

 Open access • Journal Article • DOI:10.1111/J.1468-2958.1980.TB00133.X

Talk and silence sequences in informal conversations ii — Source link

Joseph N. Cappella

Institutions: University of Wisconsin-Madison

Published on: 01 Sep 1979 - Human Communication Research (Blackwell Publishing Ltd)

Topics: Silence

Related papers:

- [Talk and Silence Sequences in Informal Conversations III: Interspeaker Influence](#)
- [Rhythms of dialogue](#)
- [A simplest systematics for the organization of turn-taking for conversation](#)
- [The chameleon effect: The perception–behavior link and social interaction.](#)
- [Speech style and social evaluation](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/talk-and-silence-sequences-in-informal-conversations-ii-3s15a6h200>



University of Pennsylvania
ScholarlyCommons

Departmental Papers (ASC)

Annenberg School for Communication

December 1980

Talk and Silence Sequences in Informal Conversations II

Joseph N. Cappella

University of Pennsylvania, jcappella@asc.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/asc_papers

Recommended Citation

Cappella, J. N. (1980). Talk and Silence Sequences in Informal Conversations II. *Human Communication Research*, 6 (2), 130-145. <https://doi.org/10.1111/j.1468-2958.1980.tb00133.x>

NOTE: At the time of publication, author Joseph Cappella was affiliated with the University of Wisconsin. Currently (April 2008), he is a faculty member of the Annenberg School for Communication at the University of Pennsylvania.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/asc_papers/105
For more information, please contact repository@pobox.upenn.edu.

Talk and Silence Sequences in Informal Conversations II

Abstract

Three models describing the structure of talk and silence sequences within and across conversations presented in a previous report (Cappella, 1979) are tested. The Markov model, describing talk and silence sequences within a conversation, is found to be a valid representation on a dyad-by-dyad basis. The Independent Decision (ID) model shows some predictive validity between conversations, although its "fit" within the conversation is less than the Markov model. The Incremental model in relaxing the consistency-across-conversation assumption of the ID model finds differences due to switching of partners in the probability of breaking or continuing mutual silences and in the probability of continuing to hold the floor. The implication for deriving dyadic interaction patterns from individual interaction styles are explored.

Comments

NOTE: At the time of publication, author Joseph Cappella was affiliated with the University of Wisconsin. Currently (April 2008), he is a faculty member of the Annenberg School for Communication at the University of Pennsylvania.

TALK AND SILENCE SEQUENCES IN INFORMAL CONVERSATIONS II

JOSEPH N. CAPPELLA
University of Wisconsin

Three models describing the structure of talk and silence sequences within and across conversations presented in a previous report (Cappella, 1979) are tested. The Markov model, describing talk and silence sequences within a conversation, is found to be a valid representation on a dyad-by-dyad basis. The Independent Decision (ID) model shows some predictive validity between conversations, although its "fit" within the conversation is less than the Markov model. The Incremental model in relaxing the consistency-across-conversation assumption of the ID model finds differences due to switching of partners in the probability of breaking or continuing mutual silences and in the probability of continuing to hold the floor. The implication for deriving dyadic interaction patterns from individual interaction styles are explored.

In an earlier report (Cappella, 1979), several models of talk and silence sequences within and between conversations were posited on the basis of previous literature, especially the work of Jaffe and Feldstein (1970). To recapitulate briefly, a two-person conversation can be described in terms of the four-state description (all combinations of talk and silence for both persons) or in terms of the six-state description (which keeps track of who is holding the floor). In addition, the individuals in each dyad can be characterized by four (or six) individual parameters describing the probability of talk given the prior dyadic state. The individual parameters are obtained directly from the four- or six-state dyadic transition matrices.

Four models, applicable to both the four- and six-state descriptions, were offered to account for the within- and between-conversation sequences of talk and silence. A simple first-order Markov Chain model was presented to describe the within-conversation sequences on a dyad-by-dyad basis. The Independent Decision (ID) model was posited as the simplest representation of between-conversation sequences with partner switching.

This model is simple because of its strong assumptions: (1) persons are perfectly consistent in their individual parameters across conversations and (2) persons act independently of the actions of others. The other two models treated, the Incremental and Regulation models, are variants of the ID model which relax the "consistency across conversations" assumption and the consistency and independence assumptions, respectively.

The research reported here presents evidence on the within-occasion Markov model and the between-occasion Independent Decision and Incremental models. The Regulation model will be evaluated in a subsequent report.

While the study reported here is hoped to be the first in a program of research on talk and silence sequences in informal conversations, it is in several respects a replication of sequential analyses done by Jaffe and Feldstein (1970, p. 30) and Feldstein, Jaffe, and Cassotta (1966, 1967). I believe that replication is important because

1. There are few laboratories capable of producing the data which Jaffe and Feldstein have reported from their own facilities.
2. Certain of their results have become controversial (Hayes, Meltzer, & Wolf, 1970), so that independent replication on a separate but structurally similar facility is desirable.

Joseph N. Cappella (Ph.D., Michigan State University, 1974) is associate professor of communication arts at the University of Wisconsin, Madison, Wisconsin 53706. This study accepted for publication June 14, 1979.

3. The beginnings of programmatic research which will rely heavily on a specialized data acquisition facility (such as that used in this study) should be based upon established results.
4. Replication is that which distinguishes scientific knowledge from other forms of knowing.

PROCEDURES

Subjects, Task, and Design

In the summer of 1975, an initial study was undertaken to replicate the work of Jaffe, Feldstein, and their associates on the Markov and ID models discussed previously. The study was modeled after Jaffe and Feldstein's study (1970, p. 30), also discussed in Feldstein, Jaffe, and Cassotta (1966, 1967).

A volunteer sample of eight summer-school students enrolled in an introductory group discussion course was obtained. The subjects were told that this was an experiment about how well people understood one another. The four males and four females were given a lengthy questionnaire on changing sex roles and sex role stereotypes upon volunteering. The questionnaire was designed to provoke thought about sex-role issues so that differences in topic familiarity might be reduced during initial conversations with peers. The subjects were permitted to take the questionnaire home with them and fill it out at their leisure. Discussion of the questions with peers and friends was encouraged. The questionnaires were returned to the experimenter at their first meeting.

Subjects were randomly assigned to a same-sex partner on their first occasion and asked to discuss the issues raised on the questionnaire which they had completed. Their task was simply to discuss the issues fully and to aim to understand not only the point of view of their interlocutors but also the reasons why their partners held that view. Each conversation lasted about 30 minutes. Subjects returned for three such conversations at approximately one-week intervals. On the second occasion, each person was randomly assigned to an opposite-sex partner and given the same instructions and task. On the third occasion, each person was randomly assigned to another same-sex partner (but

different from the first) for the third conversation.

Ideally, each discussant would have had three discussions with three different partners. However, despite numerous rescheduling efforts, one subject would not return for a second conversation. As a result, a ninth volunteer was obtained to fill the gaps. Overall there were 12 dyadic conversations for nine subjects distributed across three occasions (i.e., different partners). Due to the person who dropped, one subject had only one partner, one had four different partners, two had two different partners, and the remainder had the required three different partners.

After the third discussion subjects were debriefed but no postdiscussion questionnaires were distributed.

Data Acquisition

Subjects were seated in comfortable chairs in face-to-face position approximately 1.2 m apart in a small experimental room. Each person was fitted with a vibration-conducting throat microphone which was snug but permitted freedom of movement. A small low table between the pair also held a nondirectional microphone. Each microphone fed into a separate channel of a four-track tape recorder in an adjacent room where the three channels were preserved. Throat microphones are necessary because ordinary microphones are not unidirectional and will pick up talk in the channel of person A when person B is speaking loudly. This "spillover" problem addressed by Jaffe and Feldstein (1970, Appendix A) is especially problematic when one is concerned with the independent contributions to the conversation of each person. Since the throat microphone responds to vibrations in the speaker's vocal chords, the spillover problem is effectively solved. However, the throat microphones do not produce easily intelligible content reproductions, and so it was necessary to record the content of the conversation on a third channel.

Once the conversations are stored on audiotape, they can be analyzed for the presence or absence of sound. The electronic and computer hardware and software systems necessary to carry out the transformation of the audio recordings to a digital representation of talk or silence are described in detail in

Cappella and Streibel (in press). In brief, the process includes four steps: (1) the audio signals are amplified, rectified, and electronically smoothed, (2) the smoothed signal is sampled at 50 millisecond intervals, digitized, and stored on tape, (3) the stored data is statistically smoothed, averaged over six 50-millisecond samples to yield a 300-millisecond sample, and converted to a 1 (for talk) and a 0 (for silence) if the amplitude exceeds a certain threshold, and (4) the data are stored as an interleaved string of 0s and 1s for later analysis.

In the data set reported here, the conversations lasted 25-30 minutes. The first 4-5 minutes in each conversation were ignored to allow the subjects to become comfortable in their new roles, and exactly 20 minutes of conversation were subjected to talk-silence analysis. This produces four thousand 300-millisecond observations for each subject.

The strings of 0s and 1s can be analyzed by a series of programs¹ (Cappella & Streibel, in press), but primarily ones which convert the 0-1 sequences into the four-state and six-state representations and from these to the four-state and six-state transition matrices with each dyad being described by $4000 - 1 = 3999$ transitions. The present data were summarized into 10 two-minute transition matrices for each dyad in both the four-state and the six-state versions. These are the primary data forms necessary for testing the models.

Reliability and Validity

Studies aimed at establishing the reliability and validity of machine-coded talk and silence data, as carried out in this study, are reported fully in Cappella and Streibel (in press). Both reliability and validity were quite high. Act-by-act machine reliability showed 92 percent agreement between runs on the same data and act-by-act amplitude correlations of .987. Transition matrices were not different from the composite in adjacent runs ($X^2=6.01$, $p>.90$, $df=12$). Overall automatic analysis of talk and silence in conversation is highly reliable.

Furthermore, human coders showed 81 percent agreement with machine coding of a six-minute stretch of conversation, indicating similarity between machine judgments and naive judgments of

conversational talk and silence. Cappella and Streibel (in press) also report similarity to previously reported data in the means and the distributional shapes of pauses, switching pauses, vocalizations, and simultaneous speech. They also report data indicating that the typical 300-millisecond sampling interval provides a stable unit for the analysis of conversational sequences.

RESULTS

In the results which follow, both the four-state and six-state conclusions will be presented. First, the within-occasion Markov model will be tested. Second, the between-occasion ID model will be tested. Third, a predictive version of the Incremental model will be presented and tested.

The Within-Occasion Markov Model

To determine if a set of data fits a Markov Chain, a researcher has two choices: test each of the assumptions and conclude that the predictions will match observation, or presume that the assumptions are met until predictions fail to match observations. We will try both strategies. There are usually three assumptions to be tested in a Markov Chain: stationarity, homogeneity, and order. The order assumption is not tested here because two independent laboratories using two different sampling rates (.166 seconds and .300 seconds) have concluded that talk-silence sequences exhibit statistically significant but negligible second-order dependencies (Jaffe & Feldstein, 1970, p. 74, note 7; Hayes, Meltzer, & Wolf, 1970, p. 266). My own biases toward simplicity suggest that the first-order model be retained.

The stationarity assumption maintains that the 10 transition matrices within a conversation are all equal to the expected matrix and, hence, to one another. Using a X^2 test for stationarity (Hewes, in press), each of the 12 dyads were tested for nonstationarity in each of the 10 transition matrices. The averaged results are presented in Table 1. Using an $\alpha=.01$ significance criterion, Table 1 shows that only one of five time periods for the four-state and none of the five time periods for the six-state are

TABLE 1
Average Stationarity Values Across 12 Experiments as Computed by the Chi-Squared Statistic: Four- and Six-State Descriptions

<u>Four State</u>		<u>Six State</u>	
<u>Time Units</u>	<u>X²</u>	<u>Time Units</u>	<u>X²(+)</u>
1 & 2	21.4	1 & 2	27.7
3 & 4	24.9	3 & 4	28.5
5 & 6	18.6	5 & 6	23.7
7 & 8	25.1	7 & 8	30.4
9 & 10	28.7*	9 & 10	32.3

* Significant at $p < .01$, $df=12$.

+No X² values significant at $p < .01$, $df=6 \cdot 5 - 12 = 18$.

N.B. Any X² value greater than 21.0 is significant at $p < .05$ for $df=12$;
 any X value greater than 28.9 is significant at $p < .05$ for $df=18$.

significantly different from the composite transition matrix. Similarly, the individual dyad data show 11 of 60 (four-state) X²s and 13 of 60 (six-state) X²s significantly different from the composite at $\alpha = .01$. Together, the average X² and the individual X² values suggest that the within-occasion data is stationary at the preassigned $\alpha = .01$ level.

The choice of confidence level at $\alpha = .01$ deserves some discussion since if α were set at .05, four of five values for the four-state and two of five values for the six-state would be considered statistically significant. The problem is this: with huge amounts of data, small discrepancies between probabilities in the observed and expected matrices are magnified by the expected row frequency, which is usually large. The numerator of the X² stationarity statistic is the square of an observed minus expected probability times the row frequency of the expected matrix. Since the expected matrix for any dyad consists of 400 observations, any row frequency could easily be in the 50-100 range as a multiplying factor of the discrepancy. Thus, what appear to the researcher's eye as negligible discrepancies are greatly magnified because of the large data base upon which the statistics are calculated. The problem is further compounded because the degrees of freedom for the

X² test take into account *only* the number of independent *cells* in the transition matrix, and not the number of observations upon which they are based.

In order to compensate for the bias toward finding statistical significance, α was set as small as possible, while still maintaining reasonably high power, so that the probability of making a Type II error is not prohibitively high. Table 2 summarizes some calculations of power used in setting $\alpha = .01$ (Cohen, 1977, chapter 7). As Table 2 shows, the power available to pick up a moderate effect (.30) across all levels of α is very good (at least .93) for both the four- and six-state models. The power drops precipitously as the assumed effect size is made smaller. An effect size of .2 corresponds to a deviation of 10 percent per cell of the observed from the expected, using Cohen's formula (1977, p. 221), and an effect size of .3 corresponds to a 15 percent deviation. Setting $\alpha = .01$ or even to .001 will produce high power to detect an average 15 percent deviation per cell or larger. In general, such discrepancies are not present and stationarity is upheld.

The results for homogeneity are more clear-cut. Each of the 12 dyads is very different from the composite transition matrix by the X² test (Hewes,

TABLE 2
Comparison Among Critical and Observed Chi-Squared Values and Power Estimates for Stationarity
Analyses: Four- and Six-State Descriptions

<u>Four State Model</u>					
Critical X^2	α	df	Observed Significant/ Total	Assumed Effect+ Size	Power* (N=400)
32.9	.001	12	10/60	.20	.49
				.30	.965
26.2	.01	12	11/60	.20	.54
				.30	.97
21.0	.05	12	28/60	.20	.76
				.30	.99
<u>Six State Model</u>					
Critical X^2	α	df	Observed Significant/ Total	Assumed Effect+ Size	Power* (N=400)
42.3	.001	18	7/60	.20	.37
				.30	.93
34.8	.01	18	13/60	.20	.43
				.30	.94
28.9	.05	18	22/60	.20	.68
				.30	.99

*Power values at $\alpha = .001$ are not listed in Cohen (1977) so linear extrapolation was employed to estimate the power values at $\alpha = .001$; similarly $df = 18$ is not listed but $df = 16$ and 20 provide endpoints for a linear interpolation and Cohen quotes .30 as a "moderate" and .1 as a "small" effect size.

in press) ($p < .0001$ in all cases). The dyads cannot be pooled but must be analyzed separately. A slightly different set of homogeneity comparisons was carried out for the four-state matrices. Within occasion 1, the two male-male dyads were compared and the two female-female dyads were compared. Both comparisons were significantly differ-

ent from their expected matrices (X^2 ($df=12$) = 102.9 and 143.8, $p < .0001$). Similar results obtain for the mixed-sex occasion (two dyads) and for the same-sex occasion (three dyads). Thus, the heterogeneity among transition matrices gets no better (i.e., more homogeneous) as occasion or sex of partner is taken into account.

A final bit of evidence that the talk-silence process is basically Markovian is obtained from the distributions of sequences of states. These distributions should be exponentially decreasing in shape (Singer & Spilerman, 1974) so that the frequency of shorter strings of states is very high and that of longer strings is very low. Visual inspection indicates that all distributions are exponentially decreasing (Cappella & Streibel, in press).

The strictest test for any process supposed to be Markovian is its ability to predict. In testing the fit of these data to a Markov chain within a conversation, the initial transition matrix based upon the first two minutes of interaction, $P(1,k)$, was raised to the 10th power and that powered matrix compared to the observed matrix $\prod_{i=0}^9 P(t+i,k)$ for each conversa-

tion separately. Discrepancies between observed and predicted were calculated and average discrepancies obtained only after individual predictions were made. These discrepancies are reported in Tables 3 and 4. Two types of discrepancies are reported: the average algebraic and the average absolute discrepancy for each cell. Each index has its uses and drawbacks as a measure of the fit of observed to predicted. The average algebraic discrepancy takes the sign of the difference between predicted and observed into account. If there are any systematic discrepancies of over- or underprediction due to given rows or columns, the pattern of signs would show this. There do not appear to be such discrepancies in either the four- or six-state data. The average algebraic discrepancy does, however, tend to understate the size of the discrepancy in any one cell on the average since the positive and negative values tend to cancel one another. The average absolute discrepancy ignores the sign of the difference between predicted and observed and takes any discrepancy into account. However, in doing so, the average absolute discrepancy overstates the size of the discrepancy in any cell. The reason is that the transition matrix is constrained to have row sums equal to one. Consequently, if some cell in the predicted matrix is larger than the corresponding cell in the observed matrix, then some other cell in the same row of the predicted matrix must be smaller than the corresponding cell in the

TABLE 3
Average Algebraic and Average Absolute Discrepancies Between Markov Predicted and Observed Matrices: Four-State Description

	Discrepancy Matrix			
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
<u>1</u>	-.005*	-.006	-.018	.042
	.094**	.040	.055	.082
<u>2</u>	-.029	.000	-.014	.043
	.093	.046	.062	.086
<u>3</u>	-.021	-.005	-.010	.021
	.089	.042	.062	.060
<u>4</u>	-.026	-.006	-.021	-.053
	.094	.041	.053	.071

*The first entry in each cell is the average algebraic discrepancy.

**The second entry in each cell is the average absolute discrepancy.

observed matrix. The absolute discrepancy counts both discrepancies into its sum without any correction for this inherent upward bias in any row. A more valid index of discrepancy on a cell-by-cell basis is probably one which lies between the average algebraic and average absolute discrepancies.

Nevertheless, it seems clear from Tables 3 and 4 that the fit of the Markov-predicted matrices to the observed matrices 20 minutes ahead in time is very tight. In the four-state case, the largest absolute discrepancy is .094 and in the six-state data it is .102. The average cell in the four-state case has an absolute discrepancy of .067 and in the six-state case it has an absolute discrepancy of .051. Of course, these averages are somewhat inflated.

Despite this obviously close fit in the transition matrices, it is useful to obtain a statistical measure of fit across time under Markov assumptions. Such

TABLE 4
Average Algebraic and Average Absolute Discrepancies Between Markov Predicted and Observed Matrices:
Six-State Description

		<u>Discrepancy Matrix</u>					
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
<u>1</u>	-	.012*	-.004	.002	.004	.029	-.004
	.	.086**	.068	.030	.047	.067	.021
<u>2</u>	-	.016	-.021	.001	.004	.031	-.000
	.	.083	.071	.031	.042	.065	.017
<u>3</u>	-	.006	.005	.002	-.008	.031	-.003
	.	.055	.060	.019	.042	.083	.026
<u>4</u>	-	.040	-.032	-.005	.027	.073	.000
	.	.057	.049	.015	.060	.096	.029
<u>5</u>	-	.035	-.028	-.004	.031	.053	.000
	.	.049	.048	.014	.061	.102	.032
<u>6</u>	-	.041	-.018	-.002	.017	.062	.010
	.	.032	.066	.019	.053	.096	.037

*The first entry in each cell is the average algebraic discrepancy.

**The second entry in each cell is the average absolute discrepancy.

a measure is obtained by calculating a X^2 measure of fit of predicted to observed data.² The measure is simply the squared discrepancies between predicted and observed transition matrices at the 10th time unit, divided by the predicted, weighted by the expected row frequency, and summed across all 16 or 36 cells.

In the four-state case, six of 12 X^2 measures do not reach statistical significance at $\alpha < .01$ ($df = 12$). In the six-state case, six of 12 X^2 measures do not achieve significance at $\alpha < .01$ ($df = 30$). These dyads are not the same in both cases. In both the

four-state and six-state cases, two dyads (dyads four and 12 in both cases) have predicted matrices based upon an initial matrix, which is grossly discrepant from the composite (see note 2). They produce very large X^2 measures of discrepancy from the predicted transition matrix. If these two dyads are not considered, then the average X^2 values are 22.61 in the four-state and 51.06 in the six-state cases. The former is not statistically significant at $\alpha = .01$ ($df = 12$), while the latter is borderline significant at $\alpha = .01$ ($df = 30$). Given the size of the data set for any one dyad, and given that we have not

taken advantage of any of the tested assumptions (see note 3), the overall fit of both the four- and six-state descriptions to a Markov Chain must be viewed as remarkable.

Interestingly, the X^2 per degree of freedom for the four- and six-state predictions across all 12 dyads is 3.77 and 2.36, respectively, showing a slightly better, but not statistically significant ($F=1.60$, $df=12,30$), fit for the six-state description.

Granting that the sequence of talk and silence is stationary and Markovian within a conversation, does this consistency in individual parameters in the conversation extend across conversations with different partners?

The Independent Decision Model

The ID model tests this assumption by using the composite individual parameters, $p_1^*(T,k,A)$ and $p_1^*(T,k,B)$, from one conversation to predict the composite dyadic transition matrix for a conversation which has not yet occurred. In the present study, individual composite parameters from the first conversation were used to predict dyadic transition matrices in the second and third conversations and, similarly, for predicting from the second to the third conversations. Backward predictions are compatible with the assumptions of the Independent Decision model, but none were undertaken. Twelve conversations were predicted in this way. When two different predictions of the same conversation were available from independent data sets (as, for example, in predicting the third conversation), both sets were included. Discrepancies between predicted and observed data are presented in Tables 5 and 6 for four- and six-state matrices.

As with the within-occasion predictions, both the average algebraic and average absolute discrepancies are provided. The largest absolute discrepancies are .118 in the four-state case and .128 in the six-state case. The average absolute discrepancy per cell is .062 in the four-state case and .065 in the six-state case. These figures are about the same or a little worse than the within-occasion discrepancies. However, one must be struck by the degree of fit between occasions since what is being predicted is

TABLE 5
Average Algebraic and Average Absolute
Discrepancies Between Observed and
"Between-Conversation" Predicted Matrices:
Four-State ID Model

	Discrepancy Matrix			
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
<u>1</u>	.026*	.013	-.057	.018
	.103**	.014	.084	.046
<u>2</u>	.016	-.018	.014	-.012
	.053	.072	.067	.060
<u>3</u>	.005	.011	-.014	-.003
	.118	.030	.101	.010
<u>4</u>	.100	-.019	-.011	-.072
	.106	.027	.013	.092

*The first entry in each cell is the average algebraic discrepancy.

**The second entry in each cell is the average absolute discrepancy.

the composite structure of a nonexistent conversation from individual components derived from other conversations at least one week prior and with a different partner.

A more sensitive test permitting a more direct comparison between the within- and between-occasion models can be achieved as follows:

1. Since the within-occasion process is stationary, a prediction of the composite matrix from the Independent Decision model is also a legitimate prediction of the initial transition matrix $P_{ID}(1,k)$ where the subscript "ID" indicates that the matrix has been predicted by the ID model.
2. This matrix raised to the 10th power can be the predicted matrix for a X^2 test of fit to the observed within-conversation transition matrix.
3. X^2 values calculated between the predicted mat-

TABLE 6
Average Algebraic and Average Absolute Discrepancies Between Observed and "Between-Conversation"
Predicted Matrices: Six-State ID Model

	Discrepancy Matrix					
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
<u>1</u>	.036*	-.040	.006	0 ⁺	-.002	0
	.110**	.078	.011	0	.030	0
<u>2</u>	.014	-.007	.000	0	-.006	0
	.114	.087	.028	0	.012	0
<u>3</u>	.024	.023	-.048	0	.001	0
	.059	.086	.054	0	.051	0
<u>4</u>	0	-.055	0	.016	.031	.008
	0	.078	0	.093	.058	.011
<u>5</u>	0	-.007	0	.137	-.111	-.019
	0	.012	0	.128	.113	.029
<u>6</u>	0	.035	0	.034	.022	-.046
	0	.067	0	.066	.071	.105

*The first entry in each cell is the average algebraic discrepancy.

**The second entry in each cell is the average absolute discrepancy.

+The single digit zeroes are zero by definition of the appropriate transitions.

rix from the Independent Decision model and the observed transition matrices can be compared to the X^2 s obtained in the previous section from the simple within-conversation Markov predictions.

4. If there is no difference, then there is reason to believe that there is between-occasion consistency.

We find that the average X^2 s across the 12 predicted conversations derived from the Independent

Decision model to be 174.2 ($14.5=X^2/df$) for the four-state case and 249.4 ($8.31=X^2/df$) for the six-state case. These values are much greater than the average X^2 values of 67.5 ($5.62=X^2/df$) and 74.89 ($2.50=X^2/df$) obtained with the within-occasion, four- and six-state Markov predictions.³ The differences are, in fact, statistically significant in the six-state case ($F=3.32$, $df=30,30$, $p<.01$) and close to significance in the four-state case ($F=2.58$, $df=12,12$, $p<.07$). It is clear that using the Inde-

pendent Decision model to predict the initial transition matrix for powering under the Markov assumption is predictively less adequate than using the initial observed transition matrix for powering. The implication follows that individuals are more consistent within a conversation than across conversations with a different partner. The questions to be raised next are: Can we determine what types of factors the inconsistency across conversations is due to? Can we establish which, if any, of the 6 six-state and 4 four-state parameters are most responsible for the observed across-conversation inconsistency?

The Incremental Model: A Predictive Version

The previous results establish two important conclusions for the Incremental model. First, the finding of stationarity within a conversation suggests that the individual parameters are stable within that conversation. Second, the finding that the predictions of the Independent Decision model across conversations are less accurate than the within-conversation Markov model implies that differences in the individual parameters between conversations ought to be investigated. This investigation may be described as follows: variation in individual parameters occurs both between and within conversations, although the within-conversation variability should not be appreciable, as the stationarity finding shows. According to the Incremental model, this variation may be due to individual predispositions, fixed partner influences (occasion differences), or sequential or practice effects as the experiment progresses. The purpose is to account for as much variation as possible in each of the 10 individual parameters (6 six-state and 4 four-state),⁴ with dummy variable surrogates for individual differences, partner or occasion effects, and sequence or practice effects. The results are presented first for a person-by-person analysis across time and second for a pooled data set of persons and time. For convenience, the 6 six-state parameters will be labeled SIX-1, SIX-2, etc. and the four-state parameters as FOUR-1, FOUR-2, etc. (see note 4).

Individual-level results. These analyses seek to account for variation in each individual's parameters

(i.e., SIX-1, SIX-2, . . . FOUR-2) as a function of dummy variables representing the occasion of each conversation,⁵ an integer variable for time within each conversation and their interactions. These independent variables will uncover all linear and non-linear differences in the dependent variables across conversations and any linear trends in the parameters within a conversation.

Consider first the linear effects of time and time-by-occasion interactions. When all the independent predictors remained in the equations, only nine of 184 possible tests were significant at $\alpha = .05$. Using a stepwise procedure with forward selection and an $\alpha = .05$ inclusion criterion, only 12 of 184 possible time and time-by-occasion effects remained in the final regressions. As expected, there are essentially no linear effects of time on any of the eight individual parameters. On the other hand, several of the between-occasion contrasts were significantly different in both analyses, with SIX-2 exhibiting the strongest shifting from conversation to conversation.

A second and more careful analysis of the between-occasion effects alone was undertaken. Since we are performing regression analyses on time-series probability data, it is wise to check the data for violations of the homogeneity of variance and serial correlation assumptions. The previous results about time effects are not invalid since the usual impact of heterogeneous variance and serial correlation is to produce excessively liberal rather than conservative hypothesis tests (Kmenta, 1971, p. 256, 282).

The eight dependent variables were first checked for serial correlation using the Durbin-Watson test (Kmenta, 1971, p. 294). Only three of 64 (eight dependent variables \times eight persons) tests were significant at $\alpha < .05$ (two-tailed test; $df = 1, 20$ or $2, 30$ or $3, 40$ depending on whether the subject had two, three, or four conversations), 11 were inconclusive, and 58 were not significant. Clearly, each person is free of serial correlation on each dependent variable.

The variance in each dependent variable for each person across conversations did show considerable heterogeneity by a maximum likelihood test (Kmenta, 1971, p. 268), with 25 of 64 (the subject

TABLE 7
Significance of Partner Switching on Each of Eight Probabilities for Each Subject: Four- and Six-State Descriptions

	Dependent Variable (Individual Probability Parameters)							
	SIX-1	SIX-2	SIX-3	SIX-4	SIX-5	SIX-6	FOUR-1	FOUR-2
Subj 1 df=3,36	.55 (.0000)	.85 (.0000)	.17 (N.S.)	.51 (.0000)	.42 (.0002)	.21 (.036)	.61 (.0000)	.26 (.011)
Subj 2 df=2,27	.29 (.009)	.58 (.0000)	.04 (N.S.)	.35 (.003)	.05 (N.S.)	.41 (.001)	.41 (.001)	.33 (.004)
Subj 3 df=2,27	.12 (N.S.)	.27 (.013)	.06 (N.S.)	.28 (.010)	.01 (N.S.)	.01 (N.S.)	.11 (N.S.)	.02 (N.S.)
Subj 4 df=2,27	.03 (N.S.)	.32 (.006)	.21 (.041)	.18 (N.S.)	.09 (N.S.)	.02 (N.S.)	.16 (N.S.)	.13 (N.S.)
Subj 5 df=2,27	.05 (N.S.)	.66 (.0000)	.28 (.012)	.33 (.004)	.37 (.002)	.21 (.04)	.31 (.006)	.27 (.01)
Subj 6 df=2,27	.26 (.018)	.06 (N.S.)	.18 (N.S.)	.33 (.004)	.16 (N.S.)	.05 (N.S.)	.32 (.006)	.18 (N.S.)
Subj 7 df=1,18	.05 (N.S.)	.35 (.006)	.00 (N.S.)	.01 (N.S.)	.01 (N.S.)	.09 (N.S.)	.01 (N.S.)	.08 (N.S.)
Subj 8 df=1,18	.01 (N.S.)	.62 (.0000)	.04 (N.S.)	.00 (N.S.)	.10 (N.S.)	.03 (N.S.)	.01 (N.S.)	.00 (N.S.)
Ave. R^2	.169	.465	.122	.251	.151	.128	.241	.160
#Overall Significant/ Total	3/8	7/8	2/8	5/8	2/8	3/8	4/8	3/8

with one conversation could not be tested) comparisons significant at $\alpha < .05$. The ratios of the largest to the smallest variances were in many cases too large to simply be ignored. A number of transformations normally applied to heterogeneous probability data were tried (Krumbein, 1957; Bartlett, 1971; Winer, 1962, pp. 400-01), and the ARCSIN $\sqrt{P_i} \pm .05$ was found to reduce the number of significant heterogeneous variances from 25 to 10 and, more importantly, to reduce the ratios of largest to smallest variances to tolerable limits. While some heterogeneity remains, it should not be

so severe as to seriously bias the reported significance tests (Boneau, 1971; Cochran, 1971). The regressions reported hereafter have all dependent variables transformed by the ARCSIN-square root transformation above.

The between-occasion differences are completely confounded with changing partner since no person interacts with the same partner on subsequent occasions. Table 7 summarizes the occasion (or partner) effects for each subject and each dependent variable. Overall, the greatest effect on individual parameters due to partner (or occasion)

TABLE 8
Effects of Person, Partner, Occasion, and Person-By-Occasion Interaction on Each of Eight Probabilities Pooled Across Time and Subject: Four- and Six-State Descriptions

Variable	Overall R^2 (df=34, 205)	Person (df=8, 231)	Partner (df=8, 223)	Occasion (df=1, 222)	Person x Occasion (df=17, 205)	Ignoring R^2 Increments due to Insignificant Incre- ments then R^2
SIX-1	41.7% (.0000)	23.9% (.0000)	14.0% (.0000)	0% N.S.	3.6% N.S.	38.1%
SIX-2	63.2% (.0000)	17.8% (.0000)	36.6% (.0000)	0% N.S.	8.4% (.0004)	63.2%
SIX-3	18.5% N.S.	6.9% (.033)	8.9% (.004)	0% N.S.	2.7% N.S.	15.8%
SIX-4	35.8% (.0000)	13.7% (.0000)	15.6% (.0000)	1% N.S.	5.6% N.S.	30.2%
SIX-5	23.6% (.0047)	9.8% (.002)	5.8% N.S.	2% .019	5.9% N.S.	17.7%
SIX-6	21.2% (.022)	10.4% (.001)	7.3% (.014)	0% N.S.	3.4% N.S.	17.8%
FOUR-1	50.7% (.0000)	24.0% (.0000)	18.7% (.0000)	6% N.S.	7.4% .028	43.3%
FOUR-2	33.8% (.0000)	12.9% (.0001)	13.6% (.0000)	0% N.S.	7.3% N.S.	26.5%

factors is found in SIX-2, SIX-4, and FOUR-1, which exhibit the largest average variance explained and the largest ratio of overall significant R^2 s per subject. This pattern is strengthened when one focuses upon the number of between-occasion contrasts which are statistically significant at $\alpha < .05$ when the overall F for a particular equation is significant. For SIX-2, 12 of 15 contrasts are significant, for SIX-4, seven of 15 contrasts are significant. The other dependent variables exhibit a maximum of three of 15 contrasts significant.

The individual parameters most strongly affected by partners then are SIX-2 (the probability that A will talk given that A has the floor and is talking and B is silent), SIX-4 (the probability that A will talk given that B has the floor and is silent and A is silent), and FOUR-1 (the probability that A will talk given that A and B are silent). However, since there seem to be differences among persons in these effects and since partner and occasion are completed confounded, it is desirable to pool the time-series data across subjects to determine (1) how much variation in the eight individual parameters is due to individual differences and (2) if there is a true in-

crement in probability due to partner effects rather than occasions.

Cross-sectional plus time-series results. The data for nine subjects was pooled in a cross-sectional plus time-series data set (Simonton, 1977; Kmenta, 1971, chapter 12) with an $N=240$. Several different patterns of dummy variables were regressed on each of the eight dependent variables in a stepwise regression procedure under several different assumptions about order of entry. Since there were no appreciable differences among these alternatives, only one is presented here.

Each of the eight individual parameters was regressed on a set of eight person dummies, eight partner dummies, three occasion dummies, and 24 person-by-occasion interaction dummies with the four groups entered in the order listed. The results are presented in Table 8. In seven of eight cases, significant portions of the variance in individual parameters are accounted for by the set of predictors.

The occasion dummies account for none of the variation. Even when the occasion dummies are entered first in the stepwise hierarchy, they account

for essentially zero variation. This result is as it should be since positive increments from one conversation to the next are likely to be balanced by decrements when one pools across subjects. Similarly, the person-by-occasion interaction group adds a statistically significant amount to the variance only in the case of SIX-2 ($F=2.76$, $df=17,200$, $p<.01$) (Kerlinger & Pedhazur, 1973, p.70). All other increments are not significant with $\alpha>.05$. The final column in Table 8 presents the overall R^2 values with the statistically nonsignificant contributions deleted.

As in the individual regressions, SIX-2, SIX-4, and FOUR-1 show the greatest variance explained. In this case, the variance is attributable to the presence of particular partners separate from occasion and to individual differences among speakers. Contrary to the individual analyses, SIX-1 and FOUR-2 show that appreciable amounts of their variance can be attributed to partners and subjects. Only SIX-3, SIX-5, and SIX-6 have relatively paltry amounts of their variance explained by the 16 independent predictors (about one percent of the variance per predictor).

It is interesting to note that, among the five variables which have large portions of variance accounted for, only SIX-2 has the largest share of this variance attributable to nonperson variables (despite not being entered first). The others have about one-half or more of their explained variance accounted for by individual differences.

Summary. The individual and pooled results from the Incremental model permit some general and some particular conclusions. Generally, the individual parameters show no linear trends in conversations but do show shifts from conversation to conversation, especially in the probability of breaking mutual silences (SIX-1, SIX-4, and FOUR-1) and in the probability of continued talking (SIX-2, which is identical with FOUR-3). On the other hand, the probability of breaking or continuing simultaneous talk (SIX-3, SIX-6, and FOUR-2) and the probability of talking when the other has the floor (SIX-5, which is identical to FOUR-4) show consistencies across conversations.

More particularly, the above inconsistencies

seem to be due to the effects of changing partners rather than merely changing conversations. When there are differences in individual parameters due to changing partners, there is also considerable individual difference in the amount of such change. However, only SIX-2 exhibits a significant interaction between person and partner effects, implying that it is something about the dyad, in addition to the partner and the person, which has effected a change.

CONCLUSIONS

Let us review what we have found in representing the sequence of talk and silence and interpret these findings in light of the research reviewed earlier.

Within-Conversation Sequences

A first-order Markov Chain model was proposed to account for within-conversation talk and silence sequences for each dyad. This model was shown to fit observed talk and silence sequences. Substantively, fit requires that the transition matrices are stationary, which in turn implies that the dyad's probabilities of changing states are consistent over the conversation. In addition, since the individual parameters are obtained from the dyadic transition probabilities (Equations 4-9 in Cappella, 1979), then dyadic consistency implies individual consistency within the conversation.

The literature on individual talk and silence consistency reviewed in the earlier report leads to the expectation of individual consistency in the present data. However, the present data are an even stronger claim to individual consistency since they show consistency across two-minute intervals in conversation, whereas previous data showed consistency between segments of conversation which were aggregated to create stable mean response levels.

While it is difficult to explain how the above regularity arises, the most likely alternatives are biological or cognitive (Jaffe, 1977) and personality (Lustig, 1977; McCroskey, 1977). Whatever the reason, this strong consistency is an important structural baseline for conversations, but not one which is unmodifiable.

Between-Conversation Sequences

The attempt to account for conversational sequences as persons switch partners is really nothing more than the question of how to get dyads from individuals. I take this question to be a fundamental question in any theory of two-person interaction. It is a simple matter to characterize individual or dyadic interaction alone, but it is no small matter to describe how dyadic interaction processes derive from the styles of the individuals who make up the dyad.

Following the lead of Jaffe and Feldstein (1970), three answers are given to this question in the form of the Independent Decision, Incremental, and Regulation models. The ID model boldly states that dyadic transitions are derived from individual probabilities by assuming (1) that individuals are perfectly consistent across partners and (2) that they act independently of one another in choosing to talk or be silent.

Despite its assumptions, which are contrary not only to the available evidence but also to common sense, the ID model shows some predictive utility. Predictions from one conversation to a future one show an average absolute discrepancy of about .06 per cell in the transition probability matrix. While its predictive utility is not as strong as the within-occasion Markov model, one must be impressed with the ability of relatively weak and simplistic assumptions to predict the structure of a nonexistent dyad from individuals' previous behavior.

The ID model provides a baseline set of predictions for generating dyads from individuals which must be exceeded by subsequent models of greater complexity. If the Incremental and Regulation models cannot improve the predictions of the ID model, then they must fall under the force of Occam's razor, despite their more realistic assumptions.

The Incremental model tested here relaxes the consistency-across-partners assumption which previous literature on talk and silence would lead one to label as suspect. On the one hand, the Incremental model finds certain individual parameters to be relatively consistent (i.e., not systematically different

from partner to partner) despite partner switches. These are the probabilities of breaking or continuing simultaneous speech (SIX-3 and SIX-6) and the probability of talking when the other is and has the floor (SIX-5). On the other hand, SIX-1, SIX-4, and SIX-2 do "increment" or change as partners change. The probabilities of continuing or breaking mutual silences (SIX-1 and SIX-4) show individual differences as well as partner effects, but no interactions between partner and person factors. The probability of continuing the floor, given that you have it and the other is silent (SIX-2), shows individual differences, strong partner influences, and interaction between person and partner effects.

Together the above results suggest that the consistency-across-partners assumption (that is, $p_i^*(T,k,A) = p_i^*(T,k+1,A)$ for all k and i) should be replaced by the following tentative predictive equations:

$$p_1^*(T,k+1,A) = p_1^*(T,k,A) + \sum \beta_{1f} O_j + \sum \lambda_{1f} P_j \quad (1)$$

$$p_2^*(T,k+1,A) = p_2^*(T,k,A) + \sum \beta_{2f} O_j + \sum \lambda_{2f} P_j + \sum \delta_{2f} (O_j \times P_j) \quad (2)$$

$$p_3^*(T,k+1,A) = p_3^*(T,k,A) + \sum \lambda_{3f} P_j \quad (3)$$

$$p_4^*(T,k+1,A) = p_4^*(T,k,A) + \sum \beta_{4f} O_j + \sum \lambda_{4f} P_j \quad (4)$$

$$p_5^*(T,k+1,A) = p_5^*(T,k,A) + \sum \lambda_{5f} P_j \quad (5)$$

$$p_6^*(T,k+1,A) = p_6^*(T,k,A) + \sum \lambda_{6f} P_j \quad (6)$$

where the β s, λ s, δ s are regression weights, the O_j are exogenous partner variables, and the P_j are person differences. These equations make SIX-3, SIX-5, and SIX-6 consistent across partners, as the data requires, while permitting individual differences. SIX-1, SIX-4, and SIX-2 incorporate partner effects as well as individual differences, while only SIX-2 involves an interaction between person and partner effects.

In retrospect, one should have expected the individual probabilities SIX-1, SIX-2, and SIX-4 to exhibit the greatest between-conversation differences since these probabilities primarily account for pauses, switching pauses, and vocalization length.

The research reviewed in the earlier report showed these aspects of conversation to be susceptible to modification through factors like perceived partner warmth, anxiety, ambiguity, and cognitive load, as well as influence from the partners' own levels of response on these behaviors. The results of the Incremental model show by an indirect route that the probabilities most responsible for pause, switching pause, and vocalization duration are subject to the influences expected.

Future Directions

The next step in describing across-conversation sequences is to test the independence assumption of the ID model by testing for the mutual influence posited by the Regulation model. If the Regulation model is strongly upheld, then the ID model and the within-conversation Markov model would be opened to serious question. What is more likely is that the mutual influence between partners, if any, will be small, helping to account for residual variation within and between conversations.

Secondly, substantive predictor variables must be incorporated in Equations 1-6 above. Candidates for inclusion have been suggested in the previous literature review. Without this step, one cannot test the predictive adequacy of the Incremental model relative to the ID model. Dummy variable surrogates provide useful guides to which types of predictors are necessary, but the predictors themselves are required if comparisons are to be made. Research underway in our laboratory is focusing on personality variables such as dominance, affiliation, and self-monitoring and interpersonal factors such as attraction, previous task satisfaction, and perception of the other's personality.

Finally, the models presented and tested in these papers are actually quite general. They are appropriate to any two-person interaction system with categorical states. More importantly, they suggest alternative procedures through which dyadic interaction structure might be derived from individual behaviors. Until our research can answer this question, studies of interaction process must continue to make the unnatural choice between studying individual behaviors or studying dyadic (or group) behaviors.

NOTES

The author wishes to thank Joseph Folger, Daniel Fogel, and Michael Streibel for their assistance in data analysis and gathering, Sally Planalp for reading and critiquing a version of the manuscript, and Dean Hewes for invaluable statistical consultation. Unfortunately, I cannot blame them for the final product. The research was made possible through a grant of the Graduate Research Committee of the University of Wisconsin-Madison.

1. The types of data reduction and summarization are described in Cappella and Streibel (in press). The programs are available from the author.
2. There is a significant difference between "true prediction" and "ordinary prediction" in Markov analysis. When the observed initial transition matrix is raised to the n th power and compared to the observed, then whatever statistical fluctuations exist in the data are carried through the predictions. Consequently, fit is less good. The alternative is to test the assumptions and, finding that they are satisfied, use the *composite* matrix as the best estimate of the initial transition matrix. When the composite is so used, fluctuations in the data are smoothed, fit is better, but true prediction is lost. While we have tested the assumptions, we have also opted for true prediction, which is the more conservative predictive test. I am indebted to Dean Hewes for pointing out the above distinction.
3. Note that all conversations were predicted using the ID model. Only dyads 5, 6, 7, 9, 10, 11, and 12 could be so predicted, since backward predictions were disallowed. The X^2 values reported in this section for the within-conversation Markov predictions are obtained only from the above dyads.
4. Actually there are only 8 independent parameters, since p_3^* and p_5^* for the four-state description and p_2^* and p_5^* for the six-state description are identical by definition. The data presented will list the 6 six-state and first 2 four-state parameters as dependent variables.
5. The dummy scheme identified the first occasion for each person as the comparison. All contrasts then are contrasts from later conversations to the first conversation. Which coding of independent variables one uses is irrelevant when the focus is primarily on explained variance.

REFERENCES

- BARTLETT, M.S. The use of transformations. In J.A. Steger (Ed.), *Readings in statistics*. New York: Holt, Rinehart, & Winston, 1971.
- BONEAU, C.A. The effects of violations of assumptions un-

- derlying the *t* test. In J.A. Steger (Ed.), *Readings in statistics*. New York: Holt, Rinehart, & Winston, 1971.
- CAPPELLA, J.M. Talk-silence sequences in informal conversations I. *Human Communication Research*, 1979, 6, 3-17.
- CAPPELLA, J.N., & STREIBEL, M. Computer analysis of talk-silence sequences: The FIASSCO system. *Behavioral Research Methods and Instrumentation*, in press.
- COCHRAN, W.G. Some consequences when the assumptions for the analysis of variance are not satisfied. In J.A. Steger (Ed.), *Readings in statistics*. New York: Holt, Rinehart, & Winston, 1971.
- COHEN, J. *Statistical power analysis for the behavioral sciences*. 2nd edition. New York: Academic Press, 1977.
- FELDSTEIN, S., JAFFE, J., & CASSOTTA, L. Mathematically predicted time patterns of dialogue. Paper presented at the meeting of the Eastern Psychological Association, New York, 1966.
- FELDSTEIN, S., JAFFE, J., & CASSOTTA, L. The effect of mutual visual access upon conversational time patterns. *American Psychologist*, 1967, 23, 594.
- HAYES, D.P., MELTZER, L., & WOLF, G. Substantive conclusions are dependent upon techniques of measurement. *Behavioral Science*, 1970, 15, 265-268.
- HEWES, D.E. Stochastic models of communication processes. In P.R. Monge and J.N. Cappella (Eds.), *Multivariate techniques in communication research*. New York: Academic Press, in press.
- JAFFE, J. The biological significance of Markovian communication rhythms. In S. Rosenberg (Ed.), *Sentence production*. New York: Lawrence Erlbaum, 1977.
- JAFFE, J., & FELDSTEIN, S. *Rhythms of dialogue*. New York: Academic Press, 1970.
- KERLINGER, F.N., & PEDHAZUR, E.J. *Multiple regression in behavioral research*. New York: Holt, Rinehart, & Winston, 1973.
- KMENTA, J. *Elements of econometrics*. New York: MacMillan, 1971.
- KRUMBEIN, W.C. Comparison of percentage and ratio data in facies mapping. *Journal of Sedimentary Petrology*, 1957, 27, 293-297.
- LUSTIG, M.W. The relationship between communicative predispositions and sound-silence patterns in triads. Unpublished paper presented at the Speech Communication Association Convention, Washington, D.C., 1977.
- MCCROSKEY, J.C. Oral communication apprehension: A summary of recent theory and research. *Human Communication Research*, 1977, 4, 78-96.
- SIMONTON, D.K. Cross-sectional time-series experiments: Some suggested statistical analyses. *Psychological Bulletin*, 1977, 84, 489-502.
- SINGER, B., & SPILERMAN, S. Social mobility models for heterogeneous populations. In H.L. Costner (Ed.), *Sociological methodology 1973-74*. San Francisco: Jossey-Bass, 1974.
- WINER, B.J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.