*Annenberg School for Communication*

*Departmental Papers (ASC)*

University of Pennsylvania *Year* 1979

# Computer analysis of talk-silence sequences: The FIASSCO system

Joseph N. Cappella* Michael J. Streibel†

*University of Pennsylvania, jcappella@asc.upenn.edu

†University of Wisconsin

This paper is posted at ScholarlyCommons.

http://repository.upenn.edu/asc_papers/109

*Behavior Research Methods & Instrumentation*
*1979, Vol. 11 (3), 384-392*

# Computer analysis of talk-silence sequences: The FIASSCO system

JOSEPH N. CAPPELLA and MICHAEL J. STREIBEL
*Center for Communication Research, Department of Communication Arts*
*University of Wisconsin, Madison, Wisconsin 53706*

**The study of simple talk and silence indices that characterize conversation is limited by the costly, "labor-intensive" character of data collection and analysis. In the face of results demonstrating the significance of these data in interpersonal judgments (Hayes & Meltzer, 1972; Lustig, Note 1), more efficient collection, storage, and analysis methods are required. This report describes a hardware and software system, FIASSCO, that collects, stores, and analyzes two-person separate-channel audio-recorded conversations for various indices of talk and silence. Data output are both continuous and discrete measures in time sequence. Further, data on the validity and reliability of FIASSCO output are provided along with sample analyses of computer results.**

This report presents a computer-assisted method of coding and analyzing sound-silence patterns in dyadic relationships. Specifically, it details the procedures used at the University of Wisconsin's Center for Communication Research for automatically recording, digitizing, storing, and analyzing the sound-silence patterns that are found in dyadic communication. These procedures have been developed as an integrated research approach by utilizing the potential of the center's PDP-12 computer. A discussion of how the various programs analyze a sample dyad and a reliability test of these analyses are presented.

## THEORY AND RATIONALE

It is reasonable to ask if the intensive study of simple, naked behaviors such as talk and silence is likely to offer any practical insight into interpersonal communication situations. An extensive review of the literature on simple talk and silence indices by Lustig (Note 1) produces the following: Talkative persons are more productive (Norfleet, 1948), more task oriented (Knutson, 1960; Strodtbeck & Mann, 1956), more "leader-like" (Bass, 1949, 1951; Borgatta & Bales, 1956; French, 1950; Jaffe & Lucas, 1969), more influential (Bales, 1953; Riechen, 1958; Strodtbeck, 1951), more socially adept (Knutson, 1960; Muir, 1964; Philips, 1965, 1968; Steward, 1968), and better liked (Bales, 1953; Bavelas, Hastorf, Gross, & Kite, 1965; Strodtbeck & Hook, 1961) than their less verbal counterparts.

The development of the FIASSCO system was made possible by grants from the Graduate Research Committee and the College of Letters and Sciences of the University of Wisconsin. Early work on the hardware aspects of FIASSCO owes much to Michael Redmond and Ken Emmerich; early software developments were advanced by Dan Fogel.

Based upon a similar but less comprehensive review, Hayes and Meltzer (1972) conclude, "the foregoing results and theoretical positions could easily explain and justify an intensive interest . . . in the quantitative aspects of that most social of social behaviors–talk. But the strange fact is that most . . . have been curiously uninterested in these nonverbal but vocal features of interaction" (p. 5).

It is curious that the study of talk and silence behaviors has been neglected despite the avalanche of evidence attesting to its predictive utility. It is quite possible that this neglect is based upon a suspicion that judgments apparently based on overt simple behaviors are actually confounded with the supposed richer content dimensions of speech. Given the research of Hayes and Meltzer (1972), Soskin and Kaufman (1961), and Starkweather (1956, 1961), the above suspicion appears to have no basis.

A second reason for avoiding the study of talk and silence behaviors is that the research is time consuming and, in general, labor intensive. The few attempts at programmatic research have relied on various types of technological aids to assist in the recording and analysis of talk and silence activity (Cassotta, Jaffe, & Feldstein, 1964; Chapple, 1949; Matarazzo, Saslow, & Saslow, 1956). The most sophisticated of these was developed and implemented by Jaffe and Feldstein (1970) and their associates (Welkowitz & Martz, Note 2) and is currently operating in the departments of psychology at New York University and the University of Maryland, Baltimore County. This system (the AVTA) completely automated the collection, transformation, storage, and (at least partially) the analysis of talk and silence data. The AVTA frees the researcher from the necessity of having human coders present during conversations or having coders categorize talk and silence behaviors from audio- or videotapes. Consequently, problems of

Copyright 1979 Psychonomic Society, Inc.

0005-7878/79/030384-09$01.15/0

coder fatigue, coder reliability, time, and financial costs are substantially minimized. The possibility of extensive data gathering is enhanced. We report a very similar but more flexible automatic talk and silence recording system called FIASSCO (fundamental interpersonal arrangement of sounds and silences in conversational occurrences).

The FIASSCO system permits the collection, transformation, and storage of talk and silence data in both discrete and continuous form. The discrete data are in the form of frequencies and probabilities of various dyadic states. Following Jaffe and Feldstein (1970), these states include a four-state description in which State 1 = both Person A and Person B silent, State 2 = both A and B talking, State 3 = A talking and B silent, and State 4 = A silent and B talking, and a six-state description in which State 1 = A and B silent but A has the floor, State 2 = A talking and B silent, State 3 = A and B talking but A has the floor, State 4 = A and B silent but B has the floor, State 5 = A silent and B talking, State 6 = A and B talking but B has the floor. The discrete state description permits the analysis and representation of talk and silence sequences as some stochastic process, for example,. a Markov chain (Cappella, 1976; Jaffe, 1970; Jaffe & Feldstein, 1970; Cappella, Note 3).

The continuous form of the data summarizes the duration of various indices during each turn of each speaker. Jaffe and Feldstein (1970, pp. 18-21) postulate a mutually exclusive and exhaustive descriptive scheme that includes vocalizations (continuous sounds made by a speaker bounded by silence on both sides), pauses (silences made by a speaker bounded on both sides by a vocalization by the same speaker), switching pauses (silences made by speaker bounded on one side by vocalization of the speaker and on the other side by a vocalization by the other speaker), and simultaneous speech (simultaneous vocalizations by both speakers when one of them has the floor). These data make possible the kind of descriptive and predictive studies cited earlier in the Lustig (Note 1) review, as well as studies of matching and synchronization carried out by Feldstein and his colleagues (see Feldstein & Welkowitz, 1978, for a review), Natale (1975a), and Webb (1972).

A second kind of continuous data is also available from typical talk and silence analysis, namely, vocal intensity. As a simple by-product of collecting talk and silence data, the intensity of the talk response is also routinely stored. These data could be used for studies of convergence in vocal intensity (Natale, 1975b), for studies of interruption outcomes (Meltzer, Morris, & Hayes, 1971), or for studies of responsiveness or involvement.

Each of these types of data is routinely available from the analysis of any audiotape using FIASSCO. Our current research efforts are aimed at studies of sequential structure in talk and silence and studies of matching and synchronization in the various durational measures.

## METHOD

### Subjects and Design

In the typical experiment, subjects are brought into a comfortably furnished experimental room and outfitted with either throat microphones (Audiotone, Phoenix, Arizona) or miniature boom microphones (Plantronics, Santa Cruz, California). The wires from these microphones, as well as those from a room microphone (AKG, D-109 200 ohms), are connected to cables that run through the wall to the adjoining computer room. Subjects are then instructed to discuss a certain prearranged topic and are left alone without time constraints placed on their discussion. The dialog of each dyad is recorded on a four-channel tape recorder (Sony TC-366-4, quadraphonic stereo) with the room microphone recorded on one channel and each subject recorded on a separate channel. The fourth channel is left empty for the later insertion of cue pulses.

### Apparatus and Procedure

The overall flow of information and analysis proceeds from the dyad onto an audiotape and then, either immediately or later, through electronic smoothing and digitizing circuitry into a computer-stored data base. The various summaries, analyses, and calculations then proceed under the control of a researcher in an interactive fashion with the computer. The output from the programs that perform the various functions remains in the computer's data base and can be printed or transmitted to other computers at any time. Figure 1 schematizes the overall flow of information and analyses.

The audio signal from each person in the dyad can be processed (i.e., electronically rectified, smoothed, and digitized) in real-time while the discussion is going on, or it can be processed from a previously recorded audiotape. The audio signal is fed into a custom-built patch-panel circuit (PPC) where

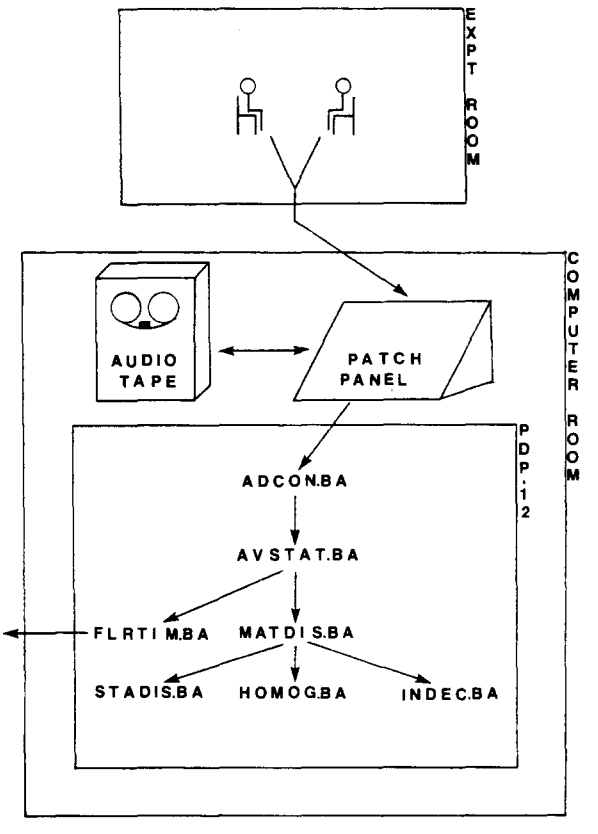


Figure 1. A schematic representation of the FIASSCO system.

it can be electronically mixed, rectified, and smoothed as desired (see Figure 1). The smoothing circuit has a rise time of 21 msec in response to a square wave. This was designed so that within-word silences would be less likely to be digitized as silences. The audio signal from the patch panel is then fed into an analog-to-digital (A/D) converter that operates under the PDP-12 computer's program control. The A/D converter senses an audio signal varying between +1 V and −1 V and converts this signal into integer numbers that range from −511 to +511 (0 to +511 when the signal has been rectified into a positive voltage). The voltage level of the audio signal can be manually adjusted by the tape-recorder output and/or PPC.

Figure 1 (bottom block) shows the various programs used to digitize and analyze the audio signal. These programs are written in BASIC and are stored by the PDP-12's disk-based operating system (DOS/12). The researcher sits at the PDP-12 console and executes each program as required. The output from each of these programs is stored as a text file by the PDP-12 computer. The output files can be printed, transmitted to other computers over the telephone, or read by other programs as input for their analyses. What follows is a brief discussion of each of these programs and how they are used with dyadic sound-silence data. Output examples from some of these programs are given in the appendices for the six-state analyses. The four-state outputs are comparable.

**Programs**

**ADCON.BA.** This program converts the incoming audio signal from each person into digital form and stores these numbers in a file for later use (see Appendix A). Its utility lies in the fact that it permits a researcher to specify the sampling rate at which the audio signal is digitized. We have settled on a 50-msec sampling rate. This means that we convert each second of dialog of each person in the dyad into 20 numbers that range from 0 to 511. This rate, in conjunction with the electronic and statistical smoothing of the audio signal, is sufficient to discriminate between-word silences for each person without picking up within-word silences (see Goldman-Eisler, 1958; Jaffe & Feldstein, 1970, Appendix A; see also data reported below). The statistical smoothing involves the calculation of a moving average typically 5 sampling units in duration. The moving average has the same effect as the electronic smoothing, minimizing substantially the presence of sharp peaks and troughs. The program ADCON.BA also permits the researcher to type an extensive header text into the output file so that the analysis is clearly identified. Another useful feature of ADCON.BA is that the digitized audio signal can be displayed on a cathode-ray tube (CRT). This permits the researcher to adjust signal levels and to choose when to store data.

**AVSTAT.BA.** This program takes ADCON.BA output and averages any number of data points into a single number; it then converts these data into a four- or six-state string of numbers. Typically, we average six data points so that the resultant string of numbers reflects the audio signal as if it had been sampled and digitized once every 300 msec. A silence between words therefore must be at least 300 msec before it is counted as a silence. AVSTAT.BA also permits the researcher to specify a cut-off level below which an audio signal is considered noise. The cut-off levels are chosen on a pragmatic and experiential basis after ADCON.BA output is examined. The same procedure is typically employed by Jaffe and Feldstein (1970). The four- and six-state analyses proceed according to the scheme described earlier for coding discrete talk and silence data. The output of AVSTAT.BA is therefore a string of numbers (1-4 or 1-6) that reflects the state of the sound-silence patterns every 300 msec for however long one chooses (see Appendix B). AVSTAT.BA output can be read by either FLRTIM.BA or MATDIS.BA.

**FLRTIM.BA.** This program performs an analysis of the six-state time-series data calculated by AVSTAT.BA and categorizes these data sequentially for each person according to the within-turn initiation time and duration of vocalization, pause, switching pause, simultaneous speech, and floor time for each floor turn. The start time and duration of each of these categories is presented in 300-msec time units. FLRTIM.BA therefore provides a complete and convenient summary of the six-state time-series data. A portion of the data presented in Appendix B is presented in Appendix C after analysis by FLRTIM.BA. This program also has an option to store and/or print only a summary of the conversation. Under this option, only total vocalization time, total pause time, and total simultaneous speech duration are printed for each speaker turn. Typically, the output from FLRTIM.BA is transferred to larger computers, where the statistical software necessary to carry out time-series analyses of the data (e.g., Cappella, in press) is available.

**MATDIS.BA.** This program takes AVSTAT.BA output and calculates a frequency and probability transition matrix for prespecified time units within the dyadic conversation, as well as a composite frequency and probability transition matrix. A researcher can therefore immediately know the frequency and probability of state sequences in the four- or six-state data (see Appendix D). MATDIS.BA also calculates the individual contingent probabilities that Person A or B will talk, given one of the prior dyadic states according to the following formulas (Jaffe & Feldstein, 1970).

$$p_i^*(t,k,A) = p_{i2}(t,k) + p_{i3}(t,k) \qquad i = 1,2,3$$

$$p_i^*(t,k,A) = p_{i2}(t,k) + p_{i6}(t,k) \qquad i = 4,5,6$$

$$p_i^*(t,k,B) = p_{i3}(t,k) + p_{i5}(t,k) \qquad i = 1,2,3$$

$$p_i^*(t,k,B) = p_{i5}(t,k) + p_{i6}(t,k) \qquad i = 4,5,6,$$

And, for the four-state description,

$$p_i^*(t,k,A) = p_{i2}(t,k) + p_{i3}(t,k) \qquad i = 1,2,3,4$$

$$p_i^*(t,k,B) = p_{i2}(t,k) + p_{i4}(t,k) \qquad i = 1,2,3,4,$$

where the $p_i^*(t,k,A)$ are the individual contingent probabilities at time t, conversation k, that Person A will talk at t + 1 given the prior state of the dyad is i and where $p_{ij}(t,k)$ is the probability of transition from state i to state j at time t and in conversation k. Typically, conversations are broken into 2-min time units that guarantee data per cell sufficient to insure nonzero row frequencies (400 observations at 300-msec time units).

The output from MATDIS.BA is stored for subsequent analysis by programs testing Markov chain assumptions and predictions (Hewes, 1975, in press; Kemeny & Snell, 1960) or is transferred to larger computers where software necessary to carry out other analyses is available.

**STADIS.BA.** This program tests the stationarity assumption of MATDIS.BA output by taking the initial transition matrix to the Mth power (i.e., predicted matrix), where M is the number of arbitrarily defined time units in the discussion, and comparing it with the product of the M transition matrices (i.e., observed matrix). A discrepancy matrix between the predicted and observed matrices indicates the degree to which these data fit a first order Markov chain (Kemeny & Snell, 1960). A chi-square measure of discrepancy for each of the M matrices from the composite matrix is calculated as follows.

$$\chi^2 = \sum_i \sum_j \frac{f_{i.}(t,k)\,[p_{ij}(t,k) - p_{ij}(T,k)]^2}{p_{ij}(T,k)} \qquad t = 1,2,\ldots,M,$$

where $p_{ij}(T,k)$ is the ijth transition element from the composite transition matrix of the kth conversation, $p_{ij}(t,k)$ is the ijth element of the tkth matrix, and $f_{i.}(t,k)$ is the row frequency of

the same matrix (Hewes, in press). This value is distributed as chi square with 12 deg of freedom in the four-state case and 18 deg of freedom in the six-state case and serves as a statistic testing the assumption of stationarity.

**HOMOG.BA.** This program takes the frequency and probability output from MATDIS.BA and tests the homogeneity assumption of a first order Markov chain of two or more composite matrices against the overall composite by a chi-square test (Hewes, in press):

$$x^2 = \sum_{g}^{G} \sum_{i} \sum_{j} f_{i.}(T,g) \frac{[p_{ij}(T,g) - p_{ij}(G)]^2}{p_{ij}(G)},$$

where $p_{ij}(G)$ is the ijth element of the composite matrix across all G groups, $p_{ij}(T,g)$ is the ijth element of the composite transition matrix of the gth group, and $f_{i.}(T,g)$ is the row frequency of the latter matrix. HOMOG.BA can test homogeneity for as many as 10 transition matrices simultaneously.

**INDEC.BA.** This program takes MATDIS.BA output of three dyadic discussions, uses the individual parameters of the same two persons in these discussions to predict a new composite transition matrix, and compares it to the observed matrix for these two persons in these discussions to predict a new composite transition matrix, and compares it to the observed matrix for these two persons (Jaffe & Feldstein, 1970, pp. 84-89, 111-112). For example, if Person A is matched with Person C in one 20-min dyadic discussion and Person B is matched with Person D in another discussion, we can use the information about Persons A and B from these two discussions to predict a transition matrix for them and compare it with an actual discussion between them. For the six-state case, the predicted transition elements on Occasion 2, $\hat{p}_{ij}$, from the individual parameters would be as follows.

$$\hat{p}_{i1} = [1 - p_i^*(T,1,A)]\ [1 - p_i^*(T,1,B)] \qquad i = 1,2,3;$$

$$\hat{p}_{i2} = p_i^*(T,1,A)\ [1 - p_i^*(T,1,B)] \qquad i = 1,2,3;$$

$$\hat{p}_{i3} = p_i^*(T,1,A)\ p_i^*(T,1,B) \qquad i = 1,2,3;$$

$$\hat{p}_{i4} = 0 \qquad i = 1,2,3;$$

$$\hat{p}_{i5} = [1 - p_i^*(T,1,A)]\ p_i^*(T,1,B) \qquad i = 1,2,3;$$

$$\hat{p}_{i6} = 0 \qquad i = 1,2,3;$$

$$\hat{p}_{i1} = 0 \qquad i = 4,5,6;$$

$$\hat{p}_{i2} = p_i^*(T,1,A)\ [1 - p_i^*(T,1,B)] \qquad i = 4,5,6;$$

$$\hat{p}_{i3} = 0 \qquad i = 4,5,6;$$

$$\hat{p}_{i4} = [1 - p_i^*(T,1,A)]\ [1 - p_i^*(T,1,B)] \qquad i = 4,5,6;$$

$$\hat{p}_{i5} = [1 - p_i^*(T,1,A)]\ p_i^*(T,1,B) \qquad i = 4,5,6;$$

$$\hat{p}_{i6} = p_i^*(T,1,A)\ p_i^*(T,1,B) \qquad i = 4,5,6;$$

INDEC.BA has been used to produce the data reported in Cappella (Note 3).

### Summary

The sequence of programs from ADCON.BA to INDEC.BA make possible the automatic collection, storage, and analysis of durational and discrete talk and silence data. The researcher can take a recorded two-channel conversation and, in a short time, have a complete description of the distribution of talk and silence parameters across time, across person, and across time and person for several different indices. What remains to be seen is whether basic talk and silence data acquired by FIASSCO are comparable to those acquired by other systems.

## COMPARABILITY OF DATA

In order to insure that the FIASSCO system produces reliable and valid data, a series of short studies was carried out. A more extensive replication and extension is reported in Cappella (Note 3). These studies attempted to establish (1) the similarity between descriptive data obtained via FIASSCO and those obtained via AVTA, (2) the reliability of machine coding of a sequence of talk and silence data, (3) the degree of difference between machine- and human-coded talk-silence data, and (4) the adequacy of the time-sampling unit (Arundale, 1977).

### Descriptive Data

There are three types of descriptive data that should be compared to existing data: mean durations on continuous indices, distributional data on continuous indices, and sequential data on categorical indices. The last comparison is reported at length in Cappella (Note 3) and will not be repeated here.

Based on 12 20-min dyadic conversations reported in Cappella (Note 3), average durations of vocalization, pause length, switching pause length, and length of simultaneous speech were calculated. These mean scores are reported in Table 1 and compared to those reported by Jaffe and Feldstein (1970) and Lustig (Note 1). Clearly, the data obtained with FIASSCO are very similar to those reported by other procedures, with vocalization being somewhat shorter and simultaneous speech somewhat longer than other reports.

If the data obtained with FIASSCO are similar to those obtained with other methods, then there is evidence that the continuous indices are comparable to identical indices gathered with other software and hardware systems. Figures 2, 3, and 4 provide such evidence. These graphs depict the relationship between the frequency of occurrence of various durations and the duration of vocalization (Figure 2), the duration of pauses (including both switching pauses and within-turn

**Table 1**
**Comparison of Descriptive Statistics on Vocalizations, Pauses, Switching Pauses, and Simultaneous Speech from Several Laboratories**

| | Brady* | Jaffe and Feldstein* | Lustig** | FIASSCO† |
|---|---|---|---|---|
| Mean Vocalization Duration | 1.17 | 1.64 | 1.48 | 1.13 |
| Mean Pause Length | .50 | .66 | | .61 |
| Mean Switching Pause Length | .40 | .77 | | .62 |
| Mean Simultaneous Speech | .25 | .40 | | .61 |

**Reported in Jaffe and Feldstein (1970, p. 128). **Data gathered with identical hardware but different software than FIASSCO at Center for Communication Research, Madison, Wisconsin; conversations were triadic. †In seconds; based on 12 20-min conversations, 48,000 300-msec samples.*

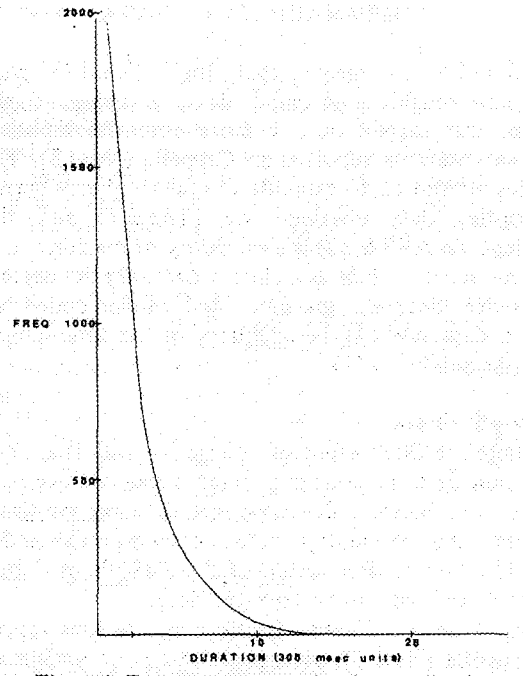


Figure 2. Frequency vs. duration for vocalizations.

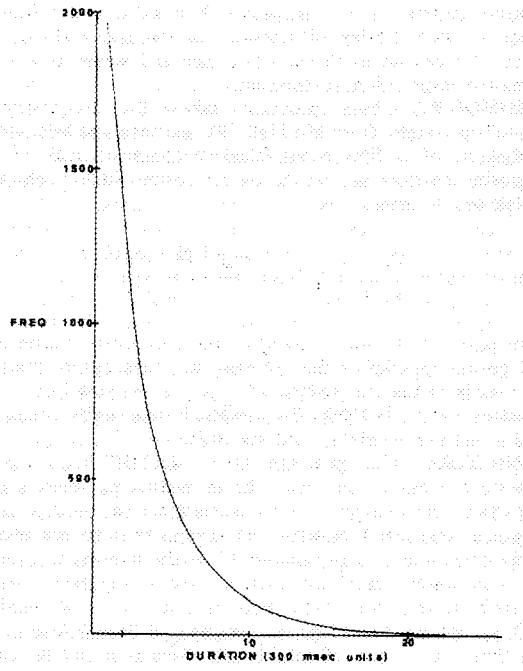


Figure 3. Frequency vs. duration for pauses.

pauses) (Figure 3), and the duration of simultaneous speech (Figure 4). The functions presented are obviously exponential and thus match the distributions reported in Jaffe and Feldstein (1970, pp. 75-77).

**Reliability Data**

Two small reliability studies were carried out to compare a machine analysis of 6 min of a randomly selected conversation to that coded by two human coders using a pushbutton procedure. The human coders entered a pulse on audiotape when they detected talk in the conversation. Each coded independently, and the machine analyzed their pushbutton responses and compared them to its own analysis. Of 1,200 coded time units, the machine and Coder A agreed on 93.5% of the cases; the machine and Coder B agreed on 69% of the cases. The average agreement was 81.2%. In general, machine coding and human coding are similar, suggesting that judges' perceptions of talk and silence are strongly related to machine indices of these data.

The second study sought to establish the reliability of the machine's analysis of a given signal. If such reliability cannot be established, then the signals typically analyzed would be so unstable as to insure different results each time the signal was studied. Such instability would call into question the ability to give an unequivocal characterization of any input signal. If the representations are equivocal, then either the phenomenon is too erratic for study or the methodology for studying it introduces random components.

Because we expect FIASSCO to be reliable, very strict criteria for reliability are necessary. A randomly sampled 4-min segment of conversation was analyzed twice using the FIASSCO system. In order to insure that the samplings in the two runs began at precisely the same time, a cue pulse on the third channel of the audiotape automatically initiated the sampling of the other two channels. The pulse was less than 50 msec wide to insure that no initial samples in the other two channels were missed. Sampling in both runs was at 50 msec. Each of the channels' signals was smoothed by a 5-point moving average procedure and later summed over six adjacent samples, yielding a sampling unit of 300 msec duration. There were 800 observations in each

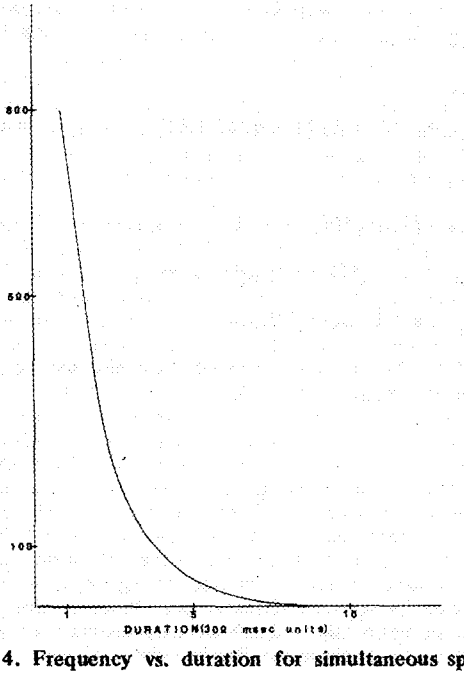


Figure 4. Frequency vs. duration for simultaneous speech.

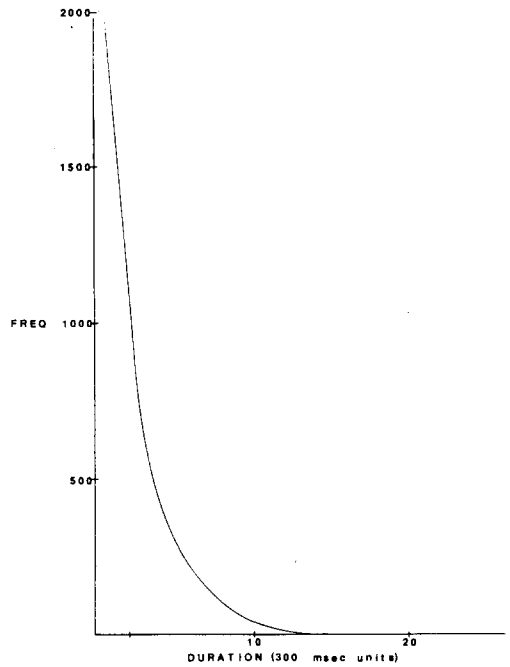


Figure 2. Frequency vs. duration for vocalizations.

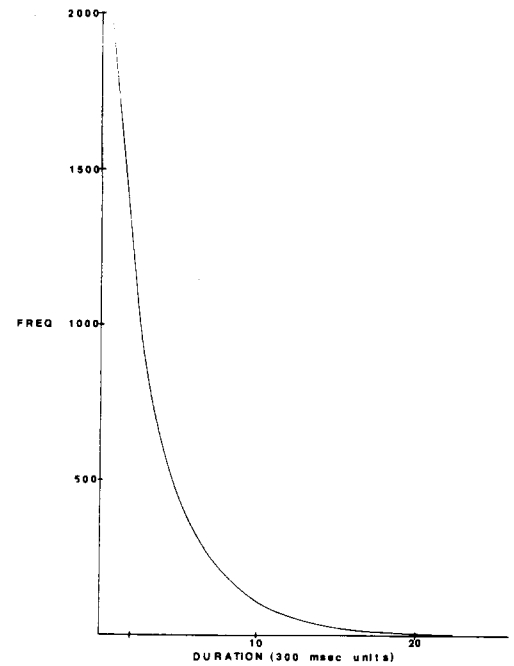


Figure 3. Frequency vs. duration for pauses.

pauses) (Figure 3), and the duration of simultaneous speech (Figure 4). The functions presented are obviously exponential and thus match the distributions reported in Jaffe and Feldstein (1970, pp. 75-77).

**Reliability Data**

Two small reliability studies were carried out to compare a machine analysis of 6 min of a randomly selected conversation to that coded by two human coders using a pushbutton procedure. The human coders entered a pulse on audiotape when they detected talk in the conversation. Each coded independently, and the machine analyzed their pushbutton responses and compared them to its own analysis. Of 1,200 coded time units, the machine and Coder A agreed on 93.5% of the cases; the machine and Coder B agreed on 69% of the cases. The average agreement was 81.2%. In general, machine coding and human coding are similar, suggesting that judges' perceptions of talk and silence are strongly related to machine indices of these data.

The second study sought to establish the reliability of the machine's analysis of a given signal. If such reliability cannot be established, then the signals typically analyzed would be so unstable as to insure different results each time the signal was studied. Such instability would call into question the ability to give an unequivocal characterization of any input signal. If the representations are equivocal, then either the phenomenon is too erratic for study or the methodology for studying it introduces random components.

Because we expect FIASSCO to be reliable, very strict criteria for reliability are necessary. A randomly sampled 4-min segment of conversation was analyzed twice using the FIASSCO system. In order to insure that the samplings in the two runs began at precisely the same time, a cue pulse on the third channel of the audiotape automatically initiated the sampling of the other two channels. The pulse was less than 50 msec wide to insure that no initial samples in the other two channels were missed. Sampling in both runs was at 50 msec. Each of the channels' signals was smoothed by a 5-point moving average procedure and later summed over six adjacent samples, yielding a sampling unit of 300 msec duration. There were 800 observations in each

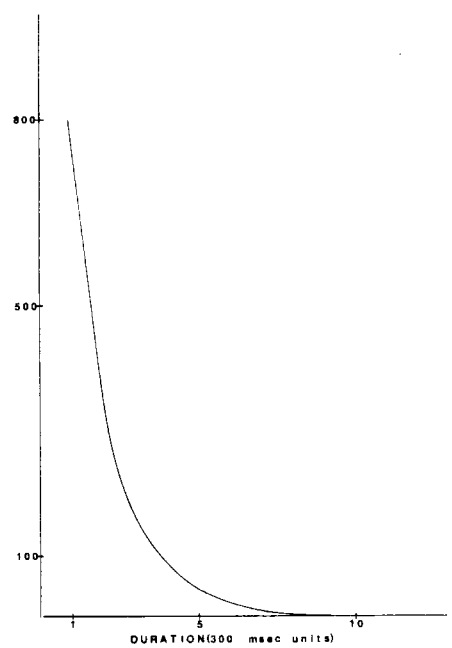


Figure 4. Frequency vs. duration for simultaneous speech.

**Table 2**
**Confusion Matrix Representing Two Separate FIASSCO Analyses of a 4-Min Segment of Conversation**

| | Run 2 | | | |
|---|---|---|---|---|
| Run 1 | 1 | 2 | 3 | 4 |
| 1 | 147<br>36.8 | 7<br>48.2 | 4<br>20.6 | 29<br>80.8 |
| 2 | 9<br>43.9 | 195<br>57.6 | 17<br>24.6 | 3<br>96.5 |
| 3 | 0<br>14.8 | 6<br>19.5 | 64<br>8.3 | 6<br>32.6 |
| 4 | 2<br>61.5 | 0<br>80.7 | 2<br>34.5 | 209<br>135.1 |

*Note–The first entry in each cell is the observed frequency, the second is the expected frequency.*

of the two runs. The first reliability test was made on the two strings of 800 state categories (one, two, three, or four) output by AVSTAT.BA. Cut-off was set at 60.

The two strings are compared in the confusion matrix of Table 2. A confusion matrix is the strictest and most revealing test of reliability, since any off-diagonal entry indicates a disagreement between the two runs on some particular act. The confusion matrix represents an act-by-act assessment of reliability. Distributional and correlational indices mask act-by-act disagreements, whereas the confusion matrix reveals each and every disagreement. The confusion matrix also indicates in its off-diagonal entries where the most likely disagreements between coding runs are to be found.

The matrix shows that the most likely sources of unreliability are in the confusions between States 1 and 4 and States 2 and 3. These confusions are probably due to the cut-off level chosen and exactly where in the sampling unit the sampling pulse itself enters. While the signals are reasonably smooth, peaks and valleys remain and the sampling pulse impinges on the signal at some random point within the sampling unit. Nevertheless, reliability is excellent, with a chi square of 1,737, which is significant. This chi square yields a contingency coefficient of .828, which is 95.6% of the maximum contingency coefficient for a 4 by 4 table. Thus, the two runs are highly associated.

In addition to the act-by-act reliability, the output of ADCON.BA is a string of numbers from 0 to 511 representing the amplitude of the signal in each 50-msec unit. The two strings from each separate run of ADCON.BA, correlated over the first 150 observations, yielded a correlation coefficient of .987.

Finally, the four-state and six-state matrices output by MATDIS.BA for each of the two runs were compared by a chi-square test. The two four-state matrices differed from the composite [$\chi^2(12) = 6.01$, $p > .90$] and the two six-state matrices differed from the composite [$\chi^2(18) = 9.40$, $p > .90$]. The transition matrices between runs were not different, and consequently, the individual row frequency distributions and the contingent probabilities that are based on the transition matrix should not differ between reliability runs.

Overall, machine reliability is very high for act-by-act criteria, for continuous amplitude data, and for transition matrices. We feel confident that the signal and the method of its analysis are sufficiently reliable to permit further study.

**Sampling Unit**

Jaffe and Feldstein (1970, pp. 116-117, 123-130) comment at some length concerning the problems of establishing the appropriate fundamental sampling unit for their sound-silence studies. Based on both theoretical considerations (Goldman-Eisler, 1958) and investigation of the mean sojourn times in the observed exponential distributions of pauses, switching pauses, and vocalizations, Jaffe and Feldstein settled on a sampling unit of 300 msec. In reviewing their choice of sampling unit according to guidelines developed in simulations of sampling unit, Arundale (1977) concludes that "Jaffe & Feldstein's research thus suggests that a value of approximately 0.3 seconds has validity as the shortest sojourn time" (p. 277). Thus there is some support for choosing the .3-sec unit as the sampling time.

Our data support this finding. Table 1 shows that the shortest mean sojourn time across 48,000 300-msec observations occurs for the simultaneous speech category and is .61. The sampling unit then should be shorter than .61 sec to effectively resolve all of the real transitions in the signal (Arundale, 1977; Cane, 1959; Darwin, 1959). However, this datum is less than conclusive, since the standard deviation for simultaneous speech is large in our sample (.465), implicating a large range of sampling units from .15 to 1.0 sec.

Consequently, we chose to take a closer look at the sampling procedure adopted in AVSTAT.BA, which averages adjacent groups of 50-msec samples (where group size varies from 1 to 10). The same 4 min of data employed in the machine reliability studies were used here, since we were confident of their stability. The data were averaged for 1 unit, 2 units, . . . , and 10 units, yielding sampling units of 50, 100, . . . , and 500 msec, in 50-msec increments. Nine four-state transition matrices were output, one for each sampling unit, and the similarity between adjacent frequency matrices was described by a chi-square measure of discrepancy using only the off-diagonal frequencies. Clearly, only the off-diagonal entries should be used, since as the sampling rate increases, the diagonals will necessarily build up for any sampling unit below the minimum. The chi-square measure is only a descriptive statistic, since the diagonal entries have been removed.

Figure 5 graphs these values at different sampling interval comparisons. While no statistical comparisons are possible, there does seem to be some stabilization in the discrepancies between adjacent frequency matrices as one moves to a sampling unit of 300 msec. Researchers with talk-silence data probably should not use sampling intervals much larger than this given the mean sojourn times of Table 1. The reader is to be

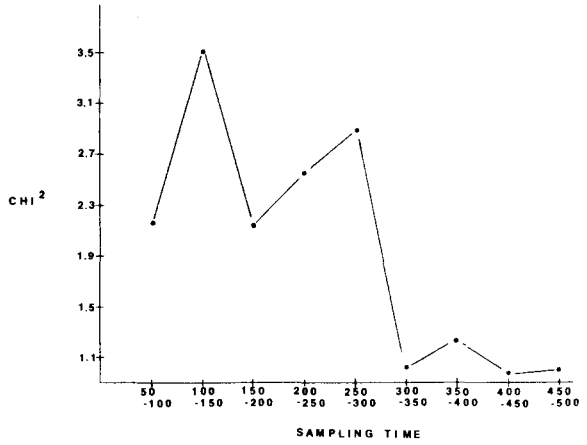


**Figure 5. Chi-square discrepancies for transition frequency matrices as sampling unit increases, based on off-diagonal frequencies only.**

cautioned concerning the data of Figure 5. In a true sampling unit study, the data in each time unit are identical and the sampling pulse randomly samples the same signal in increasingly larger sampling domains (50 msec, 100 msec, etc.). In such a case, the smallest sampling unit would contain all of the information available at that sampling duration, and larger units would possibly lose certain information. We chose to study our own averaging procedure and, consequently, the signal sampled is changing somewhat (averaging 1, averaging 2, . . . , averaging 10 values) as the sampling units (i.e., number of points averaged) change. Consequently, Figure 5 shows that at about 300 msec (six 50-msec samples averaged) the stability of results desired begins to be realized.

### Summary

The data reported in this section attest that the FIASSCO hardware and software system produces analyses of talk-silence data that are comparable to results obtained with other systems and are reliable. The sampling unit of 300 msec seems to produce stable transition matrices.

## CONCLUSION

The FIASSCO system is a usable hardware and software system for the analysis of talk and silence data in two-person conversations. It makes possible the coding and summarization of categorical and continuous talk-silence indices directly from audiotape. As a consequence, the labor-intensive and costly collection of these data is substantially reduced, and the way is opened for extensive research into talk and silence correlates and predictors, as well as research into sequential models of these basic templates upon which conversations are written.

## REFERENCE NOTES

1. Lustig, M. W. *The relationship between communicative predispositions and sound-silence patterns in triads.* Paper presented to the Speech Communication Association convention, Washington, D.C., 1977.

2. Welkowitz, J., & Martz, J. *WELMAR: Computer programs to analyze dialogic time patterns.* Unpublished manuscript, Department of Psychology, New York University.

3. Cappella, J. N. *A within- and between-situation analysis of talk-silence sequences in informal two-person conversations: Replication and extension.* Unpublished paper, Center for Communication Research, Department of Communication Arts, University of Wisconsin, Madison, 1978.

## REFERENCES

Arundale, R. B. Sampling across time for communication research: A simulation. In P. M. Hirsch, P. V. Miller, & F. G. Kline (Eds.), *Strategies for communication research.* Beverly Hills, Calif: Sage, 1977.

Bales, R. F. The equilibrium problem in small groups. In T. Parsons, R. R. Bales, & E. Shils (Eds.), *Working papers in the theory of action.* Glencoe, Ill: Free Press, 1953.

Bass, B. M. An analysis of leaderless group discussion. *Journal of Applied Psychology,* 1949, **33**, 527-533.

Bass, B. M. Situational tests: II. Variables of leaderless group discussion. *Educational Psychology and Measurement,* 1951, **11**, 196-207.

Bavelas, A., Hastorf, A. H., Gross, A. E., & Kite, R. W. Experiments on the alteration of group structure. *Journal of Experimental Social Psychology,* 1965, **1**, 55-70.

Borgatta, E. F., & Bales, R. F. Sociometric status patterns and characteristics of interaction. *Journal of Abnormal and Social Psychology,* 1956, **43**, 289-297.

Cane, V. R. Behavior sequences as semi-Markov chains. *Journal of the Royal Statistical Society,* 1959, **21**(Series B), 36-58.

Cappella, J. N. Modeling interpersonal communication systems as a pair of machines coupled through feedback. In G. R. Miller (Ed.), *Explorations in interpersonal communication.* Beverly Hills, Calif: Sage, 1976.

Cappella, J. N. Structural equation models: An introduction. In P. R. Monge & J. N. Cappella (Eds.), *Multivariate techniques in communication research.* New York: Academic Press, in press.

Cassota, L., Jaffe, J., & Feldstein, S. AVTA: A device for automatic transaction analysis. *Journal of the Experimental Analysis of Behavior,* 1964, **7**, 99-104.

Chapple, E. The interaction chronograph: Its evolution and present application. *Personnel,* 1949, **25**, 295-307.

Darwin, J. H. Note on the comparison of several realizations of a Markov chain. *Biometrika,* 1959, **46**, 412-419.

Feldstein, S., & Welkowitz, J. A chronography of conversation: In defense of an objective approach. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication.* Hillsdale, N.J: Erlbaum, 1978.

French, R. L. Verbal output and leadership status in initially leaderless group discussions. *American Psychologist,* 1950, **5**, 310-311.

Goldman-Eisler, F. Speech-analysis and mental process. *Language and Speech,* 1958, **1**, 59-75.

Hayes, D. P., & Meltzer, L. Interpersonal judgments based on talkativeness: I. Fact or artifact? *Sociometry,* 1972, **35**, 538-561.

Hewes, D. Finite stochastic modeling of communication processes. *Human Communication Research,* 1975, **1**, 271-283.

Hewes, D. Stochastic models of communication processes. In P. R. Monge & J. N. Cappella (Eds.), *Multivariate techniques in communication research.* New York: Academic Press, in press.

JAFFE, J. Linked probabilistic finite automata: A model for the temporal interaction of speakers. *Mathematical Biosciences*, 1970, 7, 191-204.

JAFFE, J., & FELDSTEIN, S. *Rhythms of dialogue*. New York: Academic Press, 1970.

JAFFEE, L. C., & LUCAS, R. L. Effects of rates of talking and correctness of decisions on leader choice in small groups. *Journal of Social Psychology*, 1969, **79**, 247-254.

KEMENY, J. C., & SNELL, J. L. *Finite Markov chains*. New York: Van Nostrand, 1960.

KNUTSON, A. L. Quiet and vocal groups. *Sociometry*, 1960, **23**, 36-49.

MATARAZZO, J. D., SASLOW, G., & MATARAZZO, R. G. The interaction chronograph as an instrument for objective measurement of interaction patterns in interviews. *Journal of Psychology*, 1956, **41**, 347-367.

MELTZER, L., MORRIS, W., & HAYES, D. Interruption outcomes and vocal amplitude: Explorations in social psychophysics. *Journal of Personality and Social Psychology*, 1971, **18**, 392-402.

MUIR, F. L. *Case studies of selected examples of reticence and fluency*. Unpublished master's thesis, Washington State University, 1964.

NATALE, M. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 1975, **32**, 790-804. (a)

NATALE, M. Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills*, 1975, **40**, 827-830. (b)

NORFLEET, B. Interpersonal relations and group productivity. *Journal of Social Issues*, 1948, **2**, 66-69.

PHILLIPS, G. M. The problem of reticence. *Pennsylvania Speech Annual*, 1965, **22**, 22-38.

PHILLIPS, G. M. Reticence: Pathology of the normal speaker. *Speech Monographs*, 1968, **35**, 39-49.

RIECHEN, H. W. The effects of talkativeness on ability to influence group solutions to problems. *Sociometry*, 1958, **21**, 309-321.

SOSKIN, W. F., & KAUFFMAN, P. E. Judgment of emotion in word-free voice samples. *Journal of Communication*, 1961, **11**, 73-80.

STARKWEATHER, J. A. The communication-value of content free speech. *American Journal of Psychology*, 1956, **68**, 121-123.

STARKWEATHER, J. A. Vocal communication of personality and human feelings. *Journal of Communication*, 1961, **11**, 63-72.

STEWARD, L. A. *Attitudes toward communication: The content analysis of interviews with eight reticent and eight nonreticent college students*. Unpublished PhD dissertation, Pennsylvania State University, 1968.

STRODTBECK, F. L. Husband-wife interaction over revealed differences. *American Sociological Review*, 1951, **16**, 468-473.

STRODTBECK, F. L., & HOOK, H. The social dimensions of a twelve-man jury table. *Sociometry*, 1961, **24**, 397-415.

STRODTBECK, F. L., & MANN, R. D. Sex role differentiation in jury deliberation. *Sociometry*, 1956, **19**, 3-11.

WEBB, J. T. Interview synchrony: An investigation of two speech rate measures. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication*. New York: Pergamon Press, 1972.

## Appendix A
## ADCON.BA Output

```
.R PIP
*TTY:<LTA1:ADC.02
/
EXPERIMENT NUMBER: 2    DATE: JULY 16, 1975
TAPE NUMBER: 1          TAPE POSITION: 275-571
PERSON A:  4            SEX: MALE
PERSON B:  3            SEX: MALE
SEQUENCE:  1            TOTAL TIME: 20.00 MINUTES
TOPIC OF DISCUSSION:    QUESTION SET 1
STRT CH:   10
# OF CH'S:  2
SAMPLING RATE: 50 MSEC
# OF OBS: 24000
/
158
8
161
  7
163
  7
176
8
207
9
187
7
154
7
118
6
.
.
.
```

## Appendix B
## AVSTAT.BA Output: Six-State Analysis

```
.R PIP
*TTY:<ADC6.02
/
EXPERIMENT NUMBER: 2    DATE: JULY 16, 1975
TAPE NUMBER: 1          TAPE POSITION: 275-571
PERSON A:  4            SEX: MALE
PERSON B:  3            SEX: MALE
SEQUENCE:  1            TOTAL TIME: 20.00 MINUTES
TOPIC OF DISCUSSION:    QUESTION SET 1
START CH:  10
# OF CH'S:  2
SAMPLING RATE: 50 MSEC
# OF OBS: 24000
6-STATE ANALYSIS
# PTS PER PERS AVGD: 6
CUTOFF LEVEL: 40
/
22222212222222222333
56
221122222222222221112222222222222222222212222222221
222222222122222222222222222122211112222211122222222
2221222222222222221222222222222111
5444444
22333
5545555545555545555455544455555455555544455
22211
5665555544
231
5445444454555555444455444444554
2
55554444444444
21221111
.
.
.
```

**Appendix C**
**FLRTIM.BA Output: Six-State Data**

PROGRAM: FLRTIM.BA

INPUT ON(DEV:NAME.EX)?ADC6.02
OUTPUT TYPE (FULL/PART)?FULL
OUTPUT ON(TTY/DEV:NAME.EX)?TTY

NOTE: 1 TIME UNIT = 300 MSEC

| # | PRS | FLOOR START | TIME LEN | VOCAL STRT | DUR LEN | W/I STRT | PAUSES LEN | SWT STRT | PAUSES LEN | SIMUL STRT | SPCH LEN |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 1 | 20 | 1 | 6 | 7 | 1 | | | 18 | 3 |
| | | | | 8 | 13 | | | | | | |
| 2 | B | 21 | 2 | 21 | 2 | | | | | 22 | 1 |
| 3 | A | 23 | 133 | 23 | 2 | 25 | 2 | 153 | 3 | | |
| | | | | 27 | 13 | 40 | 3 | | | | |
| | | | | 43 | 19 | 62 | 1 | | | | |
| | | | | 63 | 9 | 72 | 1 | | | | |
| | | | | 73 | 9 | 82 | 1 | | | | |
| | | | | 83 | 17 | 100 | 1 | | | | |
| | | | | 101 | 3 | 104 | 4 | | | | |
| | | | | 108 | 3 | 112 | 3 | | | | |
| | | | | 115 | 11 | 126 | 1 | | | | |
| | | | | 127 | 13 | 140 | 1 | | | | |
| | | | | 141 | 12 | | | | | | |
| 4 | B | 156 | 7 | 156 | 1 | | | 157 | 6 | | |
| 5 | A | 163 | 5 | 163 | 5 | | | | | 165 | 3 |
| . | | | | | | | | | | | |
| . | | | | | | | | | | | |
| . | | | | | | | | | | | |

**Appendix D**
**MATDIS.BA Output**

Six-State Transition Frequencies and Probabilities

***FREQ MATX***
TRANSITIONS: ROWS(FROM) – COLS(TO)

TIME UNIT: 1

| | | | | | | |
|---|---|---|---|---|---|---|
| 16 | 12 | 0 | 0 | 6 | 0 | 34 |
| 16 | 137 | 4 | 0 | 2 | 0 | 159 |
| 2 | 0 | 7 | 0 | 2 | 0 | 11 |
| 0 | 6 | 0 | 53 | 20 | 2 | 81 |
| 0 | 1 | 0 | 28 | 74 | 4 | 107 |
| 0 | 2 | 0 | 1 | 3 | 2 | 8 |

TIME UNIT: 2

| | | | | | | |
|---|---|---|---|---|---|---|
| 19 | 7 | 1 | 0 | 8 | 0 | 35 |
| 16 | 108 | 8 | 0 | 4 | 0 | 136 |
| 0 | 6 | 11 | 0 | 3 | 0 | 20 |
| 0 | 6 | 0 | 60 | 21 | 4 | 91 |
| 0 | 3 | 0 | 30 | 55 | 6 | 94 |
| 0 | 6 | 0 | 1 | 3 | 14 | 24 |

.
.
.

TOTAL TIME:

| | | | | | |
|---|---|---|---|---|---|
| 0.626 | 0.285 | 0.009 | 0 | 0.078 | 0 |
| 0.122 | 0.821 | 0.041 | 0 | 0.014 | 0 |
| 0.058 | 0.201 | 0.529 | 0 | 0.210 | 0 |
| 0 | 0.045 | 0 | 0.704 | 0.234 | 0.014 |
| 0 | 0.011 | 0 | 0.306 | 0.663 | 0.018 |
| 0 | 0.234 | 0 | 0.093 | 0.218 | 0.453 |

*

Six-State Contingent Probabilities

TIME UNIT: 1

| PROB(A) | PROB(B) | 1–PROB(A) | 1–PROB(B) |
|---|---|---|---|
| 0.352 | 0.176 | 0.647 | 0.823 |
| 0.886 | 0.037 | 0.113 | 0.962 |
| 0.636 | 0.818 | 0.363 | 0.181 |
| 0.098 | 0.271 | 0.901 | 0.728 |
| 0.046 | 0.728 | 0.953 | 0.271 |
| 0.5 | 0.625 | 0.5 | 0.375 |

TIME UNIT: 2

| PROB(A) | PROB(B) | 1–PROB(A) | 1–PROB(B) |
|---|---|---|---|
| 0.228 | 0.257 | 0.771 | 0.742 |
| 0.852 | 0.088 | 0.147 | 0.911 |
| 0.850 | 0.700 | 0.15 | 0.3 |
| 0.109 | 0.274 | 0.890 | 0.725 |
| 0.095 | 0.648 | 0.904 | 0.351 |
| 0.833 | 0.708 | 0.166 | 0.291 |

.
.
.

TOTAL PROB MATX

| | | | | | |
|---|---|---|---|---|---|
| 0.626 | 0.285 | 0.009 | 0 | 0.078 | 0 |
| 0.122 | 0.821 | 0.041 | 0 | 0.014 | 0 |
| 0.058 | 0.201 | 0.529 | 0 | 0.210 | 0 |
| 0 | 0.045 | 0 | 0.704 | 0.234 | 0.014 |
| 0 | 0.011 | 0 | 0.306 | 0.663 | 0.018 |
| 0 | 0.234 | 0 | 0.093 | 0.218 | 0.453 |

*

(Received for publication December 14, 1978;
revision accepted January 15, 1979.)