The voice of confidence: Paralinguistic cues and audience evaluation

SCHERER, Klaus R., LONDON, Harvey, WOLF, Jared J.

Abstract
A standard speaker read linguistically confident and doubtful texts in a confident or doubtful voice. A computer-based acoustic analysis of the four tapes showed that paralinguistic confidence was expressed by increased loudness of voice, rapid rate of speech, and infrequent, short pauses. Under some conditions, higher pitch levels and greater pitch and energy fluctuations in the voice were related to paralinguistic confidence. In a 2 × 2 design, observers perceived and used these cues to attribute confidence and related personality traits to the speaker. Both text and voice cues are related to confidence ratings; in addition, the two types of cue are related to differing personality attributes.

Reference


DOI: 10.1016/0092-6566(73)90030-5

Available at:
http://archive-ouverte.unige.ch/unige:101849

Disclaimer: layout of this document may differ from the published version.
The Voice of Confidence:
Paralinguistic Cues and Audience Evaluation

KLAUS R. SCHERER
Harvard University

HARVEY LONDON
Hunter College of the City University of New York

AND

JARED J. WOLF
Massachusetts Institute of Technology

A standard speaker read linguistically confident and doubtful texts in a confident or doubtful voice. A computer-based acoustic analysis of the four tapes showed that paralinguistic confidence was expressed by increased loudness of voice, rapid rate of speech, and infrequent, short pauses. Under some conditions, higher pitch levels and greater pitch and energy fluctuations in the voice were related to paralinguistic confidence. In a $2 \times 2$ design, observers perceived and used these cues to attribute confidence and related personality traits to the speaker. Both text and voice cues are related to confidence ratings; in addition, the two types of cue are related to differing personality attributes.

In a review of the field of person perception, Tagiuri (1969) presents a framework for research in this area, identifying the main elements in the process of person P's perception or cognition of person O: (1) O's characteristics or states, (2) the concomitants of these, (3) the distal cues or manifestations of O's characteristics available to P, (4) the proximal cues utilized by P, (5) the cognitive processes involved, and

---

1 The work in acoustic voice analysis was supported in part by U. S. Air Force Cambridge Research Laboratories (Air Force Systems Command) Contract F19628-69-C-0044. Klaus Scherer is now at the Psychological Institute, University of Kiel, West Germany. Jared Wolf is now at Bolt, Beranek and Newman, Inc., Cambridge, Massachusetts. The authors would like to thank Fred Grandy, who was the actor in the study, and Michael Grinberg, who aided in data analysis. Requests for reprints should be sent to Dr. Harvey London, 10 Forest Street, Apt. 10, Cambridge, Mass. 02140.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.
the percept or judgment by P of O's characteristics. A promising tool for research within this framework is the encoding experiment, a technique that has frequently been used for the investigation of nonverbal communication (Weick, 1969; Duncan, 1969). In this type of experiment, the “encoding subjects” are asked to manifest cues expressing a certain state or intention such as approval-seeking (Rosenfeld, 1966) or persusiveness (Mehrabian & Williams, 1969) vis-a-vis the “decoding subjects” who then yield a judgment about the encoder's states or intentions. This technique enables the investigator to study, under controlled experimental conditions, both type and frequency of cues associated with certain communicator characteristics and the utilization of these cues by the receiver in his attribution of various traits or states to the communicator. The present paper reports an investigation of the paralinguistic cues manifested by a speaker when encoding “confidence” and “doubt” and of attributions made to the speaker by observers exposed to these cues.

“Expressed confidence,” a newly identified “message” (McGuire, 1969, p. 172) variable that has been under study by London and his colleagues, seems especially useful as a vehicle for the present investigation of the person perception process, since a consistent body of evidence has accumulated with respect to the linguistic expression of confidence. Linguistically expressed confidence was originally extracted from two-person persuasive interaction (Johnson & London, 1968; London, Meldman, & Lanckton, 1970a; London, Meldman, & Lanckton, 1970b), and measured in terms of the presence of specific words and phrases, codified in a scoring manual (Yoselson, Meldman, & London, 1970). In terms of the manual, phrases such as “I am confident that . . .” are scored as expressing confidence, while phrases such as “I am unsure that . . .” are scored as expressing doubt. Within Tagiuri's framework, the words and phrases found in the scoring manual represent the linguistic cues manifested by O and utilized by P.

Furthermore, “expressed confidence” has been investigated in studies in which linguistic and kinesic channels of communication were varied independently (Maslow, Yoselson, & London, 1971; London, McSeveney, & Tropper, 1971). In both of these studies, the expression of confidence was found to affect persuasion. The paralinguistic expression of confidence has been investigated in a correlational study (London, 1973, p. 12). The latter study suggested that O's paralinguistically expressed confidence is related to persuasion.

Although it is possible, operationally, to produce paralinguistically expressed confidence (doubt) by simply asking O to speak in a confident (doubtful) tone of voice, we know little about the specific paralinguistic
cues a speaker uses to encode confidence. The first part of the present study attempts to isolate cues such as energy, pitch, and rate of speech for various conditions of "confidence encoding." Judging from the rather extensive literature on the expression of personality and affect in voice and speech (Kramer, 1963; Mahl & Schulze, 1964; Davitz, 1964; Scherer, 1970; Scherer, 1972), it seems reasonable to expect a speaker to express confidence by speaking with a strong, loud voice (i.e., high energy), with expressive intonation (i.e., high pitch variation), and with a relatively rapid, unhalting rate of speech.

**METHOD**

**Overview**

The study was conducted in two parts. First, a speaker soundtaped both "linguistically confident" and "linguistically doubtful" texts, reading each text in both a "paralinguistically confident" and a "paralinguistically doubtful" manner. This process generated four tapes, two channel-discrepant and two channel-congruent. These tapes were computer-analyzed for acoustic properties and for mean duration of pauses. Further, the four tapes were played to four different groups of 5's in the context of a jury decision-making study; ratings of various kinds were obtained.

**Experimental Manipulations of Confidence**

**Text confidence and doubt.** The text, the same as that used in earlier research carried out by London and his co-workers (Study 1 in Maslow et al., 1971), consisted of a legal argument. There were two versions of the argument, both arguing that a defendant be held not liable. Both versions made the same points in favor of the defendant. However, the argument was written so that there were "slots" throughout the text into which words and phrases expressing confidence or doubt could be placed. These expressions were drawn from the scoring manual mentioned above (Yoselson et al., 1970). To give the reader the flavor of this approach, we provide two examples. Where the Confident Text said "obviously," the Doubtful Text said "I don't know." Where the Confident Text said "I believe," the Doubtful Text said "I'm not positive." The tone of the arguments so constructed did not differ from the tone of arguments used by "real" persuaders and persuadees (cf. Johnson & London, 1968; London et al., 1970a; London et al., 1970b).

**Voice confidence and doubt.** A student with considerable experience in acting was hired to tape-record each of the two versions of the argument twice, once expressing confidence paralinguistically and once expressing doubt. The speaker indicated afterwards that he had role played gen-
eralized states of confidence and doubt and had not consciously manipulated specific paralinguistic features of his speech. The speech was recorded with $7\frac{1}{2}$ inch/sec tape speed on a UHER 8000 stereo tape recorder using an AKG 200 microphone.

**Effectiveness of Manipulations**

To test whether expressed confidence was effectively manipulated via both linguistic and paralinguistic channels, Thurstone's paired comparison technique (Edwards, 1957) was utilized. Ten female college students, paid volunteers, were tested one at a time using the paired comparison technique, and were asked, for all possible pairs of the four tapes, which tape of each pair expressed more confidence. The order of pairs and the order of tapes within each pair was randomized.

The results yielded the following scale values: Doubtful Text–Doubtful Voice, 0.00; Confident Text–Doubtful Voice, 0.45; Doubtful Text–Confident Voice, 0.70; Confident Text–Confident Voice, 3.47. These values are in line with theoretical expectations. Mosteller's test (Edwards, 1957, pp. 53–57) indicated ($0.50 < p < 0.70$) that the results do not depart significantly from theoretical values.

**Acoustic Voice Analysis**

(1) **Energy and pitch.** An electronic analysis of voice energy and pitch was performed on two kinds of speech samples, consisting of single words or short expressions: (a) expressions relevant to the linguistic manipulation of confidence (i.e., expressions to be found in the “slots” mentioned above); and (b) expressions of comparable syntactic status but irrelevant to the linguistic manipulation (i.e., expressions not to be found in these slots). In all, 19 relevant expressions and 20 irrelevant expressions were analyzed for each of the four tapes. These 39 expressions exhausted roughly half of the text.

The acoustic measurements of these samples were performed with the aid of the Speech Communication Group computer facility at the Research Laboratory of Electronics, Massachusetts Institute of Technology. This system is a Digital Equipment Corporation PDP-9 general purpose computer coupled to peripheral equipment designed to facilitate on-line speech research (Henke, 1968). A general purpose program for spectral analysis and display was available (Wolf, 1969, 1970).

The recordings were played on an on-line Ampex 401A tape deck into a computer-controlled spectrum and fundamental frequency analyzer. The spectrum analysis was performed by a 36 channel filter bank, and fundamental frequency (i.e., pitch) was extracted by means of low-pass
filtering the speech above the first harmonic and measuring the intervals between zero crossings with a subprogram. This rudimentary pitch extractor required occasional adjustment for the low-pass filter to accommodate the wide excursions of vocal pitch and amplitude of the speaker's more expressive passages. The spectrum and fundamental frequency were measured at a rate of 100 samples/sec.

The program has a capacity for storing data for 2.5 sec of speech. After the input phase, the program produces a display on a 16-in. cathode ray tube. The display contains two graphs which represent functions of time, representing the 2.5 sec of stored speech data. One graph is the sum of the outputs of the first six filters of the spectrum analyzer covering 0–750 Hz. This function is useful as a "low frequency loudness" (i.e., voice energy) map of the utterance. The other graph is fundamental frequency (i.e., pitch). Over these two graphs is displayed a manually controllable vertical cursor, which enables the operator to point to any part of the utterance represented by the display.

The data of interest were the values of the low-frequency loudness function (given in scale units monotonically related to amplitude) and fundamental frequency (given in Hz) during the sample expressions described above. These data were obtained by playing the tape into the computer and stopping it immediately after the desired expression. The analysis program stored only the most recent 2.5 sec of speech, so that the sample expression (which was always shorter than 2.5 sec) would be represented by the right-most part of the display. The cursor was positioned at the beginning of the sample expression and a subprogram was initiated to type out the values of the low-frequency loudness and fundamental frequency at 5-frame intervals until the end of the sample was reached. This process yielded a sampling of these functions at a rate of 20 per sec. The values, typed out by a printer, were later averaged for each sample expression, resulting in single values for energy and pitch for each of the 19 relevant and 20 irrelevant expressions in each of the four tapes.

(2) Number and duration of pauses and utterances. Using a Hewlett-Packard pen recorder, a continuous recording of the envelope of the speech signal was made for all four recordings of the entire text passage. The recording was made on a graph paper roll advancing at a speed of 5 mm/sec. Thus, each millimeter division of the graph paper represented 1/2 sec of speech. With the help of these pen recordings (which represented voice amplitude), the presence or absence of sound at any given position of the utterance could be ascertained and thus the number and duration of pauses could be measured. Any silent interval longer than ½ sec was
operationally defined as a pause. Using this procedure, the following measures were obtained: (a) mean number of within-sentence pauses; (b) mean duration of 14 between-sentence pauses; (c) mean duration of within-sentence pauses. Finally, the number of values printed out by the voice analysis computer in its sampling of the energy of uttered words and expressions provided an estimate of the duration of the utterance (rate of speech), since the rate of sampling was constant at 20 values/sec.

**Audience Evaluation**

The Ss were 47 undergraduate women at a small all-female teachers' college. They were paid volunteers and had signed up to participate in one of four group sessions.

The E scheduled Ss for experimental sessions by circulating sign-up sheets. Ss signed up for one session only. In all, four sessions were conducted, one per experimental condition.

At each session, it was explained via tape-recorded instructions that the point of the study was to help evaluate the ability of law students to work in legal clinics. Ss were asked to play the role of jury members in a law case. They were to read the facts of a law case and judge's instructions to the jury, and then to listen to a sound tape with comments on the case by a high-ranking law student.

While Ss were reading background materials, similar to those used in previous studies (London, Meldman, & Lanckton, 1970c), E drew a slip of paper from an envelope in order to determine which of the four sound tapes would be played.

After the sound tape had been played, Ss filled out a questionnaire, rating the speaker with respect to confidence, expertise, and legal competence on scales running from 0 to 100% and with respect to 23 personality attributes and 7 speech attributes on scales running from 0 to 9. Finally, E collected all materials, gave Ss further information about the study, thanked Ss for their participation and paid them.

**RESULTS**

**Acoustic Analysis of Tapes**

It will be recalled that differences in energy, pitch, and duration of utterances between the Confident Voice and the Doubtful Voice renderings of the two textual versions of the argument were expected. The data relevant to these hypotheses are presented in Table 1. (All p values reported in this paper are two-tailed.) The results show a number of significant differences between the Confident Voice and Doubtful Voice arguments for both Confident and Doubtful versions of the text. Com-
TABLE 1
MEAN VALUES OF ACOUSTIC SPEECH MEASUREMENTS

<table>
<thead>
<tr>
<th></th>
<th>Voice</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Confident</td>
<td>Doubtful</td>
<td></td>
</tr>
<tr>
<td>Energy, in scale units</td>
<td>Text</td>
<td>Confident</td>
<td>108.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Doubtful</td>
<td>115.2</td>
</tr>
<tr>
<td>Pitch, in Hertz (Hz)</td>
<td>Text</td>
<td>Confident</td>
<td>113.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Doubtful</td>
<td>131.8</td>
</tr>
<tr>
<td>Rate of speech (duration of utterances as measured by the number of values printed out by the computer for each sample)</td>
<td>Text</td>
<td>Confident</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Doubtful</td>
<td>17.7</td>
</tr>
</tbody>
</table>

pared to a Doubtful Voice, a Confident Voice is associated with greater energy \( F(1,148) = 38.50, p < .001 \), greater pitch \( F(1,148) = 6.65, p < .05 \), and faster rate of speech \( F(1,148) = 9.14, p < .01 \). There is a significant interaction \( F(1,148) = 9.61, p < .01 \) between voice and text for pitch, suggesting that the speaker used high pitch to express a Confident Voice under Doubtful Text conditions, but not under Confident Text conditions.

The error term in the analyses of variance posed a problem, as the word samples in the utterance resemble a time series. Thus, the variance of observations within a single rendering will be relatively small if observations are correlated with one another. To test for this possibility, matched \( t \) tests were performed for all pairs of means, and autocorrelation coefficients (Ezekiel & Fox, 1959) were computed for the observations within each condition. These coefficients were tested and, when found significant, degrees of freedom were accordingly reduced in conformity with Bartlett's formula (Ezekiel & Fox, 1959, p. 335). The pattern of results substantiates the ANOVA findings at the same or higher levels of significance.

Since there were no interactions between either voice or text and relevance, Table 1 is not broken down to show separate means for the relevance variable.

Variances for energy and pitch measurements were computed for both confidence-relevant and confidence-irrelevant words. For pairs of variances, \( F \) tests for significant differences (Edwards, 1966, pp. 271–273) were performed. Results suggest greater pitch and energy fluctuation for confidence-relevant words. The significantly greater degree of pitch and, to some extent, energy variation for confidence-relevant words in the
TABLE 2
MEAN VALUES FOR PAUSES

<table>
<thead>
<tr>
<th></th>
<th>Voice</th>
<th>Confident</th>
<th>Doubtful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pauses</td>
<td>Text</td>
<td>1.67</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>Confident</td>
<td>1.00</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>Doubtful</td>
<td>1.00</td>
<td>1.91</td>
</tr>
<tr>
<td>Within sentence pause</td>
<td>Text</td>
<td>0.27</td>
<td>0.86</td>
</tr>
<tr>
<td>duration (in sec)</td>
<td>Confident</td>
<td>0.37</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Doubtful</td>
<td>0.37</td>
<td>0.70</td>
</tr>
<tr>
<td>Between-sentence pause</td>
<td>Text</td>
<td>1.19</td>
<td>2.05</td>
</tr>
<tr>
<td>duration (in sec)</td>
<td>Confident</td>
<td>1.19</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td>Doubtful</td>
<td>1.19</td>
<td>1.53</td>
</tr>
</tbody>
</table>

Confident Voice condition may indicate that paralinguistic expression of confidence is achieved in part through selective stress of strategic linguistic units.

The means for number and duration of pauses are shown in Table 2. Significant main effects for voice (for number of pauses $F(1, 56) = 5.84, p < .05$; for within-sentence pauses $F(1, 56) = 15.54, p < .01$; for between-sentence pauses $F(1, 56) = 15.21, p < .01$) indicate that, in the Confident Voice condition, the speaker used pauses less frequently and that, when he did pause, the pauses were shorter. The means suggest that the speaker used substantially longer pauses to express Doubtful Voice under Confident Text conditions, but this interaction approaches significance only for between-sentence pauses.

There were no significant differences with respect to number or duration of pauses for the Text variable.

**Audience Evaluation**

Given differences in acoustic measures between the two voice renderings of the argument, the question arises as to whether these differences led Ss to attribute to the speaker differences in confidence, personality, and speech. Mean ratings of the speaker with respect to perceived confidence are shown in Table 3.

TABLE 3
MEAN AUDIENCE RATINGS OF SPEAKER'S CONFIDENCE IN PERCENTAGE UNITS

<table>
<thead>
<tr>
<th>Voice</th>
<th>Confident</th>
<th>Doubtful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>66.7</td>
<td>58.1</td>
</tr>
<tr>
<td>Doubtful</td>
<td>42.9</td>
<td>32.0</td>
</tr>
</tbody>
</table>
In a $2 \times 2$ analysis of variance, there were significant main effects for text ($F(1,43) = 27.31, p < .001$) and for voice ($F(1,43) = 4.19, p < .05$). No interaction was found. These results indicate that expressed confidence is perceived accurately in both the linguistic and the paralinguistic channels.

The reader will recall that the speaker was rated not only with respect to confidence, but with respect to expertise and legal competence as well. With regard to the latter two ratings, no significant differences were observed.

Mean ratings of the speaker on scales ranging from 0 through 9 for 11 of the 23 personality attributes tested are shown in Table 4. Only scales for which significant differences were found are reported.

Analyses of variance ($2 \times 2$ factorial, $df = 1/43$) yielded single main effects for either text or voice in the following cases: compared to the Doubtful Text condition, the speaker in the Confident Text condition was seen as more conceited ($p < .025$), more professional ($p < .05$), more businesslike ($p < .05$), and less personal ($p < .05$). Compared to the Doubtful Voice condition, the speaker in the Confident Voice condition was seen as more enthusiastic ($p < .001$), more forceful ($p < .001$), more active ($p < .001$), and more competent ($p < .05$). Significant main effects for both voice and text were observed for dominance (voice, $p <$...

### Table 4

<table>
<thead>
<tr>
<th>Mean Ratings of Personality Attributes of Speaker</th>
<th>Confident voice</th>
<th>Doubtful voice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Confident text</td>
<td>Doubtful text</td>
</tr>
<tr>
<td>Enthusiastic&lt;sup&gt;a&lt;/sup&gt;,&lt;sup&gt;c&lt;/sup&gt;</td>
<td>6.4</td>
<td>4.9</td>
</tr>
<tr>
<td>Forceful&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.8</td>
<td>4.9</td>
</tr>
<tr>
<td>Active&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.0</td>
<td>5.1</td>
</tr>
<tr>
<td>Competent&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.5</td>
<td>5.8</td>
</tr>
<tr>
<td>Dominant&lt;sup&gt;a&lt;/sup&gt;,&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Self-confident&lt;sup&gt;a&lt;/sup&gt;,&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.6</td>
<td>4.8</td>
</tr>
<tr>
<td>Self-assured&lt;sup&gt;a&lt;/sup&gt;,&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.3</td>
<td>4.6</td>
</tr>
<tr>
<td>Conceited&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.3</td>
<td>1.6</td>
</tr>
<tr>
<td>Professional&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.8</td>
<td>4.6</td>
</tr>
<tr>
<td>Businesslike&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.8</td>
<td>4.9</td>
</tr>
<tr>
<td>Personal&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.7</td>
<td>5.1</td>
</tr>
</tbody>
</table>

<sup>a</sup> Main effect for voice.

<sup>b</sup> Main effect for text.

<sup>c</sup> Interaction effect.
TABLE 5
MEAN RATINGS OF SPEECH ATTRIBUTES OF SPEAKER

<table>
<thead>
<tr>
<th>Voice</th>
<th>Confident</th>
<th>Doubtful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaks loudly</td>
<td>5.85</td>
<td>4.35</td>
</tr>
<tr>
<td>Fast speaking</td>
<td>2.60</td>
<td>1.40</td>
</tr>
<tr>
<td>Expressive voice</td>
<td>5.80</td>
<td>4.60</td>
</tr>
<tr>
<td>Speaks fluently</td>
<td>6.55</td>
<td>5.05</td>
</tr>
<tr>
<td>Pleasant voice</td>
<td>6.25</td>
<td>5.35</td>
</tr>
<tr>
<td>Speaks distinctly</td>
<td>6.65</td>
<td>6.40</td>
</tr>
<tr>
<td>Melodic voice</td>
<td>3.25</td>
<td>3.50</td>
</tr>
</tbody>
</table>

*a Main effect.

.001; text, p < .025), self-confidence (voice, p < .001; text, p < .001) and self-assuredness (voice, p < .05; text, p < .001). There was an interaction for enthusiasm (p < .05).

Finally, Ss were asked to rate the speaker with respect to seven speech attributes on scales running from 0 through 9. Mean ratings are found in Table 5. Since no predictions were made and no results were found for the text variable, ratings are collapsed across text conditions. Significant main effects were found in the following cases: compared to the Doubtful Voice condition, the speaker in the Confident Voice condition was seen as speaking louder (p < .025), faster (p < .025), more expressively (p < .05), and more fluently (p < .025). No interactions were found.

DISCUSSION

In terms of Tagiuri’s (1969) framework, the results suggest that some of the paralinguistic cues of confidence manifest in O’s vocal behavior and utilized by P to infer O’s state, are high voice energy (loudness), short pauses, a rapid rate of speech, and probably, under conditions that will have to be explored further, a higher pitch level. Further, large fluctuations in pitch and energy for some Confident Voice conditions seem to reflect selective stress on confidence-related expressions. While we do not intend to suggest that only confidence can be expressed by this acoustic pattern, nevertheless all of these findings are closely in line with expectations and are certainly plausible. In further replications, emphasis will have to be placed on the use of natural speech samples in order to control for the possibility that the actor in the present study used highly stereotyped, stylized expressions to encode confidence.

It should be emphasized that the present study used one speaker, one text (with the aforementioned “slots”), and one sample of raters. The
study concentrated on an in-depth acoustical analysis and a relatively unobtrusive experimental design for audience evaluation. The study differs markedly from simple encoding-decoding experiments in which judges are aware of the focus of E's interest. In such studies it is feasible to use a number of speakers or encoders. Follow-ups to the present study with a large number of encoders, texts, and raters, and using fewer and less complex measurements is clearly called for in order to assess the generality of the results.

Since the present acoustic analysis has been rather crudely centered on macroscopic variables such as pitch and energy level, more fine-grained paralinguistic analyses, possibly using transcription systems (Duncan, 1969), will have to be performed in order to isolate the more microscopic paralinguistic variables such as intonation contours, drawl, and clipping. Such analyses will be difficult and time-consuming. The present study has, however, demonstrated the usefulness and practicality of computer-based methods of voice and speech analysis which, despite crudeness, seem to be valuable tools for the objective measurement of paralinguistic variables. If the present findings are replicated, computer-based speech analyses could actually be used as unobtrusive measures of expressed confidence in ongoing interactions or in speeches having persuasion as an aim. Similarly, computer analysis may prove a valuable tool for the study of speech variables in addition to voice quality analysis (Ostwald, 1963; Hargreaves, Starkweather, & Blacker, 1965; Scherer, 1971) in personality research.

The data suggest strategies for paralinguistic impression management (Goffman, 1959, 1969). For instance, the speaker in the present study relied on pitch to express paralinguistic confidence under Doubtful Text but not under Confident Text conditions. This finding suggests that speakers may "compensate" for lack of confidence cues in the linguistic channel by making extensive use of confidence cues in the paralinguistic channel.

Expressed confidence was perceived accurately in both linguistic and paralinguistic channels. There is evidence, too, that the cues used by Ss for their perceptions are the same as those used by the speaker to encode confidence: Ss in the Confident Voice conditions correctly perceived the speech as louder, faster, more fluent, and more expressive.

Finally, we turn to the personality attributions made to the speaker on the basis of both the linguistic and the paralinguistic manipulations (cf. Table 4). It is not surprising that both confidence manipulations led to higher ratings on self-confidence and self-assuredness although it is of some interest that dominance was attributed along with these more obviously confidence-related adjectives.
Confident Text but not Confident Voice led to higher ratings on the adjectives “conceited,” “professional,” and “businesslike,” and to lower ratings on the adjective “personal.” All of these attributes suggest a single dimension: professional task orientation, which at its extreme, may well take on the semblance of conceit. In terms of impression management, the linguistic channel may be especially well suited for the communication of task-related traits.

Confident Voice but not Confident Text led to higher ratings on the adjectives “enthusiastic,” “forceful,” “active,” and “competent.” It is interesting that competence is attributed on the basis of voice, while task-orientation (as postulated) is attributed on the basis of linguistic cues, since these traits are usually seen as highly related. Perhaps congruent multi-channel behavior is required to establish a communicator as both competent and task-oriented.

Somewhat higher ratings (approaching significance) for the adjective “relaxed” in the Confident Voice conditions suggest that the paralinguistic cues associated with doubt (such as halting, hesitant speech) are interpreted as tension. Goldman-Eisler’s findings (1968) that pauses of various sorts are indicative of cognitive tension states are relevant in this respect and may indicate some of the correlates of expressed confidence and doubt.

To summarize, then, the present study suggests that confidence is expressed paralinguistically by increased loudness of voice, higher pitch level (under certain conditions), shorter pauses, and a rapid rate of speech. Observers perceive and utilize these cues to attribute confidence to the speaker. In order to isolate the essential confidence-communicating cues more reliably, direct and independent manipulation of each of these cues is called for. Recent electro-acoustical techniques developed for speech synthesis purposes (Udall, 1960; Flanagan, 1965; Wolf, 1970) point to possible directions for further research.

REFERENCES


THE VOICE OF CONFIDENCE


WEICK, K. E. Systematic observational techniques. In G. Lindzey & E. Aronson
SCHERER, LONDON, AND WOLF


