Effects of Transmission Delay on Conversational Behavior on Echo-Free Telephone Circuits

By PAUL T. BRADY

(Manuscript received August 25, 1970)

The present study sought to determine the effects of echo-free transmission delays of 600 ms and 1200 ms on conversational behavior when subjects were unaware that delay was present. Sixteen pairs of male subjects conversed 10 minutes each over an echo-free telephone circuit with zero delay and 10 minutes each with 600 ms round-trip delay; a second group of 16 pairs conversed 10 minutes each on zero delay and 1200 ms round-trip delay. Subjects noticed nothing unusual about the circuit, but the delay caused a statistically significant increase in frequency of confusions and in amount of both double talking (simultaneous speech from both speakers) and mutual silence. Analysis by means of an on-off pattern generating model revealed that subjects seem to make some adjustments in their speaking behavior; they tend to wait longer for responses and keep talking longer when interrupted. The two delay values of 600 ms and 1200 ms produced virtually identical effects in the measures made here. All of the subjects in the study were inexperienced with delay circuits; this experiment does not address itself to the possibility that experienced subjects might react differently to the delay. Further work will investigate the effects of training subjects to notice delay.

I. INTRODUCTION

1.1 *Purpose of Study*

This paper describes experiments that examined effects on the conversational behavior of subjects who talked on telephone circuits containing round-trip transmission delays of 600 ms and 1200 ms, typical of those in one- and two-hop synchronous satellites. However, satellite circuits also contain echoes and echo suppressors, which create added difficulties in conversation. The present study examines behavior on echo-free circuits, which presumably represent the circuits that would result if echoes could be eliminated from existing circuits. In particular, answers are sought to the following questions:

(*i*) Does the introduction of delay on a circuit produce any measurable effects on speaking behavior, even though the subjects notice nothing unusual about the circuit?

(*ii*) Do the subjects experience difficulty in talking over delay circuits, even though they do not notice the difficulty?

This paper reports quantitative measures of the effects of delay indicating that the answer to both of the above questions is "yes."

1.2 *Background*

For several years, studies have been directed toward measuring customer acceptability of circuits containing both transmission delay and echo suppressors. For example, these circuits were used in one study,[1] in which hundreds of customer interviews were taken just after the customers placed transatlantic calls over such a circuit. A delay of 600 ms (as opposed to the cable transmission round-trip delay of 90 ms) produced a significant increase in the number of customers who reported having difficulty, and in addition produced an increase in other complaints and comments such as "cutting on and off," and "fading." However, such customer interview studies have limited application to echo-free delay circuits because (*i*) the echo suppressors are widely believed to contribute most of the degradation and probably overwhelm the delay effects, and (*ii*) as will be shown here, many of the delay effects are subtle, unnoticed by the conversants, and hence are not reported in interviews.

Regarding delay on 4-wire* echo-free circuits, only a few studies have been conducted. These have been limited to laboratory experiments because of the difficulty of providing 4-wire service to customers in the field.[2] P. D. Bricker[3] reported a laboratory study in which subjects who were talking over a standard circuit were suddenly presented with delay, introduced during conversation but not so as to disrupt a speech burst and provide an artificial cue. Subjects were told that delay would be introduced and were asked to try to detect its presence. It commonly took more than one minute just to detect a 1.2 s round-trip delay. R. M. Krauss and Bricker[4] had subjects con-

* The terms "pure delay," "echo-free delay," and "4-wire" are equivalent terms in this study, and refer to delay circuits in which echo is prohibited by using two separate voice paths, one from speaker *A* to *B* and one from *B* to *A*. Further, each path is connected through at all times so that full duplex communication is allowed.

verse and solve puzzles over a circuit with round-trip delays of zero, 600 ms, and 1800 ms. Measurements of word count and subjective opinions showed no significant differences between the zero and 600 ms delays, but the 1800 ms delay caused an increase in reports of "difficulty in communicating due to the circuit."

E. T. Klemmer[5] summarized some experiments conducted at Bell Laboratories in which certain people used 4-wire delay circuits in their regular telephone calls to each other, and could reject the circuit if they considered it unsatisfactory. They were not told the nature of the experimental circuit. Klemmer found that users were very seldom disturbed by delays of 600 ms and 1200 ms.

The previous studies seem to indicate that subjects are generally unaware of and undisturbed by round-trip delays up to 1200 ms. However, by its nature the delay does introduce changes in the temporal patterns of two-way conversation. The present study sought to determine the effects of these changes by comparing many measurements, especially those related to temporal patterns, on conversations on both standard and delay circuits.

II. THE BASIC MECHANISM OF DELAY INTERFERENCE

One might guess that the most obvious effect of introducing delay into a transmission link would be to make the conversation more sluggish; for example, responses to questions are delayed by the round-trip delay value. In fact, this effect is rarely noticed at 1200 ms and is essentially never reported at 600 ms.

There is another delay effect that is far more important at 600 ms and 1200 ms delay. If speaker *A* momentarily pauses, especially in such a way as to suggest a response from speaker *B*, *A* will not receive this response for the normal speaker response delay plus an additional 600 ms or 1200 ms. In the meantime, *A*, not hearing *B*'s response at the expected time, may proceed on the assumption that *B* will not respond. Then, *B*'s response will arrive and interrupt *A*'s new talkspurt, thus causing an interference which at times the author has observed to momentarily disrupt normal conversation. D. L. Richards[6] has termed this an *involuntary interruption*, since from *B*'s point of view, he never interrupted *A*.

Sometimes, during a lull in a conversation, both speakers begin to talk at approximately the same time, and each person will have generated several syllables or even words before he is conscious of the other person's delayed speech. This can cause confusion.

The principal deleterious effect of delay is, therefore, to cause speech from one conversant to arrive at the other at unexpected times. This effect causes serious problems when echo suppressors are included,[7] but even on echo-free circuits the added difficulties are readily apparent, as is shown in Section IV.

III. PROCEDURE

3.1 *Experimental Circuit and Test Conditions*

Figure 1 is a diagram of the experimental circuit, which is designed to represent a standard long distance telephone circuit. The telephones were in two separate sound treated rooms; the experimenter could insert a 600 ms or 1200 ms delay from speaker B to speaker A.

It is well established, but not intuitively obvious, that the total round-trip delay can be distributed in any desired manner on the two transmission paths (A to B and B to A), as long as the total delay remains the same. Thus, the Fig. 1 configuration is identical to the subjects as a circuit in which half the delay is in each path. This point is explained in a previous paper.[8]

Sixteen pairs of male Bell Laboratories employees conversed ten minutes on the standard circuit and ten minutes on the 600 ms delay circuit, and 16 different pairs conversed on the standard and 1200 ms delay circuits for ten minutes each. In all, there were 32 different pairs of men. (Women were not used to avoid lengthening the experiment by adding a new variable.) Each pair consisted of good friends with mutual interests; the author has found that such pairs converse readily. The subjects were told to talk for 20 minutes on any topics they

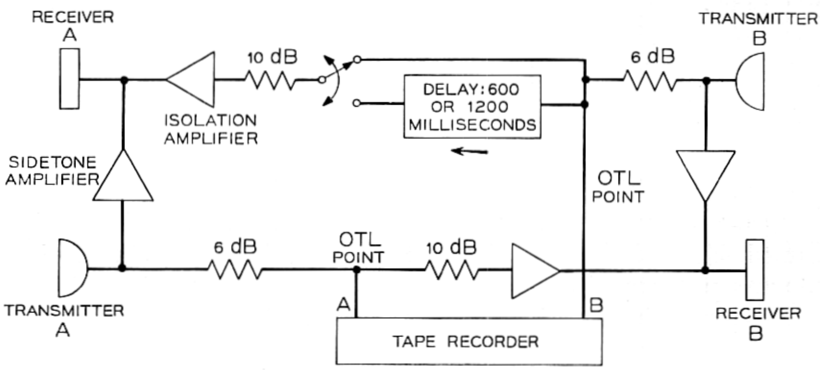

Fig. 1—Circuit over which subjects talked.

wished, but that they both should converse; monologues would be unacceptable. They were specificially instructed not to discuss the circuit, and that our interest was centered in recording a conversation for computer analysis of conversational behavior. They were told nothing about delay. Half the pairs began on the standard circuit, and half began on delay. After 10 minutes the circuit was changed, but in such a way to be undetectable to the subjects. (The delay from *B* to *A* was added or removed while *A* was talking.)

To obtain subjects inexperienced with delay, the subjects were chosen from areas of Bell Laboratories that had not been concerned with transmission delay. It was not possible to ask the subjects about delay prior to the experiment, as this would disclose the nature of the experiment. On interviews afterwards, one subject stated that he was familiar with delay circuits. His conversation was replaced with one that involved a new pair of subjects.

3.2 *Measurements Made on the Calls*

The following five measures were made on the calls.

3.2.1 *Subjective Measurements*

3.2.1.1 *Opinion of Subjects.* Immediately on completion of the call, the subjects were asked if they noticed anything unusual about the circuit. They were interviewed in separate rooms and did not hear each other's answers.

3.2.1.2 *Confusion Measures by Observers.* The author, assisted by a clerk, listened to the first five minutes of each condition. This covered 32 conditions with zero delay, 16 with 600 ms delay, and 16 with 1200 ms delay. A "confused situation" was scored whenever either speaker *reacted* to a double talk occurrence (simultaneous speech from *A* and *B*). The "reaction" had to contain a clear break in normal speech flow, such as a request for repeat ("What?"), or a self-generated repeat, or a sudden halt in the middle of an utterance. If both speakers simply continued to double talk, a confused situation was not scored, even though subsequent conversation indicated a lack of contextual understanding. Both observers had to agree on any confused situation for it to be scored.

Since the scoring instructions were imprecise, subjective judgment was used by the observers. To check validity, two other observers were given the above scoring rules and, without further instructions or practice sessions, were asked to observe four five-minute conversations.

3.2.2 *Objective Measurements*

3.2.2.1 *Speech Levels.* The equivalent peak level (epl) of a speech sample is an estimate of the peak of the instantaneous level distribution, based on a power measurement.[9,10] Epl readings were made of the first five minutes of every condition.

3.2.2.2 *Pattern Analysis.* An on-off speech pattern analysis was performed over the entire 10-minute length of each condition using a computer program written in 1963 by Mrs. N. W. Shrimpton. The program defines ten "events" related to the timing of the speech patterns, such as talkspurts, pauses, and double talk periods. The program produces considerable statistical data on the speech patterns, as illustrated in a previous paper.[11]

The patterns are strongly influenced by the design of the speech detector. In the present study, the detector was the same as that used in Ref. 11, with the threshold set to −40 dBm. (See Ref. 11 for further details on threshold settings.)

To show the effects of delay on speech patterns, the author has chosen to include data from two measurements. The first consists of the average lengths of each of the ten events. The second consists of the percent time each speaker spent in each of three states, namely, talking, double talking, and mutual silence. The talking time includes double-talking time. (These three measures are sufficient to calculate three others. A speaker is either talking or silent, hence, silent time is 100 percent minus talking time. "Solitary talking" time (no double talking) is talking minus double-talking time. "Listening" time, when a speaker is silent while the other is talking, is 100 percent minus talking minus mutual silence time.)

3.2.2.3 *On-Off Pattern Model Analysis.* The author has recently developed a stochastic model for generating on-off patterns that are statistically similar to those occurring in real conversation.[12] Using this model, only a few numbers are required to specify the conversational statistics. The model is described in Section 2.1.2 of Ref. 12. For convenience, the diagram of the model is included here in Fig. 2. Speaker A is considered to be in one of six states, as determined from the on-off speech patterns that appear at his side. The six parameters governing A's behavior are related to his probabilities for changing his talk-silence status. Numerically, these parameters, labeled α and β with suitable sub- and superscripts, specify the rates at which A departs from each state, as determined by Poisson processes. The larger the parameter, the sooner A will tend to leave the state. Thus, high values for parameters indicate short durations for the corresponding states. The subscript abbreviations stand for *sol*itary, interrup*ted*, interrup*tor*, *p*au*se*, *alt*ernate and *int*errupt.

In the present experiment, the entire 10-minute length of each condition was examined to obtain the six parameters for each speaker. These parameters are determined by measuring the frequencies of changing from state to state. The method of obtaining the parameters is described in more detail in Section 2.2 of Ref. 12.

3.3 *Technique of the On-Off Pattern Analysis*

The pattern analysis and model behavior, which are objective measures described in Sections 3.2.2.2 and 3.2.2.3 above, both use on-off speech patterns that are obtained with a speech-detector technique

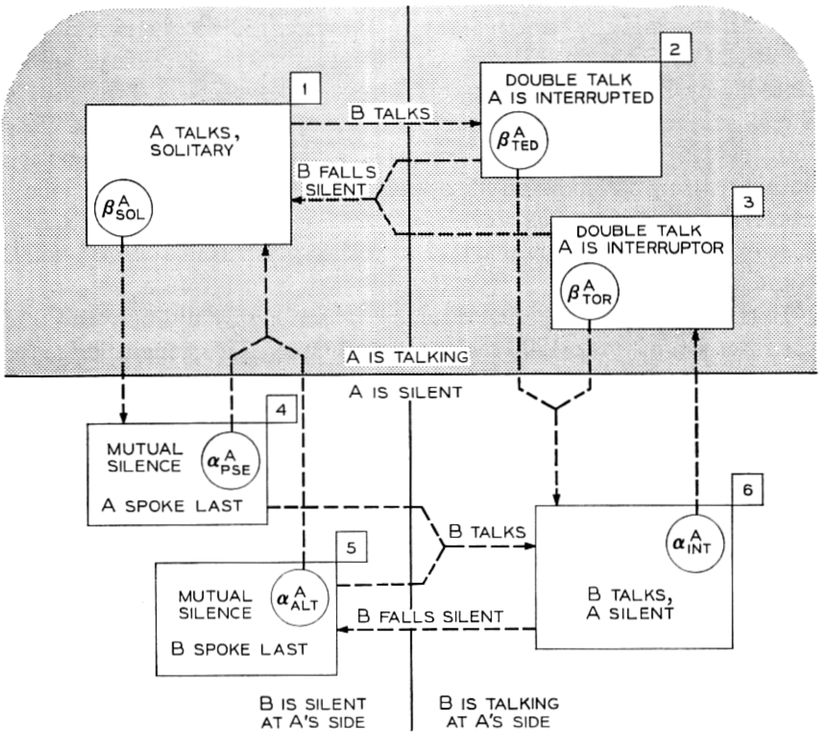

Fig. 2—The six-state model used in this study. Vertical transitions are due to Poisson processes at A's side. Horizontal transitions, due to B, are in A's external environment and are not generated by A's model.

described in previous papers.[11–13] The absolute values of any of the measures, such as average talkspurt length or the α and β values, are dependent on the speech detector parameters. For this reason, the absolute measure values were not used as an indication of the effects of delay. Instead, the *changes* in the values of the measures for each pair of subjects were noted when delay was inserted. By having each pair of subjects talk on both circuits in one session, and by balancing the pairs between delay-circuit-first and delay-circuit-second, changes in the measures from effects other than delay were minimized.

IV. RESULTS

The following is a comparison of the measurements on the standard and delay circuits.

4.1 *Opinion of Subjects*

When asked if they noticed anything unusual about the circuit, none of the 32 subjects on 600 ms delay commented on the delay. (A few stated they heard some background noise at low level.) Two pairs of the 32 subjects on 1200 ms delay made delay-related comments. These four people said they had "confused situations," but they noticed nothing wrong with the circuit. Some comments were, "I occasionally asked the other if he was still there," or "Sometimes, if we both started talking at the same time, I heard the other person but couldn't understand him."

It appears that inexperienced subjects are highly unlikely to notice the insertion of up to 1200 ms round-trip delay. This conclusion agrees with those of Bricker[3] and Klemmer.[5]

4.2 *Confusion Measures by Observers*

On listening to the conversations, both observers agreed there were many places in which confusion occurred when the delay was introduced. Table I summarizes the average number of confusions occurring in the first five-minutes of each condition.

Coincidentally, the 16 pairs of speakers on 600 ms delay had ex-

Table I—Confusions in 5 Minutes of Conversation, Averaged Over 16 Conversations

	Standard Circuit	Delay
600 ms pairs	1.2	3.7
1200 ms pairs	1.2	4.3

actly the same average number of confusions on the standard circuit as did the speakers on 1200 ms delay. Thirteen of the 16 pairs experienced more confused situations on delay than on the standard circuit, for both values of delays. This increase is significant at the 0.05 level (sign test). Note that the average number of confusions on delay for 1200 ms is not much greater than for 600 ms. The difference between 3.7 and 4.3 confusions on the two different delays is not significant (t-test).

The four five-minute segments monitored by other observers consisted of two pairs of zero and 1200 ms delay. For these four segments, the author and clerk recorded 1, 5, 1, and 6 confusions (for zero, 1200 ms, zero, and 1200 ms delay). The other two observers, as a team, recorded 0, 6, 0, and 9 confusions. Though not in exact agreement with the author's measures, these measures also indicate an increase of confusions on the delay circuits.

Summarizing this Section, when delay was added, the subjects experienced a significant increase in confused situations or difficulties, although they noticed no change in circuit quality. The 1200 ms delay effect was not significantly different from the 600 ms delay effect.

4.3 *Speech Level Measurements*

Delay had little or no effect on speaking levels. For the 600 ms delay circuits, 19 out of 32 speakers had lower epl values on the standard circuit than on delay. For the 1200 ms delay circuits, 20 out of 32 speakers had lower values on the standard circuit. Neither of these ratios is significantly different from chance.

Since half the pairs received the delay circuit first, and half received the standard first, the standard versus delay comparison does not show a possible effect of time on speaking levels. We therefore tested the difference in speaking levels between the first and second parts of each conversation. For the 600 ms delay pairs, 11 out of 32 were quieter on the first segment, and for the 1200 ms delay pairs, the ratio was 12 out of 32. Neither is significant at the 0.05 level.

The average epl for all 64 speakers was −11.6 dBm, at the OTLP point of a simulated toll circuit. (The zero transmission level point is an arbitrary reference level used to establish relative levels in a telephone circuit.)

4.4 *Pattern Analysis*

Table II lists the average lengths of the ten events, defined in Ref. 11, for all conversations. In every case, the average length is taken over 32 speakers.

TABLE II—AVERAGE LENGTH OF TEN EVENTS ON STANDARD AND DELAY CIRCUITS
(See Ref. 10 for definitions of the events.)

Event	Average Length, Seconds		Significance Level*	Average Length, Seconds		Significance Level*
	Std. Circuit	600 ms Delay		Std. Circuit	1200 ms	
1. Talkspurt	1.426	1.344	—	1.271	1.241	—
2. Pause	2.174	1.994	—	2.007	2.192	—
3. Double Talk	0.264	0.324	0.01	0.274	0.357	0.01
4. Mutual Silence	0.490	0.550	0.01	0.558	0.614	0.01
5. Alternation Silence†	0.433	0.422	—	0.496	0.505	—
6. Pause in Isolation	0.520	0.560	0.1	0.580	0.645	0.05
7. Solitary Talkspurt	1.424	1.325	—	1.230	1.127	—
8. Interruption	0.726	0.795	—	0.805	0.690	—
9. Speech after Interruption	0.852	1.001	—	0.865	0.976	0.1
10. Speech before Interruption	1.245	0.964	—	0.929	0.931	—

* Significance levels are determined from sign tests on the number of subjects that change in the indicated direction.

† In Fig. 6 of Ref. 10, the alternation silences illustrated in the sample patterns are all incorrectly labeled. The A's and B's are transposed. An alternation silence from B to A occurs when B stops talking, a mutual silence ensues, and then A talks, as determined from analysis at A's side. These are counted for all speakers A. Analysis at B's side produces the A to B alternation silences, which are counted for speakers B.

Note: Each average length represents 32 subjects. The two "Standard Circuit" columns represent two different groups of 32 subjects.

When analyzing the patterns, a separate analysis must be made at each speaker's side, because the delay, when present, causes the temporal patterns to differ at the two sides. For example, with no delay, both speakers will have an identical number of double talks of the same average length. With delay, double talking at one side does not imply double talking at the other. Thus, on the delay circuits, a separate analysis is required for each speaker, not just for each conversation.

Significance tests were made from sign tests on the lengthening or shortening of events. For example, 30 out of 32 speakers experienced an increase in average double talking length when going from a standard circuit to the 600 ms delay circuit.

Table II shows that when delay was inserted, there was a statistically significant increase in the average lengths of double talks, mutual silences, and pauses in isolation. Event 9, speech after interruption, is not very much affected since it is significant at only the 0.1 level, and then only on the 1200 ms delay.

Consider now the amount of time spent in the talking, double-talking, and mutual-silence states. These results are shown in Table III. When delay is inserted, the percent time spent in both double talking and mutual silence increases significantly.

A fairly simple mechanism can account for most of the results shown in Tables II and III. The increase in duration of double talks and mutual silences, and also of pauses in isolation, which are a subset of mutual silences, can generally be explained by noting that speech from the distant talker (e.g., *B*) will sometimes arrive at *A*'s side at inappropriate times. Specifically, responses to *A*'s speech can arrive when *A* has resumed talking rather than during *A*'s brief silences. Hence, the previous talking sequence *A*—mutual silence—*B*—mutual silence—*A* has been replaced with *A*—longer mutual silence—*A*—double talking, resulting in more time in mutual silence and in double talking.

There was little difference in the magnitude of the effect on patterns of 1200 ms versus 600 ms delay. From Table III, the double talking time for 600 ms delay was 5.9 percent versus 4.2 percent for the standard condition, a ratio of 1.40. This ratio was 1.41 for the 1200 ms case. The mutual silence ratios were 1.11 and 1.07 for 600 and 1200 ms, respectively. From Table II, the ratios of increases for average length of events for 600 ms and 1200 ms delays were 1.23 versus 1.30 for double talks, 1.12 versus 1.10 for mutual silences, and

TABLE III—PERCENT TIME IN THREE STATES ON STANDARD AND DELAY CIRCUITS, AVERAGED OVER 32 SPEAKERS

State	Standard	600 ms Delay	Significance* Level	Standard	1200 ms Delay	Significance* Level
Talking	40.5%	40.0%	n.s.	38.5%	38.4%	n.s.
Double Talking	4.2%	5.9%	0.01	4.4%	6.2%	0.01
Mutual Silence	23.0%	25.5%	0.01	27.2%	29.0%	0.01

* Significance based on sign test on number of subjects that changed in the indicated direction.

1.08 versus 1.11 for pauses in isolation. None of the differences in ratios between 600 and 1200 ms is significant.

To summarize the pattern analysis results, insertion of 600 or 1200 ms delay produced a statistically significant increase in the durations of double-talk and mutual-silence periods, and in the percent time spent in these states. There was no difference between the delays on the magnitude of the effect.

4.5 *On-Off Pattern Model Analysis*

Table IV lists the average values of the model parameters for all subjects. Two parameters change significantly. Let us first consider α_{pse}, the tendency* to resume talking after falling silent. Its value is lowered, indicating a tendency to remain silent longer, possibly in order to give the other speaker a chance to respond. The second changing parameter, β_{ted}, is also lowered, indicating a greater tendency to keep talking when interrupted.

The remaining four parameters remain unchanged. In particular, α_{int} hardly is affected; people's tendency to interrupt remains the same.

Colloquially, the downward changes in α_{pse} and β_{ted} as delay is introduced indicate that people exhibit an increased tendency to wait for replies, and have a greater tendency to ignore interruptions.

Table IV shows little difference between the effects of a 600 ms and 1200 ms delay, as noted in the ratio differences in parameters for the delay circuit versus the standard circuit.

The model concerns itself only with that aspect of human behavior reflected by a person's tendency to remove himself from a state once he enters it. The delay may rearrange speech patterns from speaker *B* so that, for example, *A* may be interrupted and thus double talk more often than he would without delay. This would increase the time *A* spends in state 2 of Fig. 2. In measuring *A*'s behavior, the model is unconcerned that with delay, *A* may spend more time in state 2. The model does indicate, however, that when *A* is in state 2, he is less likely to leave when on a delay circuit than when on a standard circuit. Thus, when the model indicates that *A*'s behavior changes with delay, it is referring to *A*'s tendencies to talk and fall silent, and not the change in interaction of *A*'s and *B*'s speech patterns, which are largely controlled by delay.

---

* The α's and β's are not probabilities in a mathematical sense; they are instead arrival rates for "pulses" to change state. However, they are proportional to probabilities for very small time increments, as defined in the Poisson process. Here, they are colloquially referred to as "tendencies."

TABLE IV—ON-OFF SPEECH PATTERN MODEL PARAMETERS ON STANDARD AND DELAY CIRCUITS

	600 ms delay subjects				1200 ms delay subjects			
	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level
β_{sol}	0.68	0.71	1.04	—	0.76	0.78	1.03	—
β_{ted}	1.93	1.41	0.73	0.01	1.88	1.33	0.71	0.01
β_{tor}	2.37	2.01	0.85	—	2.30	2.14	0.93	—
α_{pse}	2.13	1.79	0.84	0.01	1.87	1.59	0.85	(0.1)
α_{alt}	0.95	0.91	0.96	—	0.75	0.77	1.03	—
α_{int}	0.24	0.23	0.96	—	0.24	0.22	0.92	—

Note: Units of the parameters are "events per second," as explained in Ref. 11. Each parameter is averaged over 32 people. Significance level is determined from a sign test on the number of speakers that changed in the indicated direction.

The on-off pattern model can be applied to delay analysis in another way. Instead of looking at the way a speaker behaves, we can examine the way he *seems* to behave as viewed by the other speaker. For example, on the delay circuit, speaker B may talk during A's silence at B's side, but B's talkspurt, on arriving at A, may interrupt A's speech.

To perform the required analysis, all we need do is examine all the B speakers' patterns at the A terminals and vice versa. (Clearly, in the no-delay case the results will be the same as before since the patterns are identical at both sides.) Table V shows the results of the model analysis, viewing each speaker from his conversant's viewpoint.

Table V presents a striking contrast to Table IV. The first two entries β_{sol} and β_{ted}, are similar in both tables; β_{sol} remains unchanged and β_{ted} decreases. However, the parameter β_{tor} drops in Table V and remains unchanged in Table IV. This means, for example, that if, on a delay circuit, speaker A perceives B as an interruptor, B appears to A to tend to keep interrupting when in fact B is just as anxious to terminate double talking as he was on the standard circuit.

The effects on the α's are completely reversed in Tables IV and V. The drop in α_{pse} in Table IV is absent in Table V, implying that A perceives B as not tending to wait for replies when in fact he is. The drop in α_{alt} of Table V implies sluggishness; on delay, B seems much less likely to reply (or alternate) to A. In fact, his behavior is unchanged, as seen from α_{alt} in Table IV, but his replies now arrive after A has resumed talking. Thus, A perceives these as interruptions, as indicated by the rise in α_{int} in Table V. Whereas the tendency to interrupt remains the same (α_{int} is unchanged in Table IV), it appears to increase with delay.

As with all other results in this study, Table V shows no difference between 600 and 1200 ms delay on the apparent changes in the parameters.

We summarize this section by compiling a "profile" of the typical speaker conversing on a delay circuit, as expressed in terms of the model. From Table IV, his behavior changes slightly; it is as if he were more polite in letting his conversant have more time to respond (α_{pse} drops), and more persistent in continuing talking if interrupted (β_{ted} drops). But, from Table V he appears to the other speaker to be sluggish to respond (α_{alt} drops), interrupt more often (α_{int} increases about 50 percent), and once engaged in double talking, he appears reluctant to stop no matter who interrupted. This is an unflattering

TABLE V—MODEL PARAMETERS AS VIEWED FROM DISTANT SPEAKER

	600 ms delay subjects				1200 ms delay subjects			
	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level
β_{sol}	0.68	0.69	1.01	—	0.76	0.78	1.03	—
β_{ted}	1.93	1.50	0.78	0.01	1.88	1.28	0.68	0.01
β_{tor}	2.37	1.83	0.77	0.01	2.30	1.63	0.71	0.01
α_{pse}	2.13	2.12	1.00	—	1.87	1.72	0.92	—
α_{alt}	0.95	0.61	0.64	0.01	0.75	0.47	0.63	0.01
α_{int}	0.24	0.35	1.46	0.01	0.24	0.39	1.63	0.01

portrait; to the extent that the model parameters correspond to aspects of human behavior, these effects may give rise to the confused situations that are observed on the delay circuits.

V. THE EFFECTS OF AMOUNT OF DELAY

On seeing the results of this experiment, one might be puzzled by the lack of difference in the effects of 600 vs 1200 ms delay. This may be reasonable, however, in the light of the following experiment conducted with simulated conversations.

Values of all α and β parameters were measured from 8 no-delay conversations, and parameter averages were taken for the *A* and *B* speakers separately. This was done to obtain a profile of two typical speakers. (These two profiles turned out to be very similar, as they should if the *A* and *B* recording and measuring circuitry were identical.) We then had "typical speaker" *A* engage in a series of computer-simulated 20 minute conversations with "typical speaker" *B* over delays up to 2400 ms. (The simulator is described in Ref. 12.) For each delay value, a separate simulation run was made such that the simulated speakers were substituted for the speakers shown in Fig. 1. Note that it would have been meaningless to record patterns of a simulated no-delay conversation and then replay those particular patterns with one speaker's patterns delayed with respect to the other's. In this case, each model would fail to respond to delay-induced changes in the other's speech.

An analysis of the type in Table IV applied to the simulated conversations would simply yield the α and β values used to generate the patterns. That is, the "true" values of α and β were fixed for all delays. But the apparent values, as seen from the other speaker, did change, and these ratio changes are plotted in Fig. 3. Each curve represents a single parameter, and each point on each curve is the average of two values, one for speaker *A* and one for speaker *B*.

On reading Fig. 3, remember that real people make changes in their true values of β_{ted} and α_{pse} when delay is introduced, while the simulated speakers did not. In Fig. 3, all of the six apparent parameters change, while in Table V only four changed. The discrepancy may be due to holding the β_{ted} and α_{pse} parameters constant in the simulated conversations; had these parameters been varied in the same manner as observed in real subjects, Fig. 3 might be more consistent with Table V.

The most important result shown in Fig. 3 is that the results tend to

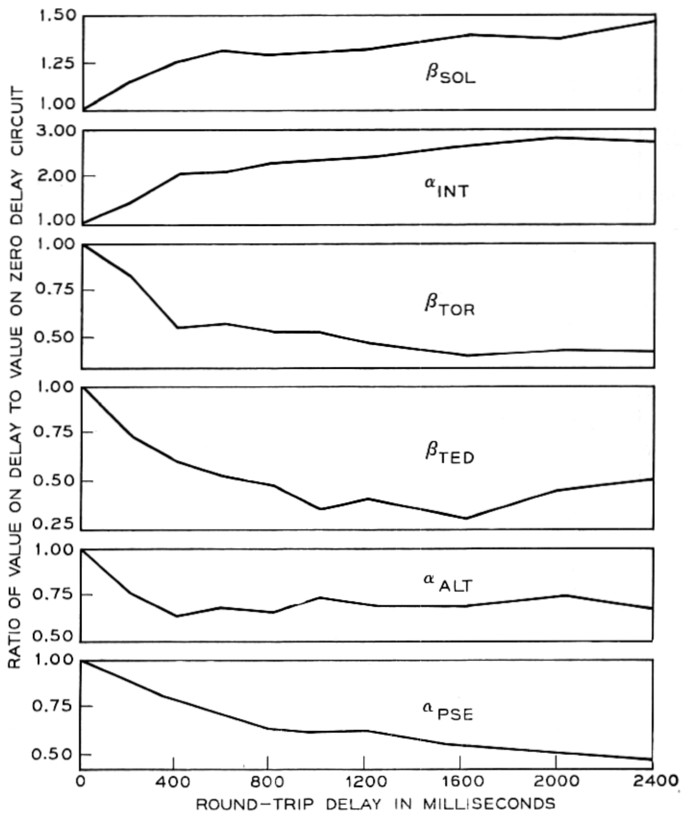

Fig. 3—Changes in the apparent values of the 6 model parameters when two simulated "speakers," with fixed parameters, "converse" over different delays. Each point is the average of the two speakers' apparent values.

asymptote with large delay values, and that they reach this asymptote in roughly 600 ms. *Therefore, a 600 ms round-trip delay is sufficient to produce all of the effects of apparent parameter change due strictly to delay.*

It is obvious that if *very* long delays are introduced, people will begin to notice the delay in responses to questions and will probably object to the delay for reasons other than confusion difficulties. Klemmer found a marked increase in dissatisfaction with 2400 ms delay.[5] The simulated speakers of Fig. 3 respond only to on-off patterns, and not contextual cues, and will never "notice" this sluggishness.

VI. CONCLUSIONS

When delay is introduced into a two-way telephone conversation between people who are inexperienced on delay circuits, measurable changes occur in their speaking behavior, even though the subjects notice no change in the circuit. The subjects become confused more often, engage in more double talking and mutual silence, and exhibit certain changes in their on-off pattern generation behavior. In addition, when a talker's speech patterns are analyzed at his conversant's side, he appears more likely to interrupt, less likely to terminate interruption, and less likely to respond to questions or pauses.

A round-trip delay of 600 ms seems large enough for all of the above effects to reach their asymptotic values. As delay becomes longer than 600 ms, additional degradation of the type measured here does not increase, but the subjects eventually may notice the sluggishness caused by very long delays.

If echoes and echo suppressors are present, other studies[1] indicate that the circuits are perceived to be unsatisfactory by some people, and that the effects get worse as delay is increased from 600 to 1200 ms. This is probably due to the mutilating effects on speech, which are caused by the echo suppressors. If, however, echoes are controlled to the point that the circuit appears to be a 4-wire circuit, inexperienced subjects or customers might behave in a manner similar to those subjects in the present study.

An important question not answered by this study is what will happen as subjects become experienced on the delay circuits. The question of behavior after acquiring experience provides a natural direction for future research that the author intends to pursue on this problem.

REFERENCES

1. Helder, G. K., "Customer Evaluation of Telephone Circuits with Delay," B.S.T.J., *45,* No. 7 (September 1966), pp. 1157–1191.
2. Emling, J. W., and Mitchell, D., "The Effects of Time Delay and Echoes on Telephone Conversations," B.S.T.J., *42,* No. 6 (November 1963), pp. 2869–2892.
3. Bricker, P. D., *Satellite Communications Physics,* Chapter 5, Foster, R. M., editor, Murray Hill, N. J.: Bell Telephone Laboratories, 1963, pp. 70–77.
4. Krauss, R. M., and Bricker, P. D., "Effects of Transmission Delay and Access Delay on the Efficiency of Verbal Communication," J. Acoust. Soc. Am., *41,* No. 2 (February 1967), pp. 286–292.
5. Klemmer, E. T., "Subjective Evaluation of Transmission Delay in Telephone Conversations," B.S.T.J., *46,* No. 6 (July-August 1967), pp. 1141–1147.
6. Richards, D. L., "Conversational Performance of Speech Links Subject to Long Propagation Times," Proc. I.E.E. (London), *116* (1969), pp. 955–963.

7. Brady, P. T., and Helder, G. K., "Echo Suppressor Design in Telephone Communications," B.S.T.J., *42*, No. 6 (November 1963), pp. 2893–2918.
8. Brady, P. T., "A Stochastic Model of Message Interchange on a Channel with Transmission Delay," I.E.E.E. Trans. on Commun. Technology, *COM-15*, No. 3 (June 1967), pp. 405–412.
9. Brady, P. T., "Equivalent Peak Level: A Threshold-Independent Speech-Level Measure," J. Acoust. Soc. Am., *44*, No. 3 (September 1968), Pages 695–699.
10. Brady, P. T., "A Comparison of the Equivalent Peak Level with Other Speech Level Measures," unpublished work.
11. Brady, P. T., "A Statistical Analysis of On-Off Patterns in 16 Conversations," B.S.T.J., *47*, No. 1 (January 1968), pp. 73–91.
12. Brady, P. T., "A Model for Generating On-Off Patterns in Two-Way Conversations," B.S.T.J., *48*, No. 7 (September 1969), pp. 2445–2472.
13. Brady, P. T., "A Technique for Investigating On-Off Patterns of Speech," B.S.T.J., *44*, No. 1 (January 1965), pp. 1–22.