A. Kendon, R. Harris + M. R. Key (eds)
Organization of Behavior in face-to-face interaction
(Mouton, 1975)

Interaction Units during Speaking Turns in Dyadic, Face-to-Face Conversations

STARKEY DUNCAN, Jr.

The problem of segmentation lies at the heart of structural research on social interaction. It is necessary to discover and to document basic structural units, having, in Scheflen's words, "(1) a given set of component parts; (2) a definite organization; and (3) specific location in a larger system" (1966: 271). As Scheflen (1966) further points out, this larger system is likely to be arranged hierarchically. A model of such a system and its constituent units would be that of phonological and syntactic units operating within a hierarchical language system.

The linguist's approach to structural units is to focus, insofar as possible, on the language behaviors (as that term is traditionally construed) of single individuals: his informants. But sociolinguists (e.g. Hymes 1969; Gumperz and Hymes 1972) have emphasized that language is typically used in a broader communication context. This broader context includes, among other things, both (a) one or more other persons with whom the language user is interacting, and (b) other behaviors, such as those in paralanguage (Trager 1958) and body motion, that are displayed conjointly with language.

The notion of units for social interaction may take into account this broader communication context. Within a given speech situation (Hymes 1972), behaviors in paralanguage and/or body motion may

This study was supported in part by Grants MH-16,210 and MH-17,756 from the National Institute of Mental Health, and by Grant GS-3033 from the Division of Social Sciences of the National Science Foundation. Susan Beekman, Jeanine Carlson, Mark Cary, Diane Martin, George Niederehe, Ray O'Cain, Thomas Shanks, Cathy Stepanek, Tichina Stubblefield, and Andrew Szasz contributed to the transcriptions and data analysis. I am indebted to Dick Jenny, Wayne Anderson, and the client, who generously consented to serve as participants in this study.

function, in an integrated fashion with those in language, to mark units. Further, it is possible that, in addition to units marked solely by the actions of an individual interactant, there may be other units that can be properly marked only through the coordinated actions of more than one interactant. It has been proposed (Duncan i.p.a.) that units of this second sort be termed "units of interaction."

Here I will discuss some findings leading to the hypothesis of units of interaction operating in dyadic, face-to-face conversations to segment speaking turns, thereby creating units on the next lower hierarchical level.

METHOD

The interviews and the methods of transcription used for this study have been detailed elsewhere (Duncan 1972). Briefly, detailed transcriptions of intonation, paralanguage, and body motion were made of both participants in two dyadic, face-to-face conversations, as recorded on videotape. The first nineteen minutes of each conversation were used.

One of the conversations was between a male of about forty years and a female of about twenty, who were previously unacquainted. The second conversation was between the male in the first conversation and a second male, also about forty. These two interactants had been good friends for a number of years. Both conversations were intrinsically motivated; that is, both would have taken place, regardless of whether or not they had been videotaped.

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 that lay in size between the phonemic clause (Trager and 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 at the ends of phonemic clauses

(though not necessarily those with rising or falling terminal junctures), which additionally were marked by the display of one or more of the following behaviors: (a) an unfilled pause; (b) the 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) an audible inhalation; (f) the use of any pitch level–terminal juncture combination other than 2 2| at the end of a phonemic clause; (g) a 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).

SPEAKING-TURN INTERACTION UNITS

The notion of "interaction units" was conceived as a result of findings on certain behaviors regularly associated with the exchange of speaking turns in dyadic, face-to-face conversations. To account for these observed regularities, a "speaking-turn system" (Duncan i.p.a, i.p.b) was hypothesized. A brief outline of the main elements of this system may provide a useful context for understanding the proposed interaction units during speaking turns.

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

Turn-System States

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.

The turn system was largely designed to explain the means by which the two participants coordinated their respective actions, so as to accomplish smooth exchanges of the speaking turn. In contrast to such smooth exchanges, both participants may simultaneously claim the speaking turn. This situation represents a breakdown of the turn system for the duration of the situation.

On the basis of research findings, a series of discrete signals and accompanying rules has been proposed (Duncan 1972; Duncan and Niederehe 1973) with respect to the exchange of speaking turns. Each of these signals is made up of a set of behavioral cues, also considered to be discrete. These signals, their constituent cues, and accompanying rules regarding the appropriate display of and response to these signals will be briefly described.

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 composed 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, g, h, i, j, 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 ($r = .96$) of the number of cues displayed, without regard to the specific cues comprising the display.

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 and 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. 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 vis-à-vis; and (b) initiation of a gesticulation. This gesticulation is the same behavior that had the effect of suppressing turn claims of the vis-à-vis, 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.

Speaking Turns as Interaction Units

It is clear that a smooth exchange of the speaking turn requires the appropriate, coordinated action of both participants. This smooth exchange involves 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, relinquishing the turn. Omission of any one of these three steps, or violation of their order, results in no exchange of the speaking turn and/or the simultaneous claiming of the turn by both participants.

This ordered sequence of actions of both participants may be said to move the conversation ahead on one hierarchical level: that of speaking turns. The exchange of the turn has therefore been hypothesized to constitute an interaction unit on that hierarchical level.

The concern of this study is with units on the next lower hierarchical level than that of the speaking turn. These lower-level units would be used to segment the speaking turn. Like the speaking turn, they will be interaction units requiring the action of both participants in dyadic, face-to-face conversations. For lack of a better term, they will be called "within-turn units."

SPEAKER-AUDITOR INTERACTION DURING SPEAKING TURNS

The most obvious behaviors observed during speaking turns are the head nods and vocalizations, such as "m-hm," "yeah," displayed by the auditor. These and related behaviors will be termed "back-channel behaviors," or more simply "back channels," after Yngve (1970).

An initial inventory was made of potential back-channel behaviors and their respective locations in our corpus. Classification of behaviors as potentially in the back channel was aided by the astute observations and research results of Dittmann and 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 355 separate auditor back channels and seventy-one separate speaker back channels.

339

Auditor Back-Channel Signal

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."

A. 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."

B. 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 . . . ."

C. 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, concerned with it?" S: "Possible that other people have . . . ."

D. 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 . . . ."

E. 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.

Auditor Back Channels and Claimings of the Turn

Within the turn system as it is presently formulated, there are two apparently different types of actions an auditor may take: he may (a) communicate in the back channel, or (b) shift to the speaker state, claiming the turn. The investigators cited immediately above were unanimous in their judgments that back-channel behaviors do not constitute a speaking turn. It was possible to lend further documentation to these judgments by exploring the differences in distribution of these two types of behavior in our corpus.

Data analysis yielded results suggesting that the various subsets of auditor back channels were similar to each other and different from claimings of the turn in four respects.

A. Auditor back channels did not tend to be marked by the speaker-state signal, whereas claimings of the turn did. Duncan and Niederehe (1973) reported that a two-by-two contingency table in which columns represented display versus nondisplay of the speaker-state signal, and rows represented auditor back channels versus beginnings of speaking turns, yielded chi-square values of 27.80, $df = 1$, $p = .00001$, for one interview, and 47.09, $p < .00001$, for the other interview.

B. Auditor back channels had a less orderly distribution than turn claims, with regard to the unit of analysis. Auditor back channels were located more frequently than expected during, and appreciably after, the units. Turn claims were predominately located immediately after such units. The chi-square associated with these distributions was 29.44, $df = 3$, $p = .00002$.

C. The auditor's display of back channels was unaffected by the speaker's display of the gesticulation signal, while turn claims were

almost entirely suppressed by that signal. For the two-by-two contingency table in which columns represented display versus nondisplay of the gesticulation signal and rows represented back channels versus turn beginnings, chi-square = 43.35, $df = 1$, $p > .00001$.

D. Turn claims and certain auditor back channels were related to two different sets of speaker cues. In Duncan (1972), turn claims were shown to be systematically related to the prior display of a set of six cues, comprising the speaker turn signal. In contrast, auditor back channels were found to be related to a set of two cues, comprising the speaker within-turn signal. This signal will be discussed in the next section.

Taken as a whole, the evidence points to multiple distinctions between auditor back channels and speaking turns. The findings in this respect are in full agreement with the observations of other investigators cited above.

Speaker Within-Turn Signal

It was found that a subset of auditor back channels was systematically related to the prior display of two speaker cues, typically occurring close to the end of units of analysis: (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 was sufficient to constitute the signal. The grammatical-completion 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 the speaking turn from instances of simultaneous claimings of the turn by the two participants.

Auditor back channels that occurred immediately after the end of a unit of analysis, or at the beginnings of the next unit (that is, slightly delayed), were related to the two speaker cues. Auditor back channels that occurred prior to the end of units of analysis were not so related. It is believed that this subset of "early" back channels was not related to the speaker cues, because these back channels too often occurred before the speaker cues were displayed.

Considering the between-unit and late auditor back channels, 88.8 percent occurred following the display of one or both of the speaker within-turn cues. A chi-square, applied to a two-by-two contingency table in which the columns represent the speaker's display versus nondisplay of the within-turn signal and the rows represent auditor's subsequent display versus nondisplay of a back channel, yielded a value of 39.31, $df = 1$, $p < .00001$. A correlation of .99 was found between the number of speaker within-turn cues displayed (0, 1, or 2), and the probability of an auditor back channel.

Under certain circumstances, the display of the speaker within-turn signal was related to the subsequent display of a speaker continuation signal, discussed immediately below.

Speaker Continuation Signal

The speaker continuation signal 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. The speaker continuation signal was composed of one of the cues of the speaker-state signal: shift away in head direction.

When no auditor back channel intervened between the end of one unit of analysis and the beginning of the next, a relationship was found between the speaker's display of within-turn cues at the end of one unit and his display of the continuation signal at the beginning of the next. It was found that 84 percent of the displays of the speaker-state signal followed the display of one or both of the speaker within-turn cues. A chi-square, applied to a two-by-two contingency table in which the columns represented the speaker's display versus nondisplay of the within-turn signal and the rows represented the speaker's subsequent display versus nondisplay of the continuation signal, yielded a value of 20.48, $df = 1$, $p = .00005$. A correlation of .96 was found between the number of speaker within-turn cues displayed (0, 1, or 2) and the probability of a subsequent speaker continuation signal.

When an auditor back channel occurred before the end of a unit of analysis, the probability of a subsequent speaker continuation signal significantly increased. This increase was observed, regardless of whether or not speaker within-turn cues preceded the back channel. The chi-square was calculated from a contingency table in which the columns represented display versus nondisplay of an early auditor back channel and the rows represented display versus nondisplay of the speaker continuation signal. In the case in which no speaker within-turn cues preceded the early back channel, the chi-square was 20.95, $df = 1$, $p = .00005$. In the case in which one or both speaker within-turn cues preceded the early back channel, chi-square = 38.10, $df = 1$, $p < .00001$.

In contrast, when an auditor back channel was located not before the

end of the unit of analysis, but rather between units, there was no accompanying increase in the probability of a subsequent speaker continuation signal. Table 1 summarizes the findings on speaker and auditor signals hypothesized to play a part in within-turn interaction units.

Table 1. 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	

DISCUSSION

The findings, briefly described above, led to the hypothesis of units of interaction serving to segment speaking turns into smaller units. These within-turn interaction units parallel speaking-turn interaction units in their structural characteristics: (a) both require the appropriate, coordinated action of both participants, and (b) both involve ordered sequences of action.

The beginning of a within-turn interaction unit appears to be marked by a speaker continuation signal, analogous to the marking of a turn beginning by a speaker-state signal. This beginning of a new within-turn unit appears to be associated primarily with either one of two preceding events: (a) a speaker within-turn signal, regardless of whether or not a between-unit auditor back channel intervened, or (b) an early auditor back channel, regardless of whether or not a speaker within-turn signal preceded it.

In view of these findings, and of the requirement that an interaction unit involve the actions of both participants, it may be hypothesized

that within-turn interaction units may be marked in either one of two alternative ways.

One of these alternative ways would involve an ordered sequence of three events: (a) a speaker within-turn signal, (b) a between-unit auditor back channel, and (c) a speaker continuation signal. In this sequence, the initial speaker within-turn signal is required, not only because of its relationship to the following continuation signal, but also because of the finding in our corpus that between-unit auditor back channels, not preceded by a within-turn signal, were never followed by a continuation signal.

The second way of marking a within-turn interaction unit would involve an ordered sequence of only two events: (a) an early auditor back channel, and (b) a speaker continuation signal. In this sequence the initial speaker within-turn signal is not required, because, when an early auditor back channel occurred, the display of a within-turn signal was found to have no appreciable effect on the probability of a subsequent display of either (a) the early back channel, or (b) a speaker continuation signal. The primary relationship was between the early back channel and the subsequent speaker continuation signal.

The within-turn interaction unit appears to provide the participants with a means by which to pace a speaking turn at a rate that takes both speaker and auditor into account. Through the back channel, the auditor may indicate his present condition with respect to what the speaker is saying. The auditor might, for example, use a request for clarification at those points at which he is not adequately following the speaker. On the other hand, the results suggest that an early back channel serves to indicate not only that the auditor is adequately 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 previously displayed a within-turn signal. This "skipping ahead" action by the speaker is not, however, automatic. It depends, apparently, both upon the auditor's early back channel and upon 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 affect, either positively or negatively, the probability of an ensuing continuation signal.

By the same logic, a late auditor back channel would indicate not only some auditor acknowledgement but also that the auditor 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.

An hypothesized interaction unit on a given hierarchical level has been discussed here. The next higher level — that of the speaking turn — has been considered elsewhere (Duncan i.p.a, i.p.b). It is hoped that continuing research will shed further light both upon units at these two hierarchical levels and upon possible units at higher and lower levels. Scheflen (1965) has proposed a higher level unit, termed the "position", that may be marked by behaviors related to posture shifts.

If research now in progress continues to confirm the findings presented here, then it would appear that interaction units may constitute an element of the structure of social interactions, at least in face-to-face dyadic conversations.

APPENDIX 1

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 the respective participants, 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 actually have been observed in the interviews subjected to analysis.

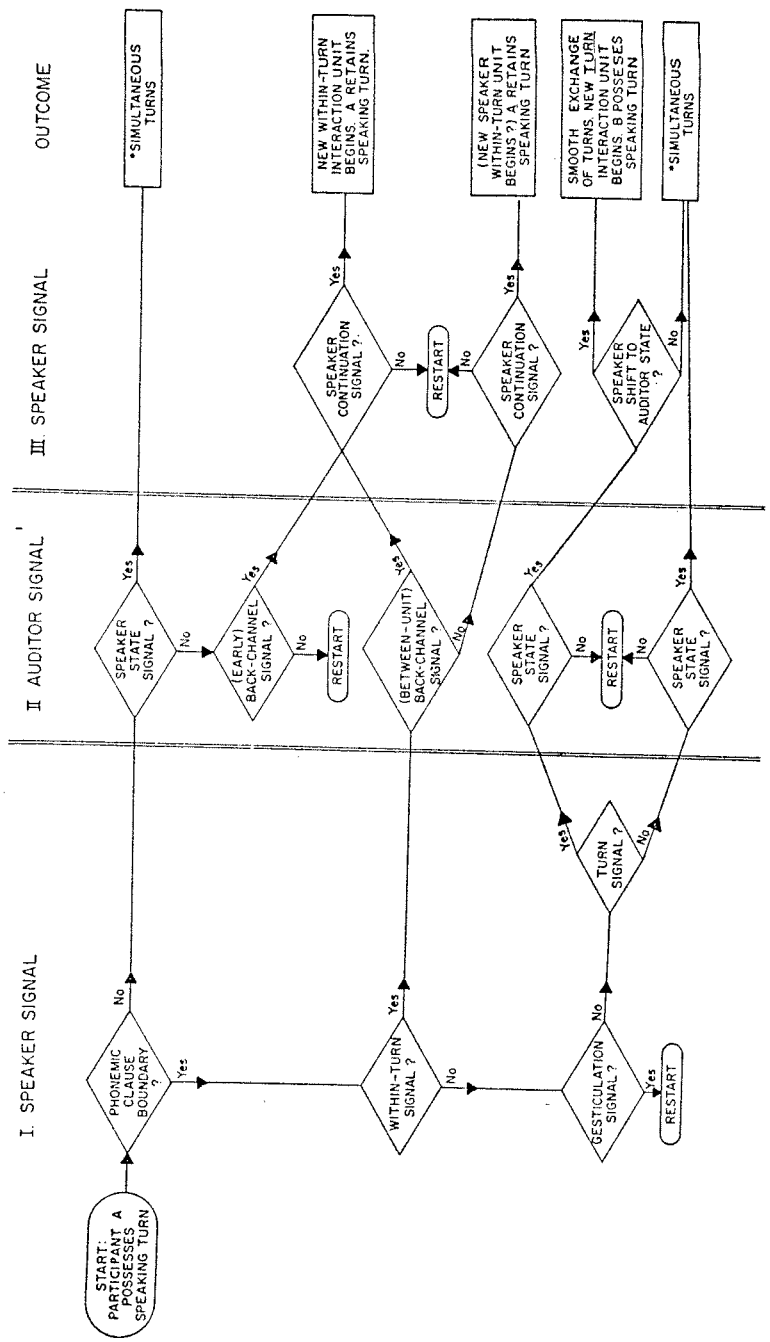

Figure 1. The turn system.

REFERENCES

BERNSTEIN, B.
1962 Social class, linguistic codes, and grammatical elements. *Language and Speech* 5:221–240.

DITTMANN, A. T., L. G. LLEWELLYN
1967 The phonemic clause as a unit of speech decoding. *Journal of Personality and Social Psychology* 6:341–349.
1968 Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology* 9:79–84.

DUNCAN, S. D., JR.
1972 Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* 23:283–292.
i.p.a Toward a grammar for dyadic conversations. *Semiotica.*
i.p.b "Language, paralanguage, and body motion in the structure of conversations," in *Language and thought.* Edited by W. C. McCormack and S. A. Wurm. The Hague: Mouton.

DUNCAN, S. D., JR., G. NIEDEREHE
1973 "A speaker-state signal within the turn system." Unpublished manuscript, University of Chicago.

FRIES, C. C.
1952 *The structure of English.* New York: Harcourt, Brace.

EKMAN, P., W. V. FRIESEN
1969 The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica* 1:49–98.

GUMPERZ, J. J., D. HYMES, *editors*
1972 *Directions in sociolinguistics.* New York: Holt, Rinehart and Winston.

HYMES, D.
1969 "Linguistics — the field," in *International encyclopedia of the social sciences,* volume 9, 351–371. New York: Crowell, Collier and Macmillan.
1972 "Models of the interaction of language and social life," in *Directions in sociolinguistics.* Edited by J. J. Gumperz and D. Hymes, 38–71. New York: Holt, Rinehart and Winston.

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

SCHEFLEN, A. E.
1973 *Communication structure: analysis of a psychotherapy transaction.* Bloomington: Indiana University Press.
1966 "Natural history method in psychotherapy: communicational research," in *Methods of research in psychotherapy.* Edited by L. A. Gottschalk and A. H. Auerbach, 263–289. New York: Appleton-Century-Crofts.

TRAGER, G. L.
1958 Paralanguage: a first approximation. *Studies in Linguistics* 13:1–12.

TRAGER, G. L., H. L. SMITH, JR.
1957 *An outline of English structure.* Washington, D.C.: American Council of Learned Societies.

YNGVE, V. H.
1970 "On getting a word in edgewise." *Papers from the sixth regional meeting Chicago Linguistic Society,* 567–577. Chicago: Chicago Linguistic Society.