appendix E
Rhythms of Vocal Sound

Paul Byers
Teachers College
Columbia University

INTRODUCTION

Mike, the youngest child in this family, has been diagnosed as disabled, and his speech is often unintelligible. Varenne has shown that he is a participant in the family's conversation, even though other members of the family do not always understand how Mike's sounds are related to their talk. Sometimes he appears to be simply ignored by the person he speaks to. Such moments have interesting features, some of which Varenne has discussed elsewhere.

When he was transcribing the taped conversation, Varenne noticed that the child's "mom" (at line 91 in the transcript) fell within a pause in Connie's speech. He asked me to examine the placement of the "mom" with my waveform analysis technology, and I found that the sound fell quite precisely *where it should be* in the conversational sequence. To say "where it should be" implies that there is an organization, that is, that there are "rules" for vocal sound placement in the conversation sequence.

In the following pages I will summarize my research into the organization (sequencing) of vocal sound placement and illustrate much of it with technologically generated "waveform" examples. More formal reports of various aspects of this research have been published (Byers, 1972, 1976, 1988). In summary I will show that:

- There are *two* superimposed rhythms (± 10 cycles per second) underlying all human speech and accompanying gestural body movement. One is an inferable, fixed biological rhythm related to brainwaves and motoric oscillations of the intercostal muscles. The other is the variable

performed sequence or rhythm of vocal units in the conversational stream. The second rhythm is a "modulation" of the first.
- In conversation the participants synchronize their underlying fixed biological rhythms.
- There are implicit "rules' governing the management of this rhythmic interplay both for individual speakers and for conversations.
- There is information significant to the conversation encoded in the relationship between these two rhythms.
- These processes are completely out of the explicit awareness of speakers, although they are accessible as "feelings about the conversation."
- There is evidence that *listening* to speech also involves rhythm-matching between speaker and listener.
- The rhythm phenomena underlying speech can also be found in the body movements and gestural relationships in human interaction.

Although a rhythmic regularity underlying speech has been suspected and efforts have been made to determine its rate or frequency (e.g., Lenneberg, 1967; Jaffe & Feldstein, 1970), these efforts did not recognize Stetson's "motor syllable," a vocal sound unit described below, or the underlying fixed rhythm and the individual (performed) modulations of that rhythm. This awaited an appropriate technology and a nonlineal multilevel structural or cybernetic paradigm with which to study patterns of relationship.

This organized behavioral netherland of rhythms underlying speech lies between physiological activity (brainwave oscillations and motor impulsing of the diaphragm and intercostal muscles), on the one hand, and the vocalization of language on the other. The organization of this vocal sequencing is independent of any particular language, but it is an infrastructure without speech or conversation as we know them would be impossible.

PARALLELS

Before I begin to describe the organization of this unfamiliar behavioral infrastructure, it will be helpful to look, first, at some analogous but more familiar organizational matters in nature.

First, there is a recognizable history of discovering organized processes that were unavailable for study until either technological or methodological advances made the discovery possible. Before the discovery and use of the microscope, it was not possible to observe and thus to recognize the significance of cells and microbiological processes as infrastructural components of body tissue (or "health"). Cells are independent of tissue, but there can be no tissue without an organization of cells.

Musical "Conversation"

The interpersonal activity of musicians playing chamber music together is a musical "conversation" that has an analogic relation to the rhythmic relationships and the "rules" I will describe. For example:

- There are two (superimposed, simultaneous) rhythms in performed music. There is the underlying regular (fixed-mechanical) beat of a particular tempo, and there is the *performed* sequence of notes. Since musicians do not play "mechanically," the performed notes vary slightly from the fixed-mechanical beat.
- No listener actually *hears* the fixed rhythm or beat, but both the musicians and their audience implicitly "know" the fixed beat and can "get into synch" with it. The regular beat-tempo is inferred from the slightly irregular note sequence.
- Each player's notes are carefully time-related to both the preceding notes of the other player(s) and the tempo-beat. It would not be musically possible, for example, to record the individual parts separately and then merge them into a musical performance.
- When a player varies the place of a note in relation to the fixed beat, the variation is called *expression*. That variation (expression) contains information about both the individual and the conjoint performance.
- Maintaining the overall tempo-beat despite the individual (expressive) variations requires continuous "self-correction." There is a parallel here with the tightrope walker who maintains his or her stability by making continuous corrections.

Radio Broadcasting

In radio or TV broadcasting there are also two kinds (levels) of frequency (rhythm is commonly used for slower perceivable trains, and frequency for faster ones).

- There are the light or sound frequencies of the studio *performance* (which enters the TV camera and microphone), and there is the broadcast frequency or carrier wave on which the performed frequencies ride as "modulation."
- Between the broadcast and the receiving, the "transmission" is both everywhere in space (the receiving aerial can be anywhere in broadcast range) and is invisible and inaudible.
- The receiver is tuned to the carrier frequency, but that channel-frequency is never heard. It is only inferable from the tuning.

- The broadcast information lies in the *relationship* between the fixed carrier frequency and the variable, unpredictable performance frequencies.

The hologram is another example of this coded informational relationship (between a reflected laser beam and a fixed reference beam) on a microscale. A very macroexample is the matter of saying, "Meet me at five o'clock," implying (without explicit awareness) a common reference to the earth's 24-hour rotation (a fixed underlying rhythm).

Language Study

Language study is concerned with the multilevel organization of articulated sounds into semantically meaningful strings (phones/phonemes, morphs/morphemes, syntactic sentences, etc.). Also in language:

- There is an invisible (and inaudible) set of shared grammatical rules that organizes the semantic content.
- The performed (i.e., spoken) utterance may or may not follow the "rules of grammar," and the utterances may not take the form of syntactic sentences, but speaker and listener will conjointly "correct for" these variations.

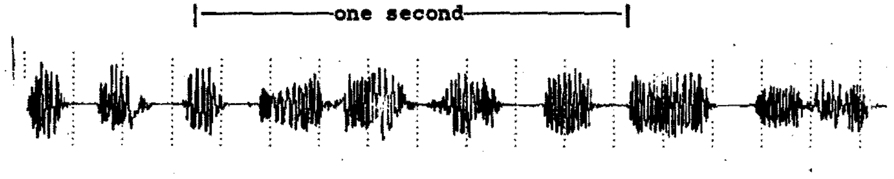
THE RHYTHMS AND THE RELATIONSHIPS

Neurologists have long understood that motor behavior is a serially organized sequence of discrete "packages," and that the temporal unit of that organization is the single "pulse" of brainwave oscillation.

During speech the diaphragm and intercostal muscles move air upward across the vocal apparatus in motoric bursts, waves, or pulses that are timed or temporally organized by the oscillations related to brainwaves. The regularity of this impulsing underlying speech was shown in Stetson's *Motor Phonetics* (1951) in which he published electromyographic records of intercostal motor activity during speech. Stetson's interest lay in rhythms of intonation, poetry, and rhythmic patterns of languages and he conceived the "motor syllable" as a minimal vocal unit. In 1903 Stetson wrote a small monograph in which he argued that rhythm was a thing apart from the content that accompanies it, and in 1905 he wrote, "The most important natural rhythm-producing apparatus is the vocal apparatus" (p. 257).

Forty-six years later Lashley (1951), after a lifetime of brain research, wrote "The Problem of Serial Order in Behavior," in which he said,

Figure 1. Motor syllables in two-man Bushman talk.



> the mechanism which determines the serial activation of the motor units is relatively independent, both of the motor units and of the thought structure. (p. 118)

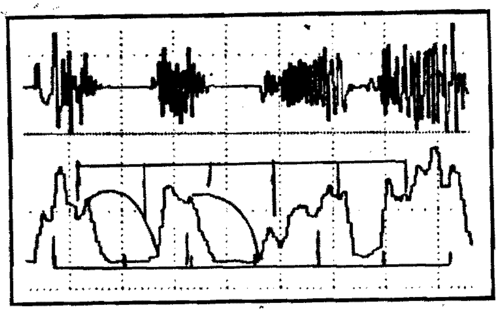
If there is a fixed rhythm underlying the motor activity of speech, one might expect that speech would emerge as a train of regularly spaced vocal packages or "motor syllables." Humans, however, do not speak in fixed mechanical rhythms. *Individuals* "expressively" vary the placement of vocal sounds, but the underlying rhythm remains stable.

The waveform below is from a Bushman conversation, and it shows, indeed, an *almost* equally spaced series of motor syllables, but a closer examination shows that they are not as regularly (mechanically) spaced as brainwaves or intercostal motor pulsing.

The waveform trace above displays the sound packages (motor syllables) of *two* speakers. The recorded event was not the usual conversation but a "storytelling" by the group's most acclaimed storytelling pair. To the Bushmen these two men perform a traditional story, and a tribal audience "listens." One of the storytellers "tells" the story with specific traditional gestures, and the other "accompanies" the story with frequent and traditional interjections and gestures. This is an oral literature that everyone knows by heart, including the gestures. These men are acclaimed for their precise "duet" skill. Below are the sequence of steps, observations, and a summary of the analysis and findings that eventually emerged from the study of the Bushman record and other recorded data from about 20 languages:

- I began this research intending to study the rhythms of *movement* in human interaction by examining film records without sound. I found that there was much synchronous movement between the two men and that points of gestural change occurred about every 10 frames (± .4 sec). (I also found this in monkey interaction.)
- If one slowed the film and counted or tapped at the visible onset or

Figure 2. Peak-to-Onsets.

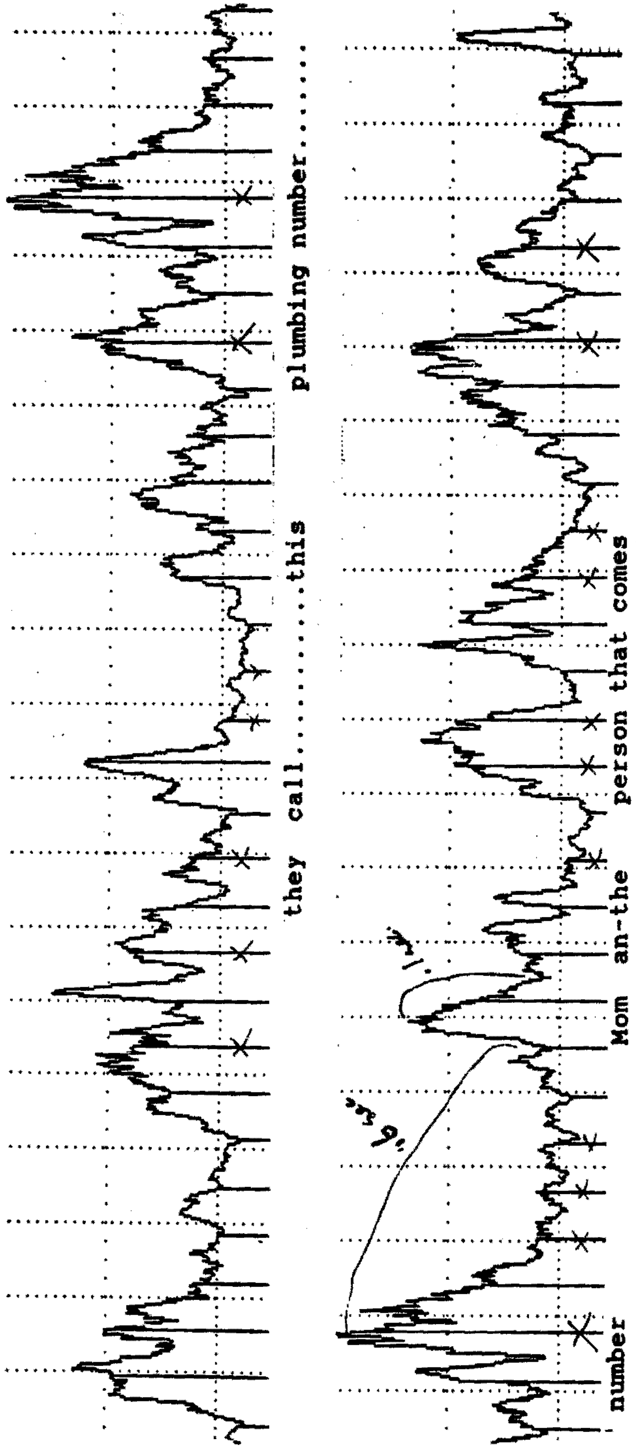


change points, there was clearly a rhythm that persisted across the entire story despite intermittent small discontinuities.

- When the recorded speech sound of this performance was fed into a chart recorder, the trace looked much like the one above. The vocal units (motor syllables) were usually about .2 secs from onset to onset. No motor syllable onset followed the preceding one by less than .2 secs, but motor syllables, according to "rule," begin (onset) *at a beat point.*
- There was clearly "rhythm" in the vocalization, but it was often briefly interrupted at each interjection or speaker switch.
- When the "listener" interjected, the onset of his interjection was .1 sec *after the preceding sound-peak.* When the first speaker continued, he also began .1 sec after the loudest sound of the other man. The chart above illustrates this in a fragment of the Bushman waveform which has been converted into a single-line trace. This segment is almost completely syncopated-overlapped.
- There was a common minimal rhythmic unit of .1 sec from the sound-peak of one speaker to the onset of the next, and a regular .2-sec interval between the onsets (motor syllables) of *uninterrupted* speakers.
- If the common ± .1 sec were projected through the entire sentence-unit, from the initial sound of one unit to the beginning of the next, the *overall* fit was quite precise despite the frequent rhythm deflections. The initial sound of the following sentence-unit began precisely on a beat point.

The capabilities of the technology, and many of the procedures and findings ("rules") in this research, can be illustrated by displaying a waveform printout and describing the successive analytic steps involved. I will do this with the waveform representing the talk in which the "Mom" occurs in Varenne's taped record. Figure 3 represents 5.46 secs of talk with the "Mom" at about 3.38 secs into the displayed chart.

Figure 3. Single-line trace of talk before and after "Mom."



The right end of the upper trace is repeated on the lower one so that the onset of "Mom" and the peak of the preceding sound can be [illegible] break. The curving lines drawn on the chart show the intervals between the peaks of one speaker and the onsets of the next.
The underlying rhythm is shown by the vertical lines below the traces.
Since motor syllables require a minimum of .2 sec between onsets, at least half the marks are expectedly unrelated to onset points. [illegible] true in Connie's speech, since she tends to "drawl" or space her onsets more than the minimal .2 sec.

WAVEFORM AND ANALYSIS

Acquiring the Waveforms

The waveform data file is acquired by using an analog-to-digital board inserted in a microcomputer slot.[1]

First, a selection of a record is made for waveform acquisition, in this case the talk surrounding the "Mom" interjection.

When the A-to-D board is activated by computer keyboard commands, the segment is played and seen scrolling across the computer screen, before acquisition begins, so that the "volume" can be adjusted (separately for each channel in case of multi-channel acquisition) and a samples-per-second rate can be selected. I typically use 400 samples per second, which displays the common .1-sec interval across about half an inch on the printout. It is possible to acquire data at 4,000 samples per second and to acquire up to 16 channel inputs simultaneously.

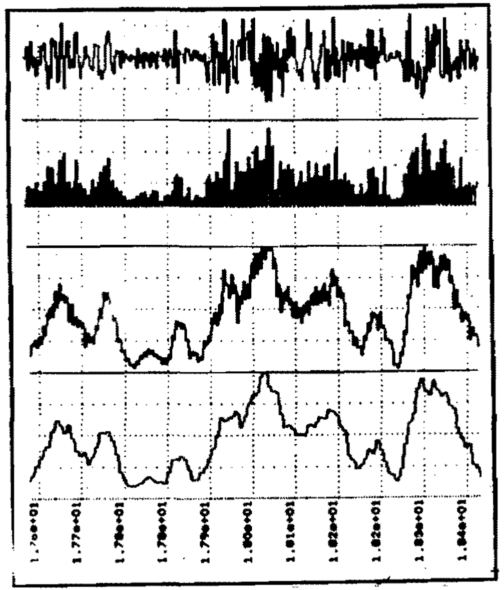
Data acquisition begins when one key is pressed and ends when another key is pressed. During acquisition, pressing the space bar will insert a mark in the record at that point. In the waveform record here, I marked the "Mom" as it was spoken during acquisition. This mark appears only on the acquired bipolar waveform trace. Markers are also useful as separators if one acquires more than one discontinuous segment in the same acquisition record.

After the waveform file is acquired, it can be retrieved on the computer screen for examination and manipulation by activating a second board in the computer, a "waveform scroller card."[2] The following are some of the possibilities:

- The initial record is a bipolar waveform similar to the top one in Figure 4. This waveform can then be rectified, integrated, and a single-line trace made with any selectable amount of detail or smoothing by specifying the number of samples to be averaged. The figure above is a much-reduced segment of a printout showing, from the top, a bipolar trace, an integrated trace, and two single-line traces with different degrees of smoothing. The numbers at the bottom show the time in seconds from the beginning of the file (the "e+01" moves the decimal point one digit to the right). These traces are created as additional but parallel records and can be examined simultaneously or separately on the screen. Up to four can be printed out simultaneously. Any segment of the originally acquired waveform file can be bracketed for printing or copied to a separate file.

When the waveform file is retrieved onto the screen, it can be enlarged vertically or compressed horizontally any selectable amount. The chart in Figure 3 was compressed by a factor of 2 for printing (and still further reduced to accommodate book-page size). The interval between any two points in the trace on the screen can be measured to the nearest thousandth of a second. On the printout there are numerical notations every half inch that show the progressive times from both the beginning of the acquired record or from the beginning of the segment being printed.

Figure 4. Waveforms.



---

[1] The "analog and digital I/O board for IBM-compatible computer systems" is a DT2811 PGL made by Data Translation, Inc. 100 Locke Drive, Marlborough, MA 01752. There are many such boards. This one is relatively low cost and compatible with the waveform scroller card. I also use a DT707 screw terminal panel which provides a direct connection for up to 16 single-ended or 8 differential simultaneous inputs. Both the converter and the waveform scroller card and software are designed for scientific work and have capabilities far beyond those required for my research.

[2] The waveform scroller card and CODAS software are from Dataq Instruments, Inc., 825 Sweitzer Ave., Akron, OH 44311. The manuals for the configuration and use of the technology and software are detailed, including information required for special programming, but they presume a knowledge of engineering concepts and language. Dataq, however, offers apparently unlimited, patient, and friendly telephone support. (Note: After these charts were made, Dataq produced an upgrade package which now supports VGA color cards and, optionally, microchannel architecture. A separate A-D board is not now required and the many added capabilities include a full- or half-page printer screen.)

Finding the Underlying Fixed Rhythm

I began work on this waveform printout by marking those points (on a strip of paper below the trace) where the trace moved suddenly upward, as they would at motor syllable onsets. This is a rough approximation, since the trace can move upward at places other than motor onsets. When the same interval, or an even multiple, begins to emerge repeatedly (this can be seen by sliding the strip with the tentative onset marks back and forth under the trace, to see which intervals on the strip fit other intervals in the trace), another strip can be made of these intervals (of about .1 sec) and this rhythm can then be marked across the entire record despite pauses or onset variations.

Since each motor syllable requires a time interval of .2 sec but could be .3 or even .4 sec, it cannot be expected that each one-tenth mark will coincide with an onset point. At least half of the .1-sec marks will fall *inside* the .2-sec motor syllables. Furthermore, the single-line trace may not accurately reflect the onset if, for example, it begins with an unvoiced or sibilant sound. Also, speakers can be expected to vary their placements "expressively." The fixed .1 sec rhythm marked at the bottom of the waveforms in Figure 3 will show the relation of the fixed rhythm train to the waveform. Careful inspection of the relation of the marks to the printed trace above shows that the underlying rhythm is not "lost." It persists even across the brief speaker switch interruption introduced by the "Mom."

After arriving at the rhythm interval inherent in the particular record and marking it across the printed segment, I marked a likely beginning onset point at the extreme left of the chart and another at the extreme right. This interval, measured on the screen, was 5.315 secs. Then I counted the number of these ±.1 sec-intervals from the first to the last. In Figure 3 there were 53 of these. By dividing the total time by the number of units (53), I arrived at a rhythm rate of .1014 sec. Then, using a waveform program that permits the placing of visible marks on the screen (and the subsequent printout), I placed a mark on the trace every .1014 sec, beginning with the left-most onset. Those marks are the vertical marks on the bottom of the trace.

Speaker Switching: The Peak-to-Onset Rule

When one speaker follows another or interjects vocal sound into the speech of another, this "next" speaker does not hear the preceding onset. He or she hears the *sound-peak* of the motor syllable, and his or her next motor syllable onset is entrained by the loudest preceding sound. The "next" or injected vocal onset will begin an exact beat interval after that sound peak. This was illustrated in Figure 2 and it was this "rule" that enabled me to say that the child's "Mom" was "where it should be." I have noted the .6 sec peak-to-onset interval (from the sound peak of Connie's "number" to the onset of "Mom") on the chart.[3]

In the Bushman storytelling performance, the "listener" quite frequently interjected sounds into the ongoing talk of the storyteller, and most of his interjected sounds (motor syllables) begin or onset exactly one .1-sec-beat unit after a preceding sound peak. The Bushmen sounds are so closely intertwined and often overlapping it was not possible to discover in detail how the underlying fixed rhythm was maintained despite the frequent rhythmic interruptions.

When Connie (the speaker before and after the interjected "mom") continued after the "Mom," she was "obliged" by the peak-to-onset rule to place her first onset in a rhythmic beat relationship to the *sound peak* of the "Mom." This could, then, "reset" the underlying rhythm. She did, indeed, begin her onset one beat unit (.1 sec) after the peak of the "Mom" sound, and the waveform chart shows a brief appearance of that new "reset" rhythm (which is out of phase with the original rhythm by about a twentieth of a second). But within a half a second her original rhythm was back in place. The waveform chart shows the relation of the "new" rhythm and the original one.

Self-Correction: The Double Onset

If the "next" speaker begins in relation to a preceding sound-peak, that speaker's onset is out of phase with the first speaker's underlying rhythm, since sound peaks are determined by the phonetic shape of the word spoken. Yet the underlying fixed rhythm or beat is not lost. This would, at first, seem paradoxical except that *conversations* are "self-correcting."

The "self," here, is not an individual. Self-correction is systemic and refers to the *conversation* as a self-correcting "system." It would not be possible for any speaker to purposefully manipulate these 25- to 50-millisecond intervals.

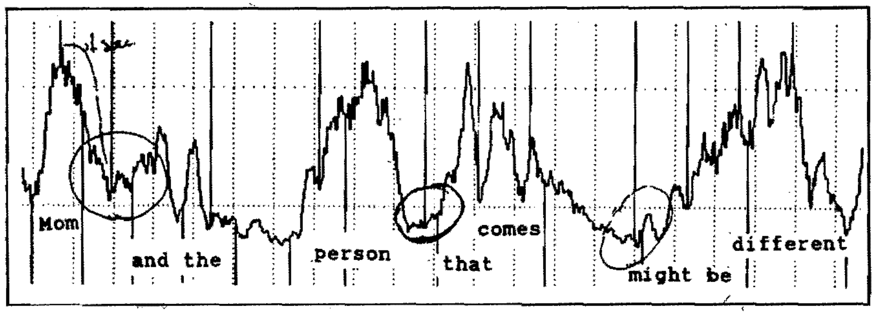
In order to "see" this rhythmic correction in greater detail, the following chart, Figure 5, displays about two seconds of the talk-sound following the "Mom."

The vertical lines at the top begin, from the left, at the peak of the "Mom." The next vertical line marks the beginning of Connie's sound .1

---

[3] Since the interval between onsets or between peaks and onsets may be any multiple of the basic .1-sec rhythmic unit, and since the perceived "rate of speech" is unrelated to the underlying rhythm, I sometimes use the analogy of the escalator to clarify the relationship. The escalator runs at a fixed speed. Riders can get on any step but not between steps, and the speed of the escalator is unrelated to how much any one step can carry.

Figure 5. Double onsets and "self-correction" after "Mom."



sec later. Subsequent marks are at .2-sec intervals (the minimal motor syllable interval). There is a noticeable upward movement of the trace at these points, suggesting that the new "reset" rhythm persists in the trace for a few beats.

The vertical lines at the bottom mark the original .1-sec fixed rhythm (as in the earlier chart). I have, in this chart, omitted the .1-sec marks that fall within motor syllable units. As one can see, Connie began her sound (motor syllable) one beat interval after the "mom" peak and *began again* on her original rhythm. This is what I have called the "double onset," which appears to be the beginning of the correction process. This double beginning after the "Mom" can, in fact, be detected in her voice, if one listens carefully, as a very brief "stutter." Then, after the motor syllable that began with the double onset (and lasted .2-sec plus .1-sec hesitation), Connie's following *two* motor syllable onsets also have a double onset appearance. At the end of the chart segment, however, Connie is clearly back on her original rhythm. The "echoes" of the double onset have disapproved.

To me this self-correction process in human speech interaction is particularly interesting, because many interactional systems in nature are self-correcting, but it is rarely possible to examine them from recorded data and at level of organization on which they are neither (a) solely biological (they involve variably performed human behavior), nor (b) matters of conscious human intention.

The Yanomamö of South America, whose ethnographer called them *The Fierce People* (Chagnon, 1968), have an unusual form of "self-correction" involving vocal sound placement. Chagnon writes,

> Yanomamö culture calls forth aggressive behavior, but at the same time provides a regulated system in which expressions of violence can be controlled. (p. 118)

If their normally aggressive talk reaches a point of incipient violence, they will move into a special form of dialogue in which each repeatedly

shouts a few syllables of challenge at the other. They almost seem to be shouting at the same time, but a close examination of the waveforms shows that each man's onset begins .1 sec after the other man's preceding onset. Since the syllables are shouted, the onsets are also the peaks. At first their vocal alternation is not very precise, and there is much "double onsetting," but as they continue these three- or four-syllable talk bursts, they become more precisely "tuned" to each other and the precise vocal interlock is an intricate vocal dance. The semantic level of talk expresses ferocity, but the underlying vocal sound relationship is (organizationally) "intimate" and violence is avoided. This is a "social" level of self-correction. These repeated bursts of precisely interlocked shouting also imply self-correction of the underlying rhythm. I have shown one of these bursts with the interlock and double onsets below in Figure 6. The marks from the top (in .2-sec intervals) are those of one man, those below the other man. There is a .2-sec pause between bursts.

A Second Rhythm

I have said, here, that there is a ±.1 rhythm underlying speech and conversation. I have, in fact, sometimes found a .133-sec (i.e., 7½ cps) underlying rhythm in some records. When I began to look at waveforms of talk recorded from radio or television broadcasts or from videocassettes of movies (because of the technical cleanliness of such records), the underlying rhythm has always been ±7½ per second. There is, in addition to the well-recognized 10 cps (alpha) brainwave, also a 7½ cps (theta) brainwave. I have found this rhythm in the speech of a Maring (New Guinea) invoking ancestors to a ceremony, in a Yanomamö (South America) shouting for others to help him, and I have found instances where "stuttered" speech will be on the expected 10 cps before and after the stuttered repetitions but with the stuttered repetitions spaced at 7½ cps. It is my present guess that

Figure 6. Yanomamö Chant-talk.

the 10 cps rhythm is found when the interaction is between a speaker and an immediate other or others, and that the 7½-cps rhythm is found when the "other" is an impersonal audience.

Observations from Other Records

In a videotape of six high school students who "spontaneously" laughed together at the successful completion of a complicated task, I discovered that there were six serial "laugh peaks" in each .2-sec interval. Laughing, however it may be regarded by conversation analysts, is a highly organized conjoint activity—like dancing at a disco. At a disco the pacer or zeitgeber is the continuous beat of the music, whereas in laughter, once begun, the conjoint behavior is mutually self-entraining.

When I examined a much-magnified trace of audience laughter recorded from the audience in a theater, I discovered that the onset of the laughter was, indeed, placed "where it should be" in relation to the last stress-peak of the preceding joke, but also that *all* the laughter peaks on the tape were triggered by (and themselves triggered) other laughter strictly according to the peak-to-onset rule. That is, everyone in the audience was "organizing" everyone else. Everyone was unknowingly participating in a conjointly organized event. Perhaps the "fun" of laughing together comes from participating in such a highly organized interpersonal event.

On the other side of the entrainment coin are occasions of rhythmic *dissonance*. I will describe three examples:

When speakers have "speech problems" that produce arrhythmias such as stuttering, stammering, or articulatory impairments, a listener experiences the discomfort of repeatedly having to reset or retune his *listening* rhythm. It is difficult to listen to speech with rhythmic discontinuity. When a blind student complained of the difficulty of listening to the mechanically generated voice of a electronic "reader" for the blind, I examined waveform printouts of the reader's "speech" and discovered that it was, indeed, constructed as a sequence of phonetic units, but that the designers had used a 16-cps interval (for which there is no human brainwave match). The student complained of the strain and fatigue engendered by listening to this "mechanical" speech.

I examined many speech tapes made of conversation in family therapy sessions with the same client couple, who complained of "communication problems." Whereas the husband, alone or in dialogue with the therapist, showed the expected ±10 cps rhythm, the wife's speech always had an underlying rhythm of about 8.7 cps in her speech *or in dialogue with the therapist*. This mismatch could clearly be identified in waveforms. The first impression, listening to dialogue with the wife, was one of conversational mismatch with much interruption and overriding. This raises a question: Are there individuals who, for whatever reason, have a rhythmic "tuning disability"? This matter certainly deserves further study.

DISCUSSION

My interest in the rhythmic infrastructure of language is only incidentally related to the study of language. My subdiscipline is *communication*, which I define broadly as the processes by which any two pieces of the universe find their relation to each other. From this point of view language, however important to our species, is an evolutionary veneer lying atop a relatively unrecognized and unstudied domain of interpersonal information management that existed long before hominids and language. My rationale for this view comes from the observation that, after centuries of research into human behavior and human relations (including the study of language), there are libraries of research reports but no significant advance in our management of human interpersonal relations. I believe that we have had neither the equivalent of a microscope nor the methodological (or paradigmatic) perspective with which to study the "infrastructural" domain in which important human information is coded in frequency or "rhythm" relationships (intervals) measurable in milliseconds.

I have explored these rhythms and their relationships in order to "deconstruct" the study of human relations, and I have reached two personal conclusions. The first is that it will be more productive to examine how humans "tune" to each other than to see humans as "doing things to each other." The Yanomamö interlock tuning described above suggests an "esthetic" quality to human relations. The second is that the *semantic* exchange in talk is a relatively unproductive domain in which to examine human interpersonal "tuning."

I believe that the most significant "finding" in the research I have described is that it points to an unexplored domain rich in possibilities for understanding an important aspect of human affairs. I do not believe that research in this domain will yield practical answers to the kinds of questions that our cultural beliefs and our past research have encouraged us to ask, but it will certainly lead us to ask new kinds of questions. For example:

- Can interpersonal "communication problems" be effectively resolved by our present language-based procedures?
- Can the present understanding of our behavioral sciences adequately describe or explain interpersonal *attraction* (or its opposite)?
- Might not the "affective domain" (the wastebasket of cognitive science) better be understood as a domain involving an interpersonal esthetic?
- Did hominid precursors have no language because the vocal apparatus

was insufficiently refined to produce the phonetic distinctions required by language, or because the carrier-wave linkage had not evolved to organize the sounds?

The vocal/interpersonal rhythms domain lies between or across the "biological," on one hand, and the "cultural" on the other, such that a clear distinction disappears. In this domain the interpersonal event, the conversation, systemically corrects itself, requiring a linkage between the autonomic and the unperceived cultural. In this domain the self-correction process is uncontaminated by cognitive intent or purpose. Here, perhaps, we can discover the process by which a flaw can sometimes turn self-correction into pathology or addiction, a matter of increasing importance in human affairs from personal self-esteem to international relations. This matter cannot easily be studied by examining live humans who are contaminated by (cultural) learning or by examining (biological) immune processes, which are largely inaccessible and only slightly understood.

I believe that the rhythm research, as I have described it, is not yet grounded by a clearly encompassing design. I have offered only bits and pieces of observation and deduced a few "rules." Often an unexpected small discovery of pattern has required a major revision in my effort to understand the design of the larger interpersonal process.

As Margaret Mead pointed out in her address as retiring president of the AAAS, we are now beginning a "hard" behavioral science in which we work with recorded data that can be archived and shared, with observations that can be precisely delineated, with a technology that permits a new order of analysis, and with results that can be replicated with the same or other data.

We are beginning a science of human behavior or human relations based on "how humans work" instead of the earlier "how we *believe* they work."

References

Byers, P. (1972). *From biological rhythm to cultural pattern: A study of minimal units.* Unpublished doctoral dissertation, Colombia University.

Byers, P. (1976). Biological rhythms as information channels in interpersonal communication behavior. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology II.* New York: Plenum. (Also in S. Weitz. *Nonverbal communication: readings with commentary* (2nd ed.). New York: Oxford, 1979.)

Byers, P. (1988). Toward a cultural epidemiology of emotion. In M. Clynes & J. Panksepp (Eds.), *Emotions and psychopathology.* New York: Plenum.

Chagnon, N. (1968). *Yanomamö: The fierce people.* New York: Holt, Rinehart and Winston.

Jaffe, J., & Feldstein, S. (1970). *Rhythms of dialogue.* New York: Academic.

Lashley, K. S. (1951). The problems of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley. (Also in S. Saporta (Ed.), *Psycholinguistics.* New York: Holt, Rinehart and Winston, 1961.)

Lenneberg, E. H. (1967). *The biological foundations of language.* New York: Wiley.

Stetson, R. H. (1903). Rhythm and rhyme. *Psychological Review Monographs, 4,* 413–466.

Stetson, R. H. (1905). A motor theory of rhythm and discrete succession I & II. *Psychological Review, 12,* 250–270, 293–350.

Stetson, R. H. (1951). *Motor phonetics: A study of speech movements in action.* (Published for Oberlin College.) Amsterdam: North-Holland.