Labov says that the second line is easily seen as an answer to the first despite the absence of syntactic information, e.g. a 'no' in line *B*, because speakers and hearers use social knowledge which Moerman (172, following Sacks) calls 'tying' and which Garfinkel calls 'common understanding' (3–6). This social knowledge needs to be included in linguistic theories of the speaker-hearer, if the model is to adequately reflect language use.

Labov contrasts 'ritual insults' with 'personal insults' to establish discourse rules. By contrast to a personal insult which is countered by a denial (*A*. Your mother drinks pee. *B*. She does not), ritual insults are answered by another 'sounding' (*A*. Your mother drinks pee. *B*. Your father eats shit) which become chained into long sequences. Although Labov's careful analysis devises a linguistic rule which goes beyond traditional linguistic theories, this rule does not explain how the speaker or hearer uses his social knowledge to 'choose' between a ritual or personal insult discourse strategy.

Cicourel and Garfinkel maintain that the apprehension of any object in the world is dependent on the methods by which that apprehension is accomplished. That is, theories of the mind, language use and data collection, including sociological conceptions of actors playing out the roles associated with certain social statuses as guided by normative rules requires an interpretive element. If actors can occupy many statuses, or play many roles, then they must have the interpretive ability ('basic rules', 250ff.) to decide which role behavior 'normative rules' is appropriate for which social situations. If there can be a plurality of normative rules guiding action, then the actor must have the ability to decide which rule applies to the situation at hand.

The conversational and discourse rules located by Sacks, Schegloff, Labov *et al.* can be understood as normative rules in the same way as social rules. They are abstracted formulations which need to be adapted to specific interactional situations. Just as the role player in sociological theory must decide which role is appropriate for social situations, the conversationalist must 'decide' which MCD, topic selection, or discourse rule, etc., applies in which situation. Both conceptions require that an underlying interpretive mechanism be attributed to the model of the actor which will instruct the actor how to relate general normative rules to specific situations.

REFERENCES

Cicourel, Aaron V. (1964). *Method and measurement in sociology*. New York: The Free Press.

Garfinkel, Harold (1967). *Studies in ethnomethodology*. Englewood Cliffs: Prentice Hall.

Hill, Richard & Stones Crittendon, Kathleen (1968). *The Purdue symposium on ethnomethodology*. Institute for the Study of Social Change. Purdue University Monograph.

Zimmerman, Don H. & Pollner, Melvin (1970). *The everyday world as a phenomenon*. In Jack D. Douglas (ed.), *Understanting everyday life*. Chicago: Aldine.

Reviewed by Hugh Mehan
Department of Sociology
University of California, San Diego
La Jolla, California 92037

(Received 12 October, 1973)

Lang. Soc. **2**, 161–180. Printed in Great Britain

On the structure of speaker–auditor interaction during speaking turns[1]

STARKEY DUNCAN, JR.

Department of Behavioral Sciences, Committee on Cognition and Communication, University of Chicago

ABSTRACT

The structure of speaker–auditor interaction during speaking turns was explored, using detailed transcriptions of language, paralanguage, and body-motion behaviors displayed by both participants in dyadic, face-to-face conversations. On the basis of certain observed regularities in these behaviors, three signals were hypothesized: (a) a speaker within-turn signal, (b) an auditor back-channel signal, and (c) a speaker continuation signal. These signals were composed of various behaviors in language and in body motion. It was further hypothesized that the display of appropriate ordered sequences of these signals by both participants, served to mark 'units of interaction' during speaking turns. (Conversational analysis; speaking turns; back-channel behaviors; interrelations of verbal and nonverbal behavior; American English (Chicago)).

It is axiomatic that language is used within a larger communication context. This context typically includes, among other things, both (*a*) other behaviors (such as those in paralanguage (Trager 1958) and in body motion) displayed conjointly with language behaviors, and (*b*) one or more other persons with whom the language user is interacting.

It is commonplace to consider language (as that term is traditionally used) to be the province of highly structured, rule-governed phenomena. One might inquire, however, as to the possibility that such structured phenomena might be found to extend beyond the bounds of language proper, to include elements of the larger communication context.

This paper reports some findings from a program of research designed to discover elements of structure in the broader communication context, not only in paralanguage and body motion, together with language, but also in the respective behaviors of both participants in dyadic, face-to-face conversations.

[1] This study was supported in part by Grants MH-16,210 and MH-17,756 from the National Institute of Mental Health, and by Grants GS-3033 and GS-3033A # 1 from the Division of Social Sciences of the National Science Foundation. Susan Beekman, Mark Cary, Diane Martin, George Niederehe, Ray O'Cain, Thomas Rossen, Thomas Shanks, Cathy Stepanek, and Andrew Szasz contributed to the transcription and data

Previously reported results from this research program (Duncan 1972; Duncan & Niederehe 1974) have dealt primarily with phenomena associated with the exchange of speaking turns in conversations. Evidence was presented leading to the hypothesis of several signals used to facilitate the smooth and appropriate exchange of the speaking turn between the two participants. These signals and accompanying rules will be briefly described below.

Speaker–auditor interaction may also be encountered during the course of a speaking turn. In dyadic, face-to-face conversations, as the speaker continues his turn, the auditor does not typically remain mute and motionless. Rather, he may frequently nod or shake his head and utter a variety of vocalizations, such as 'm-hm', and 'yeah'. (These and similar behaviors will be termed 'back-channel behaviors' in the context of this discussion. The term 'back-channel' was adopted from Yngve (1970)).

In this paper, empirical relationships found between (*a*) auditor back-channel behaviors, and (*b*) certain speaker behaviors in language and in body motion, variously occurring either before or after the back channels, will be presented. On the basis of these observed relationships, hypotheses will be advanced regarding some elements of communication structure applying to speaker–auditor interaction during speaking turns.

METHOD

Materials

The results to be reported were based on the same set of transcribed data described in Duncan (1972). Briefly, detailed transcriptions were made of speech and body-motion behaviors observed during the first nineteen minutes of two, two-person conversations, as recorded on videotape. Both conversations would have taken place regardless of whether or not they had been videotaped. All participants in the two conversations had full knowledge of the videotaping procedure and had consented to the taping prior to the conversations. All participants were native speakers of American English.

The first conversations was between a twenty-year-old female applicant for psychotherapy, and a male therapist of about 40 years. The conversation was an intake interview, in which the applicant was provided with some information about the operation of the counseling center to which she was applying and the therapist was provided with information both with regard to routine facts (age, address, etc.) about the applicant, and with regard to her reasons for applying for therapy. Thus, a variety of topics were discussed in the conversation. The applicant was at that time working as a secretary, having had about two years of college. The therapist and applicant were previously unacquainted.

The second conversation was between the therapist in the first conversation, and another therapist, also a male of about 40 years. The two men had been

good friends for a number of years. The general topic was a discussion of a client (not the one in the first conversation) whom the first therapist had seen in an intake interview, and whom the other therapist had seen in an initial therapy hour. However, the interaction was also characterized by some joking and kidding between the two men, a typical aspect of their relationship.

Transcription

Segmental and suprasegmental phonemes were transcribed, with minor modifications, in accordance with the scheme published by Trager and Smith (1957). The segmental phonemes were the least important components of the study.

Transcriptions of paralanguage followed the outline developed by Trager (1958). A number of paralinguistic behaviors were transcribed; however, only the following played a part in results obtained thus far: (*a*) intensity; (*b*) pitch height; and (*c*) extent.

In contrast to paralanguage, there was for body motion no available transcription system which could be readily adapted to our purposes. This situation led to a transcribing method based on the behaviors actually encountered in each interview. The transcription system for the first conversation was created by first making an inventory of the movements used by the two participants, and then assigning either arbitrary or descriptive labels to these movements. This system was then applied to the second conversation, after expanding it to include new movements observed.

While there is no pretense that the resulting transcription system is able to encompass all movements occurring in this culture, every attempt was made to include all movements observed in the dyads under study. Included were: (*a*) head gestures and movements (nodding, turning, pointing, shaking, etc.) and direction of head orientation; (*b*) shoulder movements (e.g., shrugs); (*c*) facial expressions, such as could be clearly seen on the videotape; (*d*) hand gestures and movements of all sorts (each hand transcribed independently); (*e*) foot movements (each foot independently); (*f*) leg movements; (*g*) postures and posture shifts; and (*h*) use of artifacts, such as pipe, kleenex, papers, and clip board. Each observed occurrence of the various behaviors included in each of these categories was carefully located in the stream of communication and transcribed. None of these transcribed behaviors was ruled out of consideration on *a priori* grounds as potential components of phenomena treated in this paper.

Coordination of body motion and speech transcriptions

Speech syllables were used to locate all transcribed events. Thus, the movements of both participants in a conversation were located with respect to the syllables emitted by the participant who happened to be speaking at the time, or to the pause between two syllables.

Unit of analysis

A unit of analysis was selected which lay in size between the phonemic clause (Trager & Smith 1957) and the speaking turn. The unit was defined in terms of behaviors displayed by the participants. Boundaries of the unit were defined as being: (*a*) at the ends of phonemic clauses (though not necessarily those with rising or falling terminal junctures), (*b*) which additionally were marked by the display of one or more of the following behaviors: (*a*) an unfilled pause; (*b*) turning of the speaker's head toward the auditor; (*c*) a drop in paralinguistic pitch and/or loudness in conjunction with a phonemic clause, either across the entire clause, or across its final syllable or syllables; (*d*) a relaxation of the foot or feet of the speaker from a marked dorsal flexion; (*e*) audible inhalation; (*f*) the use of any pitch level–terminal juncture combination other than 2 2| at the end of a phonemic clause; (*g*) paralinguistic drawl on the final syllable or on the stressed syllable of a clause; (*h*) the termination of any hand gesticulation used by the speaker, or the relaxation of a tensed hand position (e.g., a fist) by the speaker; (*i*) the use by the speaker of one of a set of stereotyped expressions, such as 'but uh', 'or something', or 'you know', termed sociocentric sequences by Bernstein (1962); (*j*) a drop in paralinguistic pitch and/or intensity, in conjunction with a sociocentric sequence; and (*k*) the completion of a grammatical clause, involving a subject-predicate combination. More detailed definitions for these behaviors are given in Duncan (1972).

Extent of corpus

There were in the transcriptions of the two conversations: 2,481 phonemic clauses; 1,043 in Conversation 1, and 1,438 in Conversation 2. In the two transcriptions there were 885 units as defined above, yielding an average of 2.8 clauses per unit.

THE SPEAKING-TURN SYSTEM

On the basis of previous research, a 'speaking-turn system' has been hypothesized to account for findings on structured phenomena related to the exchange of speaking turns in dyadic conversations. Because of the apparently close relationship between the phenomena to be considered here and those of speaking turns, some aspects of the speaking-turn system will be briefly sketched. More extended accounts may be found elsewhere (Duncan 1973, 1974).

In general, the speaking-turn system, as presently developed, is conceptualized as consisting entirely of discrete elements, arranged hierarchically.

For the purposes of the system, two, mutually exclusive, discrete states are posited for each participant in a dyadic conversation: speaker and auditor. A 'speaker' is defined as a participant who claims the speaking turn at any given

moment. An 'auditor' (Kendon 1967) is a participant who does not claim the speaking turn at any given moment.

Speaker turn signal

The auditor may claim the turn when the speaker displays a turn signal. In proper operation of the system, if the auditor so claims the turn in response to the signal, the speaker is obliged to relinquish immediately his claim to the turn. When the speaker is not displaying the turn signal, however, auditor claims of the turn are inappropriate, leading to simultaneous turns.

The turn signal is permissive, not coercive. The auditor is not obliged to claim the speaking turn in response to the display of the signal by the speaker. The auditor may alternatively communicate in the back channel, or remain silent.

The turn signal is comprised of a set of six behavioral cues, found variously in intonation, content, syntax, paralanguage, and body motion. These cues were mentioned above as behaviors *f*–*k*, used in the definition of units of analysis.

The display of any single cue was sufficient to constitute a display of the signal. However, the probability of an auditor turn-claiming response to a signal display was found to be a linear function ($r = 0.96$) of the number of cues displayed, without regard to the specific cues comprising the display.

Each cue of the turn signal was considered to be activated only at the end of phonemic clauses, even if the cue was displayed prior to the end of the clause. For example, a drawl is appropriately responded to by the auditor only when the speaker has completed the phonemic clause in which the drawl is displayed.

Speaker gesticulation signal

The display of this signal appears to negate any turn signal concurrently being displayed. It was found that display of the gesticulation signal virtually eliminates claims to the turn by the auditor.

The gesticulation signal is composed of a single, discrete cue: the hands being engaged in a gesticulation, as opposed to being engaged in a self-adaptor (Ekman & Friesen 1969), or to being at rest.

Speaker-state signal

This signal is hypothesized to mark a participant's shift from the auditor to the speaker state. In our corpus, speaker-state signals were typically displayed at the beginning of speaking turns, and were rarely displayed in association with back channels.

The speaker-state signal is defined as the display of at least one of a set of two discrete cues: (*a*) a shift in head direction, away from one pointing directly toward the partner; and (*b*) initiation of a gesticulation. This gesticulation is the same behavior that had the effect of suppressing turn claims of the partner, as mentioned above.

The cues comprising the speaker-state signal were observed, not only at the beginnings of speaking turns, but also during the course of speaking turns. These cues, displayed during turns, will be further considered below.

SPEAKER-AUDITOR INTERACTION DURING SPEAKING TURNS

In the course of analysis of the organization of interaction associated with the exchange of speaking turns, the question arose concerning the possibility that similar phenomena might be found during speaking turns. Accordingly, exploratory analysis was begun, in an attempt to discover patterns of speaker-auditor interaction as the speaker was pursuing his turn. In due course, evidence was found for such patterns, leading to the development of hypotheses concerning the operation of certain signals during speaking turns. For clarity of presentation, and to avoid terminological confusion, the signals and their constituent cues will be described first. Then evidence that led to these hypotheses will be presented in the Results section.

Auditor back-channel signal

The auditor back-channel signal does not constitute a speaking turn or a claim of the turn. It appears, however, to be a way that the auditor may provide the speaker with useful information as the turn progresses.

The following behaviors were observed in our corpus and, on the basis of the results reported below, were considered to be various forms the auditor back-channel signal took in the corpus. In the examples that follow, '*S*' stands for 'speaker'; and '*A*' for 'auditor'.

(1) *m-hm.* This expression is used to stand for a group of readily identified, verbalized signals. Included in the group are such expressions as 'm-hm', 'yeah', 'right', and the like, and Kendon's (1967) examples of 'yes quite', 'surely', 'I see', and 'that's true'. Most of the m-hm signals may be used singly or in repeated groups, as in 'yeah, yeah'.

(2) *Sentence completions.* Not infrequently in our materials an auditor would complete a sentence that a speaker had begun. In such a case he would not continue beyond the brief completion; the original speaker would continue with his turn as if uninterrupted. Sentence completions have been independently reported by Yngve (1970). Example: *S*: '... eventually, it will come down to more concrete issues ...'; *A*: 'As she gets more comfortable.' *S*: 'and I felt that ...'

(3) *Request for clarification.* Contrasting with sentence completions are brief requests for clarifications. Such requests were usually accomplished in a few words or a phrase. Example: *S*: '... somehow they're better able to cope with it.' *A*: 'You mean these anxieties, concern with it?' *S*: 'Possible that other people have ...'

(4) *Brief restatement.* This back-channel behavior is similar to the sentence completion, except that it restates in a few words an immediately preceding

thought expressed by the speaker. Example: *S*: '... having to pick up the pieces'; *A*: 'the broken dishes, yeah'; *S*: 'but then a very ...'

(5) *Head nods and shakes.* Head nods and shakes may be used alone or in company with the verbalized back-channel signals. Head nods may vary in duration from a single nod to a rather protracted, continuous series of nods.

We may, on the basis of form, distinguish the vocal back-channel behaviors (subclasses 1-4 above) from the visual back channels (subclass 5).

Speaker within-turn signal

The speaker within-turn signal was typically displayed at the end of units of analysis. The display of this signal was found to be related to the subsequent display of auditor back channels and of speaker continuation signals (described below).

This signal is composed of a set of two cues, (*a*) completion of a grammatical clause, and (*b*) turning of the speaker's head toward the auditor. The display of either one of these two cues is sufficient to constitute the signal. The grammatical-clause cue is identical to that mentioned as a speaker-turn cue. Thus, this cue is common to the two signals. (The head-turning cue was not included in the speaker turn signal, because it failed to differentiate smooth exchanges of speaking turns from instances of simultaneous turns.)

Speaker continuation signal

The speaker continuation signal was typically displayed at or very near the beginnings of units of analysis. As will be described below, the speaker continuation signal appears to operate during speaking turns in much the same way that the speaker-state signal operates at the beginnings of turns: both signals appear to mark the beginnings of new units within the conversation.

TABLE 1a. *Speaker and auditor signals hypothesized for within-turn interaction units*

Name of signal	Constituent cues	Related to subsequent display of
Speaker within-turn	1. Completion of a grammatical clause 2. Turning of head towards auditor	1. Between-unit auditor back channel 2. Speaker continuation signal
Between-unit auditor back channel	(5 different types, both audible and visible, observed)	
Early auditor back channel	(same as between-unit auditor back channel)	1. Speaker continuation signal
Speaker continuation signal	1. Turning of head away from auditor	

TABLE 1. *Location and observed frequencies of five types of auditor back-channel behaviors and of auditor turn claims*

Location	m-hm	Sentence completion	Brief restatement	Clarifying question	Head nod	*N* back channel	*N* Speaker turn claim	Grand total
No gesticulation signal displayed								
Early	20	3	2		18	43	7	50
Between units	29	3	1	3	19	55	77	132
Sociocentric sequence	9				12	21	6	27
Late	8				8	16	1	17
Total	66	6	3	3	57	135	91	226
Gesticulation signal displayed								
Early	11	2	2		20	35	3	38
Between units	32	4		2	28	66	5	71
Sociocentric sequence	4				4	10		10
Late	3		1	1	6	9	1	10
Total	50	6	3	3	58	120	9	129
Grand total	116	12	6	6	115	255	100	355

The speaker continuation signal was composed of one of the cues of the speaker-state signal: shift away in head direction. Display of this cue was related to the preceding display of both (*a*) the speaker within-turn signal, and (*b*) the auditor back-channel signal, in a manner to be described below.

Table 1a summarizes the above description of speaker and auditor signals hypothesized to be operating during the course of speaking turns. More detailed discussion of findings leading to these hypotheses will follow.

RESULTS

AUDITOR BACK-CHANNEL SIGNAL

An initial inventory was made of potential back-channel behaviors and their respective location in our corpus. Classification of behaviors as potentially in the back channel was aided by the astute observations and research results of Dittmann & Llewellyn (1967, 1968), Fries (1952), Kendon (1967), and Yngve (1970).

While back-channel behaviors by both speaker and auditor were observed in the corpus, the results to be described are based on analysis of auditor back channels exclusively. There was a total of 71 separate back channels displayed by speakers in the corpus.

Table 1 shows the observed frequencies of the five types of auditor back-channel behaviors and of auditor turn claims. The occurrence of each of these types of behavior is differentiated according to (*a*) whether or not it was displayed while the speaker was gesticulating, and (*b*) its location with respect to the units of analysis. The data shown on speaking-turn claims are elaborated from those shown in Table 1 in Duncan (1972).

Some tabulations in addition to those shown in Table 1 may be of research interest. It is possible for the auditor to display both a vocal and a visual back channel in the same unit of analysis. This happened 69 times. It is also possible for the auditor to display a sequence of two vocal back channels, such as 'yeah, yeah,' in the same unit of analysis. This happened 31 times.

The auditor's claiming the speaking turn may be preceded by his display of a back channel, either vocal or visual. Of such claims resulting in a smooth exchange of the speaking turn, seven were preceded by vocal back channels, and five by visual back channels. One occasion of simultaneous turns was preceded by a vocal auditor back channel. These back channels located at the beginnings of speaking turns were considered in Duncan & Niederehe (1972). In each case, these back channels were marked by the speaker-state signal.

In the early stages of investigating back-channel behaviors, some basic issues require consideration. (*a*) To what extent may the various types of behaviors tentatively identified as back channels be regarded as a single class of behaviors? (*b*) To what extent may the auditor back channels be regarded as a class of behaviors distinct from auditor turn claims (Duncan, 1972)? (*c*) Are auditor back channels related in any systematic way to behaviors displayed by the speaker? These issues are deeply interrelated; most of the findings to be reported are empirically relevant to more than one of the issues.

Auditor response and speaker gesticulation

In view of the sharp suppression of auditor turn claims by this speaker gesticulation signal, it is of interest to consider back-channel behaviors in this regard. Gesticulating does not appear to affect the display of back channels. Roughly 50 per cent of all back channels, and of each subset of back channels, are displayed during the speaker's gesticulation. Considering only back channels, and summing the cells for completions, brief restatements, and clarifying questions to avoid small expected frequencies, $\chi^2\ (2)=1.34$, $p=0.52$. Thus, display of the gesticulation signal by the speaker has no significant effect upon display of the back-channel signal by the auditor.

In view of this finding, it is not surprising to find that auditor back channels and auditor turn claims contrast sharply in their rates of display in the presence of the speaker gesticulation signal. If we collapse all back-channel signals into the single set: 'back channel,' a 2×2 contingency table can be created, using back channels *vs.* turn claims for the rows, and gesticulating *vs.* not gesticulating for the columns. For this table $\chi^2\ (1)=43.35$. When $df=1$, a χ^2 of 10.83 has an associated probability of 0.001. This result indicates that, if there is an auditor response of some kind (either back-channel or turn claim) in the presence of a speaker gesticulation signal, there is a highly significant tendency for that response to be a back-channel.

Location of auditor behaviors with respect to units of analysis

Back channels and turn claims can be compared in terms of their distribution across units of analysis. Inspection of the transcriptions suggested that more back channels seemed to be distributed widely throughout units, while turn claims seemed to cluster predominantly at the 'proper' place: immediately after the conclusion of the unit.

For purposes of comparison, each back channel and each turn claim was identified as having occurred in one of four positions in a unit of analysis: (*a*) early: prior to the end of the unit; (*b*) between units: immediately after the final syllable of a unit, and before the speaker had begun the first syllable of the next unit; (*c*) sociocentric sequence: when a sociocentric sequence accompanied the unit, in the brief pause, if any, between the substantive part of the unit and the sociocentric sequence and (*d*) late: on or soon after the first syllable of the next unit.

Before comparing the distribution of back channels as a whole with turn claims, it is appropriate to check the relative distribution of various subsets of back channels. Within the subset of vocal back channels, the distribution of 'm-hm' back channels was compared with completions, brief restatements, and

clarifying questions considered together: $\chi^2\ (3)=3.25$, $p=0.36$. Next, vocal back channels as a class were compared with visual back channels. In this case, $\chi^2\ (3)=4.06$, $p=0.25$. Inspection of this second contingency table showed that visual back channels tended to occur more frequently than expected, and vocal back channels less frequently than expected, in the between-unit position. This finding corresponds to subjective impressions derived from inspecting the transcriptions. However, the chi-square does not support the interpretation of a nonchance difference between vocal and visual back channels in their respective distributions across speaker units.

Table 2 shows the obtained and expected frequencies for auditor back channels (vocal and visual now combined) and for auditor turn claims. For this table $\chi^2\ (3)=39.13$, $p<0.00001$. Inspection of the cells of Table 2 indicates that back channels tended to occur more frequently than expected in the early, sociocentric sequence, and late positions, and less frequently in the between-unit position.

TABLE 2. *Observed and expected frequencies of auditor back channels and of turn claims in four locations with respect to units of analysis*

	Auditor back-channels		Turn claims		
Location	Observed	Expected	Observed	Expected	Total
Early	78	63.21	10	24.79	88
Between unit	121	145.82	82	57.18	203
Sociocentric sequence	31	26.58	6	10.42	37
Late	25	19.39	2	7.61	27
Total	255		100		355

Auditor back channels and speaker cue display

Units of analysis were defined in terms of a relatively large number of behavioral cues. We may ask if there is any subset of these cues which is closely associated with the subsequent display of auditor back channels. There was such a subset systematically related to auditor turn claims, as described above.

The association between (*a*) each speaker cue or combination of cues, and (*b*) subsequent auditor back channels, was evaluated by the following three criteria:

(1) *Percentage of back-channel behaviors accounted for.* This percentage was calculated by simply dividing the number of back channels which followed the display of the cues in question by the total number of back channels in the corpus.

(2) χ^2 *for display or absence of auditor back channel versus display or absence of speaker cues.* While the first criterion – percentage of back channels accounted for – is informative, it does not take into account the overall frequency of occurrence of the speaker cues in the corpus. Cues which occur very frequently may, by chance alone, be frequently followed by back channels. It is desirable to find cues or combination of cues, for which it is true that, if the cue is displayed, it tends to be followed by the display of a back channel; and if the cue is not displayed, there tends to be no back channel. For any given cue subset, this property was evaluated by χ^2 performed on a 2×2 contingency table, in which the columns represented (*a*) number of units in which there was a display of a back channel, and (*b*) number of units in which there was no such display; and the rows represented (*a*) number of units in which at least one of the cues under consideration was displayed, and (*b*) number of units in which none of the cues was displayed. The X^2 on such a table provides a routine statistical estimate of the probability that the observed association between cues and back-channels might have occurred by chance.

(3) *Relationship between number of speaker cues displayed and probability of back channels.* Because a strong linear relationship was found between number of speaker turn cues displayed and corresponding percentage of auditor turn-claiming responses ($r = 0.96$), the same relationship was tested for back channels. If a group of three speaker cues was being considered, for example, there were four data points for the correlation: the display of 0, any 1, any 2, and all 3 of the cues. For each one of these points, the percentage of back channels was calculated by dividing the number of back channels displayed for that point, by the total number of units represented by that point. From these data, a correlation was computed.

In calculating the correlation and χ^2, units followed by an auditor turn claim were eliminated from the analysis, because turn claims and back channels were considered, for present purposes, to be two alternative auditor tactics.

On applying these three criteria to the cues used to mark the end of the units of analysis, two of these cues were found to be highly associated with auditor back channels: (*a*) completion of a grammatical clause, and (*b*) turning of the speaker's head toward the auditor. These cues were found to be associated with both visual and vocal back channels, when these back channels were displayed after the end of the speaker units. These findings are shown in the top half of Table 3.

Considering vocal and visual back channels together, 88.8 per cent occurred following the display of at least one of the two cues. The χ^2 (1) for the associated 2×2 contingency table was 39.31. An almost perfect linear relationship was found between the number of speaker cues displayed and the probability of an auditor back channel, either vocal or visual. Similar relationships were not found for the remaining speaker turn cues.

TABLE 3. *Relationships between display of speaker within-turn cues and subsequent display of auditor back-channel behaviors*

N cues displayed	*N* units having display	Proportion of units with back channel following display	Percentage of back channels following display of ⩾1 cue	χ^2	Correlation
Between-unit and late vocal and visual back channels					
0	286	0.049			
1	326	0.172	88.8	39.31*	0.99
2	173	0.318			
Early and sociocentric-sequence vocal and visual back channels					
0	304	0.128			
1	327	0.107	50.6	3.70†	−0.95
2	154	0.032			

* Corrected for continuity; $df = 1$, $p > 0.00001$.
† Corrected for continuity; $df = 1$, $p = 0.05154$.

Separate analyses of vocal and visual back channels yield highly similar results. None of the findings with respect to speaker cue display suggested that vocal back channels should be considered as phenomena distinct from visual back channels.

Results shown in the bottom half of Table 3 indicate that auditor back channels classified as early or as sociocentric sequence were not associated with the speaker cues in question. Nor were they associated with any of the other speaker cues examined.

SPEAKER CONTINUATION SIGNAL

As mentioned above, the display of cues comprising the speaker-state signal was observed, not only at the beginnings of speaking turns, but also at various points during the course of turns. It was therefore possible to examine these latter displays as potential elements of speaker–auditor interaction during speaking turns.

Specifically, analysis was directed at possible relationships between the display of speaker-state cues during turns, and the prior display of (*a*) all cues, both in the turn signal and in the within-turn signal, and (*b*) auditor back-channel signals. Analyses revealed that one of the two speaker-state cues – shift away in head direction – was significantly related, both to within-turn cues and to auditor back channels. This speaker-state cue, when displayed during turns, was designated the speaker continuation signal. Details of these relationships are presented below.

It may be noted that a relationship between within-turn cues and subsequent speaker-state cues involves a sequence of two events and requires the action of only one individual: the speaker. In contrast, because an auditor back channel was found to be related to the prior display of the within-turn signal by the speaker, a relationship between auditor back channels and a subsequent speaker continuation signal potentially involves a sequence of three events and both participants: (*a*) speaker within-turn signal, (*b*) auditor back-channel signal, and (*c*) speaker continuation signal.

In these analyses, units of analysis that were immediately followed by an auditor claim of the speaking turn were not considered. Also, auditor back channels classified as 'late' were not included in the analysis, because their position made an ensuing display of a speaker-state signal difficult. Based upon the findings on auditor back channels, those classified as 'sociocentric sequence' were considered to be 'early'.

Speaker continuation signal and speaker within-turn signal

The effect of display of cues from the speaker within-turn signal upon the subsequent display of the speaker continuation signal may most directly be assessed through analysis of the cases in which no auditor back channel intervened. The relationship between the two speaker signals was evaluated on the basis of the three criteria (percentage accounted for, χ^2, and product-moment *r*) outlined above.

Table 4 summarizes the relevant findings. Comparable results were not obtained for the other speaker-state cue: initiation of a gesticulation.

When no auditor back channel was displayed in association with a unit of analysis, 84 per cent of the displays of the speaker-state signal followed the

TABLE 4. *Relationships between display of speaker within-turn cues, and subsequent display of the speaker continuation signal (no auditor back channels intervening)*

N cues displayed	*N* units having display	Proportion of units with continuation signal following display	Percentage of continuation signals following display of ≥1 cue	χ^2	Correlation
0	256	0.047			
1	264	0.159	84	20.48*	0.96
2	108	0.194			
Total	628				

* Corrected for continuity; $df=1$, $p=0.00005$.

display of at least one within-turn cue. The associated χ^2, pitting the display or non-display of the speaker continuation signal, against the prior display or non-display of a within-turn signal was 20.48. In addition, a correlation of 0.958 was found between (*a*) number of within-turn cues displayed (0, 1, or 2), and (*b*) percentage of units having a speaker continuation signal following that display.

Speaker continuation signal and auditor back channels

The effect of auditor back channels upon the subsequent display of the speaker continuation signal was assessed by applying χ^2 to contrast (*a*) the frequency of continuation-signal display preceded by a back channel, with (*b*) the frequency of continuation-signal display not preceded by a back-channel. Separate contrasts were made for when the preceding back channel was classified as 'early', and when the preceding back channel was classified as 'between unit'. A further differentiation was made on the basis of whether or not the back channel was preceded by the display of the speaker within-turn signal. Table 5 summarizes the relevant findings.

When the preceding auditor back channel was classified as 'between unit', no significant difference was found in the frequency of speaker-continuation-signal display following an auditor back channel, as contrasted with the frequency of display following no back channel. In fact, when no within-turn signal was displayed prior to the back channel, the opposite hypothesis received support: that the two rates of display are significantly similar.

TABLE 5. *Relationships between* (*a*) *display of speaker within-turn cues followed by auditor back channels, and* (b) *subsequent display of the speaker continuation signal*

N cues displayed	*N* back channel displays	Proportion of units with continuation signal display	χ^2 Contrast with no-back-channel display*	*p*
		Early auditor back channel		
0	36	0.278	20.95	=0.00005
>0	35	0.629	38.10	<0.00001
Total	71			
		Between-unit auditor back channel		
0	12	0.000	0.00	
>0	74	0.243	1.80	=0.18
Total	86			

* Corrected for continuity: $df=1$.

Different results were obtained from analysis of speaker continuation signals preceded by early auditor back channels. These early back channels had the effect of significantly increasing the probability of a subsequent speaker continuation signal. This increased probability of display was found, both when a within-turn signal preceded the early back channel and continuation signal, and when it did not.

DISCUSSION

The findings presented above suggest certain regularities in the display of specific behaviors by the speaker and by the auditor during speaking turns. These findings, based on data gathered from two dyadic, face-to-face conversations, led to the hypothesis of three signals: (*a*) the speaker within-turn signal, (*b*) the auditor back-channel signal, and (*c*) the speaker continuation signal.

Before the relationships between the three hypothesized signals could be analyzed, it was necessary to show a clear differentiation between the two actions an auditor might take in the turn system (as presently constituted): (*a*) shift to the speaker state, claiming the speaking turn, and (*b*) communicate in the back channel, an action that did not appear to involve claiming the speaking turn.

Auditor response to speaker turn and within-turn signals

It is interesting to note that the nature of the relationship between (*a*) the speaker within-turn signal, and (*b*) auditor back channels, appears to be identical to that found between (*a*) the speaker turn signal, and (*b*) auditor turn claims.

In both cases of speaker signal display and auditor response, the display of any single cue from the cue set was sufficient to constitute a signal. In both cases, the display of a speaker signal did not inevitably result in an auditor response: the auditor appears to retain distinct response options subsequent to the speaker's signal display. There was also found in both cases a strong linear relationship between the number of speaker cues displayed, and the probability of an auditor response. With regard to these three phenomena, the findings on speaker turn signals and on speaker within-turn signals appear to be mutually corroborating.

Speaker continuation signal

When no auditor back channel followed the display of a speaker within-turn signal, a relationship was found between this signal and subsequent display of one of the speaker-state cues (shift away in head direction). This finding led to the hypothesis that this cue, when displayed during turns, functioned as a speaker continuation signal.

The nature of the relationship between the speaker within-turn signal and a subsequent speaker continuation signal was identical to that found between

speaker signals and auditor response, as discussed immediately above. (In the case of the continuation signal, the relationship was, of course, between the speaker's within-turn cues and his own subsequent action.)

An auditor display of an early back channel had the effect of significantly increasing the probability of a subsequent speaker-state cue display, regardless of whether or not that early back channel had been preceded by a within-turn signal.

Auditor back channels occuring between analytic units had no apparent effect on the subsequent display of the speaker-state cue.

Segmenting of individual messages and of interaction

Joint consideration of the findings on the three signals hypothesized above, permits some further hypotheses concerning the organization of speaker–auditor interaction during speaking turns. These further hypotheses concern units that may serve to segment these turns.

Previous papers (Duncan 1973, 1974) have suggested the existence of two distinct, but interlocking, types of units operating in dyadic conversations: (*a*) units marked by the speaker alone, segmenting his speech, and (*b*) units of interaction, requiring the joint, coordinated participation of both participants.

The findings reported above may be interpreted to suggest the existence of units used by the speaker to segment his speaking turn. The speaker within-turn signal appears to mark the end of one such unit, and a subsequent speaker continuation signal appears to mark the beginning of the next unit. This unit, whose boundaries are marked by the successive display of these two signals by the speaker alone, may be termed the 'speaker within-turn unit'.

In contrast, the marking of interaction units would require the action of both participants. Evidence has been presented elsewhere (Duncan 1972, 1973) relevant to the speaking turn as a unit of interaction. As presently understood, a smooth exchange of the speaking turn requires the following ordered sequence of three events: (*a*) the speaker displays a turn signal (and does not conjointly display a gesticulation signal), (*b*) the auditor switches to the speaker state, displaying a speaker-state signal, and (*c*) the previous speaker switches to the auditor state, thereby relinquishing the turn. Omission of any one of these three steps results in no exchange of the speaking turn and/or simultaneous turns.

The notion of interaction units may be applied to the findings on speaker – auditor interaction during speaking turns. In this case, auditor back channels, as opposed to auditor turn claims, provide the auditor with a means by which to participate in the interaction. Through the back channel he may acknowledge his receipt and understanding – or lack thereof – of the speaker's message.

Thus, the following sequence of behaviors: (*a*) speaker within-turn signal, (*b*) between-unit auditor back channels, and (*c*) speaker continuation signal, may be hypothesized to constitute an interaction unit. This action sequence may

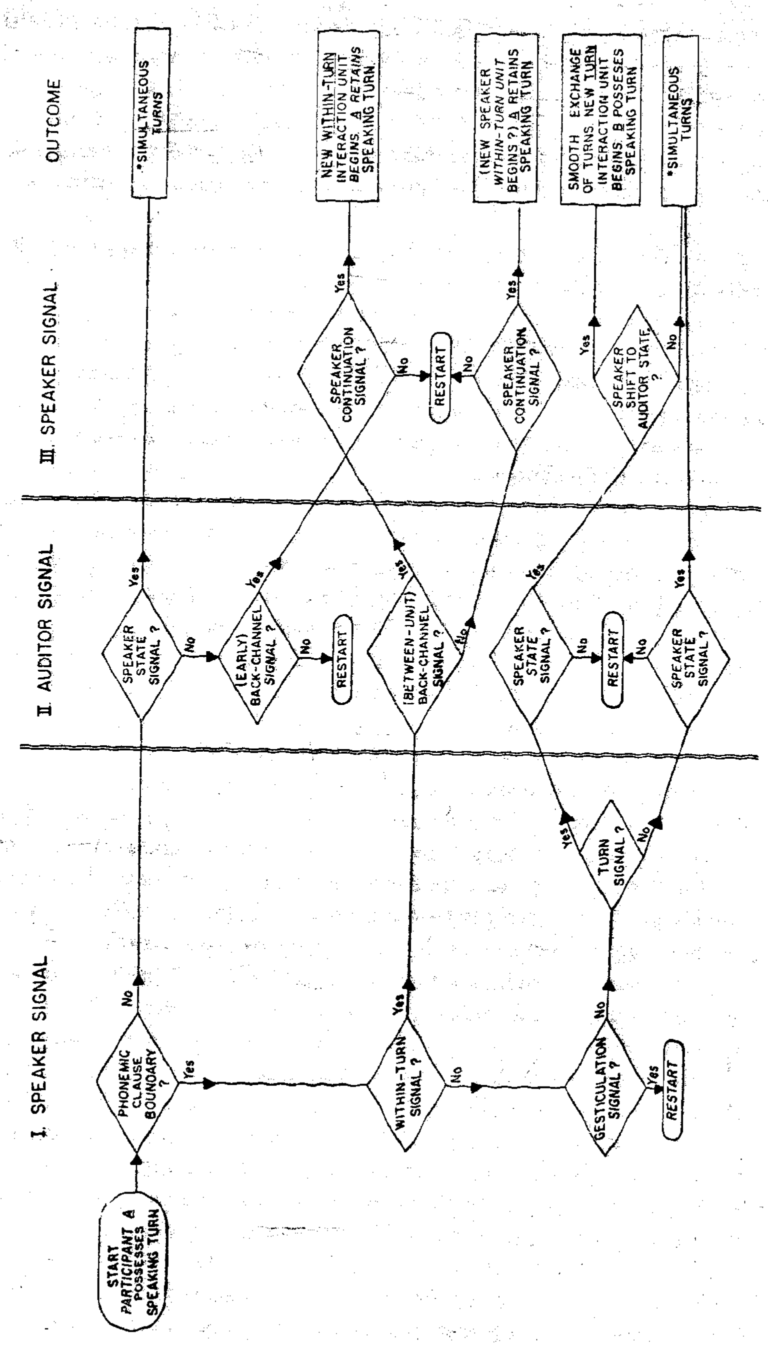

Figure 1. Logical model of the turn system.

be hypothesized as one means by which speaker and auditor may segment the conversation on an hierarchical level immediately lower than that of the speaking turn, creating what may be termed a 'within-turn interaction unit'.

The results suggest that a within-turn interaction unit may be created in at least one other way. This alternative sequence would involve only two steps (but still involve both participants): (*a*) an early auditor back channel, and (*b*) a speaker continuation signal.

The hypothesis of a two-step within-turn interaction unit is based upon the finding that early auditor back channels had the effect of significantly increasing the probability of display of a subsequent speaker continuation signal, whereas a between-unit auditor back channel did not. This increased probability of display was found, regardless of whether or not the early back channel was preceded by a speaker within-turn signal. (In contrast, Table 5 shows that, in our corpus, between-unit auditor back channel not preceded by a speaker within-turn signal, were never followed by a speaker continuation signal.)

It appears that both the display of an auditor back channel, and its location, may play a part in speaker–auditor interaction. That is, an early back channel may not be merely misplaced, but rather it may carry significant information for the interaction.

On the basis of the results, one may speculate that an early auditor back channel may indicate, not only that the auditor is following the speaker's message, but also that the auditor is actually ahead of it. Accordingly, it would be appropriate for the speaker to proceed directly to the next unit, regardless of whether or not he had displayed a within-turn signal. This 'skipping ahead' action by the speaker is not, however, automatic. It depends, apparently, both on the auditor's early back channel, and on the speaker's assessment of the situation.

In contrast, a between-unit auditor back channel would indicate that the auditor is following the speaker's message as it is developing. Therefore, it would not effect, either positively or negatively, the probability of an ensuing continuation signal.

By the same logic, a late auditor back channel would indicate some auditor acknowledgement, but also that he is not quite following the speaker's message. No analyses were brought to bear on this possibility, because late auditor back channels occupy the same position at which continuation signals are typically displayed. It was not possible, therefore, to test the probability of ensuing speaker-state signals.

Figure 1 presents a logical model of the turn system, as it is presently understood. No claim is made that the diagram represents a model of the actual perceptual and decision processes employed by the interactants.

Drawn to be read like a computer flow chart, the figure is designed to show graphically the hypothesized organization of signals and rules, in terms of the development of a single speaking turn. Each diamond represents a decision as to whether or not the indicated discrete signal is being displayed. As a result of that binary decision, a path is chosen leading from the diamond. Paths crossing vertical boundaries (shown as double lines) drawn to separate the actions of one participant from those of the other, connect signal displays that must occur in the indicated ordered sequence. Paths not crossing such boundaries carry no implication of ordered sequences.

The starting point, shown on the left, assumes that one of the participants (*A*) holds the speaking turn, uncontested. The outcomes shown at the right are based on those phenomena that have been actually observed in the interviews subjected to analysis.

This program of research was undertaken in an attempt to determine the extent to which behaviors in paralanguage and in body motion, together with those in language, might participate in highly structured, rule-governed social phenomena. The findings of definite regularities in the coordinate display of certain behaviors by the two participants in dyadic, face-to-face conversations, led to the hypothesis of a set of signals that operate within a coherent structure of rules. These results suggest the possibility of extending the analytic approach, so as to develop in time a body of systematic information regarding the organization of the larger communication context, comparable in specificity to that presently held regarding the organization of language.

REFERENCES

Bernstein, B. (1962). Social class, linguistic codes, and grammatical elements. *Language and Speech* **5**, 221–40.

Ekman, P. & Friesen, W.V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica* **1**. 49–98.

Duncan, S.D., Jr. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* **23**. 283–92.

—— (1973). Toward a grammar for dyadic conversations. *Semiotica* **9**. 29–46.

—— (1974). Language, paralanguage, and body motion in the structure of conversations. In W. C. McCormack & S. A. Wurm (eds.), *Language and thought*. The Hague: Mouton (1975).

—— & Niederehe, G. (1974). On signalling that it's your turn to speak. *Journal of Experimental Social Psychology*, **10**. 234–47.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica* **26**. 22–63.

Trager, G. L. & Smith, H. L., Jr. (1957). *An outline of English structure*. Washington, D.C.: American Council of Learned Societies.

Trager, G. L. (1958). Paralanguage: A first approximation. *Studies in Linguistics* **13**. 1–12.

Yngve, V. H. (1970). On getting a word in edgewise. *Papers from the sixth regional meeting Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 567–77.