S. Nonverbal communication. Psychological Bulletin, 1969, 72, 118-137.
- Duncan, S., Rice, L. N., & Butler, J. M. Therapists' paralanguage in peak and poor psychotherapy hours. Journal of Abnormal Psychology, 1968, 73, 566-570.
- Eisenberg, P., & Zabowitz, E. Judging expressive movement. III. Judgements of dominance-feeling from phonograph records of voice. Journal of Applied Psychology, 1938, 22, 620-631.
- Eldred, S. H., & Price, D. B. A linguistic evaluation of feeling states in psychotherapy. Psychiatry, 1958, 21, 115-121.
- Fairbanks, G. Recent experimental investigations of vocal pitch in speech. Journal of the Acoustical Society of America, 1940, 11, 457-466.
- Fairbanks, G., & Hoaglin, L. W. An experimental study of the durational characteristics of the voice during the expression of emotion. Speech Monographs, 1941, 8, 85-90.
- Fairbanks, G., & Provonost, W. An experimental study of the pitch characteristics of the voice during the expression of emotion. Speech Monographs, 1939, 6, 87-104.
- Fant, G. Analysis and synthesis of speech processes. In B. Malmberg (Ed.), Manual of phonetics. Amsterdam: North Holland Publishing Co., 1968.
- Fay, P. J., & Middleton, W. C. Judgment of occupation from the voice as transmitted over a public address system and over a radio. Journal of Applied Psychology, 1939, 23, 586-601.

- Fay, P. J., & Middleton, W. C. Judgment of Kretschmerian body types from the voice as transmitted over a public address system. Journal of Social Psychology, 1940, 12, 151-162.
- Fay, P. J., & Middleton, W. C. The ability to judge sociability from the voice as transmitted over a public address system. Journal of Social Psychology, 1941, 13, 303-309. (a)
- Fay, P. J., & Middleton, W. C. The ability to judge truth-telling, or lying, from the voice as transmitted over a public address system. Journal of General Psychology, 1941, 24, 211-215. (b)
- Fay, P. J., & Middleton, W. C. Judgment of introversion from the transcribed voice. Quarterly Journal of Speech, 1942, 28, 226-228.
- Fay, P. J., & Middleton, W. C. Judgment of leadership from the transmitted voice. Journal of Social Psychology, 1943, 17, 99-102.
- Fry, D. B. Prosodic phenomena. In B. Malmberg (Ed.), Manual of phonetics. Amsterdam: North Holland Publishing Co., 1968.
- Fujimura, O., & Lindqvist, J. Sweep-tone measurements of vocal tract characteristics. Journal of the Acoustical Society of America, 1971, 49, 541-558.
- Glass, G. V., Willson, V. L., & Gottman, J. M. Design and analysis of time series experiments. Boulder, Co.: Colorado Associated University Press, 1975.
- Gold, B. Computer program for pitch extraction. Journal of the Acoustical Society of America, 1962, 34, 916-921.
- Goldman-Eisler, F. The measurement of time sequences in conversational behavior. British Journal of Psychology, 1951, 42, 355-362.
- Goldman-Eisler, F. On the variability of the speed of talking and on its relation to the length of utterances in conversations. British Journal of Psychology, 1954, 45, 94-107.
- Goldman-Eisler, F. Psycholinguistics. London: Academic Press, 1968.
- Goldstein, A. P. Psychotherapeutic attraction. New York: Pergamon Press, 1971.
- Grigg, A. E. Experience of clinicians, and speech characteristics and statements of clients as variables in clinical judgment. Journal of Consulting Psychology, 1958, 22, 315-319.

- Hargreaves, W. A., & Starkweather, J. A. Voice quality changes in depression. Language and Speech, 1964, 7, 84-88.
- Harris, C. M., & Weiss, M. R. Pitch extraction by computer processing of high resolution Fourier analysis data. Journal of the Acoustical Society of America, 1963, 35, 339-343.
- Holmer, N. -G., & Rundqvist, H. E. Ultrasonic registration of the fundamental frequency of a voice during normal speech. Journal of the Acoustical Society of America, 1975, 58, 1073-1077.
- Kanfer, F. H. Verbal rate, content, and adjustment ratings in experimentally structured interviews. Journal of Abnormal and Social Psychology, 1959, 58, 305-311.
- Kanfer, F. H. Verbal rate, eyeblink, and content in structured psychiatric interviews. Journal of Abnormal and Social Psychology, 1960, 61, 341-347.
- Kiesler, D. J. The process of psychotherapy. Chicago: Aldine Publishing Co., 1973.
- Knower, F. H. Analysis of some experimental variations of simulated vocal expressions of the emotions. Journal of Social Psychology, 1941, 14, 369-372.
- Kramer, E. Elimination of verbal cues in judgments of emotion from voice. Journal of Abnormal and Social Psychology, 1964, 68, 390-396.
- Leitenberg, H. The use of single case methodology in psychotherapy research. Journal of Abnormal Psychology, 1973, 82, 87-108.
- Lieberman, P. Perturbations in vocal pitch. Journal of the Acoustical Society of America, 1961, 33, 597-603.
- Lieberman, P. Intonation, perception, and language. Cambridge, Mass.: MIT Press, 1967.
- Lieberman, P., & Michaels, S. B. Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. Journal of the Acoustical Society of America. 1962, 34, 922-927.
- Lindsay, P. H., & Norman, D. A. Human information processing: an introduction to psychology. New York: Academic Press, 1972.
- Luft, J. Differences in prediction based on hearing versus reading verbatim clinical interviews. Journal of Consulting Psychology, 1951, 15, 115-119.

- Mahl, G. F. Disturbances and silences in patients' speech in psychotherapy. Journal of Abnormal and Social Psychology, 1956, 53, 1-15.
- Marascuilo, L. A. Statistical methods for behavioral science research. New York: McGraw-Hill, 1971.
- Markel, N. N. The reliability of coding paralinguistic: pitch, loudness and tempo. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 306-308.
- Markel, N. N., Bein, M. F., & Phillis, J. A. The relationship between words and tone of voice. Language and Speech, 1973, 16, 15-21.
- Markel, N. N., Meisels, M., & Houck, J. E. Judging personality from voice quality. Journal of Abnormal and Social Psychology, 1964, 69, 458-463.
- Matarazzo, J. D., Saslow, G., & Matarazzo, R. G. The interaction chronograph as an instrument for objective measurement of interaction patterns during interviews. Journal of Psychology, 1956, 41, 347-367.
- Matarazzo, J. D., & Wiens, A. N. Speech behavior as an objective correlate of empathy and outcome in interview and psychotherapy research. Behavior Modification, 1977, 1, in press.
- Matarazzo, J. D., Wiens, A. N., Matarazzo, R. G., & Saslow, G. Speech and silence behavior in clinical psychotherapy and its laboratory correlates. In J. Shlien, H. Hunt, J.D. Matarazzo, and C. Savage (Eds.), Research in psychotherapy. (Vol. 3) Washington, D. C.: American Psychological Association, 1968.
- McQuown, N. A. Linguistic transcription and specification of psychiatric interview materials. Psychiatry, 1957, 20, 79-86.
- Miller, R. L. Performance characteristics of an experimental harmonic identification pitch extraction system. Journal of the Acoustical Society of America, 1970, 47, 1593-1601.
- Moses, P. J. The voice of neurosis. New York: Grune and Stratton, 1954.
- Natale, M. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. Journal of Personality and Social Psychology, 1975, 32, 790-804.

- Natale, M. Effects of induced elation-depression on speech in the initial interview. Journal of Consulting and Clinical Psychology, 1977, 45, 45-52.
- Newman, S., & Mather, V. G. Analysis of spoken language of patients with affective disorders. American Journal of Psychiatry, 1938, 94, 913-942.
- Noll, A. M. Short-term spectrum and "cepstrum" techniques for vocal pitch detection. Journal of the Acoustical Society of America, 1964, 36, 296-302.
- Ostwald, P. F. Soundmaking. Springfield, Ill.: Charles G. Thomas, 1963.
- Ostwald, P. F. Acoustic methods in psychiatry. Scientific American, 1965, 212 (3), 82-91.
- Pittenger, R. E., Hockett, C. F., & Danehy, J. J. The first five minutes. Ithaca, N. Y.: Paul Martineau, 1960.
- Pope, B., Blass, T., Siegman, A. W., & Rahe, J. Anxiety and depression in speech. Journal of Consulting and Clinical Psychology, 1970, 35, 128-133.
- Provonost, W. An experimental study of methods for determining natural and habitual pitch. Speech Monographs, 1942, 9, 111-123.
- Reik, T. Listening with the third ear. New York: Farrar, Strauss, & Co., 1949.
- Rice, L. N. Therapists' style of participation and case outcome. Journal of Consulting Psychology, 1965, 29, 155-160.
- Rice, L. N. & Wagstaff, A. K. Client voice quality and expressive style as indexes of productive psychotherapy. Journal of Consulting Psychology, 1967, 31, 557-563.
- Roessler, R., & Lester, J. W. Voice predicts affect during psychotherapy. Journal of Nervous and Mental Disease, 1976, 163, 166-176.
- Sapir, E. Speech as a personality trait. American Journal of Sociology, 1927, 32, 892-905.
- Shlien, J. M., & Zimring, F. M. Research directives and methods in client-centered therapy. In J. T. Hart & T. M. Tomlinson (Eds.), New directions in client-centered therapy. Boston: Houghton Mifflin, 1970.

- Skinner, E. R. A calibrated recording and analysis of the pitch, force, and quality of vocal tones expressing happiness and sadness. Speech Monographs, 1935, 2, 81-137.
- Soskin, W. F., & Kauffman, P. E. Judgment of emotion in wordfree voice samples. Journal of Communication, 1961, 11, 73-80.
- Stagner, R. Judgments of voice and personality. Journal of Educational Psychology, 1936, 27, 272-277.
- Starkweather, J. A. The communication-value of content-free speech. American Journal of Psychology, 1956, 69, 121-123. (a)
- Starkweather, J. A. Content-free speech as a source of information about the speaker. Journal of Abnormal and Social Psychology, 1956, 52, 394-402. (b)
- Starkweather, J. A. Vocal communication of personality and human feelings. Journal of Communication, 1961, 11, 63-72.
- Stevens, S. S. Psychophysics. New York: John Wiley & Sons, 1975.
- Taylor, H. C. Social agreement on personality traits as judged from speech. Journal of Social Psychology, 1934, 5, 244-248.
- Thompson, C. W., & Bradway, K. The teaching of psychology through content-free interviews. Journal of Consulting Psychology, 1950, 14, 321-323.
- Trager, G. L. The typology of paralanguage. Anthropological Linguistics, 1961, 3, 17-21.
- Traux, C. B., & Mitchell, K. M. Research on certain therapist interpersonal skills in relation to process and outcome. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change, New York: Wiley & Sons, 1971.
- Wexler, D. A. A scale for the measurement of client and therapist expressiveness. Journal of Clinical Psychology, 1975, 31, 486-489.
- Wexler, D. A., & Butler, J. M. Therapist modification of client expressiveness in client-centered therapy. Journal of Consulting and Clinical Psychology, 1976, 44, 261-265.
- Williams, C. E., & Stevens, K. N. Emotions and speech: some acoustical correlates. Journal of the Acoustical Society of America, 1972, 52, 1238-1250.