Perceptions of personality from speech: effects of manipulations of acoustical parameters*

Bruce L. Brown, William J. Strong, and Alvin C. Rencher

Language Research Center, Brigham Young University, Provo, Utah 84601

(Received 17 October 1972; revised 5 January 1973)

A speech analysis–synthesis system was used to manipulate variance of fundamental frequency and a mechanical rate changer was used to manipulate speech rate. The synthesized and altered voices were tested for realism. Synthesized voices were mistaken for normal 50% to 58% of the time and rate-changed voices were mistaken for normal 78% of the time. Additional studies were conducted to test the effects of these acoustical manipulations on the adjective ratings judges made of speakers. Variance of intonation was increased and decreased by 50% for eight speakers. There was a significant trend for increased intonation to cause voices to be rated more "benevolent" by judges and decreased intonation to cause them to be rated less "benevolent." In two additional studies, rate was decreased and increased by varying amounts. Slowing the voices caused them to be rated less "competent." Speeding the voices caused them to be rated less "benevolent." Results were more consistent over speakers for rate manipulations than for intonation manipulations.

Subject Classification: 9.5.

INTRODUCTION

There has been a long history of interest in the ways in which personality and emotion are reflected in speech and the inferences judges make of personality and emotion from listening to speech. Kramer,[1] Diehl,[2] Mahl and Shulze,[3] and Hymes[4] have reviewed the history of hypotheses and research regarding the relationships between voice and personality and also voice and emotion. Duncan[5] reviewed some of the more recent studies in the context of a more general review of all kinds of nonverbal communication. These studies employ almost exclusively subjective judgments of speech.

The development in the last decade of advanced techniques of acoustic analysis is only now beginning to reach these areas. Williams and Stevens[6] have made spectrographic analyses of recordings of acted emotional utterances and emotional utterances in natural situations.

In the area of personality even less has been done in finding acoustic parameters. Ostwald[7] has done some preliminary work in observing the spectrographic characteristics of various categories of neurosis and psychosis. A great deal of work has been done in recent years on speaker identification and recognition from various analyses of acoustic data,[8] but the concern has been to correctly match acoustic patterns with the specific person rather than to match acoustic patterns with general personality characteristics or categories. This paper describes a beginning attempt to find and measure acoustic parameters that are indicative of various personality characteristics. In later work, behavioral and verbal measures of personality will be compared with acoustic properties of the person's voice, but the studies discussed here report only relationships between acoustic properties of the speech and judgments listeners make of personality from that speech. Brown[9] has attempted to study the relationship by correlating personality judgments with subjective judgments of the linguistic properties of utterances from which the personality judgments were made. This correlative approach precludes any statements about causation and doesn't allow one to isolate the effects of particular linguistic variables. Kjeldergaard[10] had speakers consciously alter their voices one dimension at a time and then evaluated the effects of each manipulation on trait ratings. Although one can get some idea of the effects of various manipulations from this kind of approach, the problems in uniformity or accuracy of manipulations are obvious. The studies reported in this paper attempt to avoid these problems by altering voices one parameter at a time by means of a rate changer and a speech synthesizer and then assessing the effects of such manipulations. The first study in this paper is a preliminary assessment of the realism of synthesized and rate-changed voices and Studies II, III, and IV present the results of the personality judgments of various manipulated voices.

I. SPEECH ANALYSIS–SYNTHESIS AND RATE-CHANGE PROCEDURES

The voices of Studies III and IV were manipulated in rate by means of an Eltro Automation Rate Changer, a mechanical tape-recorder device which can alter either pitch or speaking rate while the other is held constant. The device has a cylindrical head that can be rotated either forward or backward at variable rates, thus altering frequency while rate is controlled by the tape speed. By coordinating head rotation and tape speed, either rate or pitch can be altered while the

TABLE I. Newman–Keuls analysis on group means of 11 "voice" types with unequal numbers of replications in each group.

Groups	N	Means (in mm)	Normal	Rate changer (rate decreased)	Computer generated (intonation decreased)	Computer generated (pitch increased)	Rate changer (rate increased)	Computer generated (intonation increased)	Computer generated (untouched)	Computer generated (pitch decreased)	Laryngectomy	Bell by rule
Computer synthesized (degenerated)	4	76.8	61.3[a]	52.4[a]	45.6[a]	45.2[a]	43.2[a]	42.8[a]	41.5[a]	38.7[a]	8.9	5.8
Bell by rule	4	71.0	55.5[a]	46.6[a]	39.8[a]	39.4[a]	37.4[a]	37.0[a]	35.7[a]	32.9[a]	3.1	
Laryngectomy	4	67.9	52.4[a]	43.5[a]	36.7[a]	36.3[a]	34.3[a]	33.9[a]	32.6[a]	29.8[a]		
Computer generated (pitch decreased)	4	38.1	22.6[a]	13.7	6.9	6.5	4.5	4.1	2.8			
Computer generated (untouched)	8	35.3	19.8[a]	10.9	4.1	3.7	1.7	1.3				
Computer generated (intonation increased)	8	34.0	18.5[a]	9.6	2.8	2.4	0.4					
Rate changer (rate increased)	8	33.6	18.1[a]	9.2	2.4	2.0						
Computer generated (pitch increased)	2	31.6	16.1[a]	7.2	0.4							
Computer generated (intonation decreased)	8	31.2	55.7[a]	6.8								
Rate changer (rate decreased)	8	24.4	8.9[a]									
Normal	24	15.5										

[a] Significant at 0.01 level.

other is held constant. The rate manipulations by this method are rather believable, but pitch manipulations are less believable because the machine alters formant frequencies as well as the fundamental frequency, giving the effect of a lengthened vocal tract for lowered pitch and a shortened vocal tract for raised pitch.

The voices of Study II were manipulated in fundamental frequency by means of an automatic computer-based speech analysis–synthesis scheme.[12,13] In the analysis, new parameters are calculated each 10 msec. A spectrum is computed by means of a 512-point fast Fourier transform. A Hamming window of 40-msec duration is used when the speech is judged to be voiced by a gross time analysis of the waveform. A 10-msec window is used for unvoiced speech. Fundamental frequency was measured by using the cepstrum method. A smoothed spectrum was computed from the 30 low-order cepstral coefficients and peak-picked to determine five formant frequencies and amplitudes. Some smoothing was done on the fundamental-frequency and formant-frequency contours to eliminate gross discontinuities. All of the analysis operations were completely automatic.

The synthesis was accomplished with a five-pole parallel synthesizer which was simulated on a Librascope L-3055 computer. Eleven input parameters (*FØ*, five formant frequencies, and five formant amplitudes) were input to the synthesizer at 10-msec intervals and

TABLE II. Confusion matrix of judges' accuracy in identifying the 11 voice types.

	Judged											
Actual	Normal	Rate increased	Rate decreased	Synthesized untouched	Pitch raised	Pitch lowered	Intonation increased	Intonation decreased	Laryngectomy	Bell by rule	Synthesized degenerated	
Normal	610	12	88	17	23	23	20	27	70	14	8	912
Rate increased	10	239	4	2	6	2	25	3	5	4	4	304
Rate decreased	76	7	60	34	12	6	27	28	24	17	13	304
Synthesized untouched	36	27	10	49	7	11	45	43	21	19	36	304
Pitch raised	6	4	3	18	5	1	7	8	7	9	8	76
Pitch lowered	19	4	13	24	3	24	17	17	9	9	13	152
Intonation increased	35	35	12	48	15	9	53	22	15	36	24	304
Intonation decreased	39	32	23	37	7	14	34	35	17	26	40	304
Laryngectomy	0	0	2	24	3	9	0	6	37	38	33	152
Bell by rule	0	0	18	26	0	24	3	11	10	32	28	152
Synthesized degenerated	0	0	2	22	5	23	4	12	5	45	34	152
Total	831	360	235	301	86	146	235	212	220	249	241	3116

the input parameters were linearly interpolated as needed in the synthesizer. The speech was output by means of a 12-bit digital-to-analog converter and recorded on audio tape for testing.

II. STUDY I: REALISM OF SYNTHESIZED AND RATE-CHANGED VOICES

A. Method

A testing tape was made with 82 male voices of the following types: (1) 30 synthesized voices, (2) 16 rate-changed voices, (3) 24 normal voices, and (4) 12 "unnatural" voices for comparison. Twenty-four of the 30 synthesized voices were the ones used in Study II (details of speaker characteristics, manipulations, and text are given there). The other six were some of the same speakers altered in average FØ, two with FØ increased by 20 Hz, and four with it decreased by 20 Hz. The 16 rate-changed voices are the four voices at four rate manipulation levels which were used in Study IV. The 24 "normal" voices were some of the male university students and trade school students recorded as described in the first part of Study II, which were not synthesized. Four of the unnatural-sounding comparison voices were taken from a tape of voices generated by rule (rather than by the parameter extraction method of this study) at Bell Laboratories. Another four of them were recordings of adult males who had undergone laryngectomy and who spoke either with an external vibrator or with sound generated by pharyngeal air expulsion. The final four were voices processed by the parameter extraction method but which were degenerated in the synthesis phase by either deleting some of the formants or by exciting the system with noise instead of a pulse train.

The "normal," rate-changed and synthesized voices could have all been presented speaking the same passage (the four "Little Prince" sentences presented in Study II), but unfortunately the unnatural-sounding comparison voices could not. There is a different passage for each of the 12 "unnatural" voices. For this reason the "normal," synthesized and rate-changed voices were also presented using a variety of passages, with each passage being three sentences long.[13] Voices were assigned with approximately equal numbers of each type of voice within each quarter of the tape, with a random order of voices within each quarter. The first two voices were extreme ones—a "normal" voice and one of the Bell Labs synthesis-by-rule voices—in order to give judges anchor points, at the beginning, of the most and least natural voices they would hear.

The tape was judged in three different ways:

(1) *Rating Scale*. Seventeen college-age male and female judges were given 100-mm scale lines with "realistic" on the left end point and "unrealistic" on the right. They were instructed to rate each voice by placing a mark on the line. The marks were measured with a ruler and the ratings were averaged over judges for each of the voices.

(2) *Binary Judgments*. Thirty-eight college-age male and female judges listened to the tape with instructions to categorize each voice as either "natural" or "synthetic." They were given no indication of the actual proportion of natural or synthetic voices in the tape.

(3) *Categorization of Voices*. This is a very stringent test of voice realism. Judges were given a matrix numbered from 1 to 82 across the top and with the 11 voice categories together with the percentage of the 82 voices that fell in each listed down the side. Percentages were used rather than actual numbers of each category so that judges wouldn't worry about getting the right number of each, but simply have an idea in advance of what proportion of each to expect. As they listened to each voice they placed a mark in the box below the voice number and in line with the category they thought it belonged to.

B. Results

In analyzing the *rating-scale* judgments of realism the voices were grouped into the 11 categories (shown in Table I), and a one-way analysis of variance was computed on the averaged ratings (averaged over judges) with the 11 categories as treatments. The F value was 40.085 ($p<0.01$). Table 1 shows the results of a Newman–Keuls analysis[14] of the means of the 11 groups as well as the actual means and the number of voices in each group. From the pattern of the significant comparisons it is obvious that the voices fall into three major groupings in terms of realism. The natural voices were, of course, the most realistic of the groupings. The second grouping includes the rate-changed voices and the computer-generated voices. The third and least realistic-sounding grouping is the combination of the Bell Laboratories voices generated by rule, the laryngectomy voices, and the degenerated synthesized voices. From the means given in Table I it can be seen that although the voices used in Studies II, III, and IV are not as realistic-sounding as natural voice, they are much more realistic-sounding than voices of those three categories.

The confusion matrix showing the distribution of correct guesses and of errors for the *categorization* task is given in Table II. Information theory statistics[15] were computed on the matrix. The uncertainties for actual and judged are 3.1381 and 2.9251, respectively. The transmission is 0.3537 and the coefficient of constraint of actual on judged is 0.1209. This is very low, indicating that judges cannot accurately identify type of voice (without previous experience) when all of the categories are considered as a whole. However, an examination of the row for normal voices shows that they are guessed to be normal most often, which isn't very surprising. It is important to note that they were categorized as something other than normal about

TABLE III. Information theory statistics and confusion matrix comparison of normal voices with each of six other categories of voices.

I.	Judged			II.	Judged		
Actual	Natural	Synthetic		Actual	Natural	Synthetic	
Normal	390	18	408	Normal	390	18	408
Rate changed	212	60	272	Synthesized (altered)	215	159	374
Total	602	78	680	Total	605	177	782
$U(X)=0.9710$	$T(XY)=0.0529$			$U(X)=0.9986$	$T(XY)=0.1650$		
$U(Y)=0.5139$	$C_{y.x}=0.1029$			$U(Y)=0.7716$	$C_{y.x}=0.2138$		

III.	Judged			IV.	Judged		
Actual	Natural	Synthetic		Actual	Natural	Synthetic	
Normal	390	18	408	Normal	390	18	408
Synthesized (untouched)	68	68	136	Laryngectomy	29	39	68
Total	458	86	544	Total	419	57	476
$U(X)=0.8113$	$T(XY)=0.1840$			$U(X)=0.5918$	$T(XY)=0.1646$		
$U(Y)=0.6297$	$C_{y.x}=0.2922$			$U(Y)=0.5285$	$C_{y.x}=0.3114$		

V.	Judged			VI.	Judged		
Actual	Natural	Synthetic		Actual	Natural	Synthetic	
Normal	390	18	408	Normal	390	18	408
Bell by rule	10	58	68	Synthesized (degenerated)	7	61	68
Total	400	76	476	Total	397	79	476
$U(X)=0.5918$	$T(XY)=0.3224$			$U(X)=0.5918$	$T(XY)=0.3576$		
$U(Y)=0.6336$	$C_{y.x}=0.5120$			$U(Y)=0.6485$	$C_{y.x}=0.5500$		

one-third of the time. The figures in the normal voices column show that a little over 10% on the average of the computer-synthesized voices are categorized as normal, while 25% of the rate-changer speed-decreased voices and a little over 3% of the rate-changer speed-increased voices are so judged. Under such a stringent test, where even one-third of normal voices are categorized as other than normal, it is encouraging that any of the rate-changed or computer-synthesized voices are mistaken for normal. None of the Bell Laboratory's computer-degenerated or laryngectomy voices are ever mistaken for normal.

The results of the *binary judgments* of realism (Table III) are probably more applicable than the results of the *categorization* to studies of person perception utilizing these speech altering methods, since judges in such a study would not be told the proportion of voices that are synthesized. If anything, even the binary judgments underestimate how real the voices sound to judges in such a study, since subjects in this realism study were asked to look for synthesized voices whereas it would not be suggested in a person perception study that the voices are not "real." (When we ran Study II, we found most judges to be quite surprised when we told them that some of the voices were computer generated.) For analyzing the data of the binary judgments, the voices were collapsed from the 11 categories into seven: (1) normal voices, (2) rate changed with both increased and decreased combined, (3) computer synthesized with all manipulations combined, (4) computer synthesized with no manipulations, (5) laryngectomy voices, (6) Bell voices by rule, and (7) computer-synthesized degenerated voices. Table III gives the confusion matrices[16] and information theory statistics for the comparisons of normal voices with each of the other six categories. This analysis shows rate-changer voices to be more easily mistaken for normal voices (78% of the time) than synthesized voices are (50% to 58% of the time). This finding in favor of greater realism for rate-changer voices didn't show up in the two other kinds of realism test (Tables I and III). Even at that, computer-synthesized voices appear to be fairly convincing in that more of them were judged in this analysis to be natural than were judged to be synthesized.

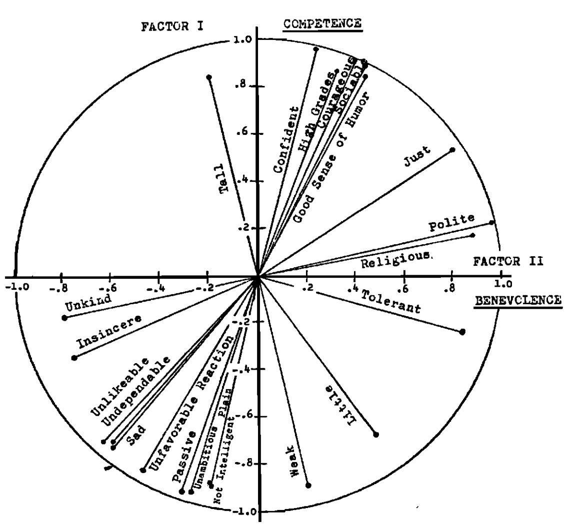

FIG. 1. Factor pattern for the adjective ratings of voices with intonation manipulated.

III. STUDY II: MANIPULATION OF VARIANCE OF FØ

A. Method

Twenty-four male university students and 24 male trade-school students (in order to give a diverse sample of young male speakers) were recorded as they recited the following four sentences:

The second planet was inhabited by a conceited man.

"Ah! Ah! I am about to receive a visit from an admirer!" he exclaimed from afar, when he first saw the little prince coming.

For, to conceited men, all other men are admirers.

"Good morning," said the little prince, "that is a queer hat you are wearing."

Fifty-six male and female judges rated the 48 speakers on 22 paired–opposite adjectives with a seven-point scale between poles. (The left pole of each is shown in Fig. 1). The averages of their ratings were factor analyzed by the principle axes method with a varimax rotation. The two resulting factors were labeled "competence" and "benevolence." To some extent, the factors obtained are dependent upon the adjectives used. These adjectives were selected to be comparable with other studies[7,18] which led up to the ones reported here. Studies using a variety of other adjectives and other techniques,[19,20] including the more open-ended method of multi-dimensional scaling, have found person perception factors closely related to "competence" and "benevolence."

From the plottings of 48 speakers' factor scores, eight speakers were selected to be representative of the 48. The eight samples were analyzed and synthesized with the computer scheme described previously. Each voice was synthesized in three forms: (1) unmanipulated control, (2) variance of *FØ* from the mean increased by 50%, and (3) this variance decreased by 50%. Only the first two sentences of the four sentence "Little Prince" passage were synthesized. A tape with a "practice voice" at the beginning[21] and the 24 synthetic voices interspersed with 12 other "filler" voices with each voice saying the two sentences twice was then played to 50 college-age male and female judges who rated the voices on the adjectives of Fig. 1.

B. Results

Figure 2 is a plotting of each of the 24 synthesized samples (three forms each for eight voices). Each speech sample is plotted according to the factor scores received on competence and benevolence. The three forms for each speaker ("I" for increased variance of *FØ*, "N" for normal, and "D" for decreased) are connected with broken lines. These three letters when enclosed in squares in Fig. 2 show were the centroids or mean vectors for each manipulation fall. A randomized-block multivariate analysis of variance was computed on the factor scores with individual speakers as blocks and the three forms for each voice (two manipulated and one normal) as treatments. The Wilks' lambda[22] value is 0.3865, which is easily significant beyond the 0.05 level with two degrees of freedom for treatments and 14 degrees of freedom for error. The individual F values for the competence factor and the benevolence factor are 2.902 (not significant) and 6.819 ($p<0.05$), respectively. It is obvious from the relative positions of the centroids in Fig. 2 as well as from the statistics that the main effect of variance of *FØ* increase upon a speaker is to make him sound more benevolent while decreased makes him sound less benevolent. A secondary trend (which is not statistically significant) can be observed from Fig. 2, for increased variance of *FØ* to make a speaker sound more competent and decreased, less competent. This is somewhat consistent with the findings of Brown[23] of a 0.53 correlation between judgments of *FØ* variance and competence scores, with a correlation of 0.50 between judgments of *FØ* variance and benevolence scores.

IV. STUDY III: MANIPULATIONS OF RATE

A. Method

The voices of 13 adult males were recorded as they recited the four sentences of Study II. The voices of four of the speakers were increased between 56% and 57% in rate[24] and four were decreased by 50%–55% by means of the rate-changer. The seven[25] manipulated voices with their corresponding seven unmanipulated controls were randomly ordered with six filler voices, with one filler voice at the beginning as the practice voice. The tape of these 20 speakers was played to 30 college-age male and female judges who rated the speakers on the adjectives of Fig. 1.

B. Results

The ratings received by each speech sample averaged over the 30 judges were factor analyzed. Figure 3 is a plotting of each speaker's normal and manipulated voice according to the factor scores of his received ratings. The Wilks' lambda for a MANOVA of the rate decreased manipulations is 0.0244 ($p<0.05$ with one degree-of-freedom treatment and three-degree-of-freedom error). The F values for competence and benevolence are 110.46 ($p<0.005$) and 0.536 (not signif-

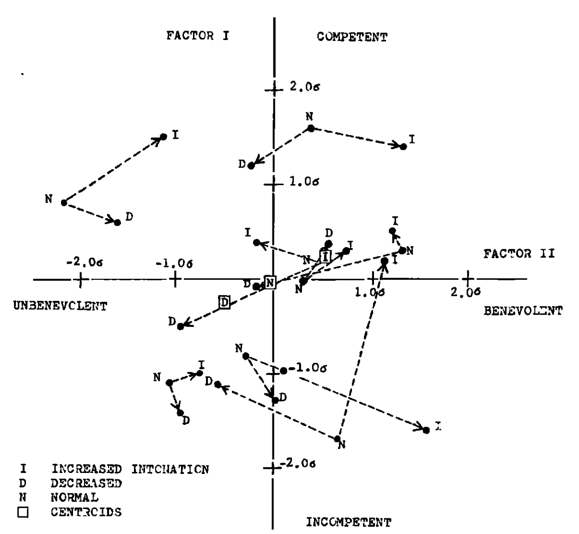

FIG. 2. Factor scores of adjective ratings of voices with intonation manipulated.

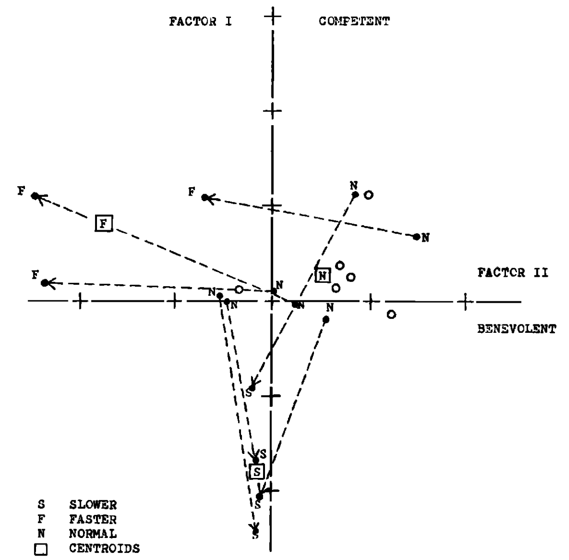

FIG. 3. Factor scores of adjective ratings of voices with rate manipulated.

icant), respectively. Slowing the voices, then, had no effect on benevolence ratings but a highly significant and consistent effect in decreasing competence ratings. The Wilks' lambda value for the rate increase manipulation is 0.0033, which isn't low enough for significance at the 0.05 level (0.0025 needed for the one-degree-of-freedom hypothesis and two-degree-of-freedom error in the randomized blocks MANOVA with two treatments and three blocks) but is significant at the 0.10 level. However, the individual F value for benevolence is highly significant ($F=224.29$, $p<0.005$) on this manipulation. The F value for competence is 2.920, which is not significant. It appears, then, that speeding the voice makes it sound less benevolent and slowing it makes it sound less competent. Since slowing the voice makes it sound less competent, one might expect speeding it to make it sound more competent, but it seems from these results that such is not the case. In every case the voice rated most positively, when both competence and benevolence are considered simultaneously, was the natural voice.

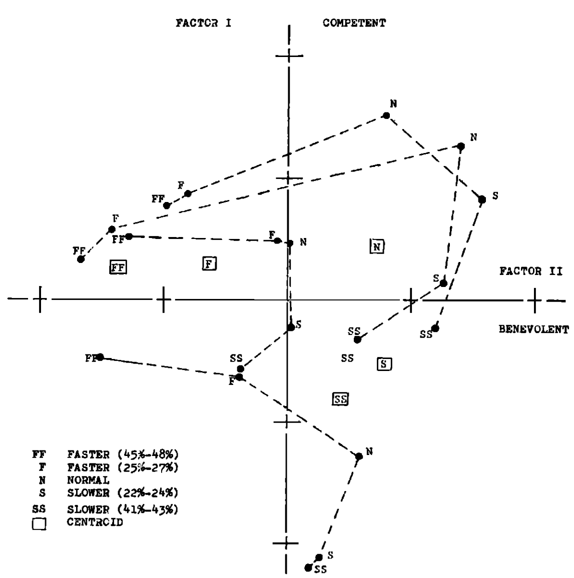

FIG. 4. Factor scores of adjective ratings of voices with five levels of manipulated rate.

V. STUDY IV: COMPARISON OF FIVE LEVELS OF RATE[26]

A. Method

Four of the voices used in Study III were also used in this study. A tape was constructed with five rates (normal, 45%–48% increase, 25%–27% increase, 22%–24% decrease, and 41%–43% decrease) for each of the four voices, making a total of 20 voices randomly ordered on the tape plus the practice voice at the beginning. The tape was played to 30 college-age female judges who rated the voices on the adjectives of Fig. 1.

B. Results

Figure 4 gives the plottings of the factor scores of the five manipulations for each of the four speakers from the factor analysis of the ratings averaged over the 30 judges. Again the general trend is for decreased rate to cause the speaker to be rated less competent and for increased rate to cause him to be rated less benevolent. In Study III, four of the voices were sped up and four different ones were slowed down; therefore, two separate MANOVA's were computed. In this study the same four were both sped up and slowed down, so only one MANOVA with five treatments was computed. The overall lambda value for rate manipulations is 0.048137. Since a lambda of 0.1965 is required for 0.01 level for four and 12 degrees of freedom, this is easily significant beyond the 0.01 level. The individual F's for competence and benevolence are 9.250 and 11.082, respectively, which are both also easily significant beyond the 0.01 level. The effects of rate manipulations are so consistent over speakers that even with a small number of speakers the effects of manipulations are highly significant. Even in Study II, the intonation study, where effects are not quite so consistent, eight speakers (with corresponding 16 manipulated and eight normal voices) were enough to detect a significant relationship.

C. Conclusions

Although the manipulations of variance of $F\emptyset$ have a statistically significant effect in increasing benevolence as variance increases and decreasing benevolence as variance decreases, the results of Study III and Study IV are more encouraging than those of Study II in that the rate manipulations had the same effect upon every speaker to whom they were applied. If this were true of other manipulations, it would be feasible to generate simple models that would predict personality judgments from certain acoustic properties. Since intonation manipulations seem to have somewhat different effects upon each speaker, future work will be directed toward many simultaneous manipulations (on two or three dimensions) of a small homogeneous sample of voices in order to determine the interactive effects of manipulations.

Of the two methods of altering one parameter of speech at a time, rate changer and computer analysis–synthesis, the former is by far the easiest, quickest, and cheapest. The latter requires a computer that can gather and store analog data at a high rate and some rather sophisticated programming. The rate changer is limited in that it can only alter rate and pitch, and even the pitch manipulations have severe limitations in that all formant frequencies are raised or lowered by the same proportion. For any manipulations other than rate, then, the computer method is to be preferred, if cost is not a factor. Even for rate manipulations the computer method is potentially superior in that the rate changer speeds or slows all segments of the utterance the same, whereas in actual speech speeding is primarily accomplished in the vowels with the consonant duration remaining constant.

Voice manipulation by the rate changer and computer synthesis methods could have broad application in social psychology and a number of applied areas. The effect of voice parameters upon attitude change and upon experimenter effects could be studied with these methods, and there are obvious applications to studies of interpersonal perception. The method could be used in applied settings to determine effective vocal qualities for advertisers on the air. Such procedures could also lead to useful methods of personality diagnosis by machine from the acoustic properties of speech. Perhaps there could even be some applications in developing immediate feedback devices for correcting undesirable vocal patterns.

*This research was supported by a grant from the National Institutes of Mental Health.

[1]E. Kramer, "Judgment of Personal Characteristics and Emotions from Nonverbal Properties of Speech," Psychol. Bull. **60** (4), 408 (1963).

[2]C. F. Diehl, "Voice and Personality: An Evaluation," in *Psychological and Psychiatric Aspects of Speech and Hearing,* edited by D. A. Barbara (C. Thomas, Springfield, Ill., 1960).

[3]G. F. Mahl and G. Schulze, "Psychological Research in the Extralinguistic Area," in *Approaches to Semiotics*, edited by T. A. Sebeok, A. S. Hayes, and M. C. Bateson (Mouton, The Hague, Netherlands, 1964).

[4]B. D. Hymes, "Linquistic Aspects of Cross-Cultural Personality Study," in *Studying Personality Cross Culturally,* edited by B. Kaplan (Harper-Row, Evanston, Ill., 1961).

[5]S. Duncan, "Nonverbal Communication," Psychol. Bull. **72** (2), 118 (1961).

[6]C. E. Williams and K. N. Stevens, "Emotions and Speech: Some Acoustical Correlates," J. Acoust. Soc. Am. **52,** 1238 (1972).

[7]P. E. Ostwald, "Acoustic Methods in Psychiatry," Sci. Am. **212** (3), 82 (1965).

[8]J. J. Wolf, "Efficient Acoustic Parameters for Speaker Recognition," J. Acoust. Soc. Am. **51,** 2044 (1972).

[9]B. L. Brown, "A Social Psychology of Variations in French Canadian Speech Styles," Unpublished Doctoral Dissertation, McGill Univ. (1969).

[10]P. M. Kjeldergaard, "Changes in Perceived Personality Traits as a Function of Manipulation of Vocal Characteristics," Unpublished Manuscript, Harvard Univ. (1968).

[11]W. J. Strong, "Machine-Aided Formant Determination for Speech Synthesis," J. Acoust. Soc. Am. **41,** 1434 (1967).

[12]W. J. Strong and R. B. Purves, "Scheme for Automatic-Formant Analysis of Speech," J. Acoust. Soc. Am. **51,** 110 (1972).

[13]Since we only had the four "Little Prince" sentences rate-changed and the first two of them synthesized, these passages sometimes consisted of one of those sentences repeated three times, sometimes another, sometimes two of the sentences with one repeated and the other not, and sometimes with three of the four. Eleven of the "normal" speakers also used these combinations of "Little Prince" sentences. These precautions were taken so that sentence spoken would not become a cue to the category of the voice.

[14]B. J. Winer, *Statistical Principles in Experimental Design* (McGraw–Hill, New York, 1962), pp. 80–85.

[15]B. F. Attneave, *Application of Information Theory to Psychology: A Summary of Basic Concepts, Methods and Results* (Holt, Rinehart and Winston, New York, 1959).

[16]G. A. Miller and P. E. Nicely, "An Analysis of Perceptual Confusions Among Some English Consonants," J. Acoust. Soc. Am. **27,** 338 (1955).

[17]B. L. Brown, Ref. 9.

[18]W. E. Lambert, "A Social Psychology of Bilingualism," J. Social Issues 23 (2), 91 (1967).

[19]K. H. Smith, D. M. Pedersen, and R. E. Lewis, "Dimensions of Interpersonal Perception in a Meaningful On-Going Group," Perceptual and Motor Skills 22 (3), 867 (1966).

[20]R. J. Hart and B. L. Brown, Personality Information Contained in the Vocal Qualities and in Content Aspects of Speech," Unpublished Manuscript, Brigham Young Univ. (1968).

[21]In previous work we have found that ratings of the voice in the first position differ markedly from ratings given the same voice later in the tape. Ratings given the second and subsequent voices have been found to be very reliable when compared with ratings of the same voice later in the tape. For this reason we begin each testing tape with a "practice voice."

[22]W. W. Cooley and P. R. Lohnes, *Multivariate Data Analysis* (Wiley, New York, 1971), pp. 223–242.

[23]B. L. Brown, Ref. 9, p. 91.

[24]Although one can alter rate continuously over a wide range with the rate-changer, the rate selector is not precise. We timed the voices after they were altered and found the slight variations indicated.

[25]A mistake was made on one of the voices increased in rate which made it unusable, leaving only three voices with rate increased.

[26]This study was planned and carried out by Mrs. Erma Argyle under the direction of Brown and Rencher as part of a master's thesis.