Communication researchers would like to know if the quality of a speaker's voice is a reliable indicator of an effective communicator and if voice quality can reveal information about the speaker's intelligence, leadership ability, and personality type. Experimental studies involving over 400 undergraduates at the University of Nebraska sought to determine the way subjects in groups of different levels of intelligence judge a neutral message and the personality of the speaker based on a positive paralinguistic structure. Results showed that there were no significant differences among the groups in rating the topic, but a difference was observed in the way each group evaluated the speaker's personality. A possibility exists that subjects of superior intellect, in paying less attention to the paralinguistic message, tend to listen more critically to the linguistic message and use the message to rate the speaker. However, since grade point averages were the only criterion for classifying intelligence, it is possible that the data may be incomplete and that more complex research is needed. (DS)
A POSITIVE PARALINGUISTIC STRUCTURE AND ITS
EFFECTS ON ATTITUDE CHANGE AND JUDGMENT
OF PERSONALITY DIMENSIONS

Earl E. McDowell
University of Minnesota
St. Paul Campus
October, 1973
The two purposes of this study are (1) to review literature that is related to the effects of voice quality and the judgment of the following personality dimensions of the speaker: intelligence, introversion, extroversion, and leadership, and (2) to report an experimental study concerned with neutral subjects of various intelligence groups, determined by previous grade point averages, and the way the subjects judge a neutral message and the personality dimensions of the speaker based on a positive paralinguistic structure. The results indicate that there were no significant differences in attitude change among the three intelligence level groups. Subjects from the average and the above average groups rated the personality dimensions of the speaker significantly higher than the subjects from the superior intelligence group.
During the past 2000 years speech scholars have contended that the speaker's voice quality is an important attribute to determine whether or not he will be a successful communicator and will be accepted by the audience. Only recently, during the past forty years, have psychologists, sociologists, linguists, semanticists, and speech scholars subjected voice quality to experimental research. Researchers are concerned with the effects voice quality has on the evaluation of the personality dimensions of a speaker.

The purposes of this study are (1) to review literature that deals with the effects of voice quality on the judgment of personality variables of the speaker and (2) to report a recent study completed by this researcher which presents additional information about the role of voice quality in the speech-communication process. Specifically, this researcher is interested in the effects that voice quality has on the following personality dimensions: intelligence, leadership, introversion, and extroversion. In reference to the above variables the researcher will answer the following question: What effects does the voice quality of the speaker have on the evaluation of his personality dimensions? As this relationship is traced, the researcher will point out whether or not the conclusions drawn are reliable and valid. Upon the completion of the review of literature, a report of the study recently completed will be presented. In addition, suggested future research which will extend past research that needs to be completed in order to determine the exact role of the voice in the total communication process will be presented.
Allport and Cantril (1931) conducted the first significant research, which included a total of eight experiments. Two experiments were conducted by radio in which listeners were asked to send in judgments concerned with such personality variables as age, body characteristics, dominance, submissiveness, extroversion, and introversion. Six additional experiments were conducted at Harvard Psychological Laboratory. Allport and Cantril drew the following conclusions: (1) the voice does not convey correct information concerning inner personality characteristics, although judges were able to match voices with photographs; (2) for various features of the personality of a speaker such as age, there is associated in the minds of judges some perception of the type of voice quality to which this feature corresponds; (3) more highly organized and deep-seated traits and dispositions for some speakers are judged more consistently and more accurately than the more specific features of the physique and appearance; and (4) if the voice arouses a stereotype of the speaker, it is likely that several features of personality will be subsumed under this stereotype. Allport and Cantril also indicated that the more information given about a speaker the more frequently is the voice quality accurately identified.

The above studies and additional studies by Pear (1931), similar to those of Allport and Cantril, reveal that the voice quality does affect the way listeners will evaluate personality variables. It is difficult to assess the validity and reliability of these studies in that only post-test measures were used. A positive point of the research was that voice quality was isolated from other forms of nonverbal communication. No information, however, was provided about the linguistic messages. Perhaps the linguistic
messages were confounding variables in these studies.

During the early 1940's, Fay and Middleton conducted a number of experiments that dealt with the judgments of the speaker's personality based on voice quality transmitted over a public address system. The first experiment was conducted in 1940 and was concerned with judging intelligence. Twelve eleventh and twelfth grade boys with supposedly equal oral reading ability were placed in three groups of four based on I.Q. scores on the Terman Group Test of Mental Ability. All subjects in the three groups were randomly selected to record twice. A different thirty second prose selection was used each time. The recorded voices were played back with musical pieces between each reading. Eighty-four psychology students at the University of Illinois rated the perceived intelligence of each reader on a seven-point scale. The researchers concluded: (1) that listeners were only fairly reliable in their judgments of the intelligence of the speakers whose voices were heard over a public address system; (2) that the intelligence of some speakers was judged more reliable than others; (3) that listeners rated speakers of superior intelligence as somewhat more intelligent than they did speakers of average or inferior intelligence; and (4) that listeners revealed some voice stereotype of superior and of inferior intelligence.

Fay and Middleton (1942) conducted another study to determine if introversion of speakers was transmitted by voice quality. Introversion of the speaker was measured by Bernreuter Personality Inventory from the transmitted voices of several speakers. Twenty-seven sophomore men at the University of Illinois were selected from a large group on the basis of their personality inventory scores. Three subjects were selected randomly from each of the nine
introversion percentile levels. All twenty-seven participants read three selections. Fay and Middleton concluded that introversion can be judged from the transmitted voice with an accuracy of about chance expectation. There was, however, a rather high degree of social agreement in judging introversion-extroversion of certain voices. The agreement was in excess of the accuracy of estimations, and thus confirmed the phenomenon of stereotype judgments.

Fay and Middleton (1943) also conducted a study to discover if a speaker's leadership ability can be evaluated by the transmitted voice quality. Specifically, this study was an attempt to determine to what extent a group of listeners can judge accurately the leadership characteristics of several speakers based upon hearing their transmitted voice qualities. Fifteen freshmen from a fraternity at the University of Illinois were rated on a seven-point scale for leadership by ten seniors. The fifteen freshmen were asked to read thirty seconds of interesting material over a public address system. The results indicated that some voices are stereotypes and that interjudge agreement is valid to some extent.

The above studies by Fay and Middleton encompass the four variables discussed in this paper. In the first two studies the researchers used standardized tests which were reliable and valid. In the reported studies, the researchers failed, however, to indicate what type of linguistic materials were read by the speakers. There were also no reports of pilot tests to classify listeners or to establish inter-rater reliability. In addition, because Fay and Middleton used a mechanical devise to transmit the voice, the results of the study might not be realistic.

This researcher did not find any studies in the journals from 1943 to 1962 that dealt specifically with the relationship of voice quality and the
four personality variables. A study by Anisfield, Bogo, and Lambert (1962) dealt with the judgment of group personality rather than individual personality. In this experiment both Jewish and Gentile students were asked to judge leadership on the basis of voice quality which displayed standard pronunciation and accented voice. There was no significant difference between the two groups' judgments of leadership.

Bowers (1965) conducted a study to determine whether introvert delivery style and extrovert delivery style have different effects upon different types of personalities. A pre-test, Social Introversion/Extroversion Scales of the Minnesota Multiphasic Personality Inventory, was the instrument used to classify introverted and extroverted personalities. On the basis of the scores of the tests, subjects were placed into homogeneous groups and exposed to the two different styles of delivery. The results of the study revealed that subjects rated the personality of the speaker using extroverted speaking style more positively than they did the personality of the speaker using introverted speaking style. There was, however, no relationship found between style of delivery and the extroversion level of subjects. Moreover, a positive delivery did not produce significant change in subjects when the topic was evaluated. In short, this study was an investigation of the relationship between two styles of delivery and changes in attitude toward both speaker and concepts when the content of the two argumentative speeches was held constant. Although the two styles produced significantly different changes in attitudes toward the concepts, there was no relationship between delivery style and extroversion level of listeners.

Bowers' study differed from previous research because the subjects were
pre-tested to determine the degree of extroversion. Bowers attempted to discover if a listener's degree of introversion-extroversion would determine how the listener evaluates the speaker's personality in terms of introversion-extroversion. Because the subjects were pre-tested and the terms were operationally defined, the experiment has more credibility than previous research. The results of the experiment, however, may not be reliable or valid as the researcher used a persuasive linguistic message.

Addington (1968) conducted a more complete study on voice cues and argued that empirical evidence of various voice attributes will pinpoint the specific nature of the stereotypes of the voice. Two hundred tape-recorded samples of the "Rainbow Passage" were obtained from two female and two male trained speakers who were instructed to simulate several qualities (breathy, tense, thin, flat, throaty, nasal, and orotund) and three variations of speaker's rate (normal, fast, and slow) and pitch variety (normal, more than normal, and less than normal). Perceived personality data was obtained from sixteen sections of freshmen rhetoric at the University of Iowa. These subjects listened to the recorded voice samples and rated the personalities of the speakers. The research indicated that increased pitch variety caused subjects to rate the voice as extroverted. The study did not indicate the pitch range or the specific rates at which the speaker read. Addington failed to mention specific voice qualities that subjects evaluated as introverted or extroverted. From the results this researcher concluded that the vocal qualities of tenseness, throatiness, and orotundity were more apt to be judged extroverted, whereas the others were judged introverted.

A similar study was conducted by Pierce (1968) who argued that voice cues are proponents to assess credibility and determine how listeners will
rate the personality dimensions of a speaker. In this study one of the bipolar
adjectives was related to intelligence. Pierce concluded that the speaker
using a Type I delivery, which was designed to be in the style of a scholarly,
dispassionate, yet involved person, serious about his subject, was judged to
be intelligent. Pierce, as did Bowers, presented a description of the speaker
prior to the experiment; thus, the fact that Type I speaker was described
positively might have been the reason for the Type I voice being rated as
intelligent.

A survey of the literature from the past to the present reveals that
attempts have been made to show a relationship between voice quality and the
personality dimensions of leadership, intelligence, introversion, and
extroversion. Early research attempted to discover if, based on voice quality,
the speaker's personality characteristics can be predicted by the listeners.
In recent years Bowers attempted to discover if listeners with certain
personality characteristics evaluated the speaker's voice quality the same.
All the reported studies indicated that there is a definite relationship be-
tween voice quality and how the listener judges the personality variables.
The studies also indicate that certain voice qualities are more stereotyped
than others. For example, voice qualities of some speakers are more apt than
others to be judged intelligent.

Several criticisms concerned with voice quality research have been advanced.
Kramer (1964) pointed out that practically all studies that deal with voice
quality and personality judgment have generally ignored listener differences
in such things as personality, cultural development, and psychophysical
aspects, which may have a profound impact on how the listener judges voice
quality. As indicated before, Bowers' study was the only one that attempted
to discover background characteristics of listeners personalities prior to the experiment. Kramer contended that researchers must discover as much information as possible about listeners prior to conducting experiments. In addition, Starkweather's 1963 summary of voice quality and judgment of emotions suggested that only neutral messages be used to determine attitude change about a topic and message. If the topic and linguistic message were neutral then any change that occurred would be caused by the delivery. This initiates the following research questions: When listening to a positive paralinguistic message with a neutral linguistic message about a neutral topic will subjects of high intelligence experience more attitude change than will subjects of above average and average intelligence? When listening to a positive paralinguistic message with a neutral linguistic message about a neutral topic, will subjects of high intelligence evaluate the personality dimensions of the speaker more favorably than will subjects of above average and average intelligence?

PROCEDURES

Subjects were selected from the experimental pool established by the Speech-Communication Department, University of Nebraska, Lincoln. A total of four hundred students participated in this experiment. The initial phases of the research involved a series of pilot studies. During the first week of winter semester, 1972, students from Basic Interpersonal Speech were asked to generate a list of five topics that they considered to be neutral. The three topics that were generated most frequently were subjected to additional experimentation. Fifty subjects were randomly selected to rate the topic as neutral or not neutral. A semantic differential scale was used to determine which topic was the most neutral. A rating of seven indicated that the topic
was neutral, whereas a rating of one indicated that the topic was not neutral. Abortion was selected as the neutral topic.

In the next pilot study, unidimensional scales for the topic abortion were selected. One hundred randomly selected subjects were presented with twenty sets of semantic differential scales. Factor analysis was used to determine which scales were unidimensional. The following scales were established as unidimensional for the topic abortion: (1) foolish-wise, (2) bad-good, (3) unwarranted-warranted, (4) valuable-worthless, (5) just-unjust, and (6) fair-unfair.

In the third pilot study, two hundred subjects rated the topic abortion. Only unidimensional scales were used to determine subject's attitude toward the topic; other filler items, however, were used to mask the needed data. Subjects with a cumulative score between twenty-four and twenty-seven were operationally defined as neutral. Next, the researcher attempted to acquire these subjects' verbal SAT scores in order to group them by intelligence. Since this information was not made available by the records office, the researcher used the subjects' previous grade point averages to group them. Although this may be less credible than the SAT scores, it was a legitimate means of classification. Subjects who had grade point averages over 3.00 on a 4.00 scale were operationally defined as possessing superior intelligence. Subjects with grade point averages between 2.5 and 3.00 were operationally defined as above average in intelligence. Subjects with grade point averages between 2.0 and 2.5 were operationally defined as average. The subjects were then placed into one of three groups.

For the next pilot study, the researcher prepared a ten minute neutral message on abortion. Ten members of the Speech-Communication faculty at the
University of Nebraska, Lincoln, evaluated the linguistic message. A Likert scale was used to measure the judges' responses. The judges agreed that the linguistic message was neutral.

A positive paralinguistic delivery of the neutral message was created by a qualified graduate student. The positive paralinguistic structure was described as the normal quality of the speaker's voice. All other voice variables were held constant. That is, the speaker spoke at the same rate and did not vary pitch or force. The above ten judges listened to the tape-recorded message and rated the Likert scale to indicate that the voice quality of the tape was normal for the graduate student.

The next pilot study was the presentation of the paralinguistic message to fifty randomly selected subjects in order to acquire scales that were unidimensional for the voice quality of the speaker. The following scales were selected as unidimensional: (1) honest-dishonest, (2) intelligent-unintelligent, (3) good-bad, (4) extroverted-introverted, and (5) skilled-unskilled.

Different members from each of the three intelligence groups listened to a tape-recorded message. The written message served as a control stimulus. All subjects were told what they were to listen to a tape-recorded message on the topic abortion and that this topic had been selected because it had been in the news recently. No information was given about the speaker, as the researcher wanted to avoid sensitizing the subjects. The subjects who read the message were told that they were to read a message about the topic abortion because it had been in the news recently and then to evaluate the topic. As indicated above, there were eighty-four subjects who had a neutral attitude about the topic abortion. Because there were many more subjects from the above-average and average groups than there were from the superior intelligence
group, seventeen subjects from each group were randomly selected so that there were the same number of subjects in each intelligence group.

The first two groups consisted of subjects with superior intelligence. The nine subjects in Group I listened to the tape-recorded neutral linguistic message, and the eight subjects in Group II read the neutral linguistic message. Groups III and IV consisted of subjects with above-average intelligence. The nine subjects in Group III listened to the tape-recorded message, and the eight subjects in Group IV read the message. The last two groups consisted of subjects with average intelligence. The nine subjects in Group V listened to the tape-recorded message, and the eight subjects in Group VI read the message. After the subjects listened to a tape-recorded message or read the message, all were given a post-test. The first part of the post-test was the same as the pre-test. That is, subjects rated a set of twelve semantic differential scales. Only those scales which had been determined to be unidimensional were used to analyze the data. The second part of the test consisted of value judgments which subjects made about personality variables of the speaker.

One-way analysis of variance was used to measure subjects' attitudes toward the topic and speaker personality dimensions. If statistical significance was attained with the ANOV, further analysis was completed using Tukey H.S.D. test.

RESULTS

The .05 level of significance was required for all hypotheses tested. First, the one-way analysis of variance was established for the three groups that heard the paralinguistic message. The results indicated that there was no significant differences among the groups. (Table I) No further analysis
was completed.

Table I

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>7.03369</td>
<td>2</td>
<td>3.03387</td>
<td>0.69160</td>
</tr>
<tr>
<td>Within Groups</td>
<td>6.7554</td>
<td>24</td>
<td>4.67071</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>13.8223</td>
<td>26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOV was also run on the three groups reading the message. Again there were no significant differences found among the groups. (Table II) No further analysis was completed.

Table II

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>4.66873</td>
<td>2</td>
<td>2.98654</td>
<td>0.61554</td>
</tr>
<tr>
<td>Within Groups</td>
<td>5.34637</td>
<td>21</td>
<td>3.66375</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>10.01510</td>
<td>23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There was a significant difference between the superior subjects' ratings of personality dimensions of the speaker and the ratings of the above-average and average intelligence groups. (Table III) This analysis was completed using Tukey H.S.D.

Table III

<table>
<thead>
<tr>
<th>Difference among Means</th>
<th>$\bar{x}_1$</th>
<th>$\bar{x}_2$</th>
<th>$\bar{x}_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}_1$</td>
<td>---</td>
<td>1.1</td>
<td>1.3</td>
</tr>
<tr>
<td>$\bar{x}_2$</td>
<td>---</td>
<td>---</td>
<td>.2</td>
</tr>
<tr>
<td>$\bar{x}_3$</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
DISCUSSION

The purpose of this research was to review studies that have been done on voice quality and to present an experimental study which the researcher has recently completed. The review of literature indicates that only Bowers' article dealt with a personality dimension of the listener. In the experiment, however, listeners might have been sensitized prior to listening to the tape-recorded message, since subjects were given a description of the speaker and then listened to a persuasive linguistic message. This author's research is the first attempt to discover differences in attitude change of neutral listeners of various levels of intelligence when presented with a neutral topic, neutral linguistic message, and positive paralinguistic message. That is, this researcher wanted to discover differences in various intelligence groups in terms of how subjects rated a topic after they listened to a tape-recorded message or read a neutral message. The subjects who listened to the paralinguistic message also rated the speaker's personality dimensions. The results reveal that there were no significant differences among the intelligence level groups in rating the topic. There was, however, a significant difference between the superior intelligence group and the above-average and average groups in the way the subjects rated the speaker's personality dimensions.

In an attempt to discover within-group differences, t-tests were run between subjects' post-test scores of attitude toward the topic and attitude toward the personality dimensions of the speaker. There was not a significant difference found between the rating of the superior group. In both the above-average and average groups, there were significant differences. Subjects from these groups rated the personality dimensions of the speaker significantly higher than attitudes toward the topic.
The results reveal that superior subjects pay less attention to the paralinguistic message. Because of the few number of subjects in the experiment and the use of grade point average as the only criterion for classification of intelligence, this researcher is not ready to conclude that subjects of superior intelligence will listen more critically than will less intelligent subjects to the linguistic message and use the message to evaluate the personality of the speaker. Instead the researcher suggests that more research be completed on the subject. Replication of this study in which a large sample is used is necessary. SAT verbal scores should be used to classify subject's level of intelligence, for as past research reveals, grade point average does not have a high degree of predictive validity with intelligence tests. Additional research questions can be generated. Some examples of these are listed below:

1. When the voice quality of the speaker presenting a neutral message conveys happiness, will a person of superior intelligence, based on verbal SAT scores, judge the speaker as more intelligent than will a person of above-average or average intelligence?

2. When the voice quality of the speaker conveys happiness, will an extroverted listener, based on Social Introversion/Extroversion scales of the Minnesota Multiphasic Personality Inventory, evaluate a speaker presenting a neutral message more positively than will an introverted listener?

3. When the voice quality of the speaker conveys happiness, will a person classified as a task leader, based on three problem-solving discussions using I.P.A. system, evaluate a speaker presenting a neutral linguistic message more positively than will a social emotional leader?

Experimentation is the key to find answers to all of the above questions. In addition, questions of the above type can be tested using a variety of different emotions presented by male and female speakers. Research also needs to be completed with different voice variables; first emphasis must be placed on the variables in isolation and then in combination. This will create more complex experiments. In studies concerned with voice quality, researchers must
attempt to discover which voice quality conveys each emotion best and discover which voice quality causes listeners to evaluate the speaker as introverted, intelligent, etc.

Future researchers must conduct a series of pilot tests prior to the main study. Subjects must be given a variety of tests so that the speech communication field can expand the knowledge about the effects various voice cues have on listeners with different types of personality dimensions. Replication of past and future studies is needed to determine even a degree of accuracy. In the future, for example, researchers will be able to state indicatively that a male nasal voice will or will not be judged as not intelligent, introverted, and not possessing leadership qualities. In short, more precise experimentation is needed to determine the exact role that isolated voice cues play in the judgment of the personality variables: intelligence, leadership, introversion, and extroversion.
BIBLIOGRAPHY


