

Analysis of Communication Among Speakers in Indonesian Context

¹Ediwarman Ediwarman, ²Syafrizal Syafrizal, ³Tatu Munawaroh

^{1,2}Lecturers at Sultan Ageng Tirtayasa

²Unma Pandeglang

Abstract

Operation of the "speaker state signal" in two-person, face-to-face conversation is hypothesized. The display of this signal by the speaker seems to indicate, among other things, that he is claiming a turn of speech, distinguishing this act from "backchannel" behavior in which he acknowledges only part of the speaker's message. The signal also appears to play a role in the resolution of situations in which both participants simultaneously claim speech turn. Signals are defined as the display of at least one of four behavioral cues, two in language and two in gestures.

Keywords: Signal, Time, Talk.

I. INTRODUCTION

In any society, whenever "the physical possibility of oral interaction arises, it seems that a system of practices, conventions, and procedural rules come into play that serves as a means to guide and regulate the flow of messages." This journal discusses a search program aimed at signal discovery and formulation. Rules relevant to the organization of two-person, face-to-face conversations. In particular, signals and rules have been sought that appear to facilitate the taking and dropping of speech turns by people in the version studied. The general outcome of participants' use of these signals and rules would probably be the regular exchange of speaking turns over time and the simultaneous minimization of turn claims by both participants.

The regular exchange of speech in the conversation has suggested that taking turns speaking is a universal language. Signals and rules Two speaker signals are hypothesized to produce a signal and (b) a motion signal. These two signals are treated in the analysis as discrete, i.e., displayed or not displayed at any given moment. It complements the notion of "speaker," which denotes a participant who does not claim a turn to speak at any given moment.

Turn-Yielding Signals These signals appear to mark points in the speaker's turn at which the speaker may act appropriately to initiate a turn swap. Each time the speaker attempts to take a turn without display of the turn-yielding signal by the speaker, simultaneous rotation results; In contrast, when the attempt by the speaker immediately follows the speaker's display of the turn-yielding signal, 92% of attempts result in a smooth turn-of-speech exchange.

A signal is defined as the display of at least one of a set of six behavioral cues, also considered discrete. These cues are indicative of certain behaviors in content—syntax, intonation, paralanguage, and gestures.

The signal does not have the properties of a green traffic light. That is, the speaker is not required to take a speaking turn on each display of the turn-yielding signal. However, it was found that the probability of the speaker taking a speaking turn increased as a positive linear function of the number of cues (O-6) that the speaker shared ($T = 0.96$).

Gesticulation Signals In the analyzed interviews, this signal from the speaker manifested the effect of inhibition of the speaker's attempt to take a turn, regardless of the number of turn-yielding cues that the speaker

simultaneously displayed. That is, the movement signal appears to negate the turn-yielding signal. The speaker hardly ever tries to take a turn while a motion signal is being displayed. In our data, such attempts are so infrequent that it is impossible to judge whether or not they are likely to cause simultaneous turns.

The movement signals consist of: the speaker's hand is involved in the hand movement, provisionally defined as "the movement of the hand is generally away" from the body, which usually accompanies, and which appears to bear a direct connection with, speech" [p. 2871, or (b) tense hand position,

The adapter is self-contained and does not operate as a cue in a motion signal. (A movable self-adapter in which one or both hands come into contact with one's own body, often with a grooming appearance, such as rubbing the chin or scratching the cheek. Analogously, an object adapter is a motion in which one or both hands come into contact with an object, such as a pipe or strap. tie.)

2. THEORY BASIS

2.1 BackChannel Behavior

Consider in turn-yielding papers "backchannel behavior." This term is meant to include verbalizations, such as "M-hm," and "yeah," and head movements, such as nods and shakes, so frequently observed on the part of speakers using the term "accompaniment signal" to refer to essentially the same behavior. In addition, there are in our conversations a number of short phrases or clauses that also seem to have a common character. These longer tail lines seem to fall into the following three classes. (In the example, "S" stands for speaker, and "A" stands for speaker.)

a. Sentence completion.

Not infrequently in our material, a speaker will complete the sentence the speaker has started. In such a case, he will not continue after a brief settlement; the native speaker will continue his turn as if undisturbed. Completion of sentences has been reported independently. Example: S: ". . . eventually, it will boil down to more concrete issues. , ."; A: "Like he became more comfortable." S: "And I feel that. . . ."

b. Request for clarification.

In contrast to sentence completion, completion is a brief request for clarification. Such requests are usually accomplished in a few words or phrases. Example: S: ". . . somehow they were better able to handle it." A: "You mean this anxiety, concern for it?" S: "Maybe someone else has. . . ."

This backchannel behavior is similar to sentence completion, except that it restates in a few words immediately the previous thought expressed by the speaker. Example: S: ". . . have to take the pieces;" A: "broken plate, huh;" S: "%ut then a very. . . ." It is important to note that the definitions of the different types of backchannels have stood out. Examples have been given, of which the reader might generalize, based on his or her knowledge of the culture. This intensive approach to definition has become the basic approach to date. It may also be noted that, with the exception of head nods and shakes, all of the above definitions are based on the verbal form of speech.

BackChannel Speaking Turns Whether they show understanding, or lack thereof, agreement or disagreement, backchannel speakers seem to imply approval of the conversation paying attention to the speaker's message. They seem to provide the speaker with a mechanism to actively participate in versioning, thereby facilitating general coordination of action by both participants within the structure of the conversation.

Furthermore, in our data, many of these backchannel speeches come to points in the middle of a speaking turn where the speaker might legitimately claim the turn. These channels often give the impression that the speaker is actively avoiding the turn, keeping it in line with the current speaker.

Among those who have commented directly on the matter, there has been unanimity in the assessment that backchannel behavior, in itself, is not an alternation. When the speaker is using the backchannel, there seems to be a mutual understanding that the speaker is maintaining his turn and that he will continue as soon as the backchannel finishes. It seemed that both participants considered the speaker's backchannel as an end goal ing automatically at the completion of the relatively short speech involved. On the other hand, the turn of speech

can be continued correctly to the point where the speaker signals his willingness to give in. This overview is complicated, however, by the observation in our conversations that for some of the longer backchannels, particularly brief restatements, the boundary between backchannel and turn talk becomes uncertain. Intuitively, some of the longer rear channels seem to take a quality turn. These observations suggest that verbal form descriptions are not sufficiently strong in general to distinguish back tract from the turn of speech.

At this point, we decided to look for a signal that would serve to clearly distinguish the beginning of the speech change from the start of the channel. The use of such signals in association with the speaker's intervention at any point in the conversation will allow the speaker to immediately recognize the intervention as either the beginning of the turn or the back of the speaker. No complicated judgments

3. METHOD

The data for this study were derived from detailed transcriptions of speech and gesture behavior during the first 19 minutes of the two conversations, as recorded on videotape. Both conversations were between two people, sitting face to face at an angle of approximately 65°. This is the same transcription. Used in previous studies.

Designated conversation No. 1 in this paper is an entrance interview held at

Counseling Research Center. The therapy applicant was a woman in her early twenties who worked as a secretary and who had not yet graduated from college. The intake interviewer was a 40-year-old male, an experienced therapist, who had been conducting this kind of interview for many years. The two participants previously did not know.

Conversation designated No. 2 was between the same man who had participated in the review in the first conversation and the second male therapist, also 40 years old. These two participants are good friends and have known each other for about ten years. Their interaction was relaxed and full of jokes and laughter. Both conversations will occur whether or not they are recorded. The transcription of this conversation

has been described at length elsewhere. The notation C is made from speech intonation and one paralinguistic, and almost all observable gesture behaviors. All of the behavior recorded in the transcription was related to the syllables of speech emitted by the participant who happened to be speaking at the time or to the pauses between two syllables.

Transcription was made at least one year before this study was conducted. Similarly, all early assessments of the speaker's turn and backchannel had been encoded in transcription long before the conception of this study. As indicated above, this assessment is based on the verbal form of speech.

Analytical Procedure Exploratory Examination Conversation No. 2 was carried out to find a set of behavioral cues that could potentially serve as signal components functioning to distinguish the back channel from the start of the turn. As a result of visual inspection, several behaviors are seen to occur regularly near the start of an alternating speech. This observed behavior was then more rigorously examined by systematic analysis, described below, of each speaker's backchannel and each turn starting from Conversation No. 2. The results of this process lead to the speaker state signal hypothesis, consisting of a set of four behavioral cues. This hypothesis is then subjected to initial validation by testing the operation of the signals in Conversation No. 1. Analytical procedures and cue definition resulting from the visual inspection were applied in an identical manner to both transcriptions. It should be emphasized that, once one of these cues is defined on the basis of visual inspection, its identification in transcription for data analysis purposes is completely routine. The exact location of the start and end of each of the four behaviors has been specifically noted in the transcription and designated by its own special code. Since the start of the turn must be voiced, only the voiced back channel of the speaker is included in the analysis. Examples of simultaneous turns are not included in the analysis since it is desirable to consider only the beginnings of turns where the turn system appears to be operating in the correct way.

To achieve the desired differentiation, the speaker state signal must be displayed around the start of the turn but not directly around the backchannel of the speaker. Therefore, a speech

interval is defined, applicable to all turn start and back channels of the speaker, during which the display of possible speaker state cues will be considered.

In determining this speech interval, the same unit of analysis as previously reported was used. This unit boundary is defined as occurring at the confines of a phonemic clause that is additionally characterized by the display of at least one of a set of nine behaviors in syntax, intonation, paralinguistic, and body movement. In our conversation, these units averaged approximately eight syllables—word in length.

For purposes of the preliminary examination, speech intervals should be considered for: speaker state signal display defined as a series of four units of analysis: two units immediately prior to the initial initiation of the turn or backchannel speaker, and two units immediately following the initiation.

However, as the examination progressed, it became apparent that the prospective speaker cues were closed around the turn initiation starts. Therefore, the results reported below for both conversations are based on intervals defined as extending from one unit before turn initiation or speaker channel through the syllable that carries the main intonation stress in the first phonemic clause. (In some cases, the first phonemic clause of a turn consists entirely of phrases such as "Yah, uh." When this occurs, the interval is defined as an extension through the stressed syllable of the turn's second phonemic clause.)

Exploratory Analysis Conversation No. 2 leads to the hypothesis that signals are typically displayed with the start of the speaking turn but not with the backchannel of the speaker. The name "state signal" was chosen because the signal can be said to mark the points at which participants shift from speaker to speaker state. For exposition purposes, it will be argued in this paper that the speaker displays a signal of the speaker's state, even though the signal is interpreted as an indication that the speaker has, at that point, shifted to the state of the speaker.

Signals are defined as the display, in a particular speech interchange, of at least one of the four behavioral cues listed below. The display of these cues is considered discrete. That is, the cue is considered to be displayed or not displayed at a certain moment.

1. Shift towards the head. This gesture is considered to have occurred when the speaker shifts his head away from pointing directly at the speaker. Gestures are based on a shift in the direction of the head, not just the fact of the direction of the head. In the case of the start of a speech turn or the speaker's backchannel where the shift occurs before the specified interval (described above), the signal is not considered to be displayed.

2. Inhalation is heard. This cue is defined in terms of sharp, audible breath intake.

4. RESULT

Early Turnover Differentiation of Back-Channel Speakers presents a summary of the findings to what extent four behavioral cues distinguish rear-channel speakers from early-turnover in our data. For each of the two conversations, the findings in each cue, considered alone, are shown. The line labeled "Speaker-State Signal" shows the results for signals defined as displaying at least one of the four signals. The line labeled "Head Shift and Gesticulation" shows comparable results for signals defined as displaying at least one of the two sets of indicated cues.

In conversations subjected to exploratory analysis, (four cues)

Table 4.1 : *Display And Individual*

Cues displayed	Turn beginnings marked by cue display		Back channels marked by cue display		χ^2 *
	N	P	N	P	
Conversation no. 2: Exploratory analysis ^b					
Head shift	17	.85	10	.12	41.71
Inhalation	10	.50	0	.00	41.35
Gesticulation	10	.50	4	.05	24.96
Overloudness	11	.55	6	.07	24.00
Head shift and gesticulation	19	.95	12	.14	47.09
Speaker-state signal	19	.95	16	.19	38.92
Conversation no. 1: Validation ^c					
Head shift	21	.34	0	.00	12.33
Inhalation	3	.05	0	.00	0.43
Gesticulation	21	.34	1	.03	9.72
Overloudness	12	.20	3	.09	0.97
Head shift and gesticulation	38	.62	1	.03	27.80
Speaker-state signal	44	.72	3	.09	30.61
Total	63	.78	19	.16	

* Corrected for continuity; $df = 1$; $p(.001) = 10.83$.

^b Turn beginning $N = 20$; back channel $N = 85$.

^c Turn beginning $N = 61$; back channel $N = 32$.

The speaker status signal is displayed at the start of the turn 95% and 19% of the speaker's backchannel. The chi-square distribution associated with this portion of the signal is 33.92 ($p(1) < .00001$). Invalidation of the signal conversion is shown at 72% of the start of the turn, and at 9% of the backchannel of the speaker, with an associated chi-square of 30.61 ($p(1) < .00001$).

Settlement of Concurrent Turn Claims As noted above, examples of simultaneous turns are not included in the data shown in Table 1. However, the results suggest the possibility that cues in the speaker state signal may play a role in the resolution of simultaneous turn cases. In this context, "resolution" refers to the process by which one of the two participants "wins" and an instance of a simultaneous turn, i.e., whichever comes up by speaking the turn. This issue has not been investigated previously in this study because it is assumed that variables related to participant status may strongly influence the resolution rather than the behavior displayed in the interaction.

The research question is framed in the following way: For what? To what extent is the resolution of the simultaneous turn samples in our data predicted from the respective display by the two cue participants in the turn-yielding and speaker-state signals? To examine this question, the transcriptions of the two conversations were considered together due to the small total number (19) of the examples of simultaneous turns in our data. As with the speaker's backchannel and early turn, all examples of simultaneous turns have been recorded in the transcription, previously with the concept of this study. The scoring system for the display of cues associated with simultaneous turns is formulated as follows: Each speaker states cue and the displayed cue is assigned a score of +1; each turn-yielding cue, a score of -1. The plus and minus values are then added up for each participant in each instance of a simultaneous turn. For each such example, the sums thus obtained for the two participants are then compared. The participant with the larger number is predicted to win the conflict over the turn.

None of our exploratory transcriptional analyses with regard to simulations was performed prior to the formulation of this scoring system. This system, to be more precise, is suggested by a

similar scoring system designed in previous studies. To predict the direction of experimental bias (Rosenthal, 1966), and with the added effect of displaying a turn-generating cue on the likelihood of the speaker trying to take a turn, mentioned above. The interval to be considered for the display of the relevant signal, in this case, is defined as the extension of one unit of analysis before the start of the simultaneous turn through the last syllable of the simultaneous turn. As a result of applying the addition procedure to 19 simultaneous turn instances, it was found that there was one case (5% of the total) where the sums obtained for the two participants were the same, thus making predictions impossible. Of the remaining 18 examples (95% of the total) where predictions could be made, all 18 predictions were correct.

Applying a subset of the two cues of the speaker state cues reported above (Initiation of the gesture and shift towards the head), together with the six turn cues to 19 simultaneous turn examples, yielded identical results to the full four cue speaker status signals. These predictions can be compared with those made only on the basis of the respective speaker and speaker states of the two participants immediately prior to the start of each simultaneous turn example. Participants who have been at the speaker 19 events (63%) Inhalations seen on videotape, but not transcribed as audible, are not included. Audible and/or visible inhalation does not appear to be a physiological need to initiate turn speech.

3. ~ni~ the gestation. This cue is considered to have occurred when the speaker initiates a movement within the specified range of intervals. The definition of gestation, and its distinction from self and object adapter, is described above with respect to the signal of gestation. 4. Paralinguistic overload. This behavior corresponds to what is termed "too loud intensity." This cue is noted to have been displayed at points where the speech syllables involved in the speaker or at the beginning turn have been transcribed as having at least one degree of loudness. High-pitched over-paralinguistic transcription is not considered an example of this cue.

5. SUGGESTIONS

Talk versus Channel Like turn-yielding signals, speaker status signals is defined as displaying at least one of their constituent cues. Taken as a whole, the speaker status signal appears to mark an initial high turn proportion and is rarely displayed with the speaker backchannel.

The hypothesized signals derived from exploratory analyzes appear to have generally been validated when applied to one further conversation. The chi-square used to evaluate signal operation in each of the two conversations yields values comparable to those found for turn-yielding signals (Duncan, 2000).

Despite the fact that the supported speaker state signal hypothesis was ported at a high level of significance in both tested interviews, the results show that there is no perfect match between (a) display signal and (b) speaker backchannel classification and early turn. It is interesting to consider the possible sources of the error in the match that occurred. Assuming that ongoing research continues to support the speaker-state signaling hypothesis operating within a specifiable specification of dyadic type, face-to-face conversation, three sources of mismatch immediately come to mind.

1. The most obvious, there may be errors in the initial classification

speaker intervention, some of which are classified as backchannel (mainly specifically "short restatements") may have changed, and vice versa. These classifications were made, it will be recalled, intuitively in the early stages of research. Should the speaker state hypothesis be? The signal continues to be validated in further studies; the display or absence of signal by the speaker with respect to the intervention will be the basic criterion for classifying that intervention as turn or backchannel speech.

2. Not all signs

the signal has a very different back channel of the speaker from the initial turn (an unlikely event); it is impossible to claim that the inventory of cues for the signal is complete, either in the context of the language being studied or in the conversation that has not yet been analyzed. On the other hand, this argument cannot be used as an excuse for imperfect results. The claim that the cue list is not

exhaustive becomes interesting only at the point where cues are further documented.

3. Participants who fail to display the expected signal are committed to a performance error. For one or more of the many possible reasons, the participant simply made an error with respect to the signal display. Although such errors may occur separately from time to time for any given participant, systematic, repeated errors of this kind are expected to promote misunderstanding and possible disruption in the conversation. Indeed, as the organizational elements of the conversation are increasingly documented, it may become possible to find and describe with high specificity many sources of difficulty or communication disorders.

Each of the four proposed speaker state cues has been previously discussed with respect to the phenomenon of conversation. Analyzing a 5 to a 9-minute sample of seven dyadic conversations, it was found that 76% of "long utterances" (lasting five sets or more) were initiated with the speaker looking the other way or just before the start of the utterance. These results are approximately comparable to those reported for this cue, despite the fact that Kendon uses somewhat different definitions of (a) speech turns and (b) the interval at which the cues should be calculated.

Behaviors that are similar to our voices inhalation cues may be a signal where an interactant may "disturb" or indicate he or she wants to say something" [p. 575]. As described above, cues by speakers have been found to be an active element in suppressing the auditor's attempts to take a turn. The function of gesture cues in speaker state signals is consistent with previous findings of hand gestures.

Eltzer, Morris, and Hayes (2000) found several significant correlations "between the percentage of successful interrupts in the discussion, (a) the change in the vocal amplitude of the interrupt recipient from before to during the interrupt, and (b) the difference between the interrupt and the interrupt amplitude during speech. Simultaneously" [p. 392]. The figures in Table 4.1 show that (a) the two paralinguistic cues— inhalation are heard, and overloudness—are rarely used in validation conversations, and (b) that the body-head gesture cues shift, and the gestures appear to work almost just as effective in most respects as a full four-cue set. In light of

these results, it might be argued that, on the basis of parsimony, two gestures of gestures became hypothesized to signal the state of the speaker.

Several considerations, however, contradict this hypothesis of a more limited definition of speaker state signals. (a) Both praline guistic cues are significantly active in the exploration of paragraph No. 2. (b) Inhalation cues are perfectly discriminated against starting from the auditor's backchannel in both conversations; and increasing in its use may have greatly increased its statistical significance in Conayat No. 1. (c) Cue overloudness finds some support in the findings of Meltzer, Morris, and Hayes (2010) with regard to simulated speech tan. Although some caution must be taken in comparing Meltzer et al. study with this one because of some differences in methods, the two sets of each result are consistent. And (d) the greater behavioral diversity of the four cue sets allows both for greater flexibility in appearance and for easier adaptability to individual differences in personal conversational styles. Therefore, it would appear desirable to continue to include the four hypothesized cues in further research looking for these signals.

In general, the results show that there is less display of full speaker status signals in Conversation No. 1 than in Conversation No. 2. Encouraged by these findings, examination of the post hoc transcription of Conversation No. 1 shows that many of the turns that begin are not signaled by signals that occur during the latter part of the transcription when the conversation has taken on psychotherapy that is clearly poetic tones. During this section, the client seriously explores the personal difficulties that have led him to seek counseling, and the therapist responds in a therapeutic way. Signals appear to operate more reliably in earlier parts of transcription when conversations involve a more straightforward exchange of information.

These observations suggest that some shifts in the use of speaker status signals may be one characteristic of the organizational nature of certain types of psychotherapeutic interactions, distinguishing them from the more typical everyday conversations for which signals are sought. This is, of course, an empirical question that can be partially answered by currently continuing investigations of the more mundane kind. Simultaneous It seems that the relative balance of speaker state cues and turn cues, as

presented jointly by each participant, is associated with the means of turn case resolution. Simultaneous. Prediction based on the display of these cues is perfect, except in one case where no prediction can be made. The evaluation of these results, however, should be softened by the fact that they are based on only 19 instances of simultaneous turns. This finding, therefore, is regarded as particularly suggestive.

Fully consistent with the general function of the signal in each signal; (b) that the number of appropriate turn cues shared by a speaker was significantly related ($T = 96$) to the probability of the auditor's turn claim; and (c) that in experimental communication bias studies the summation procedure, which includes negative and positive values for intonation and paralinguistic behavior, was found to be effective in predicting object sub performance on the Duncan experimental task, Rosenberg & Finkelstein, 2000). The findings obtained at the resolution of the simultaneous turn samples are fully consistent with these considerations.

The findings reported above suggest that speaker status signals may: serve as clear behavioral markers that the auditor earlier in the conversation had shifted from auditor to speaker status and thus claimed a turn. This signal has the advantage of being independent of the verbal content of the response, thereby simplifying the process of recognizing the character of the auditor's response. Furthermore, the signal appears to contribute to the resolution of the sample turn taneous claim simulation by the two participants. Taken together, these findings suggest a central role in the organization of dyadic, face-to-face conversations.

Reference

- [1] ELTZER, L., MORRIS, W. N., & HAYES, D. P. Interruption outcomes and vocal amplitude: Explorations in social psychophysics. *Journal of Personality and Social Psychology*, 18, 392402.
- [2] MILLER, G. A. Review of J. H. Greenberg (Ed.), *Universals of language*. *Contemporary Psychology*, 2000, 8, 417418.
- [3] ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 2001.

-
- [4] SCHEGLOFF, E. A. Sequencing in conversational openings. *American Anthropologist*,
 - [5] 1999. 76, 1075-1095.
 - [6] SULLIVAN, H. S. *Conceptions of modern psychiatry*. New York: Norton, 2001.
 - [7] TRAGER, G. L.
 - [7] TRAGER, G. L., & SMITH, H. L., JR. *An outline of English structure*. Washington, D. C.: American Council of Learned Societies, 1999.
 - [8] YNGVE, V. H. On getting a word in edgewise. *Papers from the sixth regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 2000, 567