RECENT ADVANCES IN LANGUAGE, COMMUNICATION, & SOCIAL PSYCHOLOGY

edited by
Howard Giles
Robert N. St. Clair

LEA

6 Towards a Social Psychology of Voice Variations

Bruce L. Brown and Jeffrey M. Bradshaw
Brigham Young University

INTRODUCTION

A number of reviews of the personality-speech and emotion-speech literature appeared in the early 1960s (Diehl, 1960; Kramer, 1963; Mahl & Schultze, 1964). Scherer has brought us up to date with valuable recent overviews of the functions of vocal signs in conversation (1979a) emotion and speech (1979b) and personality and speech (1979c). Siegman (1978) has also contributed an important review of the speech research that bears upon clinical concerns such as anxiety, pathology, personality and emotional state. Rather than duplicate what has already been treated very competently, our purpose in this chapter will be to re-examine (and sometimes re-analyze) a select few of the older studies in order to propose remedies for the apparent artificialities of much of current paralinguistic research. In other words, we aim to demonstrate how social-psychological studies of voice variations can uncover verifiable information about personality dynamics and social processes.

Two fundamental features of vocal qualities are distinctive: they are totally quantifiable (Brown, Strong & Rencher 1975, pp. 11-12) and are central mediators of social processes and interpersonal dynamics. In this chapter, we focus primarily upon the reflections of *personality* and *emotion* in vocal characteristics. We refer to and attempt a critique of some of the major contributions and established findings from the three Scherer reviews and tie them in with additional studies not reviewed there.

As Laver and Trudgill (1979)[1] convincingly argue, there is a wealth of information within the voice concerning the intentions, dispositions and

dynamics of individuals: their physical selves as well as their inner worlds are reflected in their speech. Laver and Trudgill identify seven muscle systems which it is possible to isolate (respiratory, phonatory, pharyngeal, verlopharyngeal, lingual, labial and mandibular) and propose that:

> Speaking thus requires the most complex and skillful collaboration between the different muscle systems, whose cooperative actions all have to be precisely and intricately coordinated in time. It is not at all surprising, therefore, that in learning to control such a complex apparatus sufficiently to be able to produce auditorily acceptable imitations of speech patterns heard in one's social environment, speakers would nevertheless develop idiosyncracies of pronunciation that serve to individuate them within their own social group [p. 5].

Undoubtedly this individuation is intricately tied to the development of people's unique personalities, both through (1) the reflections of their intentions and preferences in their selection of *vocal patterns* and (2) the view they get of themselves through their biologically-fixed *vocal features* and the reactions others have to them. The problem is in unravelling the tremendous complexity of the vocal information. Substantial progress is being made in working out the ways in which anatomical variations and the functioning of the seven muscular systems of speech are realized acoustically (see, for example, the work of Laver, 1968, 1975 and 1978; and of Titze, 1973 and 1974). The major difficulties have been, and will be, in dealing with the psychological dimensions: personality and emotion.

PERSONALITY AND VOCAL PATTERNS

The widespread public interest in radio during the 1930s and 1940s was reflected in part by the vast amount of research at that time which looked at the voice as a mirror of the personality. By the next decade, however, such research effort had been almost entirely abandoned because of repeated failures to find any noteworthy correspondence between subjective judgements of personality from voice and the results of standardized measures of personality.

Since that time, most research has ignored the accuracy question and has split into two divergent directions (See Fig. 6.1).[2] These have been identified by Scherer (1979c):

---

[1]The Laver and Trudgill chapter is a very good summary of the phonological and linguistic work that has been done in categorizing, taxonomizing and defining various kinds of linguistic markers.

[2]See Scherer's (1979c) chapter where he uses an adaptation of the Brunswik "lens model" to set this distinction out in more detail.

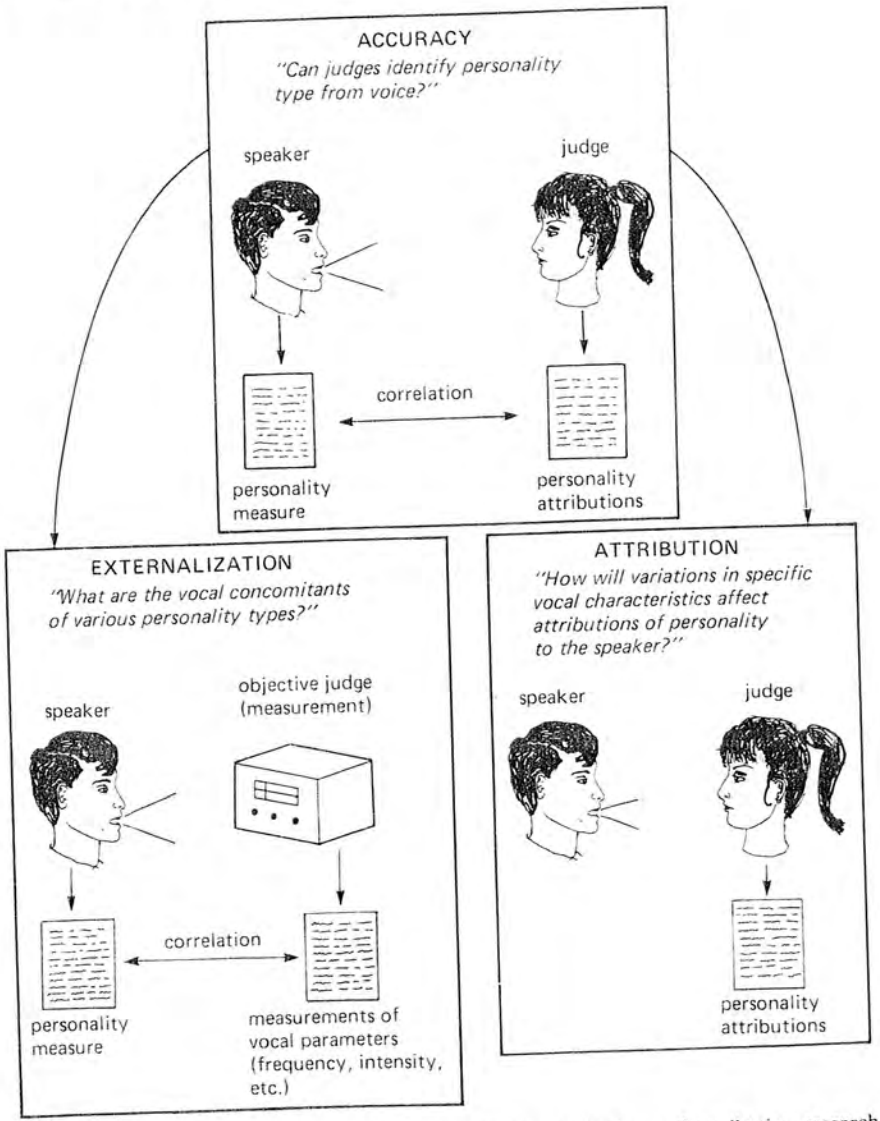

FIG. 6.1 A comparison of the standard accuracy, externalization and attribution research paradigms for studying the relationship between personality and voice.

1. *Externalization* studies, which have more or less kept the older research paradigm but have replaced naive subjective judgements with "expert ratings, systematic coding, or electro-acoustic analyses [p. 151]." These analyses are then compared with standardized personality measures.

TABLE 6.1
Potential Sources of Problems for Studies of Speech and Personality

	Inadequate Speech Measurement?	*Nonexperimental (Correlative)?*	*Inadequate Personality Measurement?*
Accuracy Studies	no	yes	yes
Externalization Studies	yes	yes	yes
Attribution Studies	no	no	no

2. *Attribution* studies, which continue to rely on subjective judgements of personality but which make no attempt to relate them to a criterion of accuracy.

Table 6.1 shows a list of potential trouble spots for studies of speech and personality. As can be seen, the accuracy and externalization studies hold the monopoly on these problems, while attribution studies are relatively problem free. Why this is so becomes clear in examining Fig. 6.2.

Figure 6.2 shows the dilemma of externalization studies — the independent variable is one based on social judgement, but the dependent variable is objective. Also, it is hard to conceive of how the independent variable could be manipulated, so externalization studies are hopelessly correlational. Attribution studies reverse this, giving the right combination for good experimental studies: An objective, manipulable independent variable and a dependent variable that is, although subjective, measureable without argument. Attributions of personality are straightforwardly social judgements, conceptually simple and quite easily measureable, compared to the conceptual difficulties of measuring personality. Accuracy studies also share the problem of the clouded conceptual status of personality measurement.

The first problem of Table 6.1, *inadequate speech measurement,* is soluble. It is possible now with Fast Fourier transforms and related acoustic analysis methods almost totally to capture numerically the information in speech (see Chapter 7). Speech that is synthesized from the numerical representation is, in the best examples, very similar to real speech (Brown et al., 1973).[3] The very important work left to be done in speech analysis is that of finding the most useful ways of summarizing the tremendous

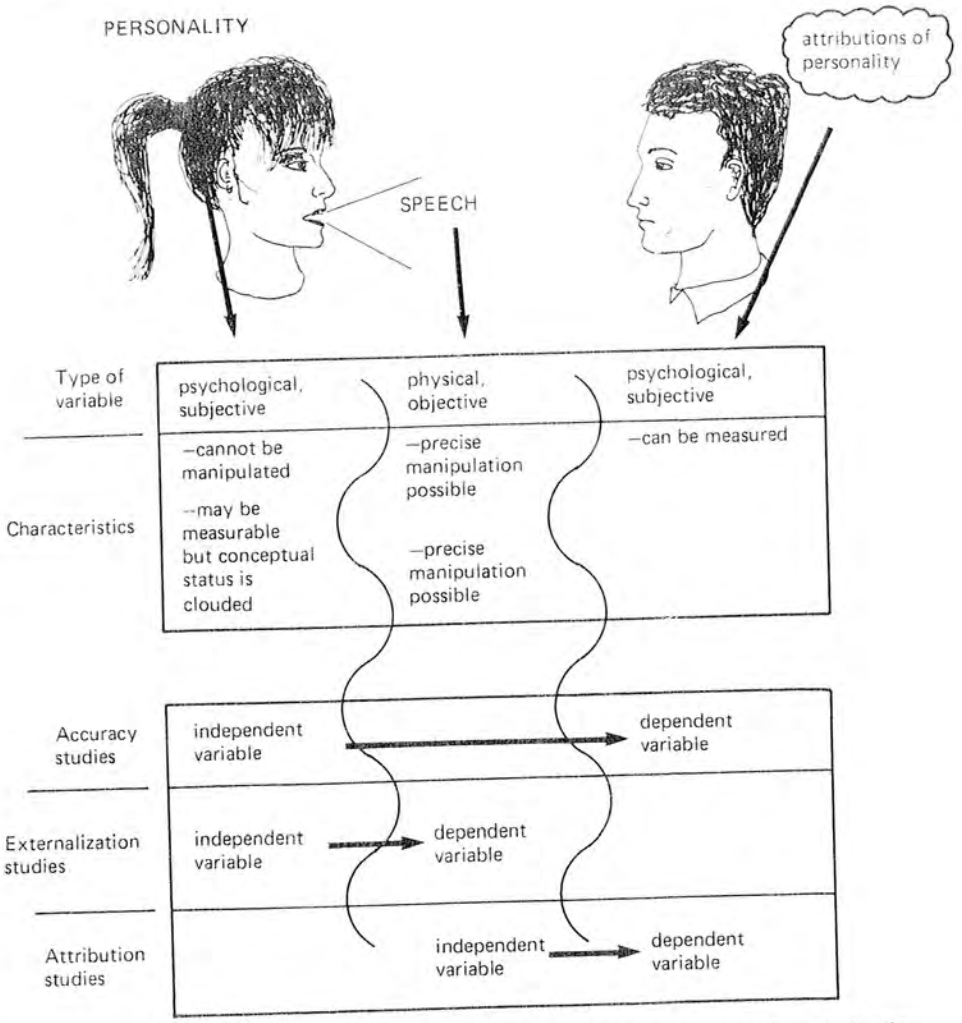

FIG. 6.2 The relative strength of attribution studies and the weaknesses of externalization and accuracy studies.

amount of data that comes from the numerical representation of pitch-tempo patterns and power spectrum, and pulling it together in a theory that relates the acoustic realization to physiological/anatomical causes.[4]

The *correlative* nature of the accuracy and the externalization studies has prevented researchers from obtaining the kind of clean results needed for the formulation of testable hypotheses and cogent theory. Even in the best studies, correlations are small and the relationships are tenuous and

unstable from one study to another. Little emerges in the way of a coherent, convincing pattern.

One way to stabilize and give some coherence to the results is to use factor analysis of the correlation matrices. There is usually considerable redundancy in both speech and personality variables, and factor score plottings show the relationships in terms of "bundles" of variables, which are more stable than single variables. (See Figs. 6.3 and 6.4 on pp. 159 and 160, for example.)

Another way to improve studies of personality reflections in speech so that they show "cleaner" results is to examine unique or extreme exemplars of various traits. In this way the chances are better of getting clear "ideal" patterns of vocal properties that differentiate one trait from another, rather than just correlating a bunch of vocal measures with a bunch of trait measures on a more randomly varying sample. Looking at the whole *gestalt* of vocal patterns characteristic of extreme exemplars of various traits will give much clearer results than endless correlational tables.

Inadequate personality measurement has also plagued accuracy and externalization studies from the very beginning. What is most surprising in examining this literature is that the quality of the paper and pencil tests used in this research has only rarely been questioned. When, for example, correlations between personality measures and personality attributions have been slight, fault has typically been assigned to the inaccurate "vocal stereotypes" which presumably bias subjective ratings. But, as we shall argue below, it is just as likely (perhaps more so) that inadequate personality measures have been largely responsible for the paucity of useful findings.

---

[3]Scherer's presentation (1979c, footnote 5, p. 184) of the results of "realism" tests is misleading. Whereas his account suggests that rate changed synthetic speech only sounds normal 50% to 58% of the time, what we in fact found was that when judges are told that some voices are synthetic and some are natural, they *mistake* synthetic for natural 50% to 58% of the time. It is important to add that even natural voices are judged to be synthetic about 5% of the time. Also, the quality of synthesized voices has improved since that time. Scherer's reference for the data is from a summary paper (Brown et al., 1975). Perhaps he did not have access to the original study (Brown et al., 1973).

[4]Titze (1973) in his doctoral dissertation used classical mechanics to mathematically model the relation between vocal anatomical functioning and its acoustic realization. Scherer in his 1979 review of emotion in speech (Scherer, 1979b) proposed a similar integration of anatomical correlates with the acoustic and psychological data, and even though he does so on the basis of the rather limited data available, the value of such an approach in producing testable hypotheses is obvious. Theoretical integration of externalization studies, giving a coherent account of why particular vocal qualities are associated with particular personal characteristics, will be heavily dependent upon the findings of such acoustical-anatomical work, exemplified in the work of Laver (1968; 1975; and 1978) and that of Titze (1973; 1974; and Titze & Talkin, 1979).

We propose, as an alternative to standard tests, to use the more direct method of nomination by social consensus to select exemplars of various traits.[5] In interpreting results, we would then say, "This vocal pattern is characteristic of those who are nominated by a number of other persons as being the most dominant person they know," rather than ". . .characteristic of those who judge themselves to be high in dominance," or ". . .characteristic of those who answer questions on a standardized scale in a way that is scored (for whatever reason, logical or empirical) as being dominant."

We would not wish, of course, to make the mistake of assuming that a social consensus can be the criterion in an accuracy study. If there is one lesson from the flurry of interpersonal-perception accuracy studies of the 1950s and 1960s, it is that we have no basis for talking about accuracy unless we have an objectively assessable reality. If we were to ask judges, for example, to identify the voice of a dominant person, we would have little basis for saying they were wrong or right according to whether their choice agreed with our social consensus nomination. If we give them all the information (e.g., that some person was nominated more than anyone else in his dormitory as being dominant) and then ask them to identify which voice it is, accuracy can be assessed because we are asking them to identify a social decision, which is real, did happen, etc., rather than appealing to some "ideal" of dominance separate from social judgement.

Accuracy Studies

The first, and largest in scale, of the accuracy studies is T. H. Pear's (1931) collection of judgements of personality from the radio voice. With the assistance of the British Broadcasting Company, he secured the judgements of 4,000 listeners concerning birthplace, occupations, age, and place of residence of each of nine speakers. Although these judgements were difficult to quantify, given their basis in the auditors' free descriptions, his results enabled him to make some tentative conclusions about their accuracy. Not surprisingly, sex (see Chapter 4) was always judged

[5] In one study (Dicks, Brown & Wells, 1977) our research group found that extreme high and low scorers on the POI, a test based upon "self-actualization" theory, did not differ in their ability to recognize portrayed emotions or their ability to portray them, but when we switched to a "nomination" strategy, comparing those who had been nominated as being unusually well adjusted with two control groups, the nominated group was substantially better at judging emotion portrayals. Scherer (1972, 1978) has also had good results with nomination strategies. He has shown that "both self and peer ratings of extroversion can be accurately inferred from voice quality for American speakers. For German speakers, peer ratings of dominance and assertiveness correlate significantly with respective attributions made on the basis of voice quality samples." (Scherer, 1979c, p. 190.)

correctly (except in the case of one eleven year-old child). Vocation and physical descriptions were sometimes judged with astonishing success. Age was estimated quite well by the listeners, although there was a strong tendency for estimates to centre in the thirty age group (see Chapter 5).

Allport and Cantril (1934), inspired by Pear's (1931) and Herzog's (1933) auspicious beginnings, set out to design a more careful study of personality judgements from voice in eight separate experiments with a total of 587 judges and 18 different speakers. Information pertaining to eight "objective" criteria (age, height, complexion, appearance in photographs, appearance in person, handwriting, vocation, and political preference) was gathered. Information was also obtained for three other criteria: extroversion/introversion was measured by the Heidbreder scale; ascendance/submission by the Allport "A-S Reaction Study"; and dominant values by the Allport and Vernon "Study of Values." In each experiment, standard passages from Charles Dickens, Lewis Carroll, and similar sources were used. Judges in each of these experiments were given descriptions of personality characteristics of speakers and asked to match them with the voices.

Age was identified quite accurately but, as in Pear's study, there was a tendency for estimates to centre in the thirty age group. Contrary to Herzog's (1933) findings, height was identified successfully in only one of four experiments. The one experiment involving complexion yielded somewhat significant results, as did the ones on appearance in photographs and in person. Vocation and political preferences were identified with moderate success, but handwriting yielded no results of significance.[6]

One of the personality measures (dominant values) gave positive and significant results in half of the studies. With the other two personality measures (extroversion and ascendance) an interesting thing happened, which Allport and Cantril (1934) state is "quite typical of our findings" — namely, that both "significantly positive" and "significantly negative" results are reported,[7] leading them to conclude that "the uniformity of opinion regarding the personality of radio speakers is somewhat in excess of the accuracy of such opinion (p.50)."

Also in 1934, Taylor published the results of his study. Recordings were

---

[6]It is unfortunate that most of these old studies do not give enough data from their results to construct a "contingency table" which shows the distribution of judged values for each measured speaker type. (This particular type of contingency table is often called a "confusion matrix" in that it displays the patterns of confusion in the judgements of actual categories.) When enough data is given, we can use Information Theory statistics to judge accuracy in a way that allows across-study comparisions (see footnote 9).

[7]The index of accuracy in this study is particularly obscure and is incomparable with other studies.

made of 20 subjects, including two groups measured high and low on neuroticism (Thurstone Personality Schedule) and a group that was randomly selected, and each recording was played to at least 20 judges. Subjects had previously filled out a self-descriptive questionnaire by putting a plus, minus or question mark next to 136 items concerning various personality traits (32 items from Freyd's list of introvert characteristics, 25 from the Thurstone Personality Schedule, and the remainder from "words commonly used to describe people"). Judges filled out the same questionnaire for each voice they heard but despite a high degree of agreement between judges, the correlations between their ratings and the subjects' self ratings indicated no relationship whatsoever. In fact, Taylor (1934) concluded that there "is a slight but statistically significant tendency for the auditors to be most consistent in their judgements when they agree least with the subjects themselves [p. 24-7]."[8]

Eisenberg and Zalowitz (1938) selected eight extremely dominant and eight extremely non-dominant women using the Maslow "Social Personality Inventory" and a personal interview. Phonograph records of their voice were played for 43 judges who were only able to judge dominance a little better than chance. This is in contrast to Allport and Cantril's positive findings with their personality measure of dominance. As in previous studies, social agreement was high despite the "inaccuracy" of the judgements.

From 1939 to 1943, a series of studies was carried out by Fay and Middleton including, among other things, judgements of Spranger personality types from voice (1939, from the Allport and Vernon "Study of Values"), Kretschmerian body types (1940b, no actual morphological measurements were made), intelligence (1940a, Terman Group Test of Mental Ability), sociability (1941, Flanagan sociability scale of the Bernreuter "Personality Inventory"), introversion (1942, also from the Bernreuter), and leadership (1943, ratings by fraternity brothers). Although the Spranger personality-type and Kretschmer body-type studies showed a better-than-chance accuracy,[9] judgements of intelligence from voice were "only fairly reliable" and sociability and leadership ratings revealed no more than chance accuracy. In each study, however, the phenomenon of significant inter-judge reliability was reported.

If one takes these studies at face value the inaccuracy of judgements of

[8]It is unfortunate that Taylor did not use an "accuracy of response prediction" paradigm in which the judges estimate what speakers would put on the questionnaire for themselves rather than putting what they think the speakers "are." In this case, there would be an "assessible reality," which would probably give much higher accuracy scores, and even if not, we would know that it was due to the judge's inability rather than the conceptual confusion of what the speaker "really is.")

personality from voice seems apparent. In every review of accuracy research only one conclusive finding emerges: *There is nearly always astonishingly high inter-judge agreement despite the poor accuracy (in the sense of personality ratings agreeing with the standardized tests of the trait in question).*

The standard way of interpreting this has been to assume that the judges are rating on the basis of inaccurate "vocal stereotypes." To quote Allport and Cantril (1934), "the various features of personality are associated in the minds of the judges (with) some preconception of the type of voice to which these features correspond [p.50]." Although judges share these preconceptions about what the voice of a certain personality type sounds like, researchers have assumed that these stereotypes have little or no basis in reality because of their disparity from assessments of personality made by paper and pencil measures.

We propose an alternative view: *That the inaccuracy of personality judgements from voice has far more to do with the inadequacy of the personality measures used than the incompetency of judges to assess traits from vocal properties.* This view is especially admissible in the light of the fact that, as Kjeldergaard (1968) points out, there is considerable accuracy whenever certain *objectively measurable* characteristics such as age, sex, social class, etc. are measured. (Social class judgements typically correlate .80+ with actual numerical social class indices, see Brown & Lambert, 1976.) Other more recent studies and reviews of the vocal realization of objective characteristics such as sex (Smith, 1979), age (Helfrich, 1979; Ryan, 1979; Ptacek & Sander, 1966), social class (Robinson, 1979; Brown & Lambert, 1976), and ethnicity (Giles, 1979) also show surprising accuracy.

In retrospect, especially with the well-established unreliability of most of the personality tests used in these studies,[10] it seems strange that it would have been assumed that the "stereotyped" judgements were in error rather than the tests. It seems particularly strange since the concept of personality is itself rooted in the perception one person has of another. The question of whether person *A* is "dominant" is ultimately a social

[9]Fay and Middleton's 1939 "Spranger personality types" study, and their 1940b "Kretschmer body types" study *did* give us enough information to make confusion matrices and to compute Information Theory statistics as suggested in footnote 6. For most of the sub-studies within the "Spranger" study, the percentage of personality information being communicated to judges by the vocal qualities was between 20% and 25% (taken from the "coefficients of constraint" in the Information Theory statistics). For the "Kretschmer" study the amount of "body type" information communicated was 5% to 8% for male judges and 14% to 17% for female judges.

[10]See for example, Tyler's (1953) critique of the Bernreuter "Personality Inventory," which formed the criterion for many of the personality and voice studies.

perception question.[11] We cannot be satisfied in any final way by agreeing that he or she is dominant if scoring high on a particular dominance test, since it is ultimately only a secondary derivative from social judgement (some particular personality theory, explicit or implied); if his or her test is "construct validated," or if consensual validation is used, we are correlatively referred to some social perception criterion. A number of personality theorists (e.g., Mischel, 1973) have articulated the dilemma: *Although personal identity seems to be a very stable thing, personality tests are notably unreliable, with little agreement among alternate measures of the supposed same trait.*[12]

Other researchers (Meehl & Hathway, 1946) have come to the conclusion that self-report scales are "notoriously susceptible to 'faking' or 'lying' in one way or another, as well as . . .even greater susceptibility to unconscious self-deception and role playing on the part of individuals who may be consciously quite honest and sincere[13] in their responses [p. 525]." More recently Sackeim and Gur (1979) concluded from their study that self-deception contributes significantly to the invalidity of self-report inventories and more substantially than does intentional lying.

We propose that if adequate alternatives to standard personality measures (such as the nomination technique described above, or selecting speakers on the basis of objective categories that are related to personality, e.g., occupation, college major or age) are utilized we will find judges to have high accuracy in identifying these from voice.

Externalization Studies

Scherer (1979c) describes the present state of research on externalization as "bleak" and "the amount of hard data negligible," and he hastens to add that this is not because of a lack of effort (there are more than 1500 references in Görlitz's 1972 review). A coherent pattern does not emerge from these "gargantuan proceedings."

---

[11]One may argue that the action pattern, the behaviour of a person, is public data, assessable by a third person observer. One way to be freed from this delusion is to carefully read Taylor's influential account (1964) of how identifying actions requires introspective attribution. There is no meaning to the action "pattern" until imposed by a human judge (either the actor himself or an observer).

[12]Could this be, for example, why Allport and Cantril (1934) in reporting judgements of dominance from voice using Allport's "A-S Reaction Study," reported findings not in accord with Eisenberg and Zalowitz's (1938), who used Maslow's "Social Personality Inventory"?

[13]Note this strange implicit hypothesis of a conscious honesty and an unconscious dishonesty that is so common in psychological literature. Compare to footnote 32 and the accompanying text and references.

The crude speech variables used in almost all of the externalization studies reviewed by Scherer are not the primary defect, since such studies could be used for preliminary hypotheses to be tested with more precise methods, and the parallel comparisions of the subjective speech judgements with the quantitative ones would be valuable. The really serious problems are (1) the correlative nature of the studies and (2) the monumental problem of personality measurement, particularly in the accuracy studies. One of the best externalization studies, Scherer's "juror study," which uses precise measurements of speech, still gives very tenuous conclusions. (The complete study is not yet in print but a partial report is given in Scherer, 1979c.) The primary data are correlations among self- and peer-ratings for extroversion and emotional stability and a set of precise measures of fluency-related variables as well as pitch and "effort". The largest correlation reported between personality ratings and speech is $-.65$ (42% common variance). Even tenuous results are better than none, and a summary of the major findings of the best of the externalization studies on specific dimensions of voice may provide some groundwork for future work. We shall concentrate on three specific dimensions of voice: fundamental frequency, vocal intensity, and articulation/timbre.

Fundamental Frequency: Scherer, in his juror study, reports that higher fundamental frequency (FØ) seems to be associated with self-ratings of competence and dominance in male Americans (and to a lesser degree in male German speakers), as well as self-ratings of discipline and dependability in male German and female American speakers.[14] He hypothesizes that a high degree of arousal may be associated with competence and dominance in males, which is in turn reflected in a heightened degree of muscle tone leading to characteristically higher FØ. However, as Scherer (1979c) points out, this hypothesis is at present highly speculative since the only evidence results from studies of an increase in transitory states of arousal under stress.

---

[14]These variables, the self-ratings on specific trait dimensions, must undoubtedly have considerable redundancy with one another, and a much clearer and coherent pattern could probably be given by summarizing them with factor analytic plots like the one in Fig. 6.4 and other figures in this chapter.

Scherer finds in his (admittedly limited) sample that the mean FØ for American speakers (128Hz) is significantly lower that that of German speakers (161Hz, $p < .001$) which supports Laver's (1975, p. 268) impression of a very low pitch range in American males. As Scherer (1979) points out, "it would be a challenging task for social psychologists to determine whether the difference is due to physiological factors—possibly related to national character or modal personality—or to differential expectations or evaluations concerning desirable pitch levels—strongly influenced by historical tradition or mass media portrayals—or the interaction of both of these factors [pp. 157-158]."

Vocal Intensity. The personality trait most consistently associated with vocal intenstity is extroversion (Allport & Cantril, 1934; Trimboli, 1973). Consistent with these findings, Mallory and Miller (1958) report a very small (.14) correlation between "inadequate loudness" as rated by members of the speech department faculty and introversion as measured by the Bernreuter Personality Inventory. Mallory and Miller (1958) also report a negative correlation (−.32) between "inadequate loudness" and "dominance," but no such correlations were found in Scherer's juror study.

Articulation and Timbre. Moses' classic work *The voice of neurosis* (1954), which discussed a number of hypotheses concerning voice changes in patients experiencing affective disturbances, inspired Rousey and Moriarty (1965, and later, Rousey, 1974) to extend and substantiate some of these ideas. These authors propose that, as well as Freud's popular ideas on the meanings of slips of the tongue as applied to words, we consider his indirect suggestion of the psychological significance of an individual's misuse of specific sounds.

Rousey's (1974) analyses of detailed case studies have led him to form inferences concerning the relationship of personality patterns to the inability to master proper speech sounds. An example of one of these inferences is the following: "The misarticulation of the vowel *r* reflects lack of impulse control and minimal neurological disfunction [p.12]." Although the strength of the evidence for his specific hypotheses and his adoption of a psychoanalytic framework may not be convincing, his work certainly represents a pioneering effort to apply our knowledge of personality and speech to the practical diagnostic setting, and the possibility of linking specific vocal characteristics to their concomitants in the development of a child's personality is an exciting one. Undoubtedly anomalies of speech are reflective of personality, and a careful working out of that relationship will be a major contribution.

Lomax (1974) has speculated that phonatory and articulatory patterns are influenced by sociological as well as psychological factors, claiming that "back, loosely enunciated and lax sounds decrease, and front, narrow, sharply enunciated sounds generally increase as culture grows more complex and laden with rules [pp. 204–205]."

Ostwald (1963, 1965) was one of the first to use spectrographic methods of voice analysis to investigate the relationship between acoustical parameters of speech and types of personality. There are now very precise methods available for looking at power spectrum of voice, such as the "fast Fourier transform" method and the "predictor coefficient" method used in Scherer's "juror study" (see Scherer, 1979c), the Brown, Strong and Rencher speech synthesis studies (1973, 1974), Smith, Brown, Strong

and Rencher (1975) and Apple, Streeter and Krauss (1979) and Stewart, Brown and Stewart (in preparation). Before these methods were well developed and became generally known, Ostwald (1965) devised a procedure that uses filters to obtain amplitude values for octave and half octave bands to provide a crude approximation to the power spectrum.

He identified certain qualitative categories of voice, such as "flat voice," associated with monotonous speech indicating depression, with sound energy spread evenly across the spectrum; "hollow voice," with only the fundamental frequency having much sound energy, the harmonics dropping off markedly in energy; "robust voice," associated with oratorical or theatrical speech. In later work, he used sound spectrograms which show the changes in power spectrum energy patterns over time, but his qualitative analysis of crude power spectrum data is much more interesting and illuminating. Even though in today's world the acoustic descriptions appear simplistic and naive, they suggest exciting possibilities for attribution studies in which experimental designs using speech synthesis methods could be used to test his hypotheses. For example, Ostwald (1965) suggests that:

> It is a fundamental fact of auditory perception that the human ear is most sensitive to tonal stimuli above 500 cycles per second. Almost all sounds calculated to arouse human beings — whistles, cries, sirens, screams — contain a large amount of sound energy above 500 cycles per second. On the other hand, the fundamental tone of the human voice, which is located at lower frequency levels, is so dispensable in ordinary conversation that telephone companies do not even bother to transmit it. We advance the hypothesis that speakers increase the amount of sound they emit at higher frequencies when they want to be listened to and reduce their high-frequency output when they prefer to be ignored. [pp. 85-86].[15]

Attribution Studies

As Scherer points out in his review (1979c), the results of attribution studies are much clearer and much more stable than externalization studies, and it is possible to approach the attribution problem experimentally.[16]

---

[15]It is interesting to note Ostwald's "medical model" approach to voice as a symptom of "emotional disturbance" (he is an M.D.), such as his statement (1965) that "one man we studied in this way entered the hospital in a state of severe depression *resulting from* manic-depressive psychosis; his voices were uniformly flat [p. 84, our italics]." He views vocal qualities as rather involuntary indices of disturbance. In contrast, Warner (1982), and Beier (1966) see both the vocal properties and the "disturbance" as primarily intentional and communicative, even insistent. (See discussion of these authors on pp. 175-176.)

[16]There is an interesting example of a very recent study in which the much weaker correlation methodology is used even in an attribution setting (Aronovitch, 1976) where an experimental method could have been used.

Stagner (1936) was the first to study the ways in which specific vocal cues are used in the process of making judgements about personality, countering Allport and Cantril's (1934) view that such an undertaking would be "absurdly atomistic." The reliability of Stagner's (1936) findings, however, are questionable, given the small number of speakers and the lack of adequate instrumentation.

One of the best of the more recent attempts at an experimental paradigm is a study by Addington (1968) in which two male and two female trained speakers simulated a number of different voice qualities, pitch patterns and speaking rates (generating 252 "voices" from four), which were then rated by a large number of judges on semantic differential type adjectives.[17] We re-analyzed his data to simplify and clarify the effects of these subjective manipulations for the two male voices and obtained some interesting hypotheses to test with computer manipulation methods. Fig. 6.4 gives the results, but Figure 6.3 helps one to understand how to read Fig. 6.4

Suppose that we were to plot 14 of Addington's 252 voices (the two male speakers doing each of seven manipulations: breathy, thin, flat, nasal, tense, throaty and orotund) on only two of his 24 paired opposite adjectives, "feminine vs masculine" and "intelligent vs. stupid." Suppose also, that we knew enough about the geometry of correlation so that, instead of putting the two axes (feminine-masculine and intelligent-stupid) perpendicular to one another, which would only be proper if they were independent ($r = .00$), we put them with their positive poles 95.5° away from one another (as shown in Fig.6.3) because they are found to correlate with one another $r = .096$, the cosine of which is 95.5°. The position of one of the voice manipulations for one of the speakers in this two-dimensional space, then, is found at the intersection of a perpendicular from his score on the masculine-feminine axis and a perpendicular from his score on the intelligent-stupid axis. Speaker 1's "breathy" voice, for example, has a score of 1.2σ on the masculine-feminine axis (this is a z score, and since the positive end of this axis is feminine, a positive score means toward the feminine end of the axis) and a score of 1.43σ on the

[17]Even though this study could in a sense be called experimental, because the same two male voices and the same two female voices are used with a number of voice dimensions manipulated in systematic combinations, it is not a precise experimental paradigm since it is doubtful that one can subjectively hold all other speech dimensions constant while altering a particular one. There is some evidence, however, that at least for the dimension of speech rate, subjective alterations give very similar patterns of judge ratings to those of computer altered voices (Stewart et al., in preparation). If such is also true of other vocal and speech dimensions, the findings of Addington may be essentially corroborated when the same speech dimensions are more carefully evaluated with the speech synthesis experimental methodology.

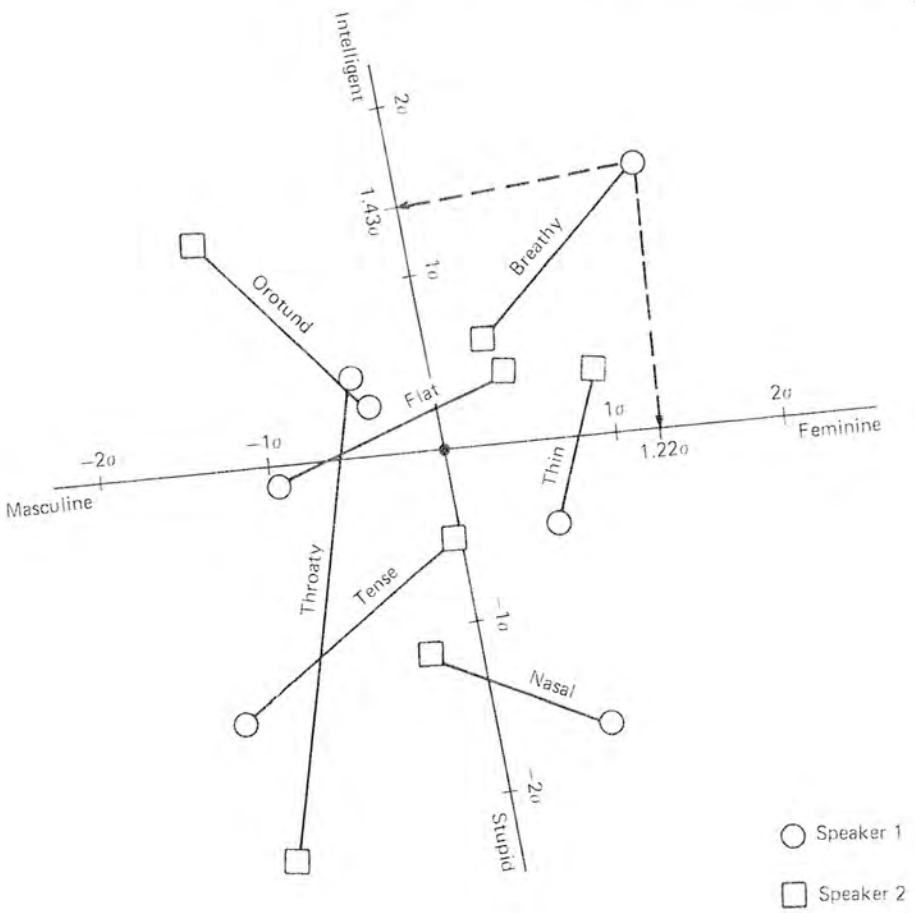

FIG. 6.3 Two male speakers in seven voice manipulations, plotted on only two of the 24 adjectives (Figure constructed from our reanalysis of the data from Addington, 1963, 1968). *Note:* The line connecting Speaker 1 with Speaker 2 for each manipulation is simply to make the two voices for each manipulation easier to find.

intelligent-stupid axis, so his point for the "breathy" voice in the two-dimensional space is located at the intersection of the two perpendiculars as shown in Fig. 6.3.

Suppose now that we wanted to see the positions of these seven manipulations for the two speakers for the other adjectives as well. If all gave independent information (all were uncorrelated with one another), this would require a 24-dimensional space representation for the 24 adjective pairs, or, more realistically, 276 graphs like Fig. 6.3 to show the plotting of all pair combinations of the 24 adjective pairs. Fortunately, there is considerable redundancy in the adjectives and they can all be

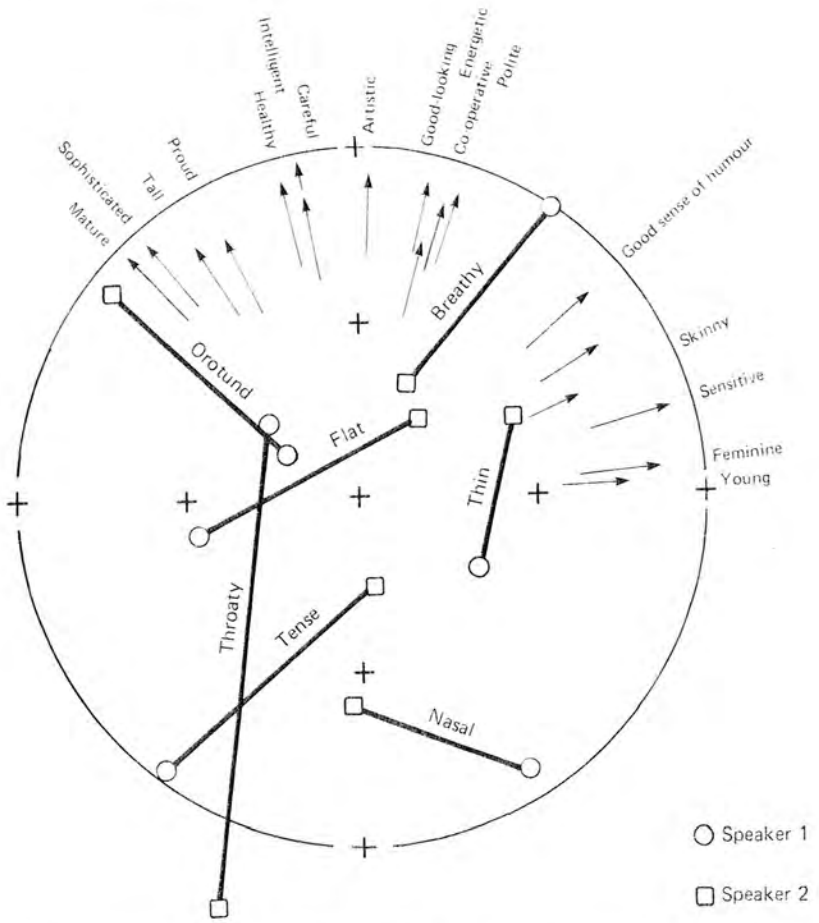

FIG. 6.4 Two male speakers in seven voice manipulations, plotted on all 24 adjectives. (Figure constructed from our reanalysis of the data from Addington, 1963, 1968: 17 of the 24 adjectives shown.)

located mostly in the same two-dimensional space of Fig. 6.3. Fig. 6.4 shows the same plotting with the other adjectives added. To know the standard score of any voice on any adjective, one simply drops a perpendicular from the plot of that voice to the adjective vector or axis in question.[18] Of course, not just any two-dimensional plane within the 24-

[18]If our factor analysis accounted for all the variance in the 24 adjectives with two factors, these projected scores on any adjective from the two-dimensional plotting would be exact. The length of the adjective vectors in Fig. 6.4 indicates the amount of variance in that adjective pair accounted for by the two factors. For those vectors that are shorter, then, the scores estimated from projections would be the least accurate.

dimensional space will do. This one was located by means of the principal factors method of factor analysis.

Although Addington also used a factor analysis to summarize his rating data, his factor structure is virtually uninterpretable, presumably because he computed correlations over both speakers and judges (thus causing a conflation of two sources of variance) and making an unnecessarily complex factorial structure. It is not clear from the paper whether this is the case, but this is a common error in analyzing data of this kind. Professor Addington graciously helped us obtain the original data from this study, his dissertation. From our factor analysis of ratings averaged over judges (found in Appendix II beginning on page 172 of Addington, 1963) we obtained the clear pattern shown in Fig.6.4. This figure gives a very terse summary of the findings of Addington's study. In examining the factor score positions of Speaker 1 and Speaker 2 (the two male speakers on the "orotund" manipulation) both voices can be seen to be relatively high on the vectors for such variables as "sophisticated," "mature," "tall," and "proud", Speaker 2 being considerably higher than Speaker 1 on these dimensions for that manipulation. Similarly, the "thin" manipulation gives rise to ratings high on "sensitive," "feminine," and "young". This is the case for both speakers. By contrast, there is considerable variance in the effect the "throaty" manipulation has on the two speakers' ratings, Speaker 2 being very low on vectors like "good looking," "cooperative," "polite," etc. for this manipulation, but Speaker 1 being about in the middle of the plotting on this manipulation. It is very difficult to pull this much information out of the original study because of the method in which the data are presented there. When one understands how to read them, factorial plots of this kind are very efficient summaries of multivariate data. Addington's study is by far the richest in information of any of the attribution studies up to the present time, and it would be very worthwhile to continue the kind of simplifying re-analysis begun in Fig. 6.4 to the remainder of the 252 speaker-manipulation combinations. The results will then have to be confirmed either by trying to do the same things using speech-synthesis technology (as employed in the studies described in the following paragraphs) to generate manipulations precisely, or by doing a careful acoustic analysis of Addington's original voice tape to determine just what the speakers were manipulating for each type of "voice."

At about the same time that Addington was doing his work, Kjeldergaard (1968), working with John Carroll at Harvard, did a similar experimental attribution study with five original voices, but with only four manipulations of each: "normal," "fast," "shouting," and "soft." Kjeldergaard's manipulations were incomplete (he had all of the manipulations for three of the speakers but only normal voices on two),

not completely crossed, and the polarity on his adjective pairs was not sufficiently clear in the version of the paper we have to enable us to put it in the form of Fig. 6.4 for comparison. Therefore, we replicated his study with the rate manipulation ("fast" vs "normal") completely crossed (all possible combinations) with the "shout" vs. "soft" vs. "normal" manipulation, using a Latin square presentation of six speakers each doing each of the six manipulation combinations and using natural speech (answers to three questions, to give three replications of the experiment), rather than a standard passage as Addington and Kjeldergaard had done. The detailed results of both this and the other Addington manipulations are being compiled, but would require much more space than the whole of this chapter, but Fig. 6.5 displays just the means for "normal" rate and "fast" placed in opposition, and the means for "shout" and "soft" placed in opposition. A "soft" voice is seen as more humble, unemotional, submissive, etc. and a "shout" voice is seen as more outgoing, energetic,

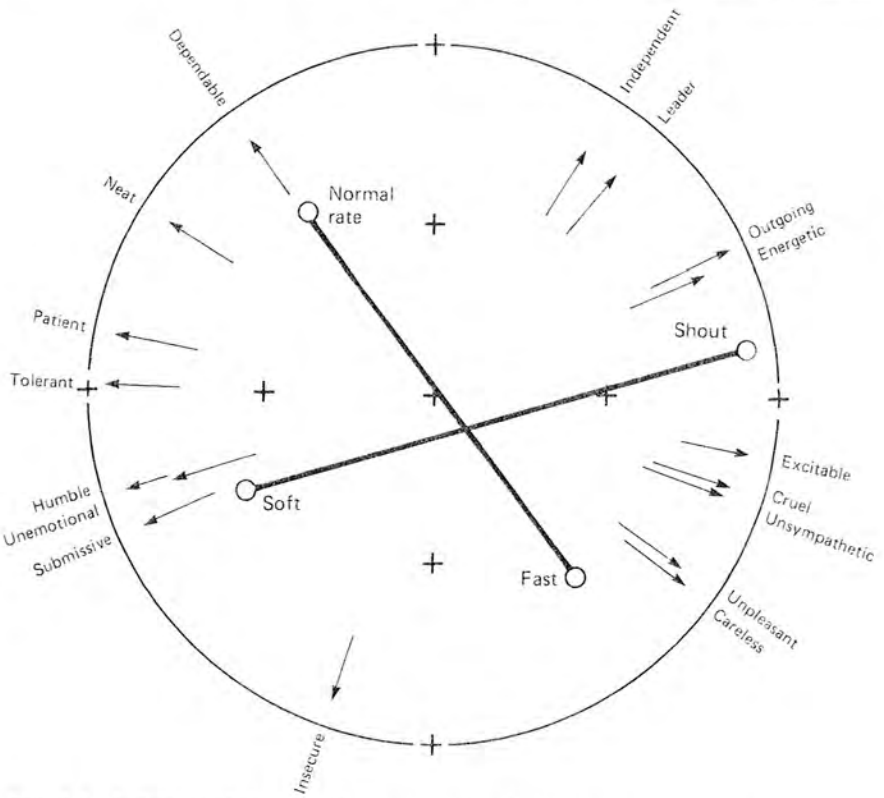

FIG. 6.5 A plotting of the mean factor scores for the "normal" rate vs. "fast" manipulation and for the "soft" vs. "shout" manipulation (Data from our replication of Kjeldergaard, 1968).

excitable, etc.[19] A "fast" voice is seen as more unpleasant, careless, unsympathetic, cruel, etc., whereas a "normal" rate voice is seen as more dependable, neat, patient, etc.

The first truly experimental studies, using a fast Fourier transforms speech synthesis method to manipulate precisely one parameter of voice at a time, were conducted in the early 1970s (Brown et al., 1973, 1974; and Smith et al., 1975) and tested the effects of rate, pitch (average FØ.) and intonation (variance of FØ.) on personality adjective judgements. The findings with respect to rate are clear cut and stable and most accurately demonstrated in the study by Smith et al. (1975).[20] Here it is shown that judgements on adjectives related to competence are a monotonically increasing function of phonemes per second. Moreover judgements on adjectives related to benevolence are an "inverted U" function of phonemes per second, with the most positive rating being given to the normal rate voice for each speaker (see Fig. 6 of Smith, et al., 1975; and the triple replication of these results under different circumstances in Brown, 1980b). The effects of intonation and pitch were much less clear cut and accounted for much less variance than rate in a three-way factorial comparison of them (Brown et al., 1974).[21]

MacLachlan (1979) prepared persuasive messages where the speaker subjectively varied his rate in three conditions: slow (111 words per minute), normal (140 wpm), and fast (191 wpm). Subjects who heard the tapes rated the speakers on the degree of their agreement with the

---

[19]To make the two-dimensional plottings more easy to read, we have only placed one pole of the bi-polar adjective pairs on the plottings, but we could just as well say that "shout", for example, is "proud", "emotional" and "dominating," since these are the opposite poles that were used for the adjective pairs "humble," "unemotional" and "submissive," which are located directly across the centre of the graph from the "shout" voice average.

[20]Scherer (1979c) questions the generalizability of these results on the basis that the manipulations may have been too extreme to sound like natural speech, but claims that one cannot determine this since the speech rates for the voices are not given. Again, this may be a problem of getting his information from a summary paper (Brown et al., 1975) rather than from the primary source (Smith et al., 1975). Fig. 6 of the primary source clearly gives the average rate of each of the nine manipulations in phonemes per second. A replication of this rate study (see Brown, 1980b) which used natural answers to questions rather than standard passages and kept the rates of manipulated speech well within bounds of believeability, gave results which are remarkably similar to the results of the Smith et al. study. It is a very reliable effect.

[21]One finding conflicts with Scherer's from correlational studies. While he finds increased competence with higher pitch (average FØ), Brown et al. (1974) found that an experimentally increased FØ, with all other vocal features controlled, resulted in a slight decrease in competence as well as benevolence ratings. However, only two "seed" voices were used in this study and the speakers were reciting a standard passage. The question will need to be settled with improved experimental methods like those discussed below (Stewart, et al. in preparation).

speaker, his intelligence and his objectivity. Although full statistics for adequate comparision with our studies are not reported, MacLachlan states that in each case the faster speaker was rated more highly. These results conflict with the findings of the Kjeldergaard replication given in Fig. 6.5 which show the fast manipulation to make a voice sound more careless, unpleasant, unsympathetic, etc. They also conflict with the many times replicated findings of the precise experimental studies (summarized in Brown, 1980b) that when a voice is increased in rate with everything else held constant, it is perceived as lower on benevolence-related adjectives generally. MacLachlan in the same paper reports an experiment in which a "time compression technique" was used to speed speech by 25%. The faster speakers were rated as "more knowledgable, intelligent and sincere than those who spoke slowly [p. 114]." These results are very consistent with those of an almost identical study reported in Brown et al. (1973) in which an Eltro-Automation Rate Changer (an electronic time compression device that holds pitch constant while altering rate) was used to both speed and slow voices. The results were presented in a similar plotting to Figs. 6.4 and 6.5 of this chapter from a factor analyisis of 22 adjective pairs, and slowed voices were found to be rated lower in all competence-related adjectives and all benevolence-related adjectives. However, speeded voices were found to be rated higher on competence-related adjectives, but lower on benevolence-related adjectives.

A recent study (Brown, Giles & Thakerar, 1985) shows clearly that making generalizations from such results should be taken with a pinch of salt, unless judges' attributions of the speaker's intent are also taken into account, i.e., the judges' perceptions of *why* the speaker is speaking that way. This was a factorial study of the interactive effects of rate ("slow," "normal," and "fast"), dialect ("standard" vs. "non-standard" in a British setting) and context (context given or not given). In the non-context or "monologue" condition, judges were unaware of the situation that led to the utterance, but in the "context" condition they were told that a psychologist was explaining some principles to a group of dentists. Fig. 6.6 gives the results for a part of the study with male judges only. For this factor analysis a three-factor solution was needed to account for over 60% of the variance. In order to plot the information for the three factors, two plottings must be given, a "front" view and a "side" view. In order to see what effect any manipulation had on all of the adjectives we must look at both plottings, Factor I by Factor II ("benevolence" by "competence") and Factor I by Factor III ("benevolence" by "social attractiveness"). The "standard, monologue" condition corresponds to all of the old studies of the effects or rate, since in those we did not have dialect or context introduced. The results for that condition are consistent with those of the older studies: Increased rate gives rise to increased

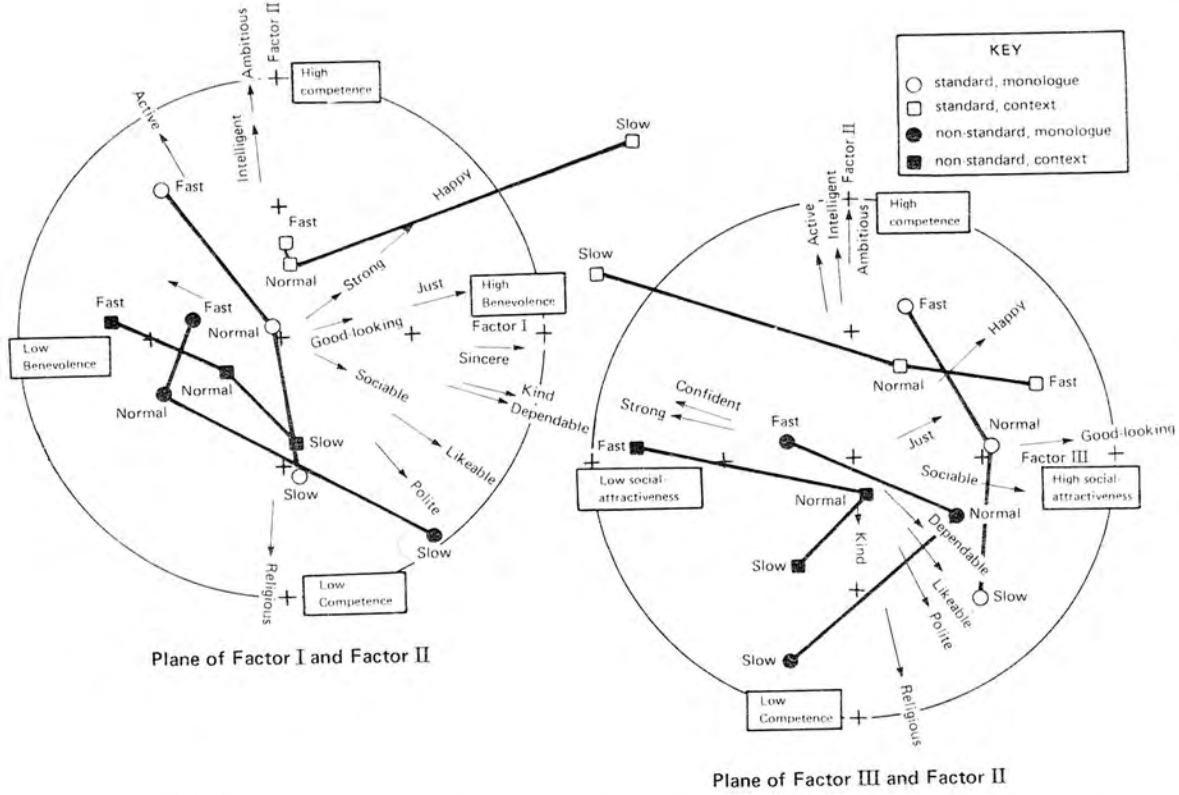

FIG. 6.6 A plotting of the factor scores for the twelve manipulations superimposed upon the factor pattern of the fifteen adjectives—male judges.

competence ratings and decreased benevolence ratings, and decreased rate gives rise to a decrease in both competence and benevolence ratings. Additionally we find in this study that both increasing and decreasing rate give rise to dereased social-attractiveness (Factor III).

The really interesting finding is the way in which context affects these old results. When judges are given context, i.e., a reason for slow speech (to make oneself intelligible to dentists), the slow rate is no longer perceived as a decreased competence but an increased benevolence, with even a slight increase in competence. A nonstandard dialect voice sounds less competent and less benevolent to judges than the standard version, but if one slows the nonstandard voice it sounds more benevolent and only slightly less competent. The study shows clearly that many factors affect judgements of voice with strong interactions among them, and, therefore that we can not know the effect of a vocal change without knowing also what meaning the judge gives to the change and, to what speaker intention it is attributed.

Perhaps the most serious defect of the early experimental studies was the artificiality of the speech samples used: recitation of standard sentences. Apple et al. (1979) tested the effects of rate and pitch manipulations using speech synthesis methods in an improved design in which the stimuli were actual spontaneous answers of subjects to one of two questions. However, there were methodological problems (see Brown, 1980b) with this study which made it difficult to evaluate whether the results agreed with those of the earlier, less natural studies. Stewart et al. (in preparation; see Brown, 1980b for a brief description of their method and results) tested the rate question with a design that also used natural speech in the form of spontanteous answers to questions, but improved upon the Apple et al. (1979) study by using a Graeco-Latin square design that uses each speaker as his or her own control. In this way they controlled for pre-manipulation characteristics of speakers, the problem that makes the Apple et al. data difficult to interpret.

The findings of the Stewart et al. study corroborated the fundamental findings of the old studies, with only slight variation across different speech content. The effects of subjective manipulations were found to be essentially the same as those of computer synthesis manipulations, suggesting that the studies that use subjective manipulation to investigate rate probably give results very close to those that would be obtained with the more precise but much more difficult and expensive speech synthesis methods.

The Graeco-Latin square paradigm worked out in the Stewart et al. study is a very sensitive and powerul one for examining the effects of even subtle vocal manipulations. In future work, it may be valuable to combine an "externalizing" methodology with this experimental method for

examining attributions. That is, after getting extreme exemplars of various traits by social consensus nomination (as described in the proceeding section), it could be valuable to test the hypotheses that come from acoustic analysis of their voices by synthesizing a voice, both in its normal form and with the hypothesized vocal index of that trait, to see if it does in fact cause an attribution of the trait.[22]

Attribution and Temporal Patterns

A serendipitous finding with respect to actual and perceived speech rate has led to some important questions related to pauses and temporal patterning. In the Stewart et al study, each voice was increased and decreased in rate both synthetically by the computer method and subjectively by the speaker. Although the two methods of manipulating rate gave almost the same results in the relative personality trait ratings, when ratings were obtained of how fast the voice sounds to judges and plotted as a function of actual rate (Fig. 6.7) it became clear that the rate of the normal spontaneous voice and the fast, medium and slow synthesized voices (which are merely mathematical transformations of the spontaneous voice) were relatively overestimated in rate. However those "acted back" (the person speaking the same thing again, as much as possible like the first time) at fast, medium and slow rates were relatively underestimated. We think this was because it is not by phonemes per second that a person subjectively judges rate, but by the rhythm of the speech (the distance between stress or accent points) independent of pauses.[23] That is, in spontaneous speech a person pauses to think, but a listener does not include the pause in a subjective assessment of rate, only reckoning it on the basis of rhythm. When one is acting back the same words, "thought pauses" are not so necessary. Although the rhythm may not be any faster, the elapsed time will be shorter and so the actual phonemes per second will be underestimated. We are now developing a method of automatically categorizing every ten millisecond slice of speech time as either "voice speech," "unvoiced speech," or "background noise." This scheme will allow a detailed micro-analysis of the temporal pattern of speech that will provide a test of these hypotheses. We expect the microanalysis of the temporal pattern of speech in natural emotion situations to be particularly revealing.

---

[22]Scherer (1979c) suggests a similar strategy of using "complementary externalization and attribution studies, to derive at least some hypotheses about the existence of further personality markers in speech [p. 191]."

[23]These, and related findings, underscore the importance of differentiating between aspects of speech and voice as they are objectively measured (distal cues) and as they are subjectively experienced (proximal percepts). (See Scherer, 1979c.) Failure to distinguish between these apsects is a serious problem in many studies.

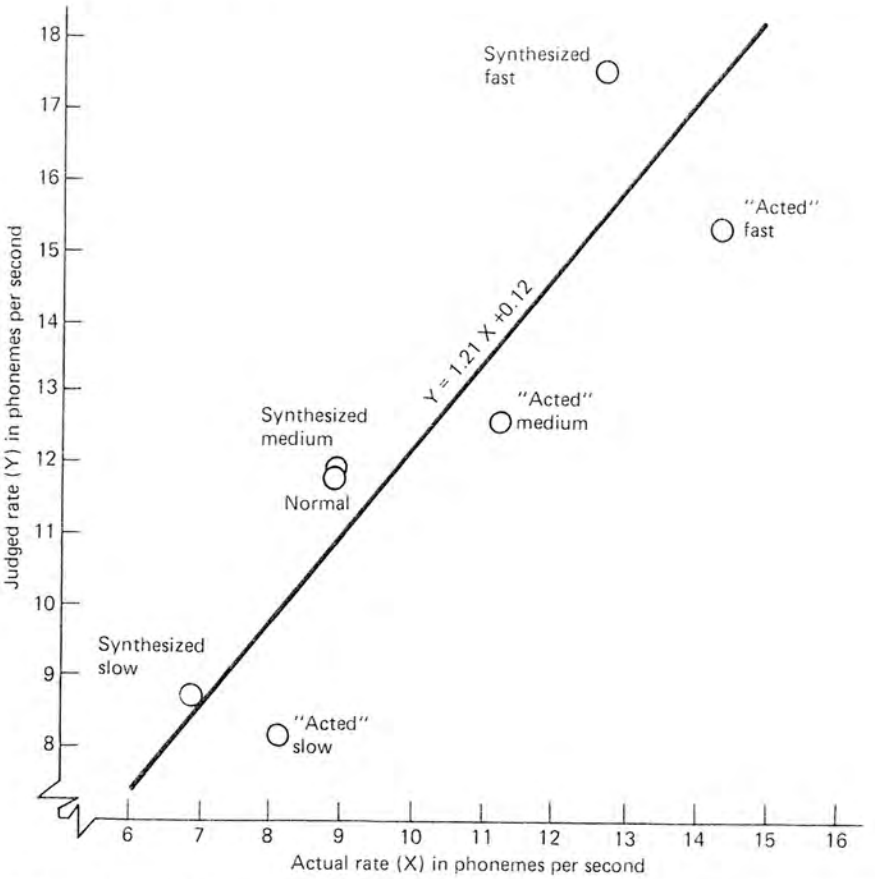

FIG. 6.7 Bi-variate scatter plot and regression line for the relationship between actual and perceived speaking rates for the average of the normal voices and for the averages of six manipulations. (from Brown, 1980b, Fig. 2, p. 296).

When Howard Giles was collaborating with our research group, he mentioned that he found it very hard to believe that slower speech is always perceived as less competent, since he knows a number of very competent slow talkers. We carefully listened to the tapes from the Stewart et al. study and made an interesting discovery: In order to subjectively slow one's speech without sounding less intelligent, one must do it with pauses. Rate decrease studies utilizing speech synthesis have slowed all aspects of the utterance uniformly. An important direction for future studies will be to test the effects of slowing through pauses. It may be possible to predict in advance the effect a pause will have from the position in which it appears. At major syntactic breaks it may be more an

"emphasis pause", but in the midst of syntactic units it may be a disruption or a "thought pause." A number of important findings have emerged from the many studies of pauses and temporal patterns in vocal interaction (see Rochester & Gill, 1973; Siegman, 1978; and Feldstein & Welkowitz, 1978, for reviews), but the development of the micro-analysis temporal patterning method mentioned above will allow more precise specification of temporal pattern. Some important advances should come from the careful analysis of the temporal patterns of voices under natural emotion conditions (which we describe in the next section) as compared with the same voices when not under emotive stress.

EMOTION AND VOCAL PATTERNS

Scherer (1979b) has also conducted a superb review of the current state of knowledge in the area of the vocal and speech reflections of emotion. Again, we do not attempt a duplication of his effort here, but instead discuss a few of the classic studies of emotion and speech in terms of their major findings and problems. Studies in this area tend to fall into one of two categories:

1. *Externalization studies:* acoustic analysis of vocal and speech cues accompanying discrete emotions (usually only portrayed rather than real) or non-specific arousal.
2. *Accuracy studies:* i.e., subjective judgements of which of several emotions is being evoked or portrayed (usually portrayed).

In principle, there is no reason why there should not also be attribution studies of emotion in speech, but these have seldom been done. For example, it would be possible to use the Graeco-Latin square design, developed for the Stewart et al. study, to make manipulations in speech to fit specific hypotheses about the vocal properties that mediate given emotive states, and to test the effects of such manipulations upon attributions of emotion from a group of judges.

Externalization Studies

Fairbanks (Fairbanks & Provonost, 1939; Fairbanks, 1940; Fairbanks & Hoaglin, 1941) was one of the first to study specific vocal indicators of discrete emotions by acoustic measurement. Since his technique was limited to the use of a modified oscillograph, the major object of his studies was to examine the relationship between FØ (fundamental frequency) and emotion.

In one study (Fairbanks & Provonost, 1939), six male actors read a standard passage while simulating emotions of contempt, anger, fear, grief, and indifference. These portrayals were played to a group of 64 "auditors" who were instructed to identify the emotions being portrayed. Their high degree of accuracy verified the correspondence between speech patterns and their intended emotions. As far as variation in pitch over time is concerned, Fairbanks noted the "few extremely wide variations in the simulation of contempt, the generally wide, rapid inflections of anger, the irregularity of the pitch changes in fear, the consistent vibrato in grief and the lack of distinguishing features in indifference [p. 458]." The mean pitch levels of the emotions in order of ascendency were, indifference (108 Hz), contempt (124 Hz), grief (136 Hz), anger (229 Hz) and fear (254 Hz)

In one of the most sophisticated approaches to the emotion-speech question, Williams and Stevens (1972) recorded the voices of professional actors reading short scenarios written to portray anger, fear, sorrow and a neutral situation. The study is carefully done and a good source for hypotheses, but much of the data are qualitative descriptions and even the quantitative results are presented in graph form or are described rather than tested for statistical significance. With such a small sample, the observed consistencies may not be applicable to a larger sample of actors, let alone the general populace. There is sufficient inter-actor inconsistency to leave the reader with a considerable burden of interpretation. The contour of fundamental frequency vs. time was found to be the strongest indicator of the emotions of the speaker. Lower fundamental frequency, with decreased range and reduced rate of articulation, were indicative of sorrow. Anger was characterized by both raised pitch and increased pitch variance. The analysis did not reveal any consistent acoustical correlates for portrayals of fear.

Accuracy Studies

One major problem in accuracy studies of emotion and speech is that of obtaining content-free speech. If we do not do this, content (*what* is being said) may be the cue for recognition of an emotion rather than vocal properties (*how* the thing is said). Davitz and Davitz (1959) dealt with this problem in a novel way by having subjects express each of ten emotions by reciting the alphabet "using letters as if they were words," after reading a brief description of a situation which might generate the feeling in question. Judges were then faced with the task of deciding which feeling situation was being expressed. They found the highest percentage of correct identification for anger (65%),[24] followed by nervousness (54.2%), sadness (49.2%), happiness (43.3%), sympathy (38.8%), satisfaction (30.8%), fear and also love (25%), jealousy (24.6%), and pride (20.8%).

The Kramer (1964) study only used five emotions but their acted portrayals used actual speech rather than letters of the alphabet and the results were displayed more completely in confusion matrix form (thus allowing more careful comparison, using Information Theory statistics, to the other studies cited in this chapter which use an accuracy paradigm). In order to hold verbal content constant so that it would not be a cue in identifying the emotion portrayed, they arranged to have three sentences that were equivalent for each of the scenarios for the five emotions. Only those three sentences were presented to judges for the identification task. The three sentences of recorded portrayals were presented to judges in three ways: untouched English portrayals, the same English portrayals filtered to pass only frequencies below 400 Hz., and untouched portrayals in Japanese. These conditions were used to test the consistency of the pattern of confusions under each method of disguising content (filtering and using a language unknown to the judges). Table 6.2 displays the results of the Kramer study presented in a format that allows maximum comparison with the other studies of this chapter. Two of the emotion names, anger and love, are the same for the Kramer study and the Davitz and Davitz study. Other emotions which are fairly comparable are grief, which was used by the Kramer study, and sadness, which was an emotion Davitz and Davitz wanted to identify. Let us see what the comparable results were for these three emotions: anger, grief/sadness, and love. In both studies, anger is most often correctly indentified (between 67% and 77% in the three Kramer conditions and 65% correct in Davitz and Davitz). Sadness/grief is next (58% to 90% in Kramer and 49.2% for Davitz and Davitz), and love is least correctly identified in both studies (38% to 56% in Kramer, and 25% in Davitz and Davitz).

One of the greatest gaps in the emotion/speech literature is the absence of studies of real, naturally occurring emotion.[25] If one wishes to be

[24]Davitz and Davitz (1959) do not report percent-correct statistics but they do give, in their Table 1, the number of correct identifications for each emotion portrayed. These percentage figures were obtained by dividing the number correct for each emotion by the total number of portrayals, which was 240 in each case.

[25]One notable exception is an externalization study included in Williams and Stevens (1972) which they carried out collateral to the main one. They analysed acoustically the recorded commentary of the radio announcer who witnessed and described the Hindenberg disaster as it happened. It is difficult to imagine a speech sample more realistic and natural or a situation more extreme. Comparisions between the actual recording and a portrayal by an actor who had not heard the recording, but only read the transcript, showed similar emotion/non-emotion differences, suggesting that, at least for this one kind of very extreme emotion, portrayals can be fairly accurate. What is missing in this data is replications to determine whether such emotion cues are consistent across persons and across time. An experimental study with a number of emotion-producing situations and a number of speakers would allow statistical inference and tests of consistency as to the relationships between emotive states and vocal properties.

TABLE 6.2

Confusion Matrices, Information Theory Statistics and Percent-correct Statistics for English Unfiltered, English Filtered and Japanese Voices Portraying Five Emotions. (The Information Used to Construct this Table is from Kramer, 1964, p. 392, Table 1, with Information Theory Statistics and Corrected Percent-correct Statistics Added.)

English Unfiltered Voices						*Percentage of Correct Judgements*	*Corrected Percentage of Judgements*	*Percentage of Information Transmitted (Coefficient of Constraint)*
Intended Emotion	Judged Emotion a	c	g	i	l			
anger	74	26	0	0	0	74%[a]	67.5%	46.4% of the
contempt	3	85	3	9	0	85%	81.25%	information
grief	2	5	58	4	31	58%	47.5%	for the five
indifference	1	16	4	76	3	76%	70%	emotions was
love	2	14	17	11	56	56%	45%	transmitted
English Filtered Voices								
Intended Emotion	Judged Emotion a	c	g	i	l			
anger	77	23	0	0	0	77%	71.25%	38.0% of the
contempt	6	48	4	36	6	48%	35%	information
grief	0	3	71	5	21	71%	63.75%	for the five
indifference	6	11	7	63	13	63%[a]	53.75%	emotions was
love	0	12	20	19	48	48%	35%	transmitted
Japanese Voices								
Intended Emotion	Judged Emotion a	c	g	i	l			
anger	67	29	0	3	1	67%[a]	58.75%	42.5% of the
contempt	0	20	17	52	11	20%	0%	information
grief	0	0	90	1	10	90%	87.5%	for the five
indifference	2	21	2	73	2	73%	66.25%	emotions was
love	1	7	3	52	38	38%	22.5%	transmitted

[a] = median

precise, neither the Kramer study nor the Davitz and Davitz study really measures judgement accuracy. In both cases the experimenters have defined the emotion that they wish the judges to portray more in the actual

scenarios used to exemplify the emotion rather than just an emotion name or label, and yet judges are only given the emotion labels to choose from. It would be much more defensible to allow the judges to read the scenarios as well so that they would have the same opportunity as the speakers to interpret the emotion labels. The investigators have not established a logically tight *assessable reality* against which to measure judge accuracy.[26]

In order to move one step closer to natural emotion in spontaneous speech, we designed an experiment (Brown, Bowen, and Hamblin, 1977, see also the partial report of it in Brown, 1980b) to obtain speech samples of five male undergraduates in reaction to six films quite diverse in emotional impact. It is not easy to define the emotion generated by each film, but having the judges also view the film provides a situational definition. That is, judges are faced with the question "which film led to this utterance?" to which there is one indisputably correct answer and five indisputably incorrect ones. The emotion involved is probably somewhere between real emotion and portrayed. It is the vicarious emotion of the spectator. We found that judges could, on the basis of only the vocal qualities (for content was obscured by having speakers in a foreign tongue) identify the film with an accuracy almost as great as portrayed emotion in the earlier Davitz and Davitz and Kramer studies.

Even though it could be argued that this is one step further removed from real emotion than portrayals, in that it is an observation of a portrayal, one crucial element makes it more comparable to real emotion: the emotion is *subsidiary* to the task the viewer is carrying out rather than being the *focus* of his attention as in portrayals.[27] A great weakness of accuracy studies which use portrayed emotion lies in the fact that it is one thing to ask whether the feelings and emotions a person experiences in natural situations are detectable in vocal qualities, and quite a different thing to ask whether a person can produce portrayals of emotion or feelings vocally that are recognizable. The former are likely to be subtle. The latter run the risk of being nothing more than stylized caricatures.

We have been carrying out additional studies (see Brown, 1980a for some of the initial data) to gather recordings of vocal emotion in natural

---

[26]Polanyi (1962, p. 139) and Taylor (1964, pp. 76–80) both argue this kind of point effectively. The Polyani reference clearly shows the motivation behind "operationism" or "operational definitions" and the Taylor reference is a demonstration of the fallacy in operationism logic.

[27]"Subsidiary" and "focal" are being used here in the same sense that Polanyi (1962, chapter 4) uses them.

situations such as: the birth of the first child;[28] patients before and after a serious operation; college athletes before and after a major event; etc. All of these voices are being judged by a small group of listeners (with the voices played backwards or low pass filtered to occlude intelligibility) to evaluate the accuracy with which they can identify which voice sample came from which situation.

The next step will be a careful acoustic analysis of the temporal patterns, pitch/amplitude patterns and long-term power spectrum associated with the voice in various situations in order to test the hypotheses of the vocal marker links with specific emotions. From these results and the ones summarized by Scherer (1979a), hypotheses will be constructed and tested with an experimental design using speech synthesis methods.

An example of a major issue in the vocal realization of emotion that could be tested effectively using this research paradigm is whether the important markers are found in the temporal patterning of FØ) and amplitude or in the characteristic timbre (long term power spectrum).[29] The bulk of the research so far has dealt with the former. Scherer, both in the emotion/speech review chapter (1979b) and the personality/speech chapter (1979c) expresses the view that power spectrum data will be a rich source of information about an individual's psychological state, proposing as an example that the relative increase of energy above 500 Hz is an indicator or marker of stress and tension in the voice.[30] However, more research is needed.

---

[28]We realize that some of these studies deal with matters of great import and personal concern, and require interpersonal respect and sensitivity. The value of even very good data (which we hope some of this will be) is very small compared to the other values involved in such an event. We have sought to act in accord with the feelings of each couple and to give them adequate knowledge of how the voices will be used, in order to allow them to decline if they wish. (Participation is voluntary in the first place, but after the tape is made they are given a chance to indicate if they prefer the tape not to be used further.) We have also been aware of the need for selectivity in the kind and size of listener groups involved in accuracy tests—it would not do to have large unselected groups of undergraduates as listener subjects.

[29]A "long term power spectrum" is simply the summation of the power spectrum over a number of time slices. It is a stabilized indicator of timbre of the voice over the utterance as contrasted with the temporary timbre characteristics of any one time slice.

[30]Scherer quotes Laver (1975) in connection with this hypothesis. Laver's clarification of it (personal communication) is that the voice of stress and tension is increased in amplitude both above and below 500 Hz, but that the relative increase is greater above 500 Hz than below. However, it can be argued that driving any resonant sound box at higher amplitudes will result in disproportionately higher amplitudes above 500 Hz, and this particular power spectrum configuration may be only an artifact of the increased amplitude, rather than a separate vocal marker for tension. (See Strong and Plitnik, 1977 for a clear demonstration of this principle in connection with the clarinet, p. 244, the oboe, p. 247, and the flute, p. 256.)

In connection with the studies of emotion, there is also a need for work on the conceptual clarification of emotion theory as it can be applied to vocal reflections. Particularly of interest is the taxonomy and very clever conceptual analysis of emotions by Solomon (1977) from within the tradition of existentialism and self-deception theory (such as Nietzsche, Sartre and Camus). Especially relevant is his emphasis on a differentiation between those emotions that are more reactive (in which a vocal accompaniment could be considered an index of an inner state) and those that can be seen as a communicative message. For example, anger can be used as a way of accusing others as Lewis Carroll expressed in *Alice in Wonderland:* " 'I'll be judge, I'll be jury,' said cunning old fury."

The shortcomings of Skinner's *Verbal behaviour* (1957) have received at least as much criticism as they deserve, but there may be some important insights to be gained from dichotomizing language functionally into "mands," verbal operants intended to bring about a certain state of affairs, and "tacts," the conveyors of information (see also Chapter 9.) Beier (1966) in his theory of "evoking messages" clearly points out the habitual ways in which vocal operants can be an important part of personality structure:

> As an illustration of the evoking message we can consider an individual who thinks of himself as lonely and without any friends. In a careful analysis of his messages, we discover the coding of very subtle cues that are likely to create an emotional climate in the people he addresses, resulting in negative and angry feelings. The sender evokes this negative response but is nevertheless able to see himself as the victim of circumstances. As he was not aware that he coded this information, he does not have a feeling of responsibility for the response he obtained. Indeed, the evoking message seems to be the type of communication which maintains the patient's present state of adjustment. With this message he helps to create those responses in his environment which confirm his view of the world. Through the responses he elicits, he constantly obtains proof that the world is exactly the place he thinks it is!. . .The evoking message is probably one of the basic tools used by individuals to maintain their consistency of personality. . .The person with emotional conflict. . .creates a world in which he typically feels victimized by others, in which he experiences great unhappiness, though he has little awareness that he is often the creator of this world [pp. 12-13].

It seems reasonable to assume that a manner of speech that evokes rejection is often because of a lack of social skill, that it constitutes a failure, but Beier's position turns the tables by viewing it as a manifestation of skill. He suggests three observations that will demonstrate that it is skillful and purposive: (1) it is frequent; (2) it is integrated with the person's total personality; and, most importantly, (3) there seems to be no learning, no

improvement associated with it.[31] We have, then, purposive messages, the intent of which are directly contradictory to the person's avowed intent—a perfect fit to the classic description of self deception (*mauvaise foi)* given by Sartre (1943/1969, pp. 96-98)[32] and others (Fingarette, 1969; Warner, 1982) i.e., purposive messages which are directly contradictory to the subject's intentions.

Although the illustrations in Beier's 1966 book of the phenomenon are mostly anecdotal accounts from clinical experience, he has, with a rotating group of students, done considerable empirical work in this area.[33] A common paradigm in these studies is to have subjects (some of them with severe adjustment problems) portray various "moods" and then to have groups of judges categorize the voice according to mood. High inter-judge reliability in an assessment that is contradictory to the person's insistence of what he or she is portraying is taken as an index of the person's problem. The logic is that it is not so much a question of who is right about the mood, but that it is a blatant demonstration of social reality—that the person is not communicating to others what he or she acknowledges as his or her intention. It is argued further that this mistake is not accidental. There is a resistance to learning. There is no increment of improvement. Therefore, it must be the social impact intended by the person.

As the concept of emotion becomes less elusive, we can begin to focus on understanding processes underlying the judge's perception of emotion. Just as the Beierian theory leads us to expect that vocal paralinguistic cues are more accurate indicators of a person's intentions than his or her own avowals, we would also expect on the basis of some recent studies (e.g. Dicks, Brown & Wells, 1977; Sackeim & Gur, 1978) that judges' involuntary "tacit" reactions (perhaps eventually measurable by EEG, EMG of facial muscles, etc.) would be more accurate than explicit judgements to the extent that the judges are themselves self-deceived we theorize that those who are more well adjusted personalities will have explicit verbal judgements more in accord with their primary tacit apprehensions and will, therefore, be more accurate in their perceptions of emotions and personality traits in others. It may be time to re-open the accuracy-of-the interpersonal-perception research tradition. One of the most important directions for future work may be a return to the accuracy issue, using a

---

[31]This is our own summary drawn from discussions with Beier.

[32]This is the example of the woman who colludes in her own seduction, though the realization of the intent that is betrayed by the total purposive pattern of her actions would horrify her.

[33]Most of that work is only in the form of dissertations and theses available at the University of Utah libary. A summary of the findings has not yet been written.

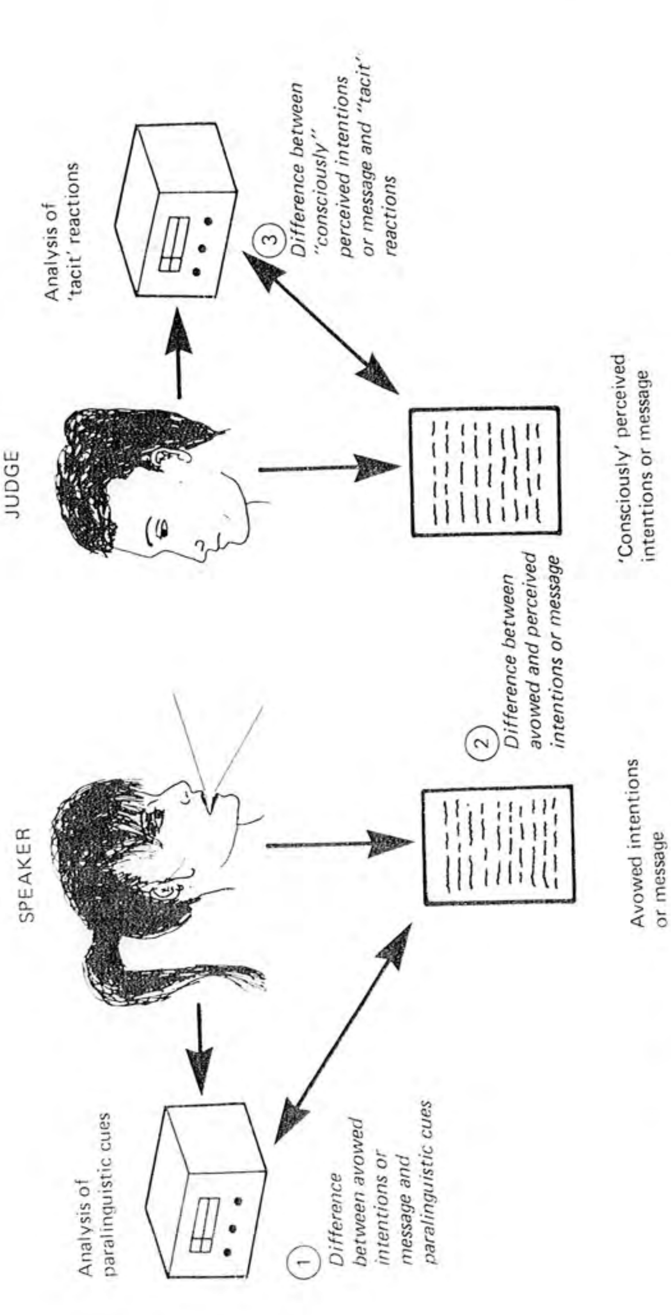

FIG. 6.8 Three possible research paradigms for a more adequate conception of emotion and speech.

logically tight accuracy paradigm to select accurate and inaccurate judges and then using the very precise attribution paradigms now in existence to uncover and analyze the attributional basis of inaccuracy and its relation to so-called "emotional disturbance" and what Szasz (1974) calls "the myth of mental illness," perhaps with research paradigms like those outlined in Figure 6.8.

REFERENCES

Addington, D. W. *The relationship of certain vocal characteristics with perceived speaker characteristics.* Unpublished dissertation, University of Iowa, 1963.

Addington, D. W. The relationship of selected vocal characteristics to personality perception. *Speech Monographs,* 1968, *35,* 492–503

Allport, G. W., & Cantril, H. Judging personality from voice. *Journal of Social Psychology,* 1934, *5,* 37–54.

Apple, W., Streeter, L. A., & Krauss, R. M. Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology,* 1979, *5,* 715–727.

Aronovitch, C. D. The voice of personality: stereotyped judgements and their relation to voice quality and sex of speaker. *Journal of Social Psychology,* 1976, *99,* 207–220.

Beier, E. G. *The silent language of psychotherapy.* Chicago: Aldine Publishing Company, 1966.

Brown, B. L. The detection of emotion in vocal qualities. In H. Giles., W. P. Robinson., & P. M. Smith (Eds.), *Language: social psychological perspectives.* Oxford: Pergamon Press, 1980. (a)

Brown B. L. Effects of speech rate on personality attribution and competency evaluations. In H. Giles., W. P. Robinson., & P. M. Smith (Eds.), *Language: social psychological perspectives.* Oxford: Pergamon Press, 1980. (b)

Brown, B. L., Bowen, D. E., & Hamblin, D. L. *The recognition of film-induced emotion from vocal qualities.* Unpublished manuscript, Brigham Young University, 1977.

Brown, B. L., Giles, H., & Thakerar, J. N. Speaker evaluation as a function of speech rate, accent and context. *Language and Communication,* 1985, *5,* in press.

Brown, B. L., & Lambert, W. E. A cross-cultural study of social status markers in speech. *Canadian Journal of Behavioural Science,* 1976 *8,* 39–55.

Brown, B. L. Strong, W. J., & Rencher, A. C. Acoustic determinants of perceptions of personality from speech. *International Journal of the sociology of language,* 1975, *6,* 11–32.

Brown, B. L., Strong, W. J., & Rencher, A. C. Fifty-four voices from two: the effects of simulations 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.

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

Davitz, J. R., & Davitz, L. J. The communication of feelings by content-free speech. *Journal of Communication,* 1959, *9,* 6–13

Dicks, R. H. Brown, B. L., & Wells, M. G. *A comparision of groups differing in mental health in their ability to judge and portray emotions.* Unpublished manuscript, Brigham Young University, 1977.

Diehl, C. F. Voice and personality: an evaluation. In D. A. Barbara (Ed.), *Psychological and psychiatric aspects of speech and hearing*. Springfield, Illinois: C. Thomas, 1960.

Eisenberg, P., & Zalowitz, E. Judgements of dominance-feeling from phonograph records of voice. *Journal of Applied Psychology,* 1938, *22,* 620-631.

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., & Pronovost, W. An experimental study of the pitch characteristics of the voice during the expression of emotion. *Speech Monographs,* 1939, *6,* 87-104.

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.

Fay, P. J., & Middleton, W. C. Judgement of intelligence from the voice as transmitted over a public address system. *Sociometry,* 1940, *3,* 186-191. (a)

Fay, P. J., & Middleton, W. C. Judgement of introversion from the transcribed voice. *Quarterly Journal of Speech,* 1942, *28,* 226-228.

Fay, P. J., & Middleton, W. C. Judgement of Kretschmerian body types from the voice as transmitted over a public address system. *Journal of Social Psychology,* 1940, *12,* 151-162. (b)

Fay, P. J., & Middleton, W. C. Judgement of leadership from transmitted voice. *Journal of Social Psychology,* 1943, *17,* 99-102.

Fay, P. J., & Middleton, W. C. Judgement of Spranger personality types from the voice as transmitted over a public address system. *Character and Personality,* 1939, *8,* 144-155.

Feldstein, S., & Welkowitz, J. A chronography of conversation: in defence of an objective approach. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behaviour and communication*. Hillsdale, New Jersey: Erlbaum, 1978.

Fingarette, H. *Self-deception*. New York: Humanities Press, 1969.

Giles, H. Ethnicity markers in speech. In K. R. Scherer & H. Giles (Eds.), *Social markers in speech*. Cambridge: Cambridge University Press, 1979.

Gorlitz, D. D. *Ergebnisse und probleme der ausdruckspsychologischen sprechstimmforschung*. Königstein, FRG: Meisenheim, 1972

Helfrich, H Age markers in speech. In K. R. Scherer and H. Giles (Eds.), *Social markers in speech*. Cambridge: Cambridge University Press, 1979.

Herzog, H. Stimme und Personlichkeit. *Zeitschrift für Psychologie,* 1933, *130,* 300-379.

Kjeldergaard, P. M. *Changes in perceived personality traits as a function of manipulation of vocal characteristics*. Unpublished manuscript. Harvard University, 1968.

Kramer, E. Elimination of verbal cues in judgements of emotion of voice. *Journal of Abnormal and Social Psychology,* 1964, *68,* 390-396.

Kramer, E. Judgement of personal characteristics and emotions from non-verbal properties of speech. *Psychological Bulletin,* 1963, *60,* 408-420.

Laver, J. Voice quality and indexical information. *British Journal of Disorders of Communication,* 1968, *3,* 43-54.

Laver, J. *Individual features in voice quality*. PhD dissertation, University of Edinburgh, 1975.

Laver, J. The concept of articulatory settings: an historical survey. *Historiographia Linguistica,* 1978, *5,* 1-14

Laver, J., & Trudgill, P. Phonetic and linguistic markers in speech. In K. R. Scherer and H. Giles (Eds.), *Social markers in speech*. Cambridge: Cambridge Univeristy Press, 1979.

Lomax, R. Social structure and sound change. In R. W. Wescott (Ed.), *Language origins*. Silver Springs, Md:, Linstock Press 1974.

MacLachlan, J. What people really think of fast talkers. *Psychology Today,* November 1979, 113-117.

Mahl, G. F., & Schulze, G. Psychological research in the extralinguistic area. In T. A. Sebeok, A. S. Hayes, & M. C. Bateson (Eds), *Approaches to Semiotics.* The Hague: 1964.

Mallory, P., & Miller, V. A. A possible basis for the association of voice characteristics and personality traits. *Speech Monographs,* 1958, *25,* 255-260.

Meehl, P. E., & Hathway, S. R. The *K* factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology,* 1946, *30,* 525-564.

Mischel, W. Towards a cognitive social learning reconceptualization of personality. *Psychological Review,* 1973, *80,* 252-283.

Moses. P. *The voice of neurosis,* New York: Grune & Stratton, 1954

Ostwald, P. F. Acoustic methods in psychiatry. *Scientific American,* 1965, *212,* 82-91.

Ostwald, P. F. *Soundmaking: the acoustic communication of emotion.* Springfield, Illinois: Charles C. Thomas, 1963.

Pear, T. H. *Voice and Personality.* London: Chapman & Hall, 1931.

Polanyi, M *Personal knowledge: towards a post-critical philosophy.* New York: Harper & Row, 1962.

Ptacek, P. H., & Sander, E. K. Age recognition from voice. *Journal of Speech and Hearing Research,* 1966, *9,* 266-272.

Robinson, W. P. Speech markers and social class. In K. R. Scherer and H. Giles (eds.), *Social markers in speech.* Cambridge: Cambridge University Press, 1979.

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

Rousey, C. L. (Ed.) *Psychiatric Assessment by speech and hearing behavior.* Springfield, Illinois: Charles C. Thomas, 1974.

Rousey, C. L., & Moriarty. A. E. *Diagnostic implications of speech sounds.* Springfield, Illinois: Charles C. Thomas, 1965.

Ryan, E. B. Why do low-prestige language varieties persist? In H. Giles & R. N. St Clair (Eds.), *Language and social psychology.* Oxford: Blackwell, 1979.

Sackeim, H. A., & Gur, R. C. Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology,* 1979, *47,* 213-215.

Sackeim, H. A., & Gur, R. C. Self-deception, self-confrontation and consciousness. In G. E. Schwartz and D. Shapiro (Eds.), *Consciousness and self-regulation: advances in research and theory* (Vol. 2). New York: Plenum Press, 1978.

Sartre, J. P. *Being and nothingness.* H. E. Barnes (trans.). New York: Washington Square Press, 1969. Originally published as *L'etre et le neant* in 1943.

Scherer, K. R. The functions of nonverbal signs in conversation. In R. N. St Clair & H. Giles (Eds.), *The social and psychological contexts of language.* Hillsdale, N.J. Lawrence Erlbaum Associates, 1979. (a)

Scherer, K. R. Inference rules in personality attribution from voice quality: the loud voice of extroversion. *European Journal of Social Psychology,* 1978, *8,* 467-487.

Scherer, K. R. Judging personality from voice: a cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality,* 1972, *40,* 191-210.

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

Scherer, K. R. Personality markers in speech. In K. R. Scherer and H. Giles (Eds.), *Social markers in speech.* Cambridge: Cambridge University Press, 1979. (c)

Siegman, A. W. The tell-tale voice: nonverbal messages of verbal communication. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behaviour and communication.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1978.

Skinner, B. F. *Verbal behaviour.* New York: Appleton-Century-Crofts, 1957.

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

Smith, P. M. Sex markers in speech. In K. R. Scherer and H. Giles (Eds.), *Social markers in speech.* Cambridge: Cambridge University Press, 1979.

Solomon, R. C. *The passions: the myth and nature of human emotion.* Garden City, New York: Anchor Books, 1977.

Stagner, R. Judgements of voice and personality. *Journal of Educational Psychology,* 1936, *27,* 272-277.

Stewart, M., Brown, B. L., & Stewart, S. *A comparison of computer manipulated speech rate with subjectively manipulated speech rate in effects upon personality attributions.* Manuscript in preparation, Brigham Young University.

Strong, W. J. & Plitnik, G. R. *Music, speech, and high fidelity.* Provo, Utah: Brigham Young University Press, 1979.

Szasz, T. S. *The myth of mental illness: foundations of a theory of personal conduct.* Revised edition. New York: Harper and Row, 1974.

Taylor, C. *The Explanation of behaviour.* New York: The Humanities Press, 1964

Taylor, H. C. Social agreements on personality traits as judged from speech. *Journal of Social Psychology,* 1934, *5,* 244-248.

Titze, I. R. The human vocal cords: a mathematical model, part I. *Phonetica,* 1973, *28,* 129-170.

Titze, I. R. The human vocal cords: a mathematical model, part II. *Phonetica,* 1974, *29,* 1-21.

Titze, I. R., & Talkin, D. T. A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *Journal of the Acoustical Society of America,* 1979, *66,* 60-74.

Trimboli, F. Changes in voice characteristics as a function of trait and state personality variables. *Dissertation Abstracts International,* 1973, *33,* 3965.

Tyler, L. E. Test review No. 77. In O.K. Buros (Ed.), *The Fourth Mental Measurements Yearbook.* Highland Park, N.J: The Gryphon Press, 1953, 135-140.

Warner, C. T. *Bondage and emancipation: a conceptual introduction to the study of motivation.* Unpublished 19-chapter manuscript, Brigham Young University, 1982.

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