

# START

## MICROFILM COLLECTION OF MANUSCRIPTS ON CULTURAL ANTHROPOLOGY

FORMERLY: MICROFILM COLLECTION OF MANUSCRIPTS ON  
AMERICAN INDIAN CULTURAL ANTHROPOLOGY  
AND  
MICROFILM COLLECTIONS OF MANUSCRIPTS  
ON THE MIDDLE AMERICAN CULTURAL ANTHROPOLOGY

Series: **XV** \_

No: **95** \_

Photographed by:

Department of Photoduplication - The Joseph Regenstein Library  
University of Chicago - Chicago, Ill. 60637

REDUCTION RATIO: **12** \_

THE NATURAL HISTORY  
OF  
AN INTERVIEW

(edited by Norman A. McQuown)

with contributions by

Gregory Bateson  
Ray L. Birdwhistell  
Henry W. Brosin  
Charles F. Hockett  
Norman A. McQuown  
Henry L. Smith, Jr.  
George L. Trager

Part I: Foreword, Chapters 1 - 5

MICROFILM COLLECTION  
OF  
MANUSCRIPTS  
ON  
CULTURAL ANTHROPOLOGY

No. 95 ~  
Series XV

University of Chicago Library  
Chicago, Illinois  
June 30, 1971

The Natural History of an Interview

Foreword	(McQuown)
Chapter 1 : Communication	(Bateson)
Chapter 2 : Vocal Activity	(Hockett)
Chapter 3 : Body Motion	(Birdwhistell)
Chapter 4 : Implications for Psychiatry	(Brosin)
Chapter 5 : The Actors and the Setting	(Bateson)
Chapter 6 : Transcript, Transcription, and Commentary	(Birdwhistell, Hockett, McQuown)
Chapter 7 : Communicative Base Lines and Symptomatic Features	(Birdwhistell, McQuown)
Chapter 8 : The Psychiatric Overview of the Family Setting	(Brosin)
Chapter 9 : Collation	(Birdwhistell, Brosin, McQuown)
Chapter 10 : Summary, Conclusions, and Outlook	(McQuown)
Appendix 1: Paralanguage: A First Approximation	(Trager)
Appendix 2: Symbology for Speech Transcription	(Hockett)
Appendix 3: Symbology for Speech Transcription	(McQuown)
Appendix 4: Microkinesic Transcription	(Birdwhistell)
Appendix 5: Macrokinesic Transcription	(Birdwhistell)
Appendix 6: Sample Kinesic Transcription	(Birdwhistell)
Appendix 7: Retranscription and Collation Charts	(McQuown)
Appendix 8: Machinery	(McQuown)
Appendix 9: Citations from the Literature	(Brosin)
Appendix 10: Bibliography	(Brosin)

### Foreword

- 2 -

In the month of November, 1955, at the Center for Advanced Study in the Behavioral Sciences, at the instigation of Dr. Frieda Fromm-Reichmann, then staff-psychiatrist at Chestnut Lodge, Baltimore, Maryland, Norman A. McQuown, participant with her and others (Alfred L. Kroeber, Gregory Bateson, Charles F. Hockett, Dr. Henry W. Brosin, Ray L. Birdwhistell, David M. Schneider) in an on-going seminar investigating the relevance of language behavior to research in other behavioral sciences, undertook to prepare a fine-grained analysis, transcription, and interpretation of the speech of participants in a tape-recorded psychiatric interview. The results of this attempt were made available in an article "Linguistic Transcription and Specification of Psychiatric Interview Materials" published in PSYCHIATRY in 1957 (Vol. 20, pp. 79-86).

Encouraged by the relative success of this initial endeavor, a team consisting of Fromm-Reichmann and Brosin, psychiatrists, Birdwhistell, kinesiologist, Hockett and McQuown, linguists, and Bateson, anthropologist, undertook, in February, 1956, to carry out a similarly fine-grained analysis, transcription and interpretation of the speech and body motion of participants in a sound-filmed (and tape-recorded) family interview. In this second effort, although one member of the filmed family was currently undergoing psychotherapy and the family environment was suspected of being schizophrenogenic, the focus was shifted from a strictly psychiatric one (psychiatrist and patient) to a less clinically oriented one (interviewer and members of the family). During a six-month period of part-time study progress was made independently by psychiatrists and linguists, and during a three-month period of intensive work on the filmed and taped materials, the kinesiologist joined these for collation. In brief meetings of the team during 1956 to 1959, work continued, as it did in the home-institutions of the team-members. Fromm-Reichmann died in April 1957 and Hockett could no longer continue as a member of the team after 1960. Bateson turned to communicational research on dolphins, although continuing to maintain contact with other team members.

In three institutions, work continued in a systematic way: At Eastern Pennsylvania State Psychiatric Institute, under the direction of Birdwhistell (with the collaboration of Dr. Albert E. Scheflen), at Western Pennsylvania State Psychiatric Institute, under the direction of Brosin (with the collaboration of a sizeable team, including Dr. Felix F. Loeb, Jr., Dr. William Charney, and Dr. William C. Condon), and in the Department of Anthropology at the University of Chicago, under the direction of Norman A. McQuown (with the collaboration of William M. Austin, Raven I. McDavid, Jr., and Dr. William Offenkrantz, and, most recently, of Starkey Duncan).

The focus of effort at EPPI was on the description of body motion and on the macro-structure of interactional behavior. The focus at WPPI was on details of body-motion and their correlation with clinical history and on the creation of adequate machinery for the efficient manipulation of sound-film materials. The focus at Chicago was on the analysis and recording of speech and on the pulling together of all the materials bearing on the analysis of the family-film materials supplied by Bateson and worked over by all the members of the team.

Each center contributed substantially to the editorial labor, and each center produced trainees who went on to become efficient collaborators. In September, 1963, Birdwhistell's group completed a typescript of the protocols of the family interview materials. In June, 1967, McQuown's group completed the editing of the explanatory materials ancillary to these protocols. In June, 1968, all materials were placed in the hands of Brosin's group (which had, since 1960, contributed substantially to their elaboration) for final revision and organization. In September, 1968, these materials were again in McQuown's hands. A training manual "The Natural History of an Interview" was ready for publication by September 30, 1968. This manual constitutes a general introduction to the theory of micro-analysis of interviews with a focus on overt behavior, to the individual systems of analysis of the speech and body-motion of participants in such interviews, to the techniques of manipulating taped and filmed materials in order to facilitate such analysis,

and to the theoretical frames suitable for the interpretation of the materials and for their use in psychotherapeutic and other practical applications. It presents a substantial body of analyzed and interpreted materials and a full record of the analytic and interpretive frames used for processing them. This manual may be used for the training of further adepts in the techniques of analysis and interpretation, and may be employed as a starting point for further research into analytic and interpretive frames and into their effective application to a wider variety of interview situations.

It has become painfully apparent, in the course of a decade of effort, that the descriptive frames currently available for undertaking microanalysis of interview behavior are in many ways inadequate to the purpose. The frame for describing and recording the varieties of the English language spoken in the United States and Canada (not to mention the rest of the English-speaking world) is in a continual state of growth and change; although in theory it is at present reasonably adequate for the description of any one regional or social dialect, in practice no full description of any one dialect is currently available, nor is there an adequate coverage of the salient points of difference characteristic of a range of dialects sufficiently wide to enable the micro-analytic interpreter to pick up, for example, the contrastive signals of differences in role and attitude, of group-adhesion and group-rejection, characteristic of participants in an interview. The frames for the description and recording of the varieties of paralinguage used in the English-speaking community are now two, that devised by Trager (1958) in the United States and that worked out by Crystal (1964) in England. Neither is complete, nor has either been widely tested on a sufficient variety of English speech. The pair has not yet been comparatively and contrastively applied to any single batch of interview materials to test for relative adequacy of fit and of coverage. The frame for the description and recording of the varieties of body-motion behavior in the English-speaking community devised by Birdwhistell (1952) has seen very considerable development since 1956, and is incorporated

into "The Natural History of an Interview." Since 1960 Birdwhistell has been engaged in perfecting the frame, and a training manual designed to teach the skills required for describing and recording body-motion in the English-speaking community is in process. Neither the original frame, nor its refined version, nor the two systems of transcription, graphic symbolic, and alphabetic, have as yet been adequately tested on a sufficient variety of body-motion behaviors in the English-speaking community. Although general theories of the structure of language are currently available, no such general theories of the structure of paralinguage or of body-motion behavior have as yet been worked out. Indeed, it is doubtful that such general theories can evolve until the currently available frames for describing and recording paralinguage and body-motion are tested in a variety of non-English-speaking linguistic communities. Likewise, cross-cultural testing of these general frames for the description and recording of language, paralinguage, and body-motion is indispensable to an adequate integration of the three frames and to a general theory of the structure of human communicative behavior.

It is suggested, therefore, that basic research into the facts of English regional and social dialect, of regional and social varieties of paralinguage and of body-motion in the English-language speaking community, needs be carried out on an appropriate variety of sound-filmed or video-taped interview-materials gathered with a view to cross-cutting regional and social dialect boundaries. A major focus of this research should be on sentence-like and sentence-sized units as these are revealed by the analysis of such cross-cutting dialect materials. It is suggested, furthermore, that basic research into the facts of regional and social dialects of other languages, of regional and social varieties of paralinguage and of body-motion in communities speaking other languages, be carried out on an appropriate variety of sound-filmed or video-taped interview materials gathered with a view to cross-cutting regional and social dialect boundaries. A major focus of this research, too, should be on sentence-like and sentence-sized units in such cross-cutting dialect materials. It is suggested, finally,

that the results of such study in the English (and other) language materials be compared and contrasted in carefully controlled frames.

It is expected that, as a result of such investigations, the frames for describing language, paralanguage and body-motion in English (and in other) language communities will be perfected, that the linguistic, paralinguistic, and body-motion markers of sentence-like units manifest in these communicative behavioral channels will be uncovered, and that the foundations of a general theory of the structure of human communicative behavior, as manifest through these channels, and in such units, will eventually be worked out. It is hoped that the materials here presented may facilitate the first steps in this on-going process.

Norman A. McQuown

September 30, 1968

CHAPTER 1

Communication

Gregory Bateson

O dieses ist das Tier, das es nicht gibt.  
Sie wusstens nicht und habens jeden Falls  
--sein Wandeln, seine Haltung, seinen Hals,  
bis in des stillen Blickes Licht -- geliebt.

Zwar war es nicht. Doch weil sie's liebten, ward  
ein reines Tier. Sie liessen immer Raum.  
Und in dem Raum, klar und ausgespart,  
erhob es leicht sein Haupt und brauchte kaum

zu sein. Sie nährten es mit keinem Korn,  
nur immer mit der Möglichkeit, es sei.  
Und die gab solche Stärke an das Tier,

dass es aus sich ein Stirnhorn trieb. Ein Horn.  
Zu einer Jungfrau kam es weiss herbei--  
und zwar im Silber-Spiegel und in ihr.

Rainer Maria Rilke, Sonette an Orpheus, II. Teil, IV.

O this is the creature that does not exist.  
They did not know that and in any case  
- its motion, and its bearing, and its neck,  
even to the light of its still gaze - they loved it.

Indeed it never was. Yet because they loved it,  
a pure creature happened. They always allowed room.  
And in that room, clear and left open,  
it easily raised its head and scarcely needed

to be. They fed it with no grain, but ever  
with the possibility that it might be.  
And this gave the creature such strength,

it grew a horn out of its brow. One horn.  
To a virgin it came hither white -  
and was in the silver-mirror and in her.

Rainer Maria Rilke - SONNETS TO ORPHEUS

(Translated by M. E. Herter Norton, W. W. Norton and Co. N.Y.)

Second Part, Sonnet 4

#### Background

At the time of the outbreak of World War II, the most promising insights in the behavioral sciences were those derivative from Freudian analysis, Gestalt psychology, and cultural relativity. Linguistics had begun to take on new life under the leadership of Sapir (1921, 1925, 1933a, 1933b) and Bloomfield (1933, 1939). Psychiatry was evolving away from the exclusive study of the individual patient towards the study of human relationships, most dramatically under the influence of Sullivan (1940); and already there were moves towards a mathematics of human relationship under Kurt Lewin (1935) and L.F. Richardson (1939).

During World War II and immediately following that period of confusion, a series of exceedingly important new approaches were evolving more or less independently in a number of different places, but the possible relevance to behavioral science of the work of George Boole (1854), Whitehead and Russell (1910-13) was still unexplored. All of these scattered advances were precipitated by the war-time development of electronic engineering. A partial list of names and locations of the principal advances will give an idea of what is happening.

Rosenblueth at Cambridge and in Mexico, and Wiener and Bigelow (1943) at the Massachusetts Institute of Technology, were laying the foundations of what has come to be called cybernetics, extending



what the engineers and mathematicians had learned about self-correcting mechanisms to the fields of biology and social organization.

Von Neumann and Morgenstern (1944) at Princeton were laying the basis of the theory of games.

Craik (1952), in Cambridge, England, before his premature death, wrote "The Nature of Explanation" raising the whole question of how messages are coded in a reticulate central nervous system.

Attneave (1959), Stroud (1949), and others at Stanford, happened to see Craik's little book and by it were inspired to a new approach to the problems of perception and adaptive action.

In Vienna, Bertalanffy (1952) was building the beginnings of systems theory with a special emphasis upon those systems (e. g., organisms) which have a continuous source of energy derived from the environment.

Shannon (1949), and others working with the Bell Telephone Laboratories, were building the structure called information theory.

Ashby (1952, 1956), in Gloucester, England, was devising new models for theories of learning and the evolution of the brain.

Other names McCulloch and Pitts (1943, 1947), Lorente de No (1922, 1933), Rashevsky (1948), Walter (1953), Tinbergen (1953), Lorenz (1952) might be mentioned as contributing to this general trend.

What has happened has been the introduction into the behavioral sciences of a number of very simple, elegant, and powerful ideas all of which have to do with the nature of communication in the widest sense of that word. The steps and sequences of logic have been coded

into the causal sequences of computing machines and, as a result, the Principia Mathematica has become a cornerstone of science.

#### The Natural History of an Interview

The present book is an attempt at synthesis. It is written by five persons who are professionally concerned with communication problems in diverse fields and we attempt a synthesis of a wide and abstract kind, starting from the most concrete data.

We start from a particular interview on a particular day between two identified persons in the presence of a child, a camera and a cameraman. Our primary data are the multitudinous details of vocal and bodily action recorded on this film. We call our treatment of such data a "natural history" because a minimum of theory guided the collection of the data. The cameraman inevitably made some selection in his shooting; and "Doris", the subject of the interview, was selected for study not only because she and her husband were willing to be studied in this way but also because this family suffered from inter-personal difficulties which had led them to seek special psychiatric aid.

These materials, then, while collected under circumstances unusual in human relationship, nonetheless provide the data for the natural history of two human beings over a brief span of time, and the data themselves are sufficiently uncorrupted by theory so that the five authors, each with a particular theoretical bias and interest, could simultaneously approach this mass of detail. Moreover, we shared something less tangible than the common data: certain theories

or preconceptions about what happens when two people interact.

#### Theoretical Premises

My major task in this preliminary chapter is to outline those theoretical premises which were engendered in us by the recent advances in the study of human communication:

#### Freudian

From Freudian theory we accept the premise (1) that only limited aspects of a part of what happens in human communication are accessible to the consciousness of the participants. Our position however differs from that of many early Freudians in two respects which are minor so far as theory is concerned, but major in their implications for method. The important corrective which the Freudian applied to man's thinking about human nature was an insistence upon the unconscious. The error to be corrected was the notion that in human beings the mental process is preponderantly or entirely conscious. This error had roots in eighteenth century culture and back into the Reformation and into earlier Judaeo-Greek philosophies of free will. But this error today seems almost fantastic.

It is now a platitude to state that mental process depends upon hierarchic organization. Whether we think of mental levels or of a brain evolved by a process of successive telencephalization, we envisage a hierarchy of both anatomy and function. And our knowledge of hierarchic function--in machines, embryology, physiology, and in human social organization--indicates as a truism that under no

circumstances can the upper echelons of any hierarchic system handle total information about the processes and events which occur at subordinate or peripheral levels. By the same token, the upper echelons can handle only limited reports--can be only partially conscious--of all that happens at their own upper level. To provide these upper echelons with total reporting would be to add to the system still higher echelons--themselves, in turn, largely unconscious. To us, then, the fact that most mental process (including, especially, the process of perception itself) cannot be inspected by consciousness is a matter of course and what is surprising, and therefore needs explanation, is the fact of consciousness. Unconsciousness is a necessity of the economics of hierarchic organization (Sapir, 1927).

This does not mean, of course, that economy of effort or the economic use of the channels of communication to avoid jamming is the only factor determining what information shall be allowed to reach the upper echelons of consciousness. The analogy of human social organizations would indicate very clearly both that upper echelons are commonly "motivated" not to receive information about certain peripheral events and that there are many events which the subordinate echelons are "motivated" not to transmit upwards. There are, therefore, many matters which remain "in the unconscious" for reasons other than those of economy, and the unconscious becomes a repository for material which is repressed in the Freudian sense.

The second difference between our position and the classical

Freudian results from our emphasis upon communication. We are interested in such questions as "what signals are emitted and what orders of awareness does the signaler show by emitting other signals about these signals? Can he plan them? Can he recall them?" And we are interested to know what signals reach the receiver and what signals he knows he has received. Our emphasis is thus upon perception and communication rather than on the internal hierarchies of mental process. From where we sit, the distinction between conscious and unconscious becomes significantly comparable to the distinction between foveal and peripheral vision.

A second premise related to Freudian theory holds (2) that everything which occurs is meaningful in the sense of being a part of the interchange as well as non-accidental. The Freudian emphasis was upon psychic determinism--that no word uttered and no detail of a dream experienced can be accidental. A man cannot "just dream." Our emphasis in this book will extend this psychological idea into the realm of inter-personal process. We shall attempt to see every detail of word, vocalization, and bodily movement as playing its part in determining the ongoing stream of words and bodily movements which is the interchange between the persons. We shall endeavor to think not only in terms of psychic determinism but in terms of a larger inter-personal determinism. Two people cannot "just agree" or "just quarrel."

Also, from Freudian theory, we accept the idea (3) that all messages, whether verbal or non-verbal, are mediated in their creation by

primary process and therefore contain, either implicitly or explicitly, all the multiple reference characteristics of dream or fantasy. If it be possible for a man to seem to talk only about the overt subject of conversation, this is achieved only by vigorous ego-function which carefully excludes or conceals the multiple under-tones of implicit content. Further, we expect that the minute analysis of speech and movement will disclose that the messages in both these modalities contain a large proportion of unconscious material with primary process characteristics, that, for example, an unconscious fingering of the dress is likely to denote (or to be a resultant of) sexual interest and/or its puritanical denial.

Also from Freudian theory, we accept (4) a generalized notion of transference: that any person emitting learned signals does so upon the (usually unconscious) assumption that the receiver of these signals will understand them "correctly"--i.e., he assumes that his vis-a-vis at the given moment will resemble psychologically some former (or even fictitious) vis-a-vis from whom he originally acquired his communicational habits.

Closely related to the notion of transference is (5) the notion of projection. This explanatory principle differs however from transference in that it does not invoke some third historical or fictitious person. When A "projects" upon B, he is merely assuming that B's signals are to be interpreted as A would interpret these signals if he himself had emitted them. That is, A assumes that B operates

according to systems of codification similar to his own. Both transference and projection may, of course, be carried into the future.

A may expect that B will exhibit meaningful action of a sort which some historical figure in A's life would have exhibited under similar circumstances (transference); or he may expect that B will behave as he himself would behave in similar circumstances (projection).

Identification must also be mentioned. This explanatory principle invokes (6) the idea "if you can't lick 'em, join 'em"--or, at least imitate them as you see them. A is said to identify with B when he starts to mold his own meaningful action in terms of what he thinks are B's principles of codification.

Notably, all of these principles--transference, projection, and identification--are likely to be unconscious in their operation and to be in some degree coercive. That is, any errors which A may make in his assumptions about B are likely to cause A to act in such a way that B is put under pressure to validate these errors by acting as if A's assumptions were true. An especially interesting case arises when A acts in a way which will coerce B into identifying with A's self-image--which may be false.

Moreover, it must not be supposed that these explanatory or descriptive principles are mutually exclusive. In a given instance, A may consciously or unconsciously assume that B is parental (transference). But A's technique for dealing with his parent may have

involved identification.<sup>1</sup> He will then adopt towards B that role which he formerly adopted towards the parent.

#### Gestalt

From Gestalt psychology, we have accepted a premise of very great importance: that experience is punctuated. We do not experience a continuum: on the contrary, our experience is broken up into what seem to us to be events and objects. In Gestalt psychology this idea is basic to the figure-ground hypothesis. And for us it is related to the premise that nothing never happens--i. e., that both sender and receiver of signals are so organized that they can and must use, for their understanding of what is going on, the fact that certain possible signals are not present. The first step in building the figure-ground hypothesis is a postulate of this kind. In order to recognize that there are stars in the night sky, we must use the fact that certain retinal end-organs are not stimulated by the darkness. In human relations, no silence is insignificant and the absence of tears may speak volumes.

More must be said concerning the punctuation of interpersonal events. Our whole procedure and, indeed, any analysis of communicational data is shaped by premises which define the units into which the stream of data is to be divided. First, in a macroscopic examination of the interview, we assumed that the 400 feet of film on which

---

<sup>1</sup>The word "identification" was perhaps an unfortunate choice for two reasons. In the phrase "A identifies B as father" is a statement of transference. And the phrase "A achieves ego-identity" suggests (as an ideal) A's escape from all the errors of transference, projection and identification.

the interview is recorded<sup>2</sup> can be punctuated into incidents or sequences with beginnings and ends psychologically meaningful to the participants. As will be seen, we have chosen certain of these incidents for microscopic study. Our macroscopic study serves to direct our more particular attention. And, while we narrow our focus from the interview as a whole to an examination of incidents within the interview and then downward to the finer and finer detail of these incidents--we work throughout with similar assumptions about punctuation of the stream of signals.

The historical basis of this assumption will make clearer what is meant. Historically, scientific linguistics has progressed most rapidly since the time when certain popular and preponderantly occidental notions about language were adopted, made rigorous, and extrapolated into the study of minute detail. In their popular form, these notions are, for example, that speech is subdivisible into sentences which in turn are subdivisible into words, which in turn can be subdivided into letters. Profound modifications have been introduced into this hierarchy by the linguists who needed to describe speech rather than written language, but the essential idea that a stream of communicative material must, of necessity, be susceptible of such multiple re-subdivision is fundamental in linguistics and in that part of communications theory which deals with coded communication--a much wider field than the conventionally linguistic. A major contribution of the

<sup>2</sup> The film also has irrelevant gaps while the camera was being reloaded after each 100 feet. The longest consecutive recording is about 3 1/2 minutes.

linguists is the demonstration that the stream of communication contains positive signals by which its units are delimited.

Moreover, Gestalt theory presupposes a hierarchy of subdivisions characteristic of the process of perception. We do not perceive the firing of unit end-organs but, from the showers of neural impulse started by that firing, we build images of identifiable and larger meaningful complexes of identifiable. We can argue from perception to communication: If an organism's perception is characterized by Gestalten and this organism is capable of emitting complex streams of communication, then these streams must be dissectible into a hierarchy of successive subdivisions. Many such analyses will be possible, and there will be one which will represent correctly the natural history of the organism.

We deal, after all, not merely with the fact that a communicational stream can be dissected but also with the question: in which of the many ways possible should this particular stream be dissected? What we know about language and communication in general indicates that there will always be one or more hierarchies of Gestalten which will be correct in the sense of describing how the message stream is created and/or how it is received and interpreted by the hearer. The Freudian findings also indicate that in any given instance several different analyses may be correct. A particular message may be simultaneously interpreted in different ways by different levels of the mind: we face problems of multiple coding.

The linguists are ahead of the other natural historians in their study of the hierarchy of Gestalten in terms of which a particular kind of behavior should be dissected. Their studies are being fortified by cross-cultural comparison, by dialectal (subcultural) comparison, and by statistics of individual variation. By way of contrast, kinesics--the study of body motion, position, and action as a modality of communication--is a relatively recent development which, like linguistics, is achieving a firm scientific base by the rigorous dissection of the kinesic stream into a hierarchy of Gestalten and subdivisions of Gestalten.

In a later chapter, Birdwhistell will outline the hierarchy of units which he is devising for kinesic description. He is proceeding in a manner comparable to but not identical with that methodology of description which has proved valuable in linguistics. The ultimate validation of this approach in kinesics will, of course, depend upon the results obtained, but there is very strong a priori argument in favor of the correctness of the approach from all that we know about communications theory in general and about human communication and perception in particular.

Returning for a moment to linguistics, other types of description which the linguists have achieved must here be mentioned. The very complex question of "meaning" is too large for discussion in this chapter but this much may be said--that a tape recording of human speech contains a great deal more than the signals correlated with the lexical meaning of what was said.

If a tape recording is transcribed into ordinary script, although some of this more-than-lexical content will be lost, some will still

survive in the transcription. Indeed, to reduce a speech to its merely lexical content would require very drastic procedure (in the course of which other and probably inappropriate non-lexical overtones will inevitably be added). It would first be necessary to strip the speech of all indications of the context in which it was uttered and by whom and to whom it was uttered. But there would still remain cadences and overtones of a non-lexical nature. To get rid of these it would be necessary to translate the speech into some other language and to use as a translator some hypothetical person (or machine) totally insensitive to non-lexical content both of the language from which he is translating and of the language into which he is translating.

As we climb the hierarchic ladder of Gestalten from the most microscopic particles of vocalization towards the most macroscopic units of speech, each step on this ladder is surmounted by placing the units of the lower level in context.

"Meaning," as this word is vulgarly used, emerges only at a very high level in this hierarchy. We discriminate the initial phoneme of the word "peter" from the initial phoneme of the word "butter" but these phonemes are in themselves meaningless apart from their setting in a stream of phonemes. Even the syllables, "pete" and "but", are in themselves either meaningless or multivalent (except insofar as their possible meaning is restricted when we know how they are placed in a stream of syllables). With each step towards a larger unit--the larger unit being always the smaller unit plus its immediate

setting--there is a more and more drastic limitation of possible referents. "Meaning," therefore, is a function of this restriction of possible meanings. Even the words "peter" and "butter" are still multi-valent. When the word "blue" is added, the hearer may be pretty sure that the referent of "peter" is a flag. But still, there is room for doubt.

The "Blue Peter" may be referred to as an actual object of action or observation in the larger context of a ship about to leave a particular port. Or the reference may only be metaphoric if the term is used on land. Or the usage of the term may be neither metaphoric nor direct but may be part of a lesson in maritime communication. Or--as here upon this page--the words "Blue Peter" may be mentioned only as an example of communicational phenomena.

Meaning approaches univalence or non-ambiguity only when very large units of the communicational stream are admitted to examination. And even then, the approach to non-ambiguity will be asymptotic. As larger and larger bodies of data are admitted, the probability of a given interpretation will be increased but proof will never be achieved. The situation is essentially the same as that which obtains in science where no theory is ever proved.

This book is concerned with trying to put together those parts of the communicational stream which the professional linguist studies (phonemes, morphemes, phrases, vocal modifiers, junctures, etc.)

with those parts of the stream which are studied in kinesics (kines, kinemorphs, etc.). A central question, therefore, which we shall have to face when we analyze the data is the extent to which there is a mutual relationship of "context" between kinesic and linguistic elements.

We face phenomena so structured that there is perhaps no definable upper limit to the size--either spatial or temporal--of the Gestalten. In practice, this would mean that no finite collection of data would confer complete non-ambiguity upon any item within the collection; that however widely "context" be defined, there may always be wider contexts a knowledge of which would reverse or modify our understanding of particular items.

#### Context

These considerations force us to a method of inquiry which will postpone the question of "meaning". When faced with a given sequence of signals, we shall delay the question "what do these signals mean" for as long as possible. We shall ask, rather, the collateral question "would the meaning be changed by a given change in the sequence or by a given change in the context?" This is a question which can be asked and answered without too much difficulty. We shall, for example, not ask whether the word "Peter" refers to an apostle or a flag, but rather whether its meaning, when the word "Peter" follows the word "Blue", is peculiarly appropriate to the new context.

In kinesic analysis, we shall similarly delay the question of

what is meant by the rapid closing and opening of an eye which is visible to the vis a vis and shall ask rather, for example, whether the meaning of this signal would be altered (a) if the other eye were blinked simultaneously and (b) if the blinked eye were one which is invisible to the vis a vis. Parenthetically, we may also ask whether the meaning of the word "Peter" is altered by winking one eye.

It is, after all, only an historic accident--a past pathway in the evolution of science--that has lead to the circumstance that linguists study data which can be heard, while the kinesicist studies data which can be seen. That the scientists have become specialized in this particular way does not indicate a fundamental separateness between these modalities in the stream of communication. It is for this reason that the work of this book starts from concrete natural history--from the recorded interaction between Doris' speech and movement and the speech and movement of Gregory. This placing of every signal in the context of all other signals is an essential discipline of our work.

A great part of the work which Birdwhistell, Hockett, and McQuown have had to do has involved a grueling process of synchronization. The audible stream for which Hockett and McQuown are specialists was recorded on tape and on film with an unsatisfactory sound track. The analysts had to work frame by frame through the film to establish the point in the audible sequence at which, for example, Doris turned her head or let her shoe fall away from her heel. I described our data loosely above as the aggregate of signals

recorded on the film. More accurately, I should have said that our data are the individual signals or messages, each in its immediate and extended context.

But the context of a signal emitted by Doris is not merely those other signals which she has recently emitted plus those which she emits soon after; it is also the room in which she is speaking, the sofa on which she is sitting, the signals emitted by Gregory with whom she is talking, and by the little boy Billy, and the inter-relationships among all of these.

#### Interaction

At this point, our concept of communication becomes interactional and our intellectual debt is to G. H. Mead (1934) and to Sullivan (1940) rather than to Freud and the Gestalt psychologists. The system which we now study is no longer merely a descriptive synthesis of Doris' body motion and speech, but the larger aggregate of what goes on between Doris and Gregory.

This larger frame determines meaning for what each person does and says. Rilke's "Unicorn" is present in every conversation between persons and this fictitious beast evolves and changes, dissolves and is recrystallized in new shapes with every move and message. Denial of the Unicorn will not prevent its existence--but only cause it to become monstrous.

This poetic fancy must be made scientifically real to the reader if he is to understand what this book is about.



For every human being there is an edge of uncertainty about what sort of messages he is emitting and we all need, in the final analysis, to see how our messages are received in order to discover what they were. For the schizophrenic this is often dramatically and conspicuously true.

Let me illustrate by an example. A schizophrenic patient tells me that he built the China Wall, rowed across the Pacific and landed in Seattle. He then walked to California where he was "affriended by those people." This narrative he offers as if it were a statement of fact. But whether it is a statement of fact for him depends upon my response. If I say, "Nonsense. You were born in California," I have thereby verified for him the notion that his narrative is to be taken literally. I have denied it as if it were a literal statement and now it exists for him as a literal statement which must be defended. From there on, we shall get into an argument not about the question "is this narrative a statement of fact?" but about the red-herring question "is this a true statement of the facts?"

The response which we get tells us about the state of the hearer after he has received the signals which we emitted. It may be evident that he misunderstood the message either grossly or subtly. The status quo, however, which obtained when we emitted the message, no longer exists and merely to repeat the message will not do. We are now communicating with a person whose relationship to us is different from what it was a moment ago. And within the framework of this new relationship we must now speak.

Of all the elements and vicissitudes of formation and re-formation of relationships, perhaps the most interesting is that process whereby people establish common rules for the creation and understanding of messages. Whatever reply I may make to the patient's delusional narrative proposes a pact to govern us both in our understanding of the message. If I deny the factual truth of the narrative, I implicitly propose that we agree to treat it as literal. If, on the other hand, I ask him whether he thinks his parents had a part in building the "China Wall" which separates him from them, I have proposed that we agree upon a different set of rules for the creation and understanding of such messages.

The possible systems of rules which two persons may share are many and complex. Among them must be mentioned a system which has been characterized as symbiotic. Such a label, as I understand it, refers to a system of non-verbalized and usually unconscious pacts in which, for example, A and B "agree" to accept each other's messages in some spirit other than that in which they were coined. By ignoring overtones and implications, or by reading in overtones which were not intended, the persons maintain a strange semblance of understanding.

#### Code Distortion

In this book we shall pay but little attention to those failures of communication which are due to the randomization of signals occasioned either by background noise or by imperfect resolving

processes in the receiving sense organ. We are concerned with a more subtle phenomenon--the distortion of messages which occurs when the persons involved differ from each other in their rules or assumptions governing the making and understanding of messages--their explicit and implicit rules of coding.

Imagine a machine which has the function of transmitting a half-tone block, (a picture formed entirely of rows of dots), over a wire to another machine. The transmitting machine will transmit over the wire a sequence of electrical impulses such that each impulse or absence of impulse is a "yes" or "no" answer to the question "is there a dot in the given space?" When the transmitting machine comes to the end of a line of spaces, it will either transmit a special signal which will cause the receiving machine to go on to the next line, or the machines will have to have been so adjusted to each other, that they operate in terms of a common pact governing how many dots there shall be in a line. A discrepancy regarding the terms of this pact will introduce code distortion. In such a case, the receiving machine will create a picture which may be an absolutely correct record of the sequence of signals emitted by the sender but which, considered as a picture, will be distortion of the original.

Figure 1

A is a picture to be transmitted, B is the distorted version which is created when the receiving machine acts upon the premise that there are only 16 squares in each line instead of 17.

In Figure 1, the effect of code distortion is shown, and it is worthwhile to stress the basic difference between such distortion and the loss of information due to entropic noise. In the case of entropic noise, the information which is lost is irretrievable, but in the case of code noise what has occurred is a systematic distortion which could conceivably be rectified.

All that is needed for this correction is that there be some means whereby the transmitter and receiver can communicate about the rules of communication. This presents special difficulties but it is a fundamental thesis of this book that at the human level such communication about the rules of communication occurs constantly. This, in fact, is the process whereby the "Unicorn" is continually created and recreated. When my patient tells his story of the "China Wall", whatever reply I make is a communication to him about how I received his message and therefore indicates to him (ideally) how he should restate it in order to have me receive that message which he wants me to receive. It tells him how to code his messages so as to elicit an appropriate response from me.

It is necessary again to insist upon the unconscious character of most communication. We are almost totally unaware of the processes by which we make our messages and the processes by which we understand and respond to the messages of others. We are commonly unaware also of many characteristics and components of the messages themselves. We do not notice at which moments in a conversation we cross and uncross our legs or at which moments we puff on our

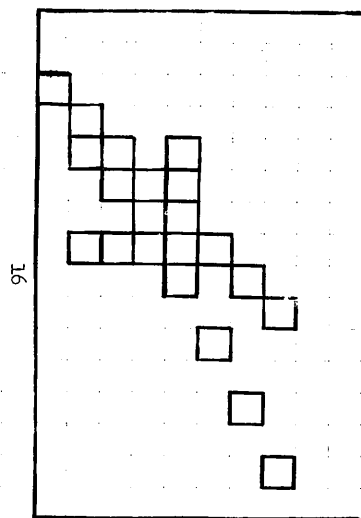


Fig. 1.B

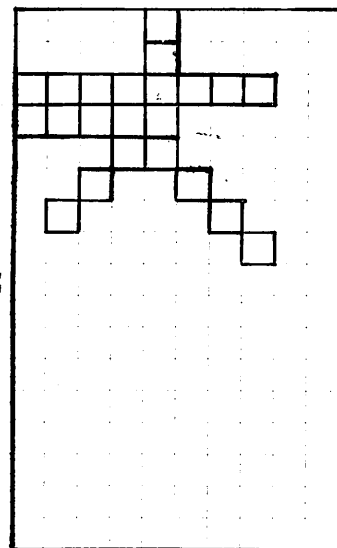


Fig. 1.A

25

29

cigarettes or blink our eyes or raise our brows. But the fact that we do not notice these things does not imply that all these details of personal interaction are irrelevant to the ongoing relationship. Just as we are in the main unconscious of the fleeting pacts which we enter into as to how messages are to be understood, so also we are unconscious of the continual dialogue about these pacts.

This dialogue is not only between persons and about the pacts which they form, it is also and more strangely a dialogue which governs what each person is. When A makes overtures which B brushes aside, this experience is to A more than a hint about how to code messages when dealing with B. In everyday language we say that a person's self-esteem is enhanced or reduced by the responses of others. Or we say that "he sees himself differently." In communicational terms, we may translate this into a statement that the very rules of self-perception, the rules governing the formation of a self-image, are modified by the way in which others receive our messages.

#### Learning and Pathogenesis

In part, this book is a study of how communication works between two persons, but it is also a study of how communication fails to work--that is, of certain pathologies of communication. Our collaborating team includes not only the two linguists and the founder of kinesics, but also two psychiatrists and the writer, whose initial training in anthropology has finally led him to study schizophrenic communication. It is therefore appropriate to examine a little more

21

closely the relation between psychiatric pathology in the individual patient and the pathologies of communication which may develop between persons. In order to keep the subject matter simple, I will exclude from consideration those psychiatric abnormalities which have an established base in organic lesion.

To build a bridge between the study of psychiatric functional pathology and the pathologies of communication, it is necessary to insist upon the existence of the facts of learning and conditioning. Two considerations become especially relevant. First, every failure of communication is painful. Second, the learning organism always generalizes from experience. Further, the business of communication is a continuous learning to communicate. Codes and languages are not static systems which can be learned once and for all. They are, rather, shifting systems of pacts and premises which govern how messages are to be made and interpreted. Every signal which establishes a new premise or pact bringing the persons closer together or giving them greater freedom may be a source of joy. But every signal which falls by the wayside is in some degree a source of pain to both. The ongoing stream of communication is thus, for each individual, a continuous chain of contexts of learning and, specifically, learning about premises of communication.

At this point, it is necessary to consider certain aspects of the learning process and to expand conventional learning theory to make it

relevant to an analysis of the interchange of signals between persons. A typical learning experiment involves two entities: an experimenter and a subject, and the theoretical conclusions derived from such experiments are commonly stated as psychological regularities descriptive of the subject. In contrast, I shall here view the experimental situation as an interaction involving two entities in whose relationship I am interested. I shall regard their relationship as formally characterized by an interchange which is repeated in successive "trials" and shall assume that not only the subject, but also the experimenter is undergoing a learning process determined, at least in part, by reinforcements which the subject provides.

As a preliminary to this, it is necessary to define a hierarchy of orders of learning. This may be done as follows:

1. It appears that the simplest learning phenomenon is the receipt of information or command. The event of perceiving a whistle may constitute, for a dog, an important piece of information or a command. Before it heard the whistle, it was in another state. The change of state I classify as the simplest learning phenomenon. It is important to remark, however, that this phenomenon is excessively difficult to investigate and has not been an immediate object of experimental study. It has, however, been a major focus of theory. What seems to have happened is that in order to arrive at a theory to describe what I call the second order of learning, the psychologists have had to provide

some description of this first order process--some verbalization of what message the dog has received. The "effect" theory proposes that this message is a promise of reward or a threat of punishment, whereas the associational theory proposes a more automatic and less purposive description of the dog's response.

2. The second order is the learning or conditioning upon which the vast mass of experimental work has been conducted. Here the word "learning" refers to a change in the dog's ability to act upon percepts or signals received. What the experimenters study is changes in the dog's behavior resulting from a sequence of trials. The phenomena studied are of a different and higher order than those discussed in 1 above. The question asked is not "what change occurs in a dog when he hears a whistle?" but "what changes have occurred in the change which a dog undergoes when he hears a whistle?" This subtle difference in the question asked by the experimenter makes it formally impossible for the theorists to deduce answers to the first question from data collected to answer the second. The behaviorists had logic on their side when they insisted that we never ask about the subjective experience of the dog. To try to deduce the dog's experience from data which could only throw light upon change in his experience is to attempt the logically impossible. From the characteristics of a class, I can make

no deductions about what a member of that class might be.

3. The third order of learning is a familiar laboratory phenomenon but has received only slight attention from the experimentalists. If we describe the second order of learning as "learning to receive signals", then the third may be described as "learning to learn to receive signals."<sup>1</sup> What happens in the laboratory is that the animal having been subjected to experiments of the second order becomes "wise". That is, when faced with an entirely new experiment of this sort, the animal requires now a smaller number of trials to achieve that learning of the second order which the experimental situation demands. The animal has acquired a knack or skill for second order learning. This phenomenon has been measured by Hull (1940) studying rote learning of nonsense syllables, and by Harlow (1949) studying problem solving in rhesus monkeys.

4. There is no theoretical reason to deny the possibility of fourth and higher orders of learning, though none of these

<sup>1</sup> Unfortunately, in an earlier theoretical paper, (Bateson, 1942) I have used for this third order of learning the term "deutero-learning." This was due to my failure to recognize the receipt of a meaningful signal or the receipt of a bit of information as an example of the simplest order of learning. To achieve any analogy between the mechanical computers and the brain it is necessary to insist that any receipt of information is, in a broad sense, learning.

have been demonstrated. The nature of the hierarchy which we are discussing is such that there is no upper limit to the series other than that set by the limitations of brain structure. The number of neurons being finite, it is certain that for any organism there is a practical upper limit to the number of orders of learning of which it is capable.

Inspection of this hierarchy of learning reveals that the difference between any order of learning and the next higher order is essentially a difference in size of Gestalt. The higher order is always documented by demonstrating change which results from a larger Gestalt, this larger Gestalt being in general built up of a multiplicity of the Gestalten characteristic of the lower order. But while this generally seems to be the case, there is no theoretical premise by which we might estimate the multiplication factor, and it is necessary to consider as at least theoretically possible the case in which this factor would be unity.

A single increment in what appears to be a context of lower order learning might conceivably precipitate major changes of some higher order, whereby all experience of the lower order would be reframed and reorganized. We face here an unpredictability of a sort which I noted earlier when discussing the indeterminacy of meaning. Larger and larger bodies of data will provide greater and greater certainty of interpretation but it is never possible to be sure that the next increment of data will not compel us to a totally new interpretation.

There is thus an analogy--perhaps amounting to identity--between those hierarchies of Gestalten which determine meaning and that hierarchy of Gestalten which we here call contexts of learning.

These abstract matters become clearer when we state that learning of the third or higher order is, in popular parlance, called "change in character." Let us suppose that an organism becomes "wise" in dealing with contexts of Pavlovian learning. The change which we here refer to may be described both as a change in the organism's expectations and as a change in its learning habits. If we speak in terms of expectations, we will say that the organism now preponderantly expects the universe of experience to be punctuated into sequences resembling the Pavlovian context; i. e., sequences in which certain percepts can be used as a basis for predicting later events. Or, if we speak in terms of learning habits, we will say that this organism will respond to the predicted certainty of that which is to come, (e. g., by salivating), but will not endeavor to change the course of events. In a word, the organism has become "fatalistic" and examination of the formal characteristics of the learning context has provided us with a formal definition of one particular sort of "fatalism."

The psychiatrist is interested largely in learning of the third or higher order. If a patient tells him that she can use a typewriter, the psychiatrist will pay but little attention. She has reported only a result of second order learning. But when the patient goes on to

describe the context in which she learned to typewrite and tells him that her teacher punished every error she made but never praised her for progress, the psychiatrist will prick up his ears. He will see in this narrative a statement of what effect the context of learning to type may have had upon the patient's habits and expectations--i.e., upon the patient's character. This enlargement of learning theory to discriminate orders of learning makes this body of experimental knowledge especially relevant to the psychiatrist. Actually the old barrier between experimentalists and clinicians seems to have grown out of this: that the experimentalists have mainly studied learning of the second order, while the psychiatrist is interested chiefly in effects of the third order. These effects he tries to evaluate in his diagnosis or to achieve in his therapy.

If this description of learning is substantially correct, that is, if there really is a hierarchy of orders of this phenomenon and the discrimination of these orders is something more than an artifact of description, then it becomes theoretically probable that there exist complex sequences of experience and action such that learning of one order will in some degree contradict the learning of some other order. We can imagine, for example, that a human subject might experience a long sequence of Pavlovian learnings but might be penalized (Bateson and Jackson, 1956) for exhibiting "fatalism." Or he might be trained towards obedience but be continually penalized for the finer detail of every obedient act. As between adults, this is familiar enough and

may make for bad "personnel relations." As between parents and small children, I believe that it is - under some circumstances - pathogenic.

#### Pathogenic Contexts

It is now clear, however, in an abstract and formal way, what patterns of interchange we should look for in our data. The discussion which preceded this reexamination of learning theory concerned the establishment of pacts and premises of communication. But evidently a premise of communication, a rule governing how messages are to be constructed or interpreted, bears the same relationship to the given message as occurs between a higher and a lower order of learning.

The acceptance of what I have called a premise of communication is the same phenomenon as the acceptance of a role--a momentary or enduring shift in habit and expectation. And "role" is only a word for some phase of character change, be it brief or enduring. It is a description of the pattern exhibited by one person in that two-person system which constitutes a context of learning.

It follows that what we have to look for in the data is sequences and, at the meta-level, sequences of sequences. The relevant units will be those segments of the stream which constitute contexts of learning. Problems of pathology within the stream will become recognizable when we see instances so constructed that learning in

a given small sequence would be contradicted by learning in some larger sequence of which the smaller is a component. Theoretically, we may expect instances in which part and whole will be identical--where the multiplication factor relating the part to the whole is unity. A single context (seen in two different ways) may propose contradictory learning at different levels.

One other peculiar phenomenon must now be mentioned--namely, that the premises of communication are commonly self-validating. By their operation they may create that consensus which will seem to validate them. He who believes that all the world is his friend--or enemy--will emit messages and act meaningfully in terms of his premise. He will meet the world in a way which puts pressure upon this very world to validate his belief, which belief he acquired in the first place by the cumulative impact of those contexts of learning which were his communication with some earlier person.

An inquiry into the functional psychopathologies thus becomes an investigation of the dynamics of past communication. But, curiously enough, because of this fact that communicational premises are self-validating, it is often not necessary to delve into the past in order to investigate their etiology. The premises are self-validating in the present and therefore the disturbed--like the normal--is continually creating around himself that environment which provides the typical etiology for his communicational habits--his symptoms. One has only to examine the present family relations of a patient to find working today the constellation which is etiologic for his symptoms. Indeed,

we may profitably examine the workings of any typical mental hospital for clues as to why its patients are mentally ill.

This broad description of the interchange between persons as a sequence of contexts of learning contains the possibilities for two kinds of psychopathological result: the learning of particular error and the disruption or distortion of the learning process itself. Historically, the first of these received most attention in the early days of psychoanalysis when emphasis was placed upon the fact that certain neuroses result from single and extremely painful experiences in childhood. In terms of what has been said above, we might rephrase this theory as a learning of error--the error being an inappropriate generalization from some terrifying, painful, or over-rewarding experience. Today, less theoretical importance is attached to this kind of pathogenesis, but its occurrence is still undoubted.

In contrast, modern psychiatric theory insists more upon those psychopathological results which derive from continual and repeated experience rather than from isolated trauma. Here the probability that simple error will be generated in the learning individual is much less, since, after all, his opinions, stemming from a multitude of instances, are to that extent validated by the repetition of instances. What is rather to be expected from such an etiology is the distortion of the learning process itself--a type of pathological result more abstract and intangible--and more difficult to correct by any therapeutic experience, since whatever



the patient learns from this experience will probably be learned by means of that process which is already distorted.

It is, however, necessary to give some substance to the phrase "distortion of learning". I have to indicate what sorts of interpersonal sequences might have this effect on one or on both of the participants.

A context of learning is a definitely structured segment of the stream of interchange between persons. We know from experimental data that while the structuring of contexts of learning is extremely variable, some structuring is always present. The events of which the context is composed--conditioned stimulus, response and reinforcement--may be variously related to each other and still constitute a structured whole. That is, we are here dealing with Gestalten (units of the interchange) and are therefore again face to face with the peculiar nature of all such units. Although they are in large part the creation of the individuals concerned, and are necessarily a product of the ways in which these individuals perceive and punctuate what is happening, their perception is inevitably guided by culture and convention. Such perception may be rigid or it may be flexible. But the essential fact is that the rules for this punctuation are a part of that system of pacts and premises upon which communication is based. For their learning communication must be viewed as a sequence of contexts.

What I am describing is a strangely retroflexed procedure;

a process which is in a way folded back upon itself. This may be said in many ways and perhaps most simply by stating that the communicational stream is a sequence of contexts both of learning and of learning to learn.

At this point, the phrase, "distortion of the learning processes" takes on meaning. It would refer to all those cases in which an individual punctuates the stream of communication in a way different from his vis-a-vis but which are reinforced nonetheless by the pain resulting from his idiosyncratic view. From the point of view of the speaker, it will seem to him that he has incurred punishment for what he thought he was communicating, whereas he is in fact being punished for what his messages seemed to be, as perceived by the other.

It is clear that this line of thought, if substantially correct, will lead to a formal theory of stability and instability in human relations. We might therefore inquire into what the engineers call criteria of stability. Is it possible to classify the degrees and orders of misunderstanding in such a way as to separate those conditions which will be corrected by the participants, so that the system continues in a steady state, from those others which lead to a progressive deterioration? At the present time such a question can only be posed in the most general terms and no meaningful answers can be imagined. One relevant matter must, however, be mentioned--that we deal with entities whose behavior is by no means describable in terms of linear equations or monotone logic. What actually seems to happen in many

instances is that when what seems to be progressive change sets in, the situation becomes more or less intolerable for one or both persons and some sort of climactic outburst occurs. Following this, the system either returns to a state which existed before the change began or entirely new patterns of communication may be evolved. There are, after all, larger and longer sequences of interchange than any which we meet with in the brief spans of data upon which this book is based.

From what little we know of the relationship between the fine details of human interaction and the longer cycles of the career line, there is reason to expect that the longer cycles will always be enlarged repetitions or repeated reflections of pattern contained in the fine detail. Indeed, this assumption that the microscopic will reflect the macroscopic is a major justification of most of our test procedures. A major function of the techniques of microanalysis is, therefore, to obtain from small quantities of data, accurately and completely recorded, insights into human relationship which could otherwise only be obtained either by long-time observation or from the notoriously unreliable data of anamnestic reconstruction.

In sum, we are concerned in this book to present the techniques for the microscopic examination of personal interaction. While, of course, the words that people say to each other have importance, the question with which we are concerned, the problem of describing the

relationship between persons, is not a question which can be answered by any summary of the dictionary meaning of their messages. There is a vast difference between the mechanical description "A gave B such and such information" and the description of the interchange "A answered B's question immediately."

The ultimate goal of the procedures outlined in this book is a statement of the mechanism of relationships. No statement of mechanism without larger context can be of long-term interest; no statement of relationship, unsubstantiated by a statement of mechanism, can warrant confidence. In order to trace the path from mechanism to validated relationship, it is first necessary to lay out for the reader some description of how the flow of linguistic and kinesic material can be systematically described.

CHAPTER 2

Vocal Activity

- 2.0. Introduction
- 2.1. The design of a language
  - 2.1.1. Introduction
  - 2.1.2. Morphemes and grammar
  - 2.1.3. Phonemes and Phonology
  - 2.1.4. Speaking as a process
  - 2.1.5. Morphophonemics
  - 2.1.6. Idiolects and dialects
- 2.2. The design of American English
  - 2.2.1. English phonology
    - 2.2.1.1. Consonants
    - 2.2.1.2. Vowels and nuclei
    - 2.2.1.3. Stress and juncture
    - 2.2.1.4. Intonation
  - 2.2.2. English grammar
    - 2.2.2.1. Functors
    - 2.2.2.2. Stems and their classification
    - 2.2.2.3. Constructions and construction-types
    - 2.2.2.4. Grammatical patterning in action
- 2.3. Nonlinguistic vocal communicative systems

Charles F. Hockett

## 2.0. Introduction

The vocal activity of a human being past earliest infancy is controlled to a large extent, though not exclusively, by a complex set of habits which we call his language. No two people have exactly identical language habits. To underscore this the total set of language habits of a single individual, at any given point in his life, is called his idiolect. It is obvious, however, that people who live together manage to understand one another's speech most of the time despite the differences. This common-sense observation leads us to speak freely of "English," of "French," of "Swahili," and so on--of different languages, each existing not monolithically but as a collection of more or less similar idiolects.

Not all vocal activity is linguistic (that is, governed by and conforming to language habits). We emit shrieks and groans, babbles and murmurs, laughs and weepings, when we are not talking--and sometimes when we are. We all stutter and stammer more or less as we speak or try to speak. Such disturbances in what would otherwise be a smooth flow of words are most easily dealt with in terms of a separate, nonlinguistic, layer of habits which have to do with the kinds of control<sup>1</sup> we exercise over our linguistic habits proper. Finally, speech in any language necessarily implies speaking

---

<sup>1</sup> See Appendix 1.

at a certain rate, and with a certain loudness, pitch, and tone quality. Some<sup>1</sup> features of these sorts constitute integral parts of a language; others<sup>2</sup>, however, are extralinguistic.

Our basic assumption is that all vocal activity audible to others--not just the linguistic segment--is communicative. That is: hearers pay attention, consciously or not, to the nonlinguistic features of vocal activity from a speaker as well as to the linguistic portion, and the former as well as the latter can be crucial in triggering responses. Our basic assumption is really broader than this, since we hypothesize that any human act detectable through any of the senses of other humans is communicative in the way described. In this chapter, however, we are concerned only with vocal activity. Our hypothesis assumes some kind of regularity and predictability in the responses of others. While we may for convenience focus on the behavior of a single person, we try always to see that person in his changing locus relative to others, triggering behavior on their part and being triggered by them.

A further assumption is that every aspect of vocal activity, nonlinguistic as well as linguistic, is learned and arbitrary, rather than innate or "natural." We must consequently expect the organization of nonlinguistic vocal activity

---

<sup>1</sup> See 2.1.3. below.

<sup>2</sup> See Appendix 1.

(and, indeed, of patterns of body motion) to vary from one society to another, just as the linguistic habits of a Frenchman are different from those of an American, and as Japanese landscape paintings do not look like those of the Western European tradition. All habits of vocal activity, and of responses thereto, are part of human culture, and, as such, are patterned, systematic, and therefore susceptible to analysis and description. We cannot be sure, of course, that these assumptions are entirely valid, but to assume the contrary would be an immediate counsel of despair--and this book would not have been written.

To make such assumptions is easy. To carry on empirical research based on them is another matter. Our scientific information is uneven. As of this writing, we know a good deal of the workings of human language; and our control of the details of some languages, including English, while not perfect, is not too lamentably bad. Nonlinguistic vocal activity, on the other hand, is a largely unexplored wilderness. The trails that have so far been blazed carry us only into its edges; we do not know how much more may lie beyond the point of deepest penetration. This imbalance is inevitably reflected both in the present chapter and in the whole investigation of which this book is the report.

## 2.1. The design of a language.

### 2.1.1. Introduction.

Here, though examples will be drawn largely from English, we shall concern ourselves only with those features of design believed to be shared by all human languages.

In my opinion--not all specialists agree--one of the most important properties of all languages is duality of patterning. To describe this property we shall turn first not to a language, but to a much simpler human communicative system--a simple substitution code of a variety often devised by American school-boys. The essence of the system is displayed in Figure 1. Here each of twenty-six simple geometrical figures is assigned an arbitrary letter-value: each figure means a certain letter of our conventional writing system. But the figures are also related, purely as to physical shape, in a systematic way. This is shown in the right half of the figure. Thus,

┘ is to ◡ as └ is to ┌;  
┐ is to ┌ as > is to < ;

and so on. Another way to point this up is to notice that each of the twenty-six figures is built out of a small stock of ingredients--straight-line segments in one of four orientations, and a dot--arranged in a strictly limited number of ways.

This systematic interrelationship in shape has nothing whatsoever to do with the meanings of the figures. If we concern ourselves only with the shapes, we can correctly assert such proportions as those given above. But these proportions do not hold between the meanings. Though one vertical segment, joining one horizontal segment at a right angle opening upwards and to the left, with a dot, means "q",

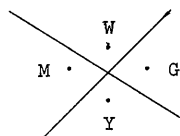
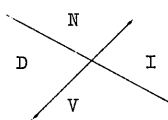
Figure 1

□	A	┌	U
◻	B	^	V
┐	C	∨	W
>	D	└	X
◻	E	^	Y
◻	F	└	Z

J	P	X
A	S	E
K	C	R

Q.	B.	. O
L:	T.	. F
Z:	H:	. U

<	G
◻	H
<	I
└	J
└	K
◻	L
>	M
∨	N
└	O
└	P
◻	Q
└	R
◻	S
◻	T



L	◻ ◻ ◻	◻ ◻ ◻	< ◻ ◻	◻ ◻ ◻	◻ ◻ ◻	∨ ∨
M	◻ ◻ ◻	◻ ◻ ◻	> ◻ ◻	◻ ◻ ◻	◻ ◻ ◻	◻ ◻
N	◻ ◻ ◻	◻ ◻ ◻	◻ ◻ ◻	◻ ◻ ◻	> ◻ ◻	<
O	> < >	^ ◻ ◻	◻ ◻ ◻	◻ ◻ ◻	◻ ◻	◻
P	◻ ◻ ◻ ◻	◻ ◻ ◻	< ◻ ◻	◻ ◻ ◻	◻ ◻	◻
Q	◻ ◻	◻ ◻	>			
R	◻ ◻	◻ ◻ ◻	◻ ◻	◻ ◻	◻ ◻	?

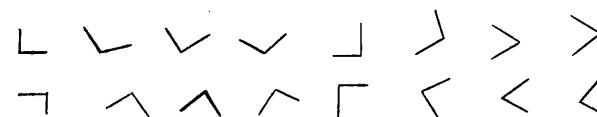
the three ingredients themselves--the vertical segment, the horizontal segment, and the dot--do not mean anything. Rather, the elementary ingredients of the figures, and their arrangements, serve only to keep the whole figure apart--they inform the recipient of a message whether one figure or another was intended by the writer. We say that the whole figures, and sequences thereof, have a semantic function, whereas the smaller ingredients have only a differential function.

The writer of a message in this system can afford to be fairly careless in his drawing of the successive figures, without giving the recipient too hard a task of decipherment. Thus if one writes

◻ < ◻ ◻ or even ◻ < ◻ ◻

the recipient will still quickly make out the intent as "SICK." In the first case, the second figure is tilted slightly counter-clockwise from the norm given in the table (Figure 1); but it is still closer to the norm for "I" than to that for any other meaning (say "X").

Figure 2



If the system included a larger number of two-segment figures (perhaps sixteen, as shown in Figure 2, instead of eight) then this much deviation from a norm would be confusing, and more care would have to be taken.

In the second case, the writing was so careless that the second figure is just as close to the norm for "X" as it is to that for "I". Here a further factor renders decipherment easy--a factor that plays a role in the first instance too, but a less crucial one: there is no ordinary written English word "SXCK," so that interpretation of the writer's intent as the figure for "X" yields an impossibility. What these examples show is that, in a communicative system of this sort, redundancy operates both within the realm of physical shapes and within the realm of meanings.

Any communicative system in which the smallest meaningful message-elements are built in a systematic way out of a stock of smaller meaningless but differentiating ingredients has duality of patterning. On one level, every whole message in such a system consists of an arrangement of minimum meaningful elements; on another level, every message in such a system consists of an arrangement of smallest differentiating ingredients.

Many systems either do not have duality, or have it in such a trivial or ultra-complex way as to be of no importance. As an example of utterly trivial duality, we can consider the system worked out by Paul Revere and his assistant for the transmission of one piece of vital information across the Charles River at night. One light was to carry one meaning,

two lights the other possible meaning. If we analyze the system with reference to meaning, we find just two minimum meaningful message-elements--namely, the two possible whole messages, one light and two. If we analyze purely with reference to physical shape, we discover, as the smallest differentiating ingredients, just the same two things: one light and two lights. The minimum meaningful message-elements and the smallest differentiating ingredients are identical, and there is thus no point in speaking of duality. Redundancy plays a role as in the system discussed earlier; but it plays that role only in one way--the light or lights can vary, without relevance, as to color, brightness, and relative placement.

Or we can consider ordinary English writing with the Latin alphabet. Roughly speaking--admittedly there are many irregularities and complexities--the minimum meaningful elements in English writing are letters, which have, as their meanings, the sounds we produce as we speak English. In a few marginal cases, we can find proportions which hold between the letters purely as to shape. Thus, in some type fonts, "p" is to "q", in shape, as "b" is to "d". But by and large there are no such simple relationships. Readers seem in the main to tell one letter from another in terms of total configurations.

The duality of language can be illustrated with English, but the reader must think in terms of spoken English in order to understand the example, even though we must necessarily present it via writing. Consider the following set of English words:

<u>pit</u>	<u>bit</u>	<u>pin</u>	<u>bin</u>
<u>pet</u>	<u>bet</u>	<u>pen</u>	<u>Ben</u>
<u>pat</u>	<u>bat</u>	<u>pan</u>	<u>ban</u>

Each of these words is a meaningful element, recurring from time to time as we speak. None of the words can be broken down into small elements that are also, on this level, independently meaningful (this is not true of all words, but it holds for the twelve listed here). Thus each word is a minimum meaningful message-element. But in physical shape--that is, in articulatory motions as we utter them, and in acoustic shape as we hear them--they are systematically related: pit is to bit as pat is to bat, bet is to pen as bat is to pan, and so on. Every utterance in a given language consists of an arrangement of minimum meaningful elements, drawn from a large but strictly finite stock of such elements. At the same time, every utterance in the language consists of an arrangement of minimum meaningless but differentiating ingredients, drawn from a much smaller stock. In describing any language, or languages in general (but not dual systems other than languages), we call the minimum meaningful elements morphemes, the smallest meaningless but differentiating ingredients phonemes (Bloomfield, 1933, p. 264).

We have seen that duality characterizes language and some, but not all, other human communicative systems. No known system used by non-human animals has been demonstrated to have duality. It appears not unlikely that among human systems, language is the earliest to achieve duality. Moreover, wherever

the property appears in simpler and more recent human systems, it may be interpreted as having developed in them directly or indirectly on the model of duality in language. All this does not necessarily imply that duality is the most crucial property of language--another property of basic importance will be described later. We have put duality first because it affords the most logical basis of organization for the further remarks that must be made about language design.

We asserted in our introductory remarks that the vocal behavior of humans is only in part governed by language habits. In analyzing vocal behavior, how are we to determine which portion is linguistic? Duality gives us the most useful criterion for the definition of language, however hard it may be to apply in practice: that segment of vocal activity which analysis shows to be dually patterned, having both morphemes and phonemes, is linguistic. The remainder, whatever its communicative relevance, is by this definition not language.

Here is a preliminary example. On two successive mornings, Jones comes into the corner drugstore and says to the clerk Coffee and doughnuts, Pete. The first morning, Pete merely fills the order; the next morning, as he is filling the order he remarks You sound tired this morning, Mr. Jones. Pete's observation may be quite accurate, in the face of absolute phonemic and morphemic identity between Jones's utterances on the two mornings. If so, then the response is triggered by nonlinguistic features of Jones's speech--a slower tempo, a "thinner" voice, less precise articulation of vowels and



consonants, and the like. Variations of these sorts are communicative, and are patterned--but they are not dually patterned, and are hence not part of language.

We are now ready to discuss, in turn, the two levels of patterning found in language and the relationships between them.

#### 2.1.2. Morphemes and Grammar.

As we have said, the smallest individually meaningful constituent signals in a language are morphemes. In the English utterance She bought a new hat, as ordinarily spoken, there are seven constituent morphemes: she, bough- (the "same" morpheme as buy, for a reason we shall see later), -t (past time), a, new, hat, and an intonation. Some acts of speech are broken off by intrusive elements--coughing, interruption by someone else, and the like. Setting aside such fragments, every utterance in a given language is composed wholly of an arrangement of an integral number of morphemes drawn from the morpheme-stock of the language.

In English, and quite possibly in all languages, at least one constituent morpheme of every utterance is an intonation; an assenting grunt Hm! or a querulous Hm? from the point of view of the language consists wholly of an intonation. Every adult speaker of a language controls several thousand morphemes; in English, about one hundred of these are intonations. The language as a whole can be said to provide an even larger stock of morphemes, since there are always some known to some speakers but not to others.

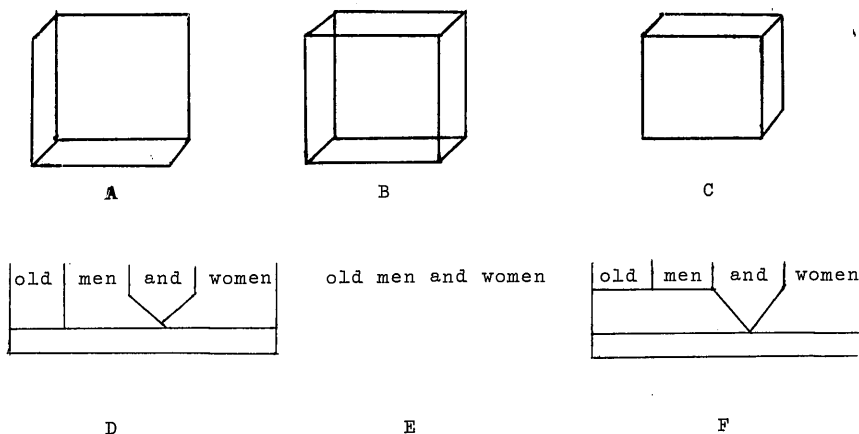
As might be expected, the matter of meanings is much more complicated for a system as complex as a language than it is for the simple substitution code discussed earlier. Some morphemes have meanings only in an indirect way. The larger combinations in which they occur have meanings different from the meanings of the combinations in which they do not occur. One can describe, at least in part, the meaning of a morpheme such as boy, girl, table, by pointing to things in the world around us; one cannot do this for such morphemes as and, or, if, a, the. Yet the meaning of men and women is different from that of men or women, and the meaning of match and book is different from that of match book, and this is enough to show us that and, or, and the like must be classed as morphemes. The meaning of a whole utterance results not alone from the meanings of the constituent morphemes, but also from their arrangement: dog bites man versus man bites dog; match book versus book match. Thus arrangements, as well as morphemes, have meanings or contribute to the meanings of larger messages.

Physically, arrangements of morphemes in utterances are linear in time and with overlap (the intonation of our sample utterance She bought a new hat is delivered simultaneously with the words). Communicatively, however, arrangements are hierarchical. In a context in which the phrase old men and women obviously means "women of all ages, and elderly men," the speaker intends the and to join the constituent parts old men and women, and his hearers so interpret it. In a context in

which the same phrase means "elderly men and elderly women," the speaker intends old to go with the whole smaller phrase men and women.

Sometimes speaker's intention and hearer's interpretation are at odds, or the hearer is uncertain; this may or may not lead to further misunderstanding. Although structural ambiguity of this sort is not uncommon, usually the specific morphemes, and the linear sequence in which they occur, allow of only one interpretation within the economy of the language. The superposition of hierarchical organization on a physically linear arrangement by a hearer is a Gestalt phenomenon, like the reading of depth into what is physically an assemblage of line segments on a plane surface. Figure 3 shows this, and also displays one device for the graphic representation of hierarchical organization.

Figure 3



Gestalt perception in vision and in the hearing of speech. A, B, and C are all assemblages of line segments on a flat surface. We see depth in all three; furthermore, B can be seen as more like A or as more like C. E stands for the physically linear arrangement of morphemes as we hear the phrase. D and F represent two different hierarchical ways in which the constituent morphemes can be organized by the hearer.

There are always stringent limitations on the arrangements in which the morphemes of a language can occur relative to one another in speech. Thus, in English, John do-es not like Mary occurs; not-es Mary John like do is nonsense. In one sense, these limitations constitute constraints on speakers. Speakers are not free to speak in any manner they please. The truth of this is not altered by the fact that the constraints are probabilistic rather than absolute. In another sense, the limitations are not so much constraints as smooth grooves along which a speaker slides with minimum effort. In the same way, a network of highways enables one to travel easily wherever one wishes to go--so long as one keeps one's car on the road and within the network.

This comparison breaks down in one respect. One can reach East Micro Junction by car only if a highway has already been built to or through that town. In a language, one can reach points that have never been reached before; for speaking a language is comparable, at one and the same time, both to road building and to highway travel. That is, a speaker may say something that neither he nor his audience has ever

heard or said before, and be perfectly understood, without any of the participants being in the slightest aware of the novelty. As an example, the reader need only consider the present paragraph: every morpheme and every way of putting morphemes together that is involved in this paragraph is familiar, and yet the specific combination is new.

This property of languages as communicative systems is called openness, and it should be obvious that openness is a trait of even more fundamental importance than duality. Yet whereas no non-human animal communicative systems seem to have duality, at least one such system shows openness: bee-communication. Thus a worker can report to her fellows the location of a source of nectar at a place where neither she nor her fellows have ever found nectar before. Yet we search apparently in vain for open systems in use by our nearest non-human relatives. Gibbon vocalization, for example, is apparently constrained to a discrete choice from a small finite set of calls: no matter how novel a situation, the communicative vocal response to it can only be one or another of these calls, not a new call consisting of two old ones in sequence or of a blending of ingredients from two or more of the old ones. The practical difference between such a closed system and an open system like human language hardly needs underscoring.

The mechanism of openness is at bottom simple: a novel utterance is produced on the analogy of various utterances of various structures which have occurred previously (in the speaker's experience) in partly comparable circumstances.

Thus the new utterance consists of individually familiar elements (morphemes), put together by familiar patterns: only the total result is new.

A result of limitations on arrangement is that any single morpheme in a language is characterized by a certain roughly definable range of privileges of occurrence relative to other morphemes. Sets of morphemes which have approximately the same privileges of occurrence constitute form-classes. In any language, some of these form-classes are very large (English man, boy, child, woman, dog, cat, etc.) and some are very small (English a and an). The very small form classes usually include only morphemes of the sort that have an indirect bearing on meaning (and, or, etc.), rather than a denotation at which one can point.

The morpheme stock of a language, and the hierarchical arrangements in which the morphemes occur relative to one another in utterances, constitute the grammar of the language. Both elements and arrangements differ from language to language. That the elements differ, not only in sound but in range of meaning, is obvious: the central meaning of English chair, for example, is broken into two ranges of meaning in French, in which one says chaise or fauteuil depending on whether the sittable is hard or stuffed; what the Cree Indians express with the single word ēskēw has to be conveyed in English with the long phrase he makes a hole in the ice to hunt beaver.

Arrangements are also incommensurable from one language to another. If only the elements were different from language

to language, then all we should have to do to master a foreign language would be to learn its vocabulary, and anyone who has tried knows that it is not that easy. To translate English I do not have time word-by-word into German yields nonsense (ich tue nicht haben Zeit!); good German Das hat alles über den Haufen geworfen yields comparable nonsense when subjected to the same sort of translation in reverse--That has everything over the heap y-thrown. Nor is it valid to say--as we may be inclined to with reference to our own language, in which everything seems reasonable and proper merely because of familiarity--that arrangements of morphemes occur if they make sense. It is nearer the truth to say that they make sense--to those who speak the language--if they occur. Even this is not the complete truth, since the perfectly natural process of coining new utterances from time to time yields something meaningful to the speaker but cryptic to others.

Words are recurrent clumps of morphemes--single morphemes in the limiting case--which manifest a degree of cohesion greater than that shown in the same language between successive morphemes in different words. In English, we can in general rely on our orthographic habits of leaving spaces at certain points as a clue to word-identity: boy is a word of one morpheme; boys and boyish are words of two morphemes each; boyishly is a single word of three morphemes. The clue is not infallible: blackboard is two words (each one morpheme) whether written that way, or with a hyphen (black-board) or with a space (black board).

Some morphemes, especially all intonations, belong to no word, but form additional non-word constituents of utterances. Conventional English orthography does a very inadequate job of representing these. Therefore the non-word morphemes supplied by a reader as he peruses ordinary written English are often different from those used by the writer as he wrote. Languages differ as to the average number of morphemes packed into single words, and as to the precise kind of intraword cohesion. English thus averages more morphemes per word than Chinese, but fewer than Spanish, which in turn averages fewer than Latin, and Latin fewer than the Algonquian languages of aboriginal North America. Such differences are relevant for purely linguistic purposes, since a language with relatively more complex words has more morphology to be described (morphology being that compartment of grammar in which words are built out of morphemes) and comparably less syntax (the building of whole utterances out of words and non-word constituents), than does a language with relatively simpler words. But these differences have never been shown to correlate with any other aspect of the life of the speakers of the languages.

Certain combinations of morphemes--sometimes exactly one word in size, sometimes smaller or larger than single words--have meanings not predictable from their constituent morphemes and the arrangements of the latter. Any such combination, and any single morpheme, is an idiom. Examples of longer idioms are white paper (governmental document), marriage of convenience, run out ("become depleted"), Statue of Liberty (in New York bay),

statue of liberty (in football). It is well to recall, too, that there are fixed longer discourses such as a poem or a cliché which are also idiomatic. Awareness and consciousness are obviously quite similar in meaning, and interchangeable in most contexts; the fact that self-awareness and self-consciousness have sharply different meanings shows that the latter two are both idioms. Idioms are learned and used as wholes; but, also, new idioms are coined in various ways, and old ones can become disused. If we measure vocabulary in idioms instead of morphemes, the count for the ordinary speaker runs to the tens of thousands: one<sup>1</sup> estimate is that the average person past infancy controls about one thousand idioms for each year of his age, though perhaps with a decline in the rate of increase past the mid forties.

### 2.1.3. Phonemes and Phonology.

Phonemes are the smaller meaningless constituents out of which morphemes, combinations of morphemes, and whole utterances are built, and which serve to keep morphemes and whole utterances apart.

The words pit and bit, said in isolation as whole utterances, differ in sound only at the beginning, and are in certain aspects the same there. In terms of articulatory movements, the lips and the nasal passages are closed for both. For pit the vocal cords are quiescent (the p is voiceless), while for bit they are in vibration (the b is voiced). There is no actual or possible third morpheme in English, differing from pit and

---

<sup>1</sup> Martin Joos in conversation.

bit only in that the initial consonant is more strongly voiced than for pit yet less strongly than for bit. Out of the continuum of physiologically possible degrees of voicing, only two ranges are selected and assigned contrastive function in English. No matter how sloppy a speaker's control of voicing may be as he says one of these two words, a hearer who know the language is absolutely constrained to interpret the result as exactly the one word, or as exactly the other, or as so badly articulated that identification is impossible. The last of these alternatives is not a third possibility within the limits of the system, but a repairable breakdown of the system. Of course, hearers may notice the production of a mumbled and undifferentiable pit/bit by a speaker, as they may note, say, a bit pronounced with an unusually strongly voiced b; and either of these may convey information of some sort. But the communicative function of such variations stands outside of language proper, within which only discrete<sup>1</sup> contrasts between quantized points or ranges of speech sound are relevant.

The phonemes of a single language are thus not speech sounds, but rather ranges of speech sound--bull's eyes, as it were, at which a speaker aims his articulation. The phonemic system of a language is not so much a stock of "things" as it is a network of distinctive articulatory-acoustic differences. A particular phoneme in a system is defined functionally not so much by what it is physically, or what it sounds like, as by

---

<sup>1</sup> Se Appendix 1, p.

what, in the same system, it differs from. This means that even if phonemes in two languages sound roughly the same, they may still differ radically in function. English makes systematic use of two degrees of voicing (voiceless, as p, and voiced, as b), but Menomini, and many other languages, make no use of this scale at all, so that the Menomini phoneme which sometimes sounds to us like p is not functionally commensurate with either English p or English b. Phonemes in different languages are never commensurate; only phonemic systems can be compared.

No known language requires articulatory motions of any part of the body except the respiratory and upper alimentary tracts, for the production of its phonemes. A few languages do not use all the manipulable parts of this region: for example, in most of the Iroquoian languages the lips do not function; and in Quileute there is no distinctive contrast between nasal passages closed and nasal passages open, so that there are no phonemically separate "nasal" sounds like English m, n, ng (come, can, king).

What we ordinarily call "vowels" and "consonants" are kinds of phonemes (segmental phonemes), but there are other kinds as well. There are phonemes involving the pitch of the voice, the general strength of articulation, differences of timing and duration, and the like. The contrast between the noun PERmit and the verb perMIT is a matter of distribution of strength of articulation, not of vowels and consonants. In Chinese, a set of syllables all of which sound something

like English you mean "oil," "have," "again," and "swing," depending on the tonal contour which accompanies the vowels and consonants. Chinese uses pitch both in the way just illustrated and also as the raw material for intonation phonemes; English uses pitch only in its intonation phonemes.

A phonemic system, then, consists of a stock of phonemes and of the arrangements in which they occur relative to one another in utterances. There are always stringent limitations on these arrangements, just as, in grammar, there are for morphemes. Phonemic systems differ from each other in two ways: (1) as to the particular quantizing of the multidimensional continuum of possible articulatory motion, and resulting sound, into contrasting ranges of speech sound, and thus as to the number of phonemes in the system; (2) as to the arrangements in which the phonemes can occur.

Systematization on the level of phonemes is, if anything, even more rigorous than on the grammatical level. Standard German has twenty-seven segmental phonemes; American English thirty-three. German has no separate phoneme acoustically similar to the English th of thick; English has none similar to the German ch of Bach. German and English share acoustically comparable phonemes t (G Tag; E time) and s (G Wasser; E sing); in German the initial sequence ts occurs (Zeit, pronounced Tseit), but not in English; in English the initial sequence st occurs (sting), but not in Standard German. Finer (and trickier) differences of this sort appear among the dialects of a single language; coarser (but also trickier) differences

appear among languages less closely allied than are English and German--say, between either of these and Chinese.

#### 2.1.4. Speaking as a Process.

When we examine an utterance via a written record, we are free to read it from beginning to end or to look back and forth from one part to another in any way we please. Of course, this is not the way utterances really happen, and our culturally imposed preoccupation with writing must not be allowed to mislead us. Language manifests itself first and foremost in the behavior of speaking. An utterance is not spoken all at once, but bit after bit in time: speech is not only unidimensional but unidirectional. Each successive bit is spoken and then vanishes into limbo, never to be recovered or unsaid.

It is convenient to pretend that a speaker "first" chooses the morphemes he wishes to transmit (perhaps very few at a time), arranging them appropriately, and "then" encodes the result into an array of phonemes for articulation--perhaps as he proceeds to the selection of the next few morphemes. Conversely, the hearer "first" interprets the incoming speech signal into an array of phonemes, and "then" decodes the phonemes into morphemes.

The convenience is twofold. First, we may thus imagine the process of morpheme emission as going on continually in any human past infancy. When not being directly guided by contemporary input from elsewhere in the person, it is driven, as it were, under its own power; but only sometimes does it "break through" to the surface to be audible to others. The progression

is probabilistic, not determinate. It is the arrival of a morpheme at "headquarters" that triggers the emission of the next morpheme, and it does not matter whether the incoming morpheme has been transmitted via feedback routes from the hearer's own "headquarters" or through the air from someone else. Thus, self-communication, either aloud or "thinking in words," mocks speech communication between people.

There is also only a probabilistic tie between the morpheme-flow within a person and other events within and around him. The tie is in both directions. Sight of a cat may or may not lead a speaker to say There's a cat: the probability that it will is conditioned by where his internal morpheme-flow is at the moment. If the self-driven probabilistic progression leads to the emission of the morpheme cat, that increases at least somewhat the probability that the speaker will look around for a cat.

Thus the "impact of language on thought and behavior," as discussed by Whorf, can be understood behaviorally; it will differ in kind, if not in degree, from language to language. So, also, can we understand one variety of Korzybskian "identification": if an experiential stimulus (say the sight of a plate of food of a certain kind) triggers the emission of a certain train of morphemes ("This stuff is Irish stew"), subsequent behavior may be controlled largely by the economy of the individual's idiolect--which, in its turn, derives from a whole history of previous experiences, of the individual and of his predecessors--rather than by the immediate idiosyncratic reality

("I don't like Irish stew"). It must never be forgotten that the generation of the stream of speech, even internally, involves work, and that it may be more work to force the succession of morphemes to match specific novel facts than it is to let the progression follow geodesic lines of highest intra-systemic probability.

#### 2.1.5. Morphophonemics.

The second convenience of the mode of description adopted above (2.1.3.) is that it helps us to specify the remaining aspect of language design: the morphophonemics of a language is the code by which arrays of morphemes are transduced into arrays of phonemes and vice versa.

In any single context of other morphemes, a given morpheme is represented by a specific arrangement of phonemes. Some morphemes are represented in the same way in all occurrences: pay in pay, pays, paying, paid, payer, etc., and in all the longer utterances which incorporate any of these words. Two morphemes may share a representation: pair "couple" and pare "to peel." A single morpheme may have more than one representation, depending on linguistic environment: wife with a final f in the singular, but wive- before the pluralizing morpheme -s; the latter is pronounced s in cats, pits, but z in cads, ribs, wives; buy is changed to bough- when followed by the past tense ending -t.

Some such alternations in phonemic shape are imposed by phonemic habits. English phonemics allows us to say -ts and -dz

at the ends of words (cats, cads), but not -tz or -ds, so that the use of both -s and -z for the pluralizing morpheme is "natural" within the language (not in any more general sense, of course: some languages allow sequences like tz and ds). Other alternations are survivals of earlier historical accidents; buy and bough-, since there is nothing in English phonemics to preclude us from saying buyed. Any of the latter is a morphophonemic irregularity.

Languages differ as to the number and importance of their morphophonemic irregularities. Russian ranks high in irregularity; German is less irregular, but still slightly more so than English; Hungarian is somewhat more regular than English, and Chinese is the most regular of the five languages we have just named. In language-learning, morphophonemic regularity is an obvious advantage. For an adult student of a second language, this advantage can be outweighed by other factors: e.g., German is easier for speakers of English to learn than is Chinese, because vocabulary and pronunciation are less alien. It has been suspected that for a child learning his first language these other factors can play no part, so that, for example, the Hungarian child should learn his language somewhat more rapidly or easily than the American or Russian child. This hypothesis seems reasonable to me, but so far as I know it has never been supported by carefully controlled observation.

#### 2.1.6. Idiolects and Dialects.

A person's idiolect is a product of a succession of different contacts: each period of residence in one or another



community, each period of interaction with a given set of friends and associates, contributes a layer, its importance depending on the length and intimacy of the contact and on the individual's receptivity at the time. Early habits can be apparently submerged by those acquired later, but they are perhaps never altogether lost. Special circumstances can bring them again to the surface. One woman, raised in the Bronx, went to the Middle West at the peak of her adaptability (age 20), and now, fifteen years later, usually speaks fairly typical Middle Western English, with a few generalized East Coast features but with no clear indications of the Bronx. But whenever this woman talks over the telephone with her parents or sister (who did not move west), typical Bronx features reappear. They also turn up when she is engaged in certain emotionally colored types of argument with her children. In the latter context, there is presumably some cue of similarity to her earlier relations with her own parents that elicits this shift in speech-pattern.

A dialect is a variety of a language characteristic of a particular geographical-social region. Dialects vary from one geographical area to another, and from one social<sup>1</sup> stratum or occupational grouping to another, sometimes slightly, sometimes radically. Regional and social dialect continuity seems to be

---

<sup>1</sup> For social concomitants of dialect variation see McDavid, 1948, 1951, 1952-3.

maintained in this country in the face of considerable mobility of individuals. If a given dialect loses speakers through emigration or social-climbing, in compensation it gains the children of new arrivals. The idiolect of a person who stayed near his birthplace indefinitely might reflect features typical of the dialect of his region and social stratum. If all speakers were like this, then we could say that a dialect consists of a set of closely similar idiolects, just as a language, in turn, includes one or more fairly similar dialects. But this assertion does not hold for American English. Many an idiolect far exceeds in dialect affiliation the complexity of that of the woman mentioned in the preceding paragraph, including within itself features from two or more dialects, the speaker's environing circumstances eliciting now the features of one dialect, now those of another.

Dialect differentiation also shows certain partial correlations with age-grading and with what we may call sex-grading. In addition to purely linguistic features, certain typical "tones-of-voice" tend to be differentially associated with different age-grades and sex-grades. Age-grading is self-explanatory. For example, within many social strata in many parts of the country, the typical voice and speech of an upper-teen-aged girl are quite different from that of a matron in her thirties, and if we hear the teen-aged voice from a matron there is an incongruity. By "sex-grading" we refer not merely to male versus female, but to the two or three different culturally transmitted manners in which the fact of being a male, or a

female, can be acted out. The "locker-room male"--taught how to be a male by other males, in a context of emphasis on physical sports--acts and speaks differently from the male who has been conducted through childhood and adolescence largely by women. Statements of male homosexuality, as a "tertiary" (i.e., culturally developed) sexual characteristic sometimes unaccompanied by any overt homosexual tendencies, can show up in certain features of voice quality and of the articulation of vowels and consonants.

Interpersonal differences in speech habits are quite like interpersonal differences in all other varieties of communicative behavior. In observing and comparing people, we must remember, however, that interpersonal differences are of two orders. Obviously, what John Jones signals on a given occasion (in words or otherwise) differs from what Jim Brown might signal in the "same" circumstances, and from what John Jones might signal in other circumstances. But there are also interpersonal differences in the systems, and it is primarily to those that we refer when speaking of dialects and of idiolects. We cannot understand what John Jones means by his signals without knowing the idiosyncratic twists given to each communicative system by John. Thus a nod means "yes" to us, "no" to an Eskimo; homely means "pretty" in Australia, "ugly" in the United States; the intonation of a polite request in British English is that of exasperation in American English; lost in some Southern speech sounds exactly like loused for some Northerners. Children whose control of their language is not yet fluent often answer yes or no for lack of the machinery for a more precise response:

parents may take this as a falsehood when it is not. Blood--the word, not merely the substance named by it--connotes laboratory work to a serologist; it may send someone else into a faint.

## 2.2. The design of American English.

A complete description of any language--one spoken by hundreds of millions of people, like English, or one spoken by a few dozen people in some New Guinea village--is a lengthy and laborious matter. A complete description of English is impossible here; fortunately, it is also unnecessary. What follows is an outline<sup>1</sup> or frame of reference, describing the basically important habits of the language, and supplying appropriate points of attachment for the multitudes of details which have to be omitted.

Grammar is always more complicated than phonology; consequently, our discussion of English phonology will be relatively less sketchy than what we say about English grammar.

### 2.2.1. English phonology.

What we shall present on phonology is not the exact system of any single idiolect or dialect, but the overall pattern of (General) American<sup>2</sup> English: a system incorporating all (or almost all) of the contrasts functional in any idiolect, so

---

<sup>1</sup> For fuller treatments, see Trager and Smith, 1951, and now, since this writing, Hill, 1958, and Sledd, 1959.

<sup>2</sup> For a fuller treatment of American dialect variation see McDavid, 1963.

that the precise system of a single idiolect is a weighted selection from the overall system. The logic of the overall pattern approach will become clear from the examples. Certain Southern and Canadian varieties of American English may require slightly different overall patterns, but this possibility has not been sufficiently explored for consideration here.

Any utterance in (General) American English is built out of a selection from a stock of forty-four phonemes. In presenting and describing these phonemes, we also supply symbols for their graphic representation. Since many of the symbols are identical with letters of ordinary English spelling, they will regularly be cited between slant lines (thus: /a/), except in tabular displays.

Segmental phonemes:

Consonants:

<u>Full Consonants:</u>	p	t	č	k
	b	d	ǰ	g
	f	θ	s	š
	v	ð	z	ž
	m	n		ŋ
		l		
		r		
<u>Semiconsonants:</u>	y	h	w	
<u>Vowels:</u>	i	ɪ	u	
	e	ə	o	
	æ	a	ɔ	

Nonsegmental phonemes:

<u>Juncture:</u>	+			
<u>Stresses</u> (written over a vowel):	á	â	à	
<u>Intonation phonemes:</u>				
<u>Pitch Levels (PLs):</u>	1	2	3	4
<u>Terminal Contours (TCs):</u>			#	

2.2.1.1. Consonants.

The first eight of the full consonants are illustrated by the following words:

/p/	<u>pull</u> /púl/	<u>lip</u> /líp/	/b/	<u>bull</u> /búl/	<u>sib</u> /síb/
/t/	<u>tell</u> /tél/	<u>let</u> /lét/	/d/	<u>dull</u> /dél/	<u>led</u> /léd/
/č/	<u>chill</u> /číl/	<u>rich</u> /ríč/	/ǰ/	<u>Jill</u> /ǰíl/	<u>ridge</u> /ríǰ/
/k/	<u>cull</u> /kél/	<u>luck</u> /lók/	/g/	<u>gull</u> /gél/	<u>lug</u> /lóg/

These eight are stops: they are articulated with closed nasal passages and with a complete closure somewhere in the oral cavity. For /p b/ the oral closure is between the lips (bilabial), for /t d/ between the tip of the tongue and the alveolar ridge (apico-alveolar), for /č ǰ/ between the blade of the tongue and the general region of the alveolar ridge (lamino-alveolar), for /k g/ between the back part of the upper surface of the tongue and the velum (dorso-velar). /b d ǰ g/ are voiced, and /p t č k/ voiceless. /č ǰ/ are released in such a way that the air passes between the blade of the tongue and the alveolar ridge to produce audible friction; for this reason these two are called affricates. The other six stops

are not affricates.

The next eight full consonants are illustrated in the following words:

/f/ full /fúl, cuff /kóf/    /v/ vat /vát/, live /lív/  
/θ/ thumb /θém/, myth /míθ/    /ð/ thus /ðés/, scythe /sáyð/  
/s/ sell /sél/, less /lés/    /z/ zip /zíp/, haze /héyz/  
/ʒ/ shell /šél/, mesh /méš/    /ʒ/ Jeanne /žán/, rouge /rúwž/,  
vision /vížán/

These eight are spirants (or fricatives, whence the term "affricate" for /ʒ ž/): the nasal passages are closed, and a partial closure is made in the oral cavity so that the air, in passing through, is forced into turbulence. For /f v/ the partial closure is made between lower lip and upper teeth (labio-dental); for /θ ð/, between the front edge of the tongue and the backs of the upper teeth, the air escaping through a transverse slit (apico-dental slit spirants); for /s z/, between tip of tongue and alveolar ridge, the air escaping through a tiny hole in the median line (apico-alveolar rill spirants); for /ʒ ž/, between blade of tongue and the region of the alveolar ridge (lamino-alveolar). /f θ s ʒ/ are voiceless, /v ð z ž/ voiced. Some speakers do not use /ž/ either initially or finally (they do not use the "French" pronunciation of Jeanne, and end rouge, garage with the affricate /ʒ/); this is why we have also given an example in medial position above.

The next three full consonants:

/m/ mull /mól/, dumb /dém/  
/n/ null /nól/, dun, done /dén/  
/ŋ/ singing /síŋŋ/.

These three are nasals: a complete closure in the oral cavity, but with the nasal passages open. For /m/ the closure is bilabial, for /n/ apico-alveolar, and for /ŋ/ dorso-velar. All three are voiced. /ŋ/ does not occur initially. Note the difference between singer /síŋər/ and finger /fíŋgər/: some speakers with a direct or indirect background of certain Central and East European languages say singer /síŋgər/, sing /síŋg/, singing /síŋgŋg/, and have no /ŋ/ except directly before /g/ or /k/ (sink /síŋk/).

The last two full consonants are sometimes classed together as liquids, a term which makes no particular reference to articulation:

/r/ red /réd/, tar /táhr/ (some speakers /táh/)  
/l/ led /léd/, dell /dél/

For /r/, the tip of the tongue is curled back, but does not touch anything (a retroflex glide vocoid). For /l/, the tip of the tongue touches the alveolar ridge, but the sides are not in contact with the sides of the mouth, and the air can pass through without any friction (apico-alveolar lateral). Both are voiced, and for both the nasal passages are closed. Many speakers, in parts of New England, the Hudson Valley and New York City, and certain regions in the old South, have /r/ only before a vowel.

The three semiconsonants occur before vowels, as follows:

/y/ yes /yés/, yacht /yát/ (or /yáht/)

/h/ hem /hém/, hut /hát/

/w/ wet /wét/, wind /wínd/

For /y/, the blade of the tongue approaches the hard palate and is pushed forward in the mouth, for /h/ the position of the tongue approximates that of the following (or preceding) vowel, and for /w/ the back part of the tongue approaches the soft palate. The lips for /y/ are spread, for /w/ rounded, and for /h/ accommodated to the lip position for the following (or the preceding) vowel. Both /y/ and /w/ are voiced, whereas /h/ is voiceless before, but voiced following, another vowel.

The full consonants and the semiconsonants occur in various sequences uninterrupted by any vowel. These combinations are called consonant clusters. Initial clusters have as many as three constituents: sigh, spy, spry /sáy spáy spráy/, well, quell, squelch /wél kwél skwélč/. Final clusters have as many as four: sing, sink, sphinx, jinxed /sín sínk sfíŋks jíŋkst/. Of the thousands of mathematically possible consonant clusters, only a few hundred actually occur, and some of these are extremely rare. Thus some speakers use initial /ts/ in a few uncommon words (tsar /tsáhr/, tsetse /tsíytsiy/, Tsimbian /tsímšiyín/), while many speakers do not use this initial cluster at all.

#### 2.2.1.2. Vowels and Nuclei.

Consonant clusters directly after a vowel do not have a semiconsonant as first constituent, because when a semiconsonant

occurs directly after a vowel the combination of vowel and semiconsonant forms a complex syllable nucleus or complex syllabic.

A vowel with no following semiconsonant constitutes a simple syllabic. Since there are nine vowels and three semiconsonants, there are thirty-six syllabics in all:

simple	complex /-y/	complex /-h/	complex /-w/
i	iy	ih	iw
e	ey	eh	ew
æ	æy	æh	æw
ɪ	ɪy	ɪh	ɪw
ə	əy	əh	əw
a	ay	ah	aw
u	uy	uh	uw
o	oy	oh	ow
ɔ	oy	oh	ow

It is not unlikely that most speakers of American English use most, if not all, of these thirty-six nuclei. But in any single idiolect some of them are commonest and a few are very rare, and the distribution of common and rare varies widely from one variety of English to another. We illustrate here all the simple nuclei and the most widespread of the complex syllabics:

/i/: hip, bit, kick /híp bít kík/

/é/: bet, neck, kept /bét nék képt/

/æ/: cap, hat, batch /káp hát báč/, except that some speakers (e.g., in Western New York State) have /əh/ or /éh/ in almost all such words.

/ɪ/: just /jɪst/ in I just got here (not in a just man, which has /jɛst/); often in this /ðís/); common in children,

silver /ʃíldrín sáilver/; in parts of the South in sister, dinner /sístə(r) díne(r)/; the commonest syllabic in unstressed syllables.

/ʊ/: but, bud, bun, buzz /bét béd bén béz/; southern British English has /á/ in such words.

/á/: cot, lock, hop /kát lák háp/, except that in Eastern New England, as in southern British English, these words have /ó/.

/ú/: book, put, puss /búk pút pús/.

/ó/: most widespread in gonna /góna/ (=going to: I'm gonna stop now), whole /hól/ (the whole thing, not make me whole again, which has /ów/), because /biykóz/ (also with /é/ or /óh/); in parts of rural Eastern New England in home, road, coat and some other household words /hóm ród kót/, elsewhere most generally /ów/.

/ó/: not common, except for the British use in cot, lock, hop, etc. (See above under /á/); some New England rural types may have this rather than /ó/ in home, road, coat; I have it before /l/ in balm, calm, psalm, alms /bólm kólm sólm ólmz/.

/íy/: beet, peep, leak, reach /bíyt píyp líyk ríyč/. In bee, bean, seem, leave most speakers have /íy/, but some in the region of Philadelphia and Baltimore have /áy/.

/éy/: pate, cape, sake /péyt kéyp séyk/.

/áy/: I, lie, tight, spike /áy láy táyt spáyk/.

Many New York City speakers have /óy/ in these words; in the South, /áh/ is common; in Central and Western Canada /éy/ is usual before a voiceless consonant, as /téyt spéyk/, but /áy/ otherwise.

/óy/: boy, hoist, coil, oil /bóy hóyst kóyl óyl/.

Some Southern speakers have no cases of this, sometimes using /óh/ instead (as before /l/: /kóhl óhl/), sometimes /áy/ (boil, "heist" /báyl háyst/), sometimes pronouncing two syllables where other speakers have one (boy /bówiy/).

/éh/: commonly in yeah /yéh/; in the Central Atlantic Seaboard in can "container" /kéhn/, in contrast with can "be able" /kán/. Many Middle Westerners have this or /éh/ or /éy/ almost to the exclusion of the /ə/ found in the East and in New England in words like cab, cad, fad, sand, ham (/éh/ or /éh/), bag, ashes, back (/éy/).

/áh/: speakers (mainly New England and the East) who distinguish between bomb /bám/ and balm /báhm/ have /áh/ also in words like father, starry (vs. sorry with /á/), Dali (vs. dolly with /á/ or /ó/). In the Middle West /áh/ is more usual before a voiced consonant, /á/ before a voiceless (e.g., bomb and balm both /báhm/, but /á/ in pot). Everywhere words like spa, fa, la, rah, pa, ma end with /áh/, except that some speakers say rather paw, maw with /óh/.

/óh/: widespread in law, saw, ought, haul /lóh sóh óht hóhl/ and the like. But in New York City these words usually have /óh/; in the Northwest including Central and Western Canada, they have /áh/ or /á/ (so that cot and caught are identical, /kát/, and cod and cawed are both /káhd/); and in many parts of the South they all have /ów/.

/aw/: house, cow, about, loud. In parts of the South one hears /əw/ (or even /əh/ or /əy/), in the Virginia Tidewater /ów/ is customary. In Central Canada and

the Northern Middle West, /šw/ is usual before a voiceless consonant (house /hšws/) but /šw/ otherwise (houses /hšwzšz/).

/úw/: boot, spook /búwt spúwk/. In too, boom, moon most speakers have /úw/, but some in the region of Philadelphia and Baltimore have /šw/.

/šw/: cope, boat, roach, poke /kówp bówt rówš pówk/. Some speakers (East?) have /šw/, also common in British English in these words; an extreme Britticism is /šw/.

From these examples we can describe the pronunciation of the vowels and semiconsonants in articulatory terms. The three vowels /i e ə/ are all pronounced with the tongue bunched towards the front of the mouth (all are front vowels): but they differ in that /i/ has the tongue high /ə/ has it low, and /e/ has it at an intermediate height (mid). As shown by the arrangement of vowel symbols on the Table of English Phonemes, the same three-way difference of tongue-height distinguishes among the members of the triad /ɪ ə a/, and also among /u o ɔ/. For all six of these, the bunching of the tongue is towards the central or back part of the mouth instead of forward: all six are back vowels. /u o ɔ/ are perhaps typically a bit further back than /ɪ ə a/; more important is the fact that the lips are slightly rounded for /u o ɔ/, but not for /ɪ ə a/ (nor for /i e ə/).

After a vowel, as part of a complex syllabic, /y/ is a glide of the tongue towards (not necessarily to) high front position; /h/ is a glide towards mid or low central unrounded position; and /w/ is a glide towards high back position, accompanied by an increase of lip rounding. As parts of complex

syllabics, all three semiconsonants are voiced, as are all vowels. Before a vowel, the three semiconsonants are glides from rather than towards the positions just described: /y/ and /w/ are voiced in this environment too (except that speakers who say hue, white /hyúw hwáyit/ may have voiceless /y/ and /w/ after /h/), but /h/ is here voiceless (compare hot /hát/ with voiceless /h/ and spa /spáh/ with voiced /h/).

#### 2.2.1.3. Stress and Juncture.

In utterances longer than a single syllable, some syllables are louder than others. Thus in our transcription of the word singing, used above to illustrate certain segmental phonemes, we placed the mark ' over the vowel of the first syllable to indicate that the first syllable is pronounced more loudly than the second: /sínŋŋ/. Some pairs of words do not differ at all in segmental phonemes, and yet are different: in my speech, the noun permit /pérmit/ versus the verb /permít/ (in some parts of the country the noun is /péhrmít/, showing a difference in segmental structure also; and for some speakers the noun is pronounced /permít/, just like the verb). This pair shows that differences of stress or loudness are distinctive, but does not show how many distinctively different levels there are. There are in fact four: from loudest /'/, called loud or primary stress, through secondary /^/, tertiary /`/, to weak (usually no mark, but ~ when needed for clarity). The following examples show the contrasts and the unpredictability of distribution of the four levels:

(He lives in a) whíte hóuse. (any house of the specified color)

(He lives in a) whíte hóuse, (not a green one). (contrasting colors)

(He's living in the) Whíte hóuse. (the house that belongs or once belonged to a family named White)

(The President lives in the) Whíte Hóuse. (a specific structure in Washington, D. C.)

Lòng Ísländ (is a) lòng Ísländ.

(He's our) nêw ùndërtákër.

(It's a) nêw ùndërtákíng.

In general (though not without exception), the pronunciation of what are ordinarily thought of as "single words" does not contain instances of the juncture phoneme /+/, and consequently our transcriptions of single words for illustrative purposes have not yet illustrated the /+/. But compare:

(That's sodium) nitrate. /...náytrèyt/

(After six you get the) night rate. /...náyt+rèyt/

(Dr. Gerald Nye discovered that, and it is called the)

Nye trait. /...náy+trèyt/

Shé's (going.) /šíyz.../

Béa's (going.) /bíy+z.../

and for many speakers:

(I want a) tomato /...taméytow/

(The Republic of) Plato /...pléy+tòw/

where the single word Plato is pronounced exactly as though it were a compound of play and toe.

The juncture phoneme /+/ consists regularly of a slight lengthening of the segmental phoneme represented in transcription directly before it--that is, that segmental phoneme is held slightly longer than it would be under the same circumstances but without the juncture. Its presence is usually marked also by certain additional features, depending on just what segmental phonemes flank it: thus in night rate the /t/ (of night) is inaudibly released, as it is at the end of an utterance, and the /r/ is fully voiced, as it is initially in an utterance; whereas in nitrate the medial /t/ is released with a slight puff of breath which also partially devoices the /r/. Any segmental phoneme directly after /+/ is pronounced as it is at the beginning of an utterance, rather than as it is medially not flanked by /+/, or finally.

#### 2.2.1.4. Intonation.

The English intonation phonemes (PLs = pitch levels and TCs = terminal contours) occur only in certain combinations called intonations. An intonation involves one and only one TC, at least two PLs, and may (at least in theory) involve as many as five PLs. An intonation occurs most usually (certain "grunts" of acceptance, rejection, and the like, may be exceptions) simultaneously with a sequence of segmental, junctural, and stress phonemes; the whole combination we shall call a macrosegment. We shall label the five positions for PLs with the letters a through e. The following examples show the positions:



<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>e</u>	
					/ <sup>3</sup> h ó n l t <sup>1</sup> #/ ( <u>Halt!</u> )
					/ <sup>2</sup> g ô w + ə <sup>3</sup> w é y <sup>1</sup> #/ ( <u>Go away!</u> )
					/ <sup>2</sup> its <sup>3</sup> fâyv+ə <sup>3</sup> k l á k <sup>1</sup> #/ ( <u>It's five o'clock.</u> )
					/ <sup>2</sup> s ô w + <sup>2</sup> fyt <sup>3</sup> šer+lênc <sup>2</sup> #/ ( <u>So eat your lunch!</u> )

(The spacing out of the symbols is designed only to show the PL positions. Each of the four "sentences" could, of course, be intoned in various other ways.)

Position c is at the syllable which bears primary stress. There is always one and only one such syllable in a macrosegment, and there is always a PL accompanying it.

Position e is at the end, where the TC also occurs. There is always a PL at this position.

Position a exists only in those macrosegments in which there is material before the syllable with primary stress. Whenever position a exists, there is a PL at it.


Position b exists only when position a does and when, in addition, at least one syllable between positions a and c carries secondary stress. Under these circumstances, some intonations have a PL at this position and some do not.

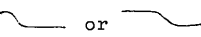

Position d falls between positions c and e. Relatively few intonations involve a PL at position d, and those that do are relatively rare.

No sure examples have yet been observed of intonations with PLs at all five positions. It is possible that the occurrence of a PL at position b precludes the occurrence of one at position d and conversely, but this is not certain.

Phonetically, the four PLs are four rather narrow ranges of pitch. They are relative, not absolute: a woman's lowest PL may be higher than a man's highest, and a single speaker may switch register from time to time. /<sup>1</sup>/ is the lowest and /<sup>4</sup>/ the highest. The actual pitch of the voice, in the intervals between the points at which the PLs occur distinctively and must be marked in transcription, varies somewhat, but is largely determined by the bounding PLs, by the distribution of stresses, and by the length of the interval. Thus, for example, if the last syllable of a macrosegment carries primary stress, so that position c is at its beginning and position e at its end, and if /<sup>3</sup>/ occurs at c and /<sup>1</sup>/ at e, the pitch will fall quite rapidly during the syllable. If, on the other hand, position c is several syllables from position e, with no PL at d and the same distribution of /<sup>3</sup>/ and /<sup>1</sup>/, then the pitch may fall fairly rapidly immediately after position c, and stay down, or it may slide down more slowly; the difference may be nonlinguistic or it may involve the presence of a PL at position d.

To illustrate:

/<sup>3</sup>hóhlt<sup>1</sup># (Halt!) contour 

/<sup>3</sup>péypər+həŋgər<sup>1</sup>#/ (Paperhanger.) contour  or 

Of the TCs, /#/ and /||/ are characterized by quite distinctive properties, while /|/ is marked largely (not exclusively) by the absence of any of the properties distinctive of the other two. /#/ involves a certain amount of stretching and fading. The segmental material directly before it is articulated somewhat more slowly than it would be with no TC, and the

force of articulation decreases. When the PL at position e is /<sup>1</sup>/, /#/ also involves a lowering of pitch below the ordinary level for /<sup>1</sup>/ in otherwise similar circumstances. It is mnemonically helpful to note that the combination /<sup>1</sup>#/ almost always carries the implication "I am done speaking, at least for the moment: your turn." This implication is strongest if the intonation ending with /<sup>1</sup>#/ has /<sup>3</sup>/ at position c. Thus:

/<sup>2</sup>hàw+<sup>3</sup>áhr+yùw<sup>1</sup>#/ (How are you?)

/<sup>2</sup>àym+<sup>3</sup>fáy<sup>1</sup>#/ (I'm fine.)

/||/ involves some stretching, as does /#/ , but without any fading; and the pitch of the voice rises to a point higher than that of the PL at position e though not so high as that of the next higher PL. The combination /<sup>2</sup>||/ or /<sup>3</sup>||/ occurs very often in questions which call for yes-or-no answers, though by no means exclusively here:

/<sup>2</sup>àhrye+<sup>3</sup>gówij<sup>3</sup>||/ (Are you going?)

/|/ involves slight stretching, but neither fading nor lowering nor raising of pitch. When /|/ occurs followed by silence, the impression is usually that of hesitation or incompleteness. Medially:

/<sup>2</sup>in+<sup>3</sup>òát+kèys<sup>2</sup>|<sup>2</sup>àhl+<sup>3</sup>gów<sup>1</sup>#/ (In that case I'll go.)

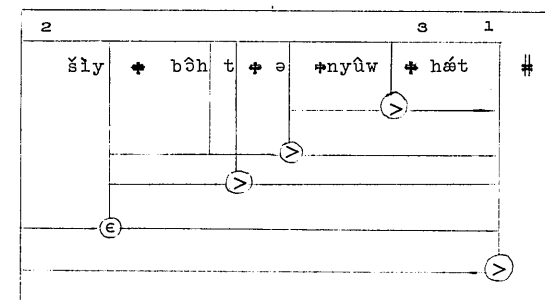
#### 2.2.2. English grammar.

In Figure 4 is displayed the sentence She bought a new hat, in a diagram that shows much of its grammatical structure. Like any other utterance, this one is built wholly of morphemes, put together wholly by recurrent patterns of combination called constructions. The ultimate constituent morphemes of the

utterance do not participate directly in the whole. Rather, they are put together two-by-two, or few-by-few, into somewhat larger forms, which in turn are put together a few at a time into still larger forms, until we reach the whole. Thus the immediate constituents of the whole sentence, as the diagram shows, are (1) its intonation and (2) the non-intonational remainder. The non-intonational portion in turn has, as its ICs (=immediate constituents), the smaller forms she and bought a new hat. She is a morpheme; but bought a new hat is composite, and has the finer-grained structure shown in the diagram.

No English morpheme is ever represented phonemically by a combination of intonational and non-intonational phonemes. Intonational phonemes occur only in recurrent combinations which represent intonational morphemes; all other morphemes are segmental.

Figure 4.



The marks "e", ">", and so on, represent types of structural relationships; see 2.2.2.4. below.

Despite their regular simultaneous manifestation in speech and despite the extent to which intonation marks the structure of the segmental material it accompanies, the intonational and the segmental parts of English constitute, in a sense, two separate communicative systems, each characterized by duality. What little we know of the intonational system so far, beyond its phonology, concerns mainly how it impinges on and marks segmental structure. Beyond this we are not ever certain of the identity of intonational morphemes; it is only a matter of convenience to assume that each intonation, as defined in 2.2.1.4., is a morpheme. So the discussion which follows will deal mainly with the segmental system. In the first two sub-sections we describe the kinds of elements (morphemes and small idiomatic combinations of morphemes) used in English; in the third sub-section we describe the constructions by which they are put together.

#### 2.2.2.1. Functors.

The segmental system of English involves morphemes, and small idioms, of two basic kinds: functors and contentives. Examples of functors are and, if, he, in, on, the -es (/z/) of goes, the -s (/z/) of umbrellas, the -ly of quickly the /'+'/ stress-and-juncture pattern of blackbird. Examples of contentives are black, bird, umbrella, quick, go, boy, boyish, boyhood.

Functors are of two kinds: markers and substitutes. Markers are elements which signal the grammatical relationships of surrounding forms. The ICs (=immediate constituents) of men and

women are the forms men and women; and is a marker. The ICs of John's hat are John and hat; 's (/z/) is a marker. These two markers (and and 's) are pure: they signal the relationships of surrounding forms, without themselves having any position in the hierarchical structure of the larger forms in which they occur. Other pure markers are the or of men or women, the either and or of either John or Bill, the neither and nor of neither John nor Bill, the both and and of both John and Bill, and the /' + `/' of blackbird.

Impure markers participate in the hierarchical structure of the larger forms in which they occur, but still mark structural relationships. The in and on of the pencil in the desk and the pencil on the desk are examples. The ICs of the former are the pencil and in the desk; the ICs of in the desk are in and the desk. In addition to standing in construction with the desk, the impure marker in also signals the fact that the whole form in the desk stands in a certain relationship (here attributive: see 2.2.2.3.) to something else--in the example, to the pencil. Other impure markers are the if of if John comes, the when of when you get there, the is of John is tall.

Cutting across their classification as pure or impure is another classification of markers: as free or bound. Free markers are separate words, called particles: and, or, either, neither, nor, both, in, on, if, when. Bound markers are less than whole words, and are called inflectional affixes: the /z/ of (he) goes, the /z/ (a different morpheme) of umbrellas, the -er of poorer, the -ed (/d/) of begged, the -ly of quickly.

The 's of John's hat is on the boundary between bondage and freedom, and can be classed either way. Stress-and-juncture patterns are neither words nor parts of words, and thus fall outside the free-bound classification and are separately designated as superfixes.

Substitutes are special forms the meanings of which are redefined on each occurrence by (1) the selection and structuring of nearby linguistic material, or (2) certain elementary features of the communications situation, or both. Thus the denotation of I and me depends entirely on who (or what) is speaking; the denotation of you depends entirely on who (or what) is being addressed. If one says John took his car, his means John's (unless special features in a larger context force his to refer to someone else); if we say Bill took his car, then the same form his refers to Bill's. In John doesn't hate her, but I do, do means "hate"; but in John doesn't like her, but I do, the same form means "like."

All English pronouns (personal, demonstrative, interrogative, relative, negative nobody etc., indefinite anybody, somebody etc., and inclusive all, everyone, etc.) are substitutes. So, probably, are the numerals one, two, etc., and the articles a/an, the; pro-adverbs here, now, there, then; the pro-verb do as illustrated above; so (a pro-adverb in He did it so, a clause substitute in if so); then also as a clause substitute, as in if he comes, then we must leave.

None of the classifications mentioned and illustrated above are airtight. Some forms appear now in one function, now in

another; or, if one prefers, some forms come in pairs identical in shape but different in function. Is is a functor in John is poor or God is good, but a contentive in God is or That which is, is. The terminal /z/ of goes is an inflectional affix, and hence a marker; but it is also a substitute, since it indicates that the subject of the verb is singular and other than speaker or addressee. Any form is used on occasion as a designation for itself; "And is a functor", as we have said earlier. When so used, the form is of course a contentive.

Allowing for such uncertainties and boundary-line cases, we define a contentive as a form which is not a functor.

#### 2.2.2.2. Stems and their Classification.

A stem is what is left of a word when all inflectional affixes, if any, are stripped from it. Thus the stem of boys is boy- (where we write a hyphen to show that we are not designating the word boy), and the stem of the word boy is also boy-, since there are no inflectional affixes to be deleted. Some stems are functors: the i- which is left when the inflectional affix /z/ is stripped from is. But most stems are contentives (e.g., boy-), and it is with the latter that we are concerned here.

English (contentive) stems fall into seven basic classes depending on their privileges of occurrence in larger forms: that is, on the inflectional affixes that are used with them, and on the constructions in which they participate.

One set of privileges of occurrence is the noun pattern: inflection for plural (boy : boys), though not invariably; use

after a or an, the, this or these, that or those, unstressed some (/sɪm/) (a boy, an elephant, the boy, this boy, these boys, some boys, some milk). English stems which follow this pattern, or much of it, but do not also follow either of the patterns of usage yet to be described, belong to class N. Examples are strength, food, action, day, friend, art, danger, music, boy, elephant.

A second pattern is the adjective pattern: inflection for degree (pretty : prettier : prettiest) or use after more and most to yield phrases (more beautiful : most beautiful) which are then used in much the same way in still larger combinations; inflection with -ly for adverbial use (prettily, beautifully). Stems which follow this pattern, but not also the noun pattern nor the third pattern to be described, belong to class A: long, false, likely, certain, icy, sleepy, short, soft, civil, beautiful.

Some stems follow both the noun pattern and the adjective pattern, but not the third pattern to be described below. These stems belong to class NA: American, sweet, savage, private, male, white, red, innocent.

The third pattern is the verb pattern: many stems have five inflected forms, as sing, sang, sung, sings, singing: most have only four, as describe, describes, described, describing; a few have even fewer, as can, could, or must and ought with only one form. Syntactically, the typical verb uses are as verb in objective constructions (saw John; the syntactical terms used here are all expounded in 2.2.2.3. below); as verb in intransitive

predicates (I see, John was singing loudly); and as connector in a connective construction (seem tired). Stems which show this pattern, but not the noun or adjective pattern, belong to class V: describe, admit, punish, bury, strengthen, falsify, penetrate, collaborate, denazify.

Stems which show both the noun and verb pattern, but not the adjective pattern, belong to class NV: walk, love, cure, change, air, eye, nose, beard, elbow, finger, cut, build. Stems showing both adjective and verb patterns belong to the sixth major class, class AV: clean, dry, thin, slow, clear, busy, idle, true. The seventh class, class NAV, includes stems which show all three patterns: fancy, faint, black, yellow, blue, brown, gray, damp.

Any stem of any of these classes, in a specific occurrence in which it is functioning by one of the three basic patterns, is, in that context, a noun, a verb, or an adjective, as the case may be. Its momentary usage does not alter its basic affiliation: fancy is of class NAV whether appearing as a noun (Her head was full of strange fancies), as a verb (She fancies herself a dancer), or as an adjective (a fancy dress, a fancier dress). Some positions of occurrence do not indicate whether a stem is functioning by one pattern or by another: Milk! is an NV stem combined with an intonation, and is ambiguously a command (= "Milk the cow now!") and thus a verb, or an exclamation (= "Why, it's milk!") and thus a noun, but nothing in the utterance tells us which.

The seven major stem classes include many smaller subclasses of stems marked by one or another peculiarity or limitation of usage. We give only a few examples. Scissors, shears, trousers,

pants, clothes are always plural. Some nouns, such as music, are almost invariably singular. Some nouns are used syntactically as either singular or plural without any overt inflection: sheep, deer, trout, bass, carp, fish, people. Such features of behavior place these stems in special subclasses of class N, NA, NV, or NAV.

A great many English stems are not single morphemes, but are built from smaller stems by the addition of certain elements called derivational affixes. Thus girlish and boyish, stems of class A, are built from the class N stems girl and boy with the derivational affix -ish; the class N stem actress is formed from the class N stem actor with the affix -ess; actor is in its turn formed from the class NA stem act with the affix -or. Some stems larger than a single morpheme do not contain a smaller stem, though derivational affixes are present: the in- and re- of infer and refer are derivational affixes, but -fer is not a stem. Derivation--the patterns of formation of stems larger than a single morpheme--constitutes a lengthy chapter of English grammar, about which, fortunately, we need here say no more.

#### 2.2.2.3. Constructions and Construction-Types.

The number of constructions in English is very large, but most of them fall into certain fundamental construction-types, and it is these that we shall describe.

A construction is either centered or uncentered. A construction is centered if the larger forms built by it have much the same privileges of occurrence, in still larger combinations,

as at least one of the ICs. Otherwise the construction is uncentered. The constructions involved in black cat, blackbird, ran quickly, quite clear, men and women, either John or I, Lake Michigan, Queen Elizabeth are centered; those involved in saw John, seems a shame, with Bill, if so, I ran are uncentered.

To show that the construction by which black cat is built is centered, we note that the form black cat can in general be substituted for the smaller form cat: I saw a cat or I saw a black cat; a big cat or a big black cat; and so on. The constituent cat is called the center (or head) of the larger form; similarly bird is the center of blackbird (and of black bird, which involves the same constituents in a different but also centered construction), clear that of quite clear, and ran that of ran quickly. If a centered construction has only one center, its other constituent is called an attribute and the construction is of the attributive type; the examples given in this paragraph are all attributive.

Some centered constructions have two (or more) centers or heads, and no attributes. Such double-headed (or multiple-headed) constructions are of two subtypes: coordinate and appositive. Examples of the former are men and women (with heads men and women; and is a marker, not a constituent), men or women, two plus five, either John or I. Examples of the latter are Lake Michigan and Queen Elizabeth, in which, though both constituents are centers, also each constituent seems to modify the other.

Centered constructions appear at all "size-levels," as it were. In the sentence If he gets here soon enough we'll all go

together, the portion If ... enough is attributive to the remainder. In any form which consists of an intonation morpheme and a string of segmental material--say our diagrammed example, She bought a new hat (Figure 4)--the construction of intonation and remainder is centered, probably with the intonation as head and the segmental part as attribute.

The major types of uncentered constructions in English are directive, connective, and predicative. Directive are the constructions of in | the box, on | the table, if | he is going, while | we were there, saw | John, asked me | a question, asked | me. Connective are the constructions of is | a big man, is | tired, became | excited, lay in the corner | motionless. Predicative are the constructions of He | is a big man, She | sings beautifully, I | saw him, That man | I just don't like, (She watched) us | cross the street. A vertical line separates the ICs in the above examples; parenthesized material is to indicate context.

Directive constructions are prepositional, conjunctive, or objective. In prepositional constructions the first IC is a preposition: in | the box, on | the table. The form built by a prepositional construction occurs most often as an attribute in a larger centered construction: e.g., the pencil | in the box; but also as predicate attributes (see below) in connective constructions: (The pencil) is | in the box. Conjunctive constructions have a subordinating conjunction as first constituent: if | he comes, when | you see him, if | so. Forms built by conjunctive constructions are also used mainly as attributes in

larger centered constructions: if he comes | we can go. Objective constructions have a verb (single word or phrase) as first constituent: saw | John, asked me | a question, asked | me. The forms built by objective constructions occur typically as predicates (second constituents) in predicative constructions: I | saw John. But if the verb is in certain forms, there are other uses. Thus to see John can be used as subject of a predicative construction (To see John | is my only desire), or as object in an objective construction (want | to see John).

Connective constructions have a verb of a special kind, or one particle (as), or, in certain contexts, zero, as first constituent. There are three subtypes based on the nature of the second constituent: a noun predicate attribute, as in (That) is | John; an adjectival predicate attribute, as in (He) grew | large; or an adverbial predicate attribute, as in (He) is | here. The forms built by connective constructions are used like those built by objective constructions: mainly as predicates (That | is John), but sometimes otherwise (wants | to be good). In (select him) as | our president, the connector (first constituent) is the particle as. In (elect him) president the first constituent is zero, but is allied to the verb be, as shown by comparison with (choose him) to be | president.

Predicative constructions have as their ICs a topic and a comment: John ran away; That new book I haven't read yet; (the man) whom you visited here yesterday. In the first of these three examples, the topic is what is traditionally called the subject, and the comment is the predicate: this is true

perhaps in the majority of cases, but not in all, as shown by the second and third examples.

Forms built by predicative constructions are clauses. Independent clauses are those to which one need only add an intonation to produce a complete simple sentence: John ran. Independent clauses can be rendered dependent with a subordinating conjunction: if John ran. Otherwise, dependent clauses are shown to be so by the verb: John is there is independent, but John be there, John were there, John being there, John to be there, John there are dependent. That is, clauses of these latter varieties occur as constituents in larger segmental forms, not, as a rule, alone. Thus a dependent clause with -ing on the verb occurs: as a subject or topic (John singing that song annoys me); as object of certain prepositions (Don't count on John singing that song); and as object of certain verbs (I heard John singing that song).

Independent clauses are also classed by order. In direct order, the subject precedes all of the verb: John is going. In inverted order, part of the verb, at least, precedes the subject: Is John going? and Has John been going?, where the verb is the interrupted phrase has been going. There are also some special orders (e.g., that of Away ran John) found under limited circumstances.

A sentence is a form which is not in construction with any other form. This requirement does not preclude certain kinds of ties between parts of successive sentences, which are not technically constructions but cross-cutting connectivities. Thus one

sentence may be a question and the next an answer. Or a word like he in one sentence may refer back to a name, say John, in the preceding sentence. Or a second sentence may begin with some loose connective, such as but or and. Thus: John did arrive, didn't he? -- Yes, but he forgot his books. Mainly, the selection and distribution of intonations marks the limits of sentences: thus any intonation ending with /<sup>s</sup> 1 # / marks the end of a sentence, and no intonation ending with / // marks the end of one.

Many sentences, at least of formal discourse, are built around an independent clause consisting of subject and predicate (or other type of topic and comment). Such sentences are called favorite. Sentences which are not of this type are of various minor types: Oh? Hey! Bring that here! Well for goodness's sake! There is a sharp difference between any sentence, favorite or minor, and a fragment which the speaker breaks off without completing. If one begins with I was going to-- and then simply stops, one has a fragment, not any kind of sentence; this is usually obvious because only part of an intonation has been produced.

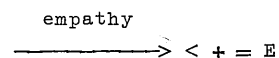
#### 2.2.2.4. Grammatical Patterning in Action.

In terms of my earlier definitions, the grammatical structure of an utterance emerges, for a hearer, as the successive elements of the utterance are received. The first element sets up certain limitations on what may come next, and certain expectations as to what is more likely or less likely to come next. The same is true of a speaker, except for those occasions on which he has



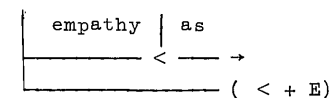
planned an utterance in detail in advance of starting to speak it aloud.

To illustrate this, let us assume that we are listening to a formal lecture, and that the first word we hear is empathy. We shall assume also that the production of the word is not accompanied by an utterance-closing intonation. Our expectation-pattern after reception of the first word, before we hear the next one, can be graphed as follows:



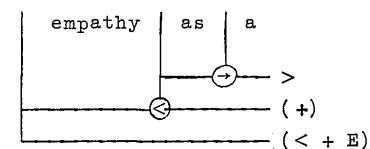
The marks have the following meaning. ">" means that the next thing spoken might be something to which the part already spoken is an attribute; empathy methods ... or the like. "<" means that the next to come might be a postposed attribute and modifying empathy: empathy of some sorts or empathy in psychotherapy. "+" means that the next to come might stand in a coordinate construction with what has already been said: empathy and intuition. "≡" means (though this is less likely) that the next to come might stand in apposition with empathy: empathy, the method used in ... Finally, "E" means that empathy may be all of a topic, the next element in the sentence beginning the comment: empathy is often used in ineffective ways. Now this is a wide range of grammatical possibilities, but it does not include everything--for example, it is entirely precluded that the next element should be the object of empathy (in an objective, prepositional, or conjunctive construction), since empathy is not a word which can occupy the first position in any such construction.

The next word we hear is as; and the expectation-pattern is restructured:

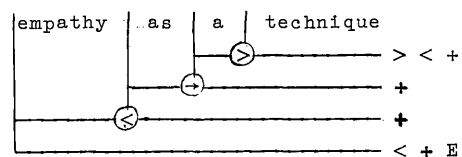


Most of the possibilities left open by the uttering of empathy have now been eliminated. As begins a grammatical form which will stand as a postposed attribute to empathy; this is indicated by the mark "<" at the junction of the boxes in the diagram for empathy and for as ... As must itself be followed by something standing to it in the relation of object; this is the meaning of the mark "→". The other three marks have the same meanings as before; they are put in parentheses because they do not indicate what can happen next--the possibility of a further postposed attribute, or of something in apposition, or of a comment, is temporarily in abeyance, until the materials demanded by as have been spoken.

The next word we hear is a:

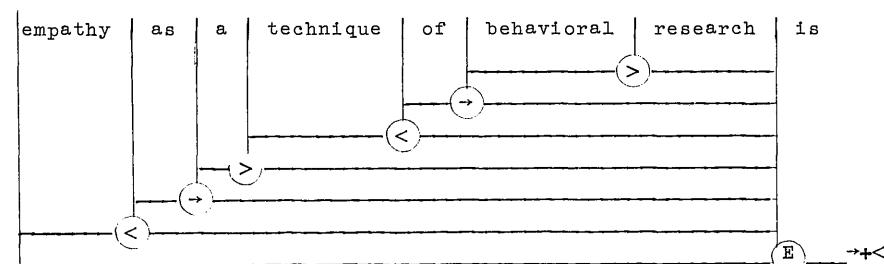
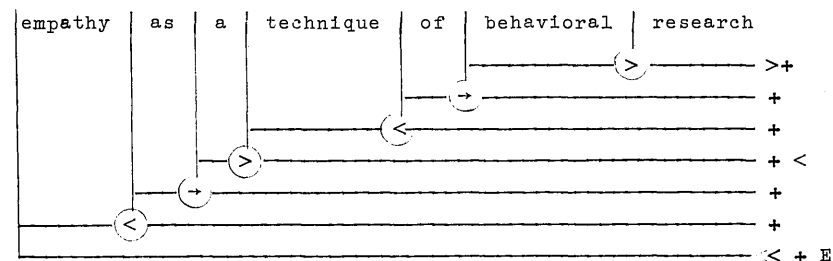
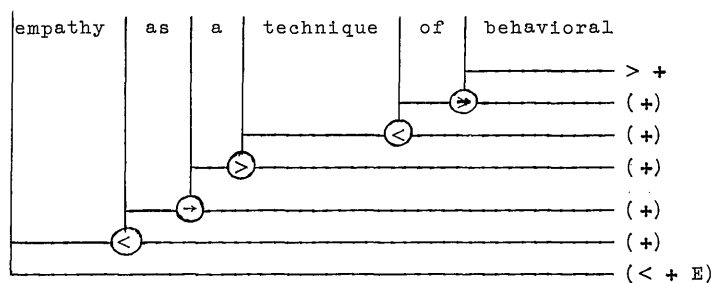
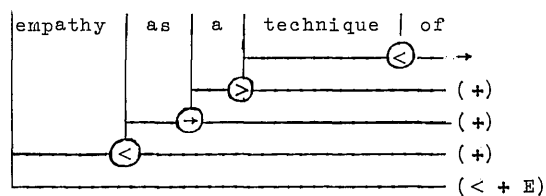


What comes next must be something modified by a. Next we hear technique:



--and this opens up the possibilities again. What comes next may stand in construction only with technique, or with a technique, or with as a technique. The first would be the case in Empathy as a technique in psychotherapy; the second in Empathy as a technique or a method; the third in Empathy as a technique has challenged many of us.

The next four forms are of, behavioral, research, and designates, and we present the expectation-diagrams for the four resulting situations without comment save about the last:



Note, particularly, the difference between the openness of the possibilities after research and the decisiveness of structure which results when is is added: the adding of is (or, of course, of seems, designates, implies, appeals, strikes or the like) finally establishes that all the forms so far constitute a unit, and that anything further in the utterance will relate to that whole composite unit, not to any of its individual components.

Such "growth diagrams" thus show some dynamics which the mere diagramming of a whole sentence, after the fact, cannot reveal.

Two related points must be added. Not infrequently, a

speaker will simply abandon an emerging structure, and start over again. This is always a possibility, though it is not represented, in the foregoing diagrams, by any positive mark. Also, the emerging structure may become so complex that the speaker forgets some of the "hang-over" irresolutions implied by his earlier words, so that the resulting whole simply does not structure. Such a "hang-over" irresolution is the need for a predicate, which is established in our example by the very first word. By the end of research, this could have been lost, the speaker going on to produce an imparsable sentence like Empathy as a technique of behavioral research developed particularly by psychiatrists, but also taken over by anthropologists, who adapt it for use with people with vastly different cultural backgrounds. Our strong habit of editing out such structural irresolutions in written material should not close our eyes to their very high frequency in oral communication, even of a technical and formal sort.

### 2.3. Nonlinguistic<sup>1</sup> Vocal Communicative Systems.

The nonlinguistic features covered in the following discussion were observed from speakers of English. Some of them are doubtless much more widespread, in connection with many languages; but insofar as we have been able to see the systematic patterning

<sup>1</sup> See Appendix 1 for a somewhat more recent (Trager, 1958) and somewhat fuller treatment, and Pittenger, 1960, 185-206, for further modifications of the system presented here.

of any of these features, our description must be interpreted as applying only to speakers of English.

Some vocal activities take place not accompanied by any linguistically organized behavior: coughs, belches, sneezes, laughing, crying, humming, and the like. Some of these, of course, occur also during linguistically organized speech: as when a cough interrupts speech, or a laugh either interrupts or is simultaneous with it. In transcribing recorded vocal behavior, it is convenient to use a symbol for laughing:

L

repeated as necessary over the transcription of the relevant sequence of words or syllables. Other interrupting effects of this kind are perhaps best taken care of via added notes. There are, of course, various kinds of laughing, and no doubt the variety is culturally patterned; but we have not yet subjected this range of phenomena to closer analysis.

Certain types of activity incorporate linguistically organized speech with certain limitations. In singing with words, the segmental portion of language is fully represented; the stresses are sometimes undistorted but sometimes modified (depending on the style of song); the intonational system of the language does not clearly manifest itself, since the raw-material in which it normally is made evident--pitch--is preempted by the conventions of singing. In whispering, of various different kinds, the intonational system is also distorted. The segmental phonemics is maintained, though the actual sound of some of the phonemes is quite different from

normal unwhispered speech. In transcription, whispered passages are marked

X- -X

For most of the above phenomena, and for those to be discussed below, there are two sets of features which can be classed under the terms vocal quality and vocal set. Without attempting to define either of these with absolute precision, we can say that both have to do with longer-termed features of the quality of the voice. Thus women's voices are different from men's (yet some men can mimic women's voices, and vice versa); the voice of a child sounds different from that of a young girl, or a woman in her twenties, or a matron of thirty-odd, or an old woman (yet young actresses can play old-woman roles). A cold in the nose affects voice quality. These "biological" or quasi-biological manifestations are matters of voice quality. Most speakers adapt to the size of the room in which they are speaking, and the size of the audience, and the amount and variety of extraneous sound, in such a way as to establish a personal norm for the specific circumstances: these adaptations are matters of vocal set.

It is within the range established for a relatively larger period of seconds, minutes, or hours, that the variations take place which we analyze in terms of vocal qualifiers. The system of vocal qualifiers consists of a number of dimensions of variability of voice quality (acoustically speaking) or style of articulation, each of which is more or less independent of all the others. On each scale, the "norm"--the most colorless placement--is roughly in the middle, while deviations from the

norm in either direction constitute significant signals. We are not certain of the exact number of dimensions; the writer has found it feasible to work with the following nine:

(1) Tempo: overslow, normal, and overfast. Symbols, the following, used like brackets through the passage effected:

overslow << - - <<  
overfast >> - - >>

(2) Syllable control: clipped, normal, and drawled:

clipped ∩ - - ∩  
drawled † - - †

(3) Volume: oversoft, normal, and overloud:

oversoft ∨ - - ∨  
overloud ^ - - ^

(4) Height of register (the bandwidth of frequencies within which the pitch levels are pronounced): overflow, normal, overhigh:

overflow † - - †  
overhigh † - - †

(5) Width of register: from overnarrow or "muted" through normal to overwide (also called "singing," but this is a questionable application of the latter term):

overnarrow † - - †  
overwide † - - †

(6) Pitch control: glissando (slurring from one pitch level of the intonational system to the next), normal, and portamento (sharp stepwise motion from one pitch level to the next):

glissando G1 - - G1

portamento Po - - Po

(7) Glottal tension: rasping, normal, and open:

rasping q - - q

open o - - o

(8) Air flow: breathy, normal, and squeezed (or "overvoiced"):

breathy B - - B

squeezed V - - V

(9) Precision of supraglottal articulation: overloose ("sloppy"), normal, and overprecise.

overloose S1 - - S1

overprecise Pr - - Pr

Each of these is, as indicated, a scale along which variation is possible. It is not yet known just how many points along each scale are communicatively distinctive. In the earliest work done on vocal qualifiers, it was believed that for each scale there were three and only three contrasting points: those assigned labels in the above listing. That is, as to tempo for example: speech was either normal, or enough faster than normal to count as distinctively overfast, or enough slower than normal to count as distinctively overslow. More recent work has convinced several investigators independently that, at least on some of the scales, more degrees are distinctive. It may, indeed, be the case that there is no exact small finite number of contrasting points along each scale, but rather a genuinely continuous range of variation. For transcription purposes, at present, however, our only practical course has been to use only the symbols

listed above.

Akin to the vocal qualifier system, yet distinct from it in one way, is a set of dimensions in which the norm is at one end of the scale instead of in the middle. Whispering can be classed as one of these dimensions: normal speech is not whispered, and whispering has no "opposite" which also deviates from the norm. A second is breaking: normal speech has no breaking, which is added as a rapid turning on and off of the vibration of the vocal cords (a sort of "chatter" or nervous "giggle"). Laughing can be interpreted as a third. Symbols:

whispering X - - X

laughing L - - L

breaking ó - - ó

Quite different from either of these systems is the system of boundary phenomena (also called vocal identifiers--a term with no mnemonic value). The items of this system appear typically at the boundaries of actual linguistically organized utterances--as one starts an utterance, as one finishes one--and during hesitations either between grammatically organized utterances or in the middle of utterances as the speaker is (as it were) deciding how to go on. The analysis of boundary phenomena is far from complete; all we can do here is to list the symbols which, it seems, are minimally necessary for transcribing them:

voiceless aspiration (sometimes a "sigh") h

voiceless intake of air (a "gasp") d

voiced aspiration through mouth e

(without the articulation of  
vowels and consonants)  
Glottal catch release into speech,  
or glottal closure breaking  
off speech

Thus a very common way of saying the assenting "grunt" (conventionally spelled M-hm) begins with the last of these and has the first of them in the middle.

Tentatively I find it convenient to distinguish between "quarter-speech," "half-speech," and full (completely linguistically organized) speech. In quarter speech, neither the segmental nor the intonational portion of language is represented: "utterances" in quarter-speech involve the system of boundary-phenomena described above and some of the system of vocal qualifiers. In half-speech, intonations are present in addition to boundary phenomena and vocal qualifiers, but not linguistic segmental material. In full speech, of course, intonational and segmental linguistic material are both present, in addition to the non-linguistic features.

### CHAPTER 3

#### Body Motion

Ray L. Birdwhistell

"Every little motion has a meaning all its own."

This line, from a song popular in the twenties, has been suggested as expressing the theme of the work of the student of body motion. Reflecting the growing self-consciousness of our culture, this line contains within it the challenge and the promise which motivates such a student. Like Bateson (p. ), we are convinced that 'everything which occurs in a social interaction is meaningful in the sense of being part of the interchange as well as non-accidental.' At the same time, however, "Every little motion has a meaning all its own," carries with it an implication which is probably the major impediment to the analysis of communicative systems: the assumption that not only does each motion have a meaning but that the relationship between that movement and that meaning is precise, integral and universal.

It is the task of this chapter to make clear that while body motion behavior is based in the physiological structure, the communicative aspects of this behavior are patterned by social and cultural experience. The meaning of such behavior is not so simple that it can be itemized in a glossary of gestures. Nor is meaning encapsulated atomistically in particular motions. It can be derived only from the examination of the patterned structure of the system of body motion as a whole as this manifests itself in the particular social situation. It should be evident to the reader that this is precisely the same point which has been made by Hockett, as a linguist, in

Figure 1

HIGHWAY SCENE

Time: Estimated 5 seconds

Characters: Soldier No. 1, Soldier No. 2 and Driver

Soldier No. 1:

		Car Passes																
Head	H> 1°	.	.	.	.	.	.	.	.	.	.	Hffbzbz	.	.	.	.	.	.
Forehead-Brows	Hfb-b	.	.	.	.	.	.	.	.	.	.	Hffbzbz	.	.	.	.	.	.
Eyes	oo driver	.	.	.	.	.	.	.	.	.	.	oo= car	.	.	.	.	.	.
Nose	Mz	.	.	.	.	.	.	.	.	.	.	oMo	.	.	.	.	.	.
Cheeks		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Mouth	L-L	.	.	.	.	.	.	.	.	.	.	t1-1	.	.	.	.	.	.
Chin		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Neck		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Shoulders	"> 1°	.	.	.	.	.	.	.	.	.	.	"	.	.	.	.	.	.
Trunk	TpTp	.	.	.	.	.	.	.	.	.	.	Txpivot	.	.	.	.	.	.
Hips		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Right Arm	RAN[RA2:45< ≥ 3:45n	N	N	.	✓	✓	✓	✓				RAN(4<5)n ^ ^ ^	RA25n	.				
Hand and fingers	R/1?4P	.	.	.	.	.	.	.	.	.	.	R/14-p	.	.	.	.	.	.
Left Arm	LAn-15'3ul[A;TA]	.	.	.	.	.	.	.	.	.	.	LAn<3n	.	.	.	.	.	.
Hand and fingers	L/1c2C3C4C5C belt	.	.	.	.	.	.	.	.	.	.	L/14c(on R. biceps)	.	.	.	.	.	.
Right Leg	Y45Y	.	.	.	.	.	.	.	.	.	.	RY<3°(Y45+30nY)	.	.	.	.	.	.
Foot		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Left Leg	Y45Y	.	.	.	.	.	.	.	.	.	.	(Y45+30'Y)	.	.	.	.	.	.
Foot		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.

Soldier No. 2:

Head	H>1°	.	.	.	.	.	.	.	.	.	.	H>1	.	.	.	.	.	.
Forehead-Brows	Hfb-b	.	.	.	.	.	.	.	.	.	.	Hffbzbz	.	.	.	.	.	.
Eyes	oo driver	.	.	.	.	.	.	.	.	.	.	oooo	.	.	.	.	.	.
Nose	Mz	.	.	.	.	.	.	.	.	.	.	oooo	.	.	.	.	.	.
Cheeks		.	.	.	.	.	.	.	.	.	.	oooo	.	.	.	.	.	.
Mouth	L-L	.	.	.	.	.	.	.	.	.	.	oooo	.	.	.	.	.	.
Chin		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Neck		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Shoulders	">1°	.	.	.	.	.	.	.	.	.	.	"	.	.	.	.	.	.
Trunk	T pTp	.	.	.	.	.	.	.	.	.	.	Txpivot	.	.	.	.	.	.
Hips		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Right Arm	RAN[RA2:45< ≥3:45n	N	N	.	<	<	.	<	.			✓ RAN(5<vn)n . .	RAnnn	.				
Hand and fingers	R/1?4p	.	.	.	.	.	.	.	.	.	.	R/?1 1-5	R/14c	P> o				
Left Arm	LAn-15'3ul [A;TA]	.	.	.	.	.	.	.	.	.	.	LAnnn	.	.	.	.	.	.
Hand and fingers	L/1c2C3C4C5C belt	.	.	.	.	.	.	.	.	.	.	L/14cP≥	.	.	.	.	.	.
Right Leg	Y45Y	.	.	.	.	.	.	.	.	.	.	RY<3°(Y45+30'Y)	Ly5≥	Ly3°				
Foot		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
Left Leg	Y45Y	.	.	.	.	.	.	.	.	.	.	(Y45+30'Y)	Ly5≥	Ly3°				
Foot		.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.

Driver\*:

Head	H>	.	.	.	.	.	.	.	.	.	.	H>	.	.	.	.	.	.
Forehead-brows	Z	.	.	.	.	.	.	.	.	.	.	Z	.	.	.	.	.	.
Eyes	oo road	.	.	.	.	.	.	.	.	.	.	oo road	.	.	.	.	.	.
Nose	Z	.	.	.	.	.	.	.	.	.	.	Z	.	.	.	.	.	.
Mouth	L/L	.	.	.	.	.	.	.	.	.	.	tL-L	.	.	.	.	.	.
Cheeks		.	.	.	.	.	.	.	.	.	.	Z	.	.	.	.	.	.



Cheeks  
Chin  
Neck  
Shoulders

---

\*Driver maintains upright, bi-manual driving position throughout scene.

(Sample)

Shoulders  $\|\geq 1^\circ$  . . . . . ||

Shoulders  $\|\geq 1^\circ$  . . . . . ||

the preceding chapter (p. ). It is the task of this chapter to review the present status of kinesics, the systematic study of body motion, in such a manner that the nature of the relationship between the two communicative systems will be adumbrated--if not entirely revealed. Based on the same assumptions, the two modes, kinesic and linguistic, have parallel, even at time analogous, structures. They are, however, infra<sup>1</sup>-communicational systems, not directly meaningful in themselves, and the reader should not be surprised to discover that their correlation brings difficulties into the analysis of the communicative process.

In the pages to follow an example of a non-verbal communicative system will be presented. This is designed to sensitize the reader to the kinesic scene. The second sub-section will contain a general review of the present status of kinesic theory and research. Finally, a discussion of the problem of systematic interpretation should prepare the reader for the experiment with which this present work is concerned.

#### Example

Just west of Albuquerque on Highway 66 two soldiers stood astride their duffle bags thumbing a ride. As a large car sped by them, the driver jerked his head back signifying refusal. The two soldiers wheeled and one Italian-saluted him while the other thumbed his nose after the retreating car.

---

<sup>1</sup> See p.

#### Macro-kinesic Translation

The two soldiers stood in parallel, legs akimbo with an intrafemoral index of 45 degrees. In unison, they raised their right upper arms to about an 80-degree angle with their bodies and, with the lower arm at approximately a 100-degree angle, moved the arm in an anterior-posterior sweep with a double pivot at shoulder and elbow; the four fingers of the right hand were curled and the thumb was posteriorly hooked; the right palm faced the body. Their left arms were held closer to the body with an elbow bend of about 90 degrees. The left four fingers were curled and the thumb was partially hidden as it crooked into their respective belts.

The driver of the car focused momentarily on the boys, raised both brows, flared his nostrils, lifted his upper lip, revealed his upper teeth, and with his head cocked, moved it in a posterior-anterior inverted nod which in its backward aspect had about twice the velocity of the movement which returned the head and face to the midline and, thus, to driving focus.

Without apparent hesitation the boys rightstepped posteriorly, one of the boys moving in echo following the movement of the other. Facing the retreating car, one of the boys raised his upper lip to expose his teeth, furrowed his forehead, lowered his brows, contracted the lateral aspects of his orbits, and flared his nostrils. His right arm swept from its posteriorly thrust position, on a shoulder

pivot, to rest, fist clenched, upper arm across the right half of the body and the lower right arm thrust up and slightly anterior to the body line. The left hand left the belt and the lower arm swept right and upward to meet the descending upper (right) arm. The left hand grasped the right biceps as, fist still clenched, the right arm moved quickly in an anterior-superior thrust in line with his shoulder and the retreating automobile.

The other boy dropped his face into "dead pan," pivoted his right arm at the elbow, flared and straightened his fingers into crooks, and, as the already hooked thumb crossed the midline of the body in the lower arm's downward sweep, the apex of the thumb made contact with the apex of the nose. Without hesitation the arm completed its sweep across the body and came to rest hanging, palms slightly forward, at his side. The left arm, on an elbow pivot, swept downward and came to rest mirroring the right.

#### Discussion

These three portrayals, the brief statement, the macrokinesic transcription Figure 1, and the kinesic description derived from the macrokinesic recording, all tell the same story with varying degrees of fullness. Some readers may feel like the little boy who received a birthday book about penguins from his aunt and felt it contained more about penguins than he ever wanted to know. However, such a record as is provided by these latter descriptions makes it possible for us to do extended analysis of the transaction.

The initial descriptive statement is totally inadequate for such purposes.

This scene contains much more than three men gesticulating at each other. In the time it takes an auto to pass a fixed point at seventy miles an hour, a communicational transaction has taken place. In five seconds a social group is established, a social ritual is performed, and, presumably, the lives of three human beings are somehow affected. This is patterned activity; its components were learned in a multiple of comparable but differing situations by the participants. Yet this is no mere mechanical performance. We cannot, for the moment, "explain" it; nevertheless, it is a piece of microculture whose natural history we may attempt to relate.

We have no way of telling how the driver felt or what he thought about as he approached the soldiers. Our only evidence comes from the driver's compressed mouth (L/L). Our experience with other American scenes suggests that this orifice compression scarcely indicates receptivity to their plea. We have for the purposes of this example elected to limit the scene to that period during which all members of the transaction could "see" each other. The "why" of the transaction may rest upon the boys' previous experience that day which occasioned the particular stance which they maintained. Perhaps as the car came into view it swerved almost imperceptibly toward the soldiers, thus alerting them to the driver's attitude. Interpersonal space

variations are in part extensions of kinesic activity and are often definitional of communication situations. Within the range of our abstracted scene, the driver's face was clearly visible to the soldiers for scarcely two seconds, and his head and face movement took less than a half second to complete. The observer has no way of finding out exactly what the soldiers "saw." Yet their unhesitating reaction indicates that the driver's analyzable act was transmitted to them. Both soldiers responded with acts of the same class as that used by the driver. Further, Boy No. 2 selected his from the same South European (post-World War II American male overlay) diakinesic<sup>1</sup> system as that expressed by the driver. This supports the conclusion that theirs was a response to the driver's activity and not simply an idiosyncratic reaction to being refused a ride.

Further questions arise from the analysis of this microcosmic scene. Was the driver initially stimulated to his insultingly rejective activity by the spread-legged stance of the boys? And/or were the left thumb in the belt combined with the spread-legged stance (often part of the pre-fight or pre-sexual advance behavior of adolescents) dominant as parts of a definitional act which challenged him, a male, into his negative response? Obviously, only by observing this driver and these soldiers in a series of contrasting scenes would such questions as these be answered. There is a strong suspicion, however, that if the driver had

---

<sup>1</sup> Analogous to a dialect system in language.

responded with a back nod of less ascending velocity, raised his eyebrows bi-laterally and lowered the corners of his lips in a "I would if I could, but if I can't, I can't," the boys would have carried out their activity in a considerably less hostile manner.

This scene is illustrative of the extent to which a human communicational event, a transaction, can be completed without recourse to verbal behavior. At the same time it demonstrates the fact that communication within even one modality is seldom a simple affair. The student of body motion behavior is not always so fortunate as to have a scene so clearly defined for him. Nor do most transactions have their interactional tempos so neatly marked, as in this case, by the explicitly conventionalized "gestures."

Notwithstanding its relative simplicity, the scene provides a useful point of departure for our present discussion. The ritual of "thumbing a ride" is familiar in American culture, yet a closer analysis of this special incident is illustrative of the hidden complexity of such scenes. In the soldiers' persuasive activity with the "thumbing a ride" gesture as the ostensible action proposition of this scene, we are provided with an excellent example of the extent to which an act can be modified by incongruent movement complexes which complete it. The spread-legged stance, congruently modified by the thumb-in-belt complex, contains two components which combine in a larger act. At the same time this act is, at one level, incongruent with the gesture of thumbing. As Bateson has suggested, such

components may modify, that is, may constitute commentaries on, each other. What they mean is another matter. At the moment we are concerned only with their relationship to each other and to the package act of "thumbing."

By careful cross-context analysis, we can derive a series of working hypotheses concerning the soldier's initial act and its incongruent components. The stereotypic "thumbing" gesture is deceptively familiar.

We must remember that the same gesture in another actional setting is conventionalized as the insulting or mock-insulting directive to "Get lost!" In fact, the complex act described above, if it took place on a street corner in Los Angeles or Chicago, could have just this explicit interpretation. Through contrast analysis, we are able to say that the "thumbing" action as produced is itself incongruent with its context--if we postulate that the dominating purpose of the boys was to persuade the driver to give them a lift. The recognition that communicational behavior can be congruent in one setting and incongruent in another should serve as a warning against any theory of meaning which suggests that the particles carry meaning in and of themselves.

Contrast analysis permits us to define this particular combination of movements in this context. We postulate the arm and thumb as an "appeal for a ride", the spread-legged stance modified by the thumb-in-belt as "male defiant," and the whole as an act conveying a "defiant appeal for specific assistance." This complex of behavior is consistent with

the role of these late adolescents, in uniform, who are avoiding "begging." These young soldiers are in no position to play the role of the college boy who "thumbs" a ride but whose college sticker and clothes belie the ingratiating stance and head cock plus smile with which he modifies his petition.<sup>1</sup> We could pursue such contrastive examination throughout the entire scene, and in the final analysis the social meaning of the individual movements, gestures, acts, and action must be phrased in terms of the entire scene. These are all susceptible of analysis if the activity is seen as a transaction, in the context provided by the various participant social roles as defined by American male sub-culture. The scene may be viewed as a role-stating ritual in which the component activity is such that it negates the central gesture. The boys must wait for another car and driver in order to get to Los Angeles. It is probably safe to say that the boys either must amend their activity or wait until a driver with a different set toward such messages comes along if they hope to get a ride.

#### Background to Kinesics

Kinesics is the systematic study of the communicational aspects of human body motion. The methodology of kinesics is still extremely crude. At its present stage of development, kinesics may claim to be a science only by virtue of

<sup>1</sup> For an interesting analysis of the complex social psychological aspects involved in such "presentations," see Goffman (1954).

the canons which dominate its operations and by virtue of the postulates upon which these operations depend. As a body of knowledge, it cannot yet be judged worthy of the appellation kinesiology. Yet five years of research which has utilized and constantly refined the methodological procedures of kinesics have been so fruitful that it is without qualms that the present investigation, employing those procedures, is attempted.

It is entirely fitting that psychiatrically oriented interview material be the subject matter for this initial attempt to apply practically the data derived from kinesic investigation. Psychiatrists and psychologists have for over a century been aware that body motion and gesture were important sources of information regarding personality and symptomatology. Allport (1), Dunlap (2), James (3), Krout (4), Lersch (5), Ombredane (6), Groddeck (7), and Wolff (8) are but a few of the students of personality who have contributed to a considerable body of literature concerning expressive movement. The brilliant observations of Felix Deutsch (9) on what he calls "posturology" must be especially noted. His is one of the clearest statements concerning the diagnostic value of body motion and posture. Kinesics, however, represents both a theoretical and a methodological departure from studies such as these which stress personal activity and individual performance. It is our hope that communicational research and, particularly, kinesic research, will provide a methodology, an annotational

system and a set of norms against which these kinds of intuitional systems can be checked. It is our conviction that significant statements concerning the behavior of particular individuals must be based on an understanding of the patterns of intercommunication of more than one actor. The significance of particular individual variation can be assessed only when the range of permissible group variation has been established.

There is nothing new about the recognition that formalized gestures play a role in communication. Theatrical performances, whether centering around dancing, drama, opera, or the mime have long emphasized the role of gesture, particularly in its stereotyped or conventional form. Integral to every religious ritual, the gesture is stressed in all novitiational training. A considerable bibliography has been collected with representation from almost every literate country and extending back in time to early India, which evidences the international character of the interest in gestures, and their proper performance. Most of these writings are of collateral interest to the kinesiologist.

The concentration upon the particular gesture and its meaningful performance leaves most of these writings of primary concern to the folklorist. Perhaps when extended research into the kinesis systems of particular areas has provided a body of background material, much of this earlier material will become relevant in a new way, just as linguistic

research consistently opens new perspectives upon old data of a verbal nature.

Of these earlier publications, most relevant have been those which have dealt with the development of systems for annotating body movement. Before the publication of the Introduction to Kinesics<sup>1</sup>, we carefully reviewed a series of annotational systems and were particularly impressed by those of Craighead (1942), Lifer (1940) and Pollenz (1949) and these doubtless influenced our system of microkinesic recording, the logic of which is presented in Appendix 3. These annotational systems are all extremely useful for recording the conventionalized patterns present in the dance, modern, classical, or folk. They are, however, somewhat too limited in scope for use as instruments of broad kinesis research. Perhaps the most complete and extensive recording system in usage today is that provided by the Laban<sup>2</sup> school. Used principally for industrial studies, this system has been used effectively both for stage and for general movement recording. The decision to develop the specialized system presented here rests finally upon the conviction that annotational conventions which signal the specific operations governing their abstraction are probably desirable. In short, recording systems should derive, in the first instance, from considerations of theory and methodology, rather than the reverse. As

---

<sup>1</sup> Birdwhistell, 1952.

<sup>2</sup> See Laban, 1940.

research and theoretical re-evaluation continues, such recording procedures must necessarily be revised.

Both the microkinesic system outlined in Appendix 3 and utilized on pp.     to pp.     , and the macrokinesic system demonstrated above in Figure 1 and outlined in Appendix 4, have been revised a number of times and must be further revised as body motion research continues. Certainly any system which is as accurate and which would permit still easier and weifter notation would be more desirable. As the annotational system for microkinesic recording now stands, only a relatively large, well-trained (and thus expensive) team could record live micro-cultural material with any degree of completeness and accuracy. Designed for the analysis of filmed material, the kinegraphs are useful only for checking kinesic research with live subjects. They are insufficiently flexible for primary microkinesic research on such subjects.

This is not the occasion for an extensive presentation of the background to kinesic research. However, to put the present work into perspective, some reference should be made to a series of men whose work bears directly upon the development of kinesic theory and methodology. The pioneer work, Darwin's Expression of Emotions in Man and Animals<sup>1</sup>, contains many lucid suggestions which foreshadow the theories emergent in the present-day fields of ethology or comparative psychology or kinesics. The rigor of his observations combined with his

---

<sup>1</sup> See Darwin, 1871.

nascent sense of the dynamics of social interaction make many of his statements seem contemporary. One cannot help but wonder how far Darwin would have gone with this work if he had had the tools of descriptive linguistics, of communication and information theory, and the technical aids to precision provided by the tape recorder, the sound camera and the time-motion analyzer.

The interdependent nature of linguistic and kinesic research is anticipated by Edward Sapir,<sup>1</sup> who, a little more than a half century later says,

"Gestures are hard to classify and it is difficult to make a conscious separation between that in gesture which is of merely individual origin and that which is referable to the habits of the group as a whole ... we respond to gestures with an extreme alertness and, one might almost say, in accordance with an elaborate and secret code that is written nowhere, known by none, and understood by all."<sup>2</sup>

Sapir did not follow up his own lead, but it is his students and other linguists strongly influenced by his work who have contributed most to the systematization of body motion research. George L. Trager and Henry Lee Smith, Jr., at the time doing research in the structure of American English at the Foreign

---

<sup>1</sup> Sapir, 1927.

<sup>2</sup> Our italics.



Service Institute, provided an atmosphere and the special guidance which encouraged the original formulation of kinesics as a science. John Broderius, another student of Sapir's, worked cooperatively with me at later stages of the refinement of kinesic principles. His constant insistence that kinesics be firmly based in pre-kinesic research and not be lost, as he phrased it, "in the thin stratosphere of intuition," helped to maintain the frame which early association with Smith and Trager had produced. The present research with the linguists is another logical step in the necessarily interdependent companionship of descriptive linguistics and kinesics.

Parallel to these influences and consistent with them, have been the writings of a series of anthropologists whose field experience, as did my own, led them to the conclusion that body motion and facial expression were strongly conditioned, if not largely determined, by the socialization process in particular cultural milieux. While affirming the ultimate biological basis for all human behavior, there seems little doubt that out of the vast range of possible combinations of muscular adjustments, perhaps a quarter of a million in the facial area alone, each society "selects" certain ones for recognition and utilization in the interaction process.

Probably the pioneer anthropological analysis of gestural activity is Efron<sup>1</sup>'s test of the hypothesis that there is a direct correlation between the previous social environment of European immigrants to America and their gestural systems.

---

<sup>1</sup> 1941.

Concentrating largely on the range of movement in the arms and hands, Efron contrasted the gestural systems of Italian and Southeastern European Jewish immigrants. Although his thesis correlating certain ecological factors with the respective gestural systems remains inconclusive, his work effectively demonstrates the social genesis of the evident variation in the gestural systems of these two groups.

While Efron's experimental approach has not been pursued by other investigators, Labarre<sup>1</sup> and Hewes<sup>2</sup>, with quite different emphases, have directed the attention of field workers to the importance of recording and analyzing the gestural behavior of human groups. However, the most important anthropological contributions to the development of the study of body motion as a communicational system have come from the work of Mead and Bateson<sup>3</sup>. Their concern with the relationship between socialization and communication, assisted by considerable skill with and appreciation for the camera as a research instrument, set the stage for the development of kinesics as a behavioral science. Not only has their field work provided a body of materials for cross-cultural study but their insights into the systemic quality of the communicational process have prevailed upon the writer to take up his profitable association with the linguists.

---

<sup>1</sup> 1947.

<sup>2</sup> 1950.

<sup>3</sup> 1942.

These few paragraphs were not intended exhaustively to relate the contributions which underlie the present status of kinesic research. The recognition that body movement contributes to interpersonal understanding (and misunderstanding) is probably as old as man's interaction with man. Yet kinesics as a systematic approach to such phenomena is still relatively untested. The brevity of these remarks does not represent a failure to appreciate the extensive scholarship from which kinesics has emerged but rather a feeling on my part that our present experiment demands an assessment of the infant discipline rather than a genealogical legitimatization of it. The scientific history of a discipline should be concerned with the tests and modifications of its theory and methodology; until investigators are trained and experiments are performed, kinesics will have no history.

In seeking to comprehend and to make intelligible those aspects of human body behavior which contribute to the communicational process, the kinesicist-anthropologist employs a set of procedures which are special only in the sense that they must be adapted to the peculiarities of the system under examination. Dealing with a universe which he has pre-defined as ordered and interdependent, his primary task is that of developing a methodology whereby units and sub-systems can be abstracted and manipulated. From the seminal insight that kinesic activity constitutes an infra-communicational system is derived a plethora of data

which, unless explicitly and methodically ordered, drowns the investigator in a myriad of shapes and sizes and orders of behavioral pieces. Having fixed his eyes upon the behavior which constitutes the human interactional scene and having adjusted himself to the outrage of the recognition that communication is continuous, he must resist a series of temptations which would shortcut and, coterminously, predetermine the results of the observational process. Some of these temptations are suggested in the discussion above but their subtle influence upon the work of those concerned with "non-verbal communication" has been such that they are probably worthy of explicit delineation.

#### Temptations

##### The "carrier" temptation

This derives from a linguistic naïveté which assumes that each gesture, whether as gross as a thumbed nose or as tiny as a first degree right lid droop, has a "real" meaning just as "words" are supposed to have. If the investigator succumbs to this, his attention is directed into a kind of dictionary wherein he draws up lists of moves and their meanings only to discover that most human beings are kinesically illiterate and move improper English. As shall be demonstrated below, even a preliminary approach to kinesic data reveals that no abstracted body motion, gesture or kine, has a precise and absolute meaning apart either from its position in the kinesic system or from the social context

in which it appears. The question "What does X mean?" can only be met by the counter questions of "At what level of analysis and in what contexts?".

The "closer to nature" temptation

This category really covers two companion but differing hidden assumptions. One of these is that body movement is somehow more primitive and thus closer to biological nature than is verbal behavior. Animals move and animals don't talk. Humans move and talk. Ergo, moving and kinesthetic-visual communication came earlier in evolutionary history than did talking and thus remain unpatterned. Depending upon the predispositions of the writer, this same assumption has permeated the work of the individualists who feel that body motion and facial expression reveal the "true" feelings of a communicant, the writings of the racists who confuse social variation in response pattern with genetically determined "stoicism", "vivacity" or even rhythmicity, and the universalists, who assume that since there is minimal biological variation in homo sapiens and since moving came early, there is species-fidelity and universality in all movements. The way in which these assumptions are expressed varies from that of some of the individualists who say that everyone is so different from everyone else as to preclude generalization at all to that of some of the universalists who optimistically anticipate a movement catalogue. Whether simply nihilistic or modern pastoralist these assumptions do not hold up as we examine the communicational situation. ~~Not~~

Not only is kinesic activity systematically patterned but this pattern varies significantly from culture to culture and even from sub-group to sub-group. While eventually we may find that the special physiological patterning of special groups may influence to a considerable degree the characteristic tone of the kinesic activity of such groups, we expect also to find a reciprocity of influence between the biological and social systems rather than any pattern of basic priority of a simple genetic nature.

More subtle and more seductive than these assumptions which deal largely with the total membership of society are those which see infantile behavior as more natural than adult behavior. Those so persuaded see maturation as somehow artificial and distortional of infantile naturalness or, accepting maturation as a natural process, these writers seem to feel that those behaviors which are characteristic of the infants of a group (or of all infants?) are somehow truer representations of the feelings of the communicant than are those more characteristic of adolescence or maturity. So long as generalizations such as these are related to the examination of individual responses and deal with the documentation of personal histories they are not of direct concern to the kinesicist. However, if they are permitted the dignity of becoming basic to all systemic interpretation it is well to point out that our knowledge of the ontogenetic development of individual kinesic systems is less than fragmentary. There exist a number of suggestive--even exciting--

studies of maturational behavior. But we lack the cross-cultural longitudinal analyses which would permit any safe generalization of "how" humans learn to become communicators or give us more than an intuitional feel for the sustaining strength of infantile response.

It must be pointed out that this does not in any way affect the validity of the regression hypotheses. It is evident to any observer that adults will in special situations behave incongruently with their level of maturation. However, to assume in an a priori manner that this proves the strength of the infantile response is to ignore the communicational function of the act.

As we shall discuss below, while a body curl or a thumb suck may on one level of analysis be incongruent with other kinesic behavior being exhibited by an actor, such behavior may be quite congruent in the total communication situation. In the sections below on body set and motion quality the differences between "ageing" and "age grading" will be discussed. Suffice it for the moment to say that it is the present premise of kinesics that considerable research on the social learning patterns of infants and children must precede any security on our part concerning "basic" behavioral manifestations.

#### The "modifier" temptation

As professionally literate members of a culture devoted to literacy, we are strongly tempted to believe that words carry the meaning and that all other non-word behavior merely

modifies it. Thus, there are those who feel that words form the natural center of the communicational universe and that all other modes of communication are to be studied as sub-systems subordinate to it. Such a decision predetermines the nature of the communicational process and I am as yet unwilling, from the situations which I have examined, to assign any such priority to any of the infra-communicational systems. For the kinesicist, silence is just as golden as are those periods in which the linguistic system is positively operative.

As shall be seen below in the discussion of kinesic markers, there are aspects of kinesic activity which have an infra-communicational function only heuristically separable from the vocal activity. Correlated with the process of verbalization, these markers, whether an aspect of the speaker's production of the message or the listener's contribution to the transaction, deserve special attention in an assessment of an interview like the one presented below. Indicating position, temporality, special emphasis, subject, object, etc., the markers, like many gestures, are often so closely bound to linguistic behavior as to seem like extensions of it. Further research may well force a special categorization of this kind of kinesic behavior. At present, however, with the recognition that during much of human interaction verbalization is absent, it seems proper to study the two systems as of comparable weight in the communicational process.

This temptation has received stress because of its implications for communicational theory and research. When do humans verbalize? Is there a correlation between intimacy, for instance, and a reduction of conversation? Is there a correlation between the culture of a group and its dependence upon one mode of the communicational process? What are we talking about when we say that one person is verbal and the other taciturn? Even such a subjective term as "good listener" may now be within the reach of objectification. It seems unlikely that such questions as these can be answered until we have considerable understanding of the nature and the role of the infra-communicational systems and their relationship to each other. To assume priority for one or the other sub-system prior to such research would be to oversimplify the problem in a manner already too familiar in so-called "content analysis."

The "central movement" temptation

Somewhat more technical than these temptations is the tendency on the part of the investigator to assume that one part of the body "carries the meaning" and other parts "modify" this central message. This is particularly seductive because we "know" intuitively as a member of a particular diakinesic system that certain movements seem to take precedence in the presentation or reception of a message. The eyes, the mouth, the face, the hands, the posture, the shoulders have all been listed by informants as being the

primary carrier of meaning. To accept such statements would be a little like accepting an informant's conviction that nouns or verbs or even consonants or vowels are the most important part of language. Further, as is true in linguistic analysis, simple particle counting does not give us a score revealing system importance. I have no doubt but that research will reveal that given cultures will, by sheer count, tend to produce more movements from one body area than from the remainder. Such counting does not, however, permit the investigator to assume a correlation between the incidence of usage of a body area and its functional importance either to the infra-communicational system or to the communicational process. Redundantly, I must again insist that only following systematic analysis of kinesic units and patterns can so-called central movements be established.

Even with the minimum of cross-cultural data at our disposal, the evidence is clear that cultures will tend to concentrate activity in certain body areas and permit the activity of others only under certain very limited circumstances. It seems evident that this will have momentous implications for students of national character. However, it does not follow that we can make statements like "Spanish women use their eyes and Russian Jewish women their hands and American stenographers their feet to say what they really mean." Such statements as these will remain at best brilliant intuitions until we comprehend the respective

kinesic systems of these women and the role of these systems in the communication processes of their respective cultures.

The "analytic informant" temptation

Kinesics, like the other behavioral sciences, uses informants as well as direct observation in gaining control of the data of the discipline. Like linguistics, however, it insists that the informant be an informant and not a fellow analyst. The young investigator is particularly prone to ask the informant what he has done or what the movement meant and to forget that the answer provides further data for analysis, not an acceptable conclusion to his analytic research. Even those investigators too sophisticated to rely on such subjective contributions may in lieu of behavioral description and analysis substitute the "multiple judge" technique. Often little more than a pooling of ignorance, such a technique is perfectly valid if the investigator is concerned with questions of establishing patterns of recall; it contributes little to the final abstraction and analysis of the kinesic system.

Kinesics is concerned with the abstraction of those portions of body motion activity which contribute to the process of human interaction. Much, if not the overwhelming proportion, of such behavior is learned by a member of any society quite out of awareness. It is my belief that not only is much of such behavior not within the range of recall but that the learning pattern may carry within it positive prohibitions on recall. Kinesics is not concerned, as such,

with the movement potential of the human species, but rather with those portions of the movement spectrum which are selected by the particular culture for patterned performance and perception. At the same time, as is true with other cultural behavior, much of what happens and which is necessary to the proper performance of a social act cannot be recalled by the actor or the untrained spectator. I said above that I had a belief that as the child is taught to move, to view and meaningfully to reproduce movement, an integral part of this education is concerned with enhancing or preventing recall of much of this activity. Preliminary observation of "flat-land" southern contrasted with New England children in Louisville from comparable socio-economic positions supports the conclusion that, even within a single culture, sub-groups may experience socialization processes sufficiently different to create misunderstanding between them. Not only was the southern "raised" child encouraged to engage in gender-identifying behavior earlier than his or her Yankee cousin, but it had far greater recall in this area both as actor and as viewer than did the northern child.

The need for skilled observers in kinesic research is evident, but even training is at times insufficient guarantee of objectivity in certain situations. One of the critical scenes discussed below of the Billy-Doris-Gregory interview contains extensive intrafemoral hand play on the part of Billy. I must confess that it was only after some thirty viewings and with the demand for micro-kinesic recording that I allowed

myself to see that his hand play was patterned. I venture to suggest that early training which precluded my "seeing" male play in the genital area contributed to my concentration of attention on the little boy's eyes and head.

Suffice it to say that an informant should be used as a window into a culture. As shall be seen below, his contribution to the research is indispensable. The investigator must constantly remind himself, however, that his informant is an adherent, not an objective interpreter, of his communicational system. Not only one but fifty million Frenchmen are likely to be wrong in their view of their own communicational system.

#### Methodology

Having determined the systematic nature of human interaction and having recognized that membership is attained in a social system only after patterned experience in this system, it is the task of the behavioral scientist to ascertain what it is that is learned which provides any particular system with its particular dynamic. It is not my task, but that of the psychologist to determine how the organism incorporates the experiences which make him a human being. Neither is it my task to map the internal relationships of the physiological systems out of which emerge the perceptible shifts in the various parts of the body. As an anthropological kinesiologist I am concerned with the learned and visually perceptible shifts in the body which contribute to the peculiar communication systems of particular societies. Kinesics is

concerned with abstracting from the continuous muscular shifts which are characteristic of living physiological systems those particular groupings of movements which are of significance to the communicational process and thus to the interactional systems of particular social groups.

The human body is capable of producing literally thousands of distinguishable positional shifts per second. Even at "rest" the body is not inactive. A high speed movie camera, the so-called slow motion camera, as it is speeded up, records more shifts or motions the faster it is set. Obviously, on some level of analysis these are of significance. The question which immediately confronts the kinesiologist is whether or not his minimal unit of activity is in the last analysis to be determined merely by the speed of his film and camera and the patience of the recorder.

There is a considerable body of data concerning the speed of neural transmission. An even larger bibliography is concerned with the psychological study of visual perception. Neither of these, unfortunately, provide us with a statement of biological potential which might in any a priori way delimit the raw material of kinesics. In short, the body of one human being produces a volley of signals, an indeterminate proportion of which may excite the optical nerve of another human being. Observation of the two over any extended period of time will reveal that, if the two were selected from a common social group, each adapts his behavior to the activity of the other. The intra-personal activity which results in

such adaptive muscular shifts and electro-chemical activity in the visual area are pre-kinesic in nature.

This is not to say that the behavior of the physiological system is isolated from the social environment. Even the most cursory examination of the cross-cultural or ontogenetic data indicates that the developing system is influenced, if not shaped, by its patterned interaction with its environment. The reverse is equally evident. In the same way, the knowledge that member X of society A will tend to be more active in one area of the body than is member Y of society B, is of obvious concern to the kinesicist, but such interest is still pre-kinesic. The data of kinesic's is not derived from the observation of intra-personal behavior. A product of systematic social interaction, the kinesic system is a social system. Out of the range of muscular adjustments produced by a human being some are utilized by the social system for communicational purposes. Thus, to say it simply, no human body produces a kine (least kinesic unit); it moves or adjusts in a set of muscular relationships. In social interaction, certain of these have demonstrably special utility in the communicational process. That is, under analysis, they emerge as kines. Every visible body movement, accordingly, is not a kine any more than every audible noise made by the vocal apparatus is a phone. Only after analysis has revealed that the presence or absence of a given movement in a particular context systematically affects the interactional process do we assert that that movement has kinesic significance.

#### The Kine

A kine is an abstraction of that range of behavior produced by a member of a given social group which, for another member of that same group, stands in perceptual contrast to a different range of such behavior. While, theoretically, within certain limits provided by the physiological structure, a given complex of muscular reactions may produce a continuous series of positions, in actuality, any social system patterns these into a discontinuous or discrete series for reception or reproduction. Thus, while, for example, the membership of culture A will report only two degrees of lid closure, culture B may recognize as many as five. As a skilled spectator under optimal conditions I can record or reproduce fifteen degrees of lid closure quite distinct from each other but most middle majority informants "see" only three. Similarly, while a finger of even the distal joint of the finger can produce a continuous arc of position in relation to the remainder of the finger or hand, four degrees of finger position on this axis are all that elicit the report of a perceptual contrast from a middle majority informant.

Thus, a kine is not a point or position of articulatory activity; it is a range which the unsophisticated informant reports as "the same". In a previous publication<sup>1</sup> points within this range were described as being in allokinic

---

<sup>1</sup> Birdwhistell, 1955.



relationship to each other. I propose now that these be called kine variants, since they may be substituted for each other and are, thus, symbolizable by a single class denoting symbol. At the risk of being repetitive I must restate the point that these equivalences are culturally defined. Each kinesic system will have differently shaped kinic classes. As a demonstration, we may use laterality as a special test of kinic significance. All indications are that, at least on the level of the kine, American movers do not necessarily, in awareness, distinguish laterality. Given individuals may favor the fingers of the right hand, the right eye or the musculature of the right side of the face. Two American middle majority movers, one favoring the right side of the body, the other the left, can, as far as we now know, interact without translating "right lid droop" into "left lid droop" or vice versa. This seems to hold for all body parts considered on the kinic level. This is not to deny the obvious fact that handedness is of social significance. What we are here concerned with is whether we can record, say, the movements of the right or left lid as variants of the same kine. We must test whether R~ is equivalent to L~ and whether they can be regarded as variants of a kine class ( <sup>3</sup> ). It is obvious that they are distinguishable on the level of articulation. The test is not, however, whether the informant tells us that the right or left lid is used. What we need to discover is whether they function interchangeably in larger kinesic contexts.

As a test let X stand for a specific brow-kine; Y stand for a specific lid kine; Z stand for a specific lateral orbit kine:

	<u>LX</u>		<u>RX</u>		<u>LX</u>		<u>RX</u>	
Are	<u>RY</u>	,	<u>LY</u>	,	<u>LY</u>	,	<u>RY</u>	equivalent
	<u>LX</u>		<u>RZ</u>		<u>LZ</u>		<u>RZ</u>	

to each other in a manner which permits us to establish a

$$\frac{X}{\frac{Y}{Z}}$$

class of  $\frac{Y}{Z}$  covering all four of these as class variants?

If it is inconsequential whether the right or left eyelid is involved in each of these structures, we have no need to establish RY and LY as members of different classes since they are variants of (Y). The fact that the difference between R and L may not be of significance on this level does not however preclude the possibility that on other levels of analysis they may function contrastively.

#### Kinemorphics

In earlier formulations of kinesics, to expedite recording, yet with the intuitive feeling that the particular division of the body "made sense", I arbitrarily divided the body into eight specific areas. Systematic investigations, utilizing contrast analysis, have since justified this body division--when applied to American movers. However, even a few hours of work with Indonesian and Bombay Indian

informants makes it clear that the specific divisions will not hold up cross-culturally. The eight areas, head and neck, face, shoulders and trunk, right arm, left arm, pelvic region, right leg and left leg, will probably be differently subdivided according to the body conception of a given social system. The particular range of such segmentations can only be determined by further research. Nevertheless, the kinemorph was defined then as an assemblage of movements (kines) in one such area.

A kinemorph is not merely an assemblage of movements in a given body area. A moving picture of such an area would not provide the investigator with a kinemorph. Such a picture or abstraction from it in the form of an exhaustive list of micro-kinographs or articulations would provide us with relatively little concerning the kinesic system of the actor. We must again use the method of abstraction and contrast analysis. As soon as we begin to contrast, with the aid of an informant, a series of kine assemblages, it becomes possible to abstract those which form unitary complexes. To return to the example which we used in the test above: We may find that we cannot set

	<u>RX</u>		<u>RX</u>		<u>RX</u>
up a single kinemorph to cover	<u>LY</u>	,	<u>RY</u>	,	<u>RY</u>
	<u>RZ</u>		<u>LZ</u>		<u>RZ</u>

<u>LX</u>		<u>RX</u>		<u>LX</u>
<u>LY</u>	.	<u>RY</u>	and	<u>LY</u>
<u>LZ</u>		<u>RZ</u>		<u>LZ</u>

Further, we may discover that

are kinemorphic variants to which informants react as substitutable for each other at this level. Similarly,

<u>RX</u>		<u>LX</u>	
<u>LY</u>	and	<u>RY</u>	
<u>RZ</u>		<u>LZ</u>	

may also be found to be substitutable for

each other. We may then conclude that we have two kinemorphs, which may be recorded as [XYZ] and [ $\dot{X}$ YZ].

While this example gives some idea of how the kinesicist deals with contrast analysis, it will be exceedingly misleading if it is not seen as over-simple. For while all of the kines which compose a kinemorph are to be found within a given time frame (while will be discussed below), they are not necessarily coterminous. I have thus far been able to abstract three kinds of kinemorphic constructions, their definition dependent upon the order behavior of the component kines:

- (1) synchronic kinemorphs; in which the component kines are simultaneous and of equal duration;
- (2) series kinemorphs, in which the kines follow one another in time; and
- (3) mixed kinemorphs which have both synchronic and series features but in which all component kines are not of the same duration.

Figure 2

Figure 2

(1)	<u>X</u>	(2)	<u>XYW</u>	(3)	<u>XY</u>
	<u>Y</u>				<u>W</u>
	W				

Each of these meets our definitional criterion of taking place within one body area and each forms a complex in which all components are necessary for the production of the unit and all are to be found within a given time frame. In the discussion of the kine we did not deal with the durational aspect of its definition, since by extended test, it is clear that performance not duration determines the kine.

A raw movement becomes classifiable as a kine at any time that its performance (of whatever duration) suffices to change the contrastive function of the complex in which it operates. The same kind of test is utilized on the kinemorphic level since the kinemorph is more than an arbitrary grouping of kines. We can establish the kinemorph, not only because the informant tells us that "these movements fit together", but also because we find transition devices which mark its initiation and terminus, and because we are able to establish its unitary function in larger contexts.

The most readily apparent kinemorph is one which begins with the body at zero (Z) and ends with it at zero (Z). Zero is defined as attention without specific movement, or, in the appropriate context as an arbitrary norm from which all kines are traced. Such kinemorphs can be described as pause-marked (=). A second type is characterized by onset of activity in one body part and is terminated by the introduction, from zero, of activity in another part. The term areal transition (X) seems useful here.

There is a third type of transition, the bound transition (+), which marks kinemorphs which can only be detected by extended contrastive research. This occurs when one kinemorph is replaced in the same body part by a different kinemorph which utilizes the same points of articulation but by rearrangement of order and/or duration establishes a complex with a meaning demonstrably different from that of the previous complex. The fact that these types of kinemorphs are differently marked by differential transitional behavior indicates that future analysis may reveal their special roles in the kinesic system. On the other hand, it is within the range of possibility that they are functionally equivalent and are merely contextual variants.

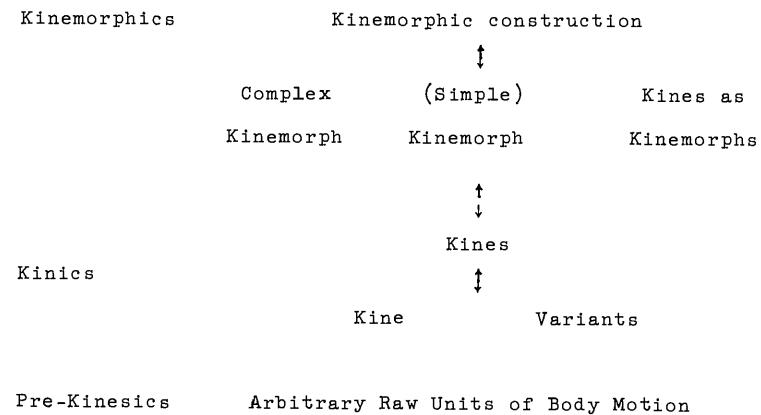
The linguist will see that the kinemorph and the morpheme are in some ways comparable. For several years I have been hopeful that systematic research would reveal a strict hierarchical development in which kines could be derived from articulations, kinemorphs from complexes of kines, and that kinemorphs would be assembled by a grammar into what might be regarded as a kinesic sentence. While there are encouraging leads in the data, I am forced to report that so far I have been unable to discover such a grammar. Neither have I been able to isolate the simple hierarchy which I sought.

While, by count, a major proportion of the kine assemblages of the American kinesic system may be meaningfully segmented in one body area, there are many occasions when the restriction of contrast analysis to one area leads only to confusion. This occurs when kines from two or more areas form a complex, which, under contrast analysis, behaves precisely like a kinemorph. These I have chosen to call complex kinemorphs. The complex kinemorph and the simple kinemorph seem to be on the same level of analysis in that they may both be directly analysed into kines. By definition, of course, the complex kinemorph differs from the simple kinemorph both in its placement and in the fact that we cannot utilize simple shift of body area as a transition marker. Letting (A), (B), and (C) stand for kines in one body area and (X), (Y), and (Z) stand for kines in another, we may specify the shape of the complex kinemorph as [A, Y, B], while the shapes of simple kinemorphs are [A, B, C] or [X, Y, Z].

To complete the description of this level of analysis, I must include those single kines which emerge as kinemorphs: thus (A) becomes [A], (X) becomes [X], and so on. The test for kinemorphic function continues to be one of abstraction and contrast analysis. Our testing context is the kinemorphic construction. The raw unit of body motion is classifiable as a kine when it is seen to have differential value in a kinemorph. Ultimately, the existence of the simple kinemorph, the complex kinemorph, and the kine as

kinemorph must all be established in the kinemorphic construction.

Figure 3



In Figure 3, the reader will note the succession of two-way arrows. This indicates that at each level of analysis a unit not only must be abstractable from non-significant variation, but must be seen to have differential meaning in the complex in which it operates.

The term kinemorphic construction is suggested to cover the next order of combination in kinesic behavior. As kines combine with other kines to make isolatable units

(the kinemorphs), or as single kines emerge as kinemorphs, these forms combine with each other in a variety of kinemorphic constructions. The simplest of these is the kinemorphic combination. A kinemorphic combination is constructed of two or more kinemorphs--either in parallel or in series. Recorded initially as [A, B, C] [X, Y, Z], if the combination can be tested for unit existence, i.e., as having differential meaning in a wider context of body activity, it can be recorded as / [A, B, C] [X, Y, Z] /. We further find that complex kinemorphs can combine with other complex kinemorphs in complex kinemorphic complex combinations / [A, X, C] [Y, N, Z] /, and with single kinemorphs to form kinemorphic compounds / [A, B, C] [X, N, Z] /. Finally, constructions of the shapes / [A] [X Y Z] / and [A] [X, N, Z] / have been abstracted.

The existence of a kinemorphic construction is determined by exactly the same procedure as has been utilized in the abstraction of the kine morphic or kinic combinations. We abstract an assemblage in which the components repetitively appear in conjunction with each other. Then by substituting behavioral events of a comparable shape within the combination we establish the kine morphic function of the components and, by extension, the reality of the morph, on one level, and that of the construction on another. Thus, the kinemorphic value of kines is revealed when we discover that / [A] [X, Y, Z] / stands in contrast to / [B] [X, Y, Z] / in exactly the same way as / [A] [X, Y, Z] / stands in contrast to

/[A] [M, N, D]/. Thus, recording /[A, B, C] Z [X, Y, Z]/ signifies that the construction has been abstracted from a larger action sequence and that during its duration the rest of the body has remained at what one of my students aptly referred to as "ready rest." The (Z), in this case, serves to remind the analyst that he is dealing with an included construction. When a full actional sequence is dealt with, (//)'s are utilized to mark the initial and terminal aspects of the sequence and all segments within the double slashes are bound constructions and form a unit on the next level of analysis.

I have not yet found any way of determining whether or not there is a conventional limitation, in terms of the number of component morphs, to the size of a kinemorphic construction. In the examples above two-part constructions were used. The reader must not be misled by this. I have seen kinemorphic constructions which contained as many as seven morphs. The test for the unitary nature of a kinemorphic construction takes place in the larger field of body movement which we call action.

In the definition of the kine, I said that zero does not stand for no behavior at all but for behavior which does not have variational kinic or kinemorphic significance.

#### Action

Analysis of the American kinesic system has led to the tentative conclusion that in the probable absence of cross-referencing systems similar to those of linguistic grammar and syntax, meaningful segmentation and binding together of kinesic construction sequences if handled, in all likelihood, parakinesically<sup>1</sup> the medium of stance. Stance is a term designed to cover a pattern of total body behavior which is sustained through time, within which one or a series of constructions take place, and which contrasts with a different stance. Stance subsumes position (p), (which is a statement of the relative position of all the body parts in space) locomotion (l), (the movement of the body through space) and velocity (v) (which covers sustained velocity of movement of the total body).

A stance change is said to occur when any one of these or combinations of these is varied to such an extent that there is a marked shift in the total message. In a major proportion of the interactions which we have observed these shifts coincide with a transition (+) (X) or (=) on the construction level. Our problem would be a good deal simpler if we could say that action-sequence transitions always coincide with inter-construction sequences. Certain of these stance changes, however, take place within what appear to be bound constructions as well as within an included construction. Such stance changes

---

<sup>1</sup> See now Birdwhistell, 1960.

may or may not be coterminous with morph transitions. The term stance shift is used to indicate this variety of stance variation which may, as research develops, turn out to be parakinesic in nature.

While our research in this area is far from exhaustive it seems probable that stance variation may serve at least a dual function. On the kinemorphic level, stance <sup>shift</sup> serves to mark the beginning and end of action sequences. In such cases (//)'s mark the action sequences and all elements included within are analysed as bound components in an interactional system. Thus, we might record a typical action sequence according to the notational logic, //[A, B, C] [N] [X, O, Z] [etc.]//. The type of stance change is marked by a small letter at the upper right of the double slash notations: (//<sup>P</sup>), (//<sup>L</sup>), or (//<sup>V</sup>). When a stance shift occurs within a sequence of bound constructions, a capital S is used with the identifying marker. Thus: <sup>V</sup>//[X, Y, Z] [N] [P, Q, R] S [A, F, C] [etc.]//<sup>V</sup> indicates an actional sequence bounded externally by two perceptible shifts in velocity and containing a stance shift (S) of one of the other two variants (position or locomotion).

Only further research can reveal the functional nature of these internal shifts for the action sequence. As shall be seen below, the gross behavior noted on the microkinesic level as stance, contains behavior which, on the macrokinesic (i.e. probably parakinesic) level, emerges as posture, demeanor, pose, and presentation. It would be desirable to

have the evidence which would give assurance that all internal stance shifts can be ignored on the microkinesic level. However, for the time being such a conclusion must be postponed. In discussing the interpretation of kinesic systems more will be said about the incidence and relative placement of stance shifts and changes. The fact that communicants react unfavorably toward "too many" or "inappropriately" placed stance shifts suggests that these are especially patterned.

#### Interaction

While it is hardly the function of this section to develop a social-psychology of human interaction, the data to follow are perhaps illuminated by establishing exactly what it is we mean by interaction. Review of the existent literature on social animals gives us some security in making a generalization which states when social animals of a common species make sustained sensory contact with each other they must engage in behavior which identifies each to the other as a species member, a group member, and as being in a particular state of readiness. Ethologists and comparative psychologists have presented us with an impressive array of behavioral data which indicates that some term like "learned" or "conditioned" or "released" must be applied to this behavior. That is, behavior of identification is not only necessary for the adaptation of the species but is apparently patterned by the particular experiences of the group. This is hardly the place to review the evidence, but it seems clear

that a member of any social group must "recognize" and "emit" certain signals in order to sustain association with that group. The data are at the present time too sparse to indicate the range of discrimination of in-group and out-group identification signals.

The fact that animals engage in species, group, and state readiness signals does not give us the right to call this complex of identifying signals animal behavior with the implication that it is somehow instinctual. I prefer to call such behavior social, since it emerges from the patterned association of species members with patterned activity systems. The fact that we use the same term to cover an aspect of kinesic behavior does not make such behavior "more biological". It rather emphasizes its functional importance to the social system.

The term encounter will be used to cover that communicational situation which occurs prior to interaction. The duration of an encounter will depend upon the nature of the communication systems exhibited by the participants. An encounter becomes an interaction when the participants become communicants. That is, the participants interact rather than counteract when they find it possible to introduce cross-referencing signals into the scene in such a way as to sustain continuing adaptive association.

This difference between an encounter and an interaction is stressed because it so clearly sets boundary lines between those situations in which mutual cross-referencing signals are

appropriately used and those in which none yet exist or if they exist, are inappropriately used by the participants. It is probably evident to the reader that with this definition few encounters are ever of sufficient duration to be recordable. It may well be that "encounters" have no real existence and that "encounter" (or non-communication) refers to the subjective feelings of distress which we have when we do not comprehend the communication situation in which we are participating. There are cross-cultural situations in which an "encounter" becomes an interaction by the introduction of the cross-referencing signal that the participants shall search for cross-referencing signals. Such a sequence may be no more complex than the joint presentation of palms followed by squatting, buttocks resting on heels, with the weight balanced on the ball of the feet and the toes. This set of signals gives evidence of the willingness to participate in some sort of sustained interaction. This simple action sequence stands in sharp contrast to a scene in which participants may not engage in a sustained encounter because one or more of the participants use only internal cross-referencing signals and thus prevents the emergence of an interaction.

Kinesics and linguistics provide recording and analytic techniques which should give new insight into the processes of acculturation and group-formation. At the same time such situations should provide micro-cultural laboratories for sharpening the tools of these disciplines. In a seminar at



the University of Buffalo, the authors of this present book working closely with Smith and Trager, who had originally suggested comparable though by no means identical classification of similar vocal phenomena, developed a methodology for the analysis of visual recognition patterns. The skeletal structure of this aspect of communicational behavior is presented below with full recognition of its crudeness. However, even in this unrefined state, such abstraction provides a tool which has proved invaluable in the establishment of actor base-lines (see Chapter 7).

#### Visual Recognition Patterns

While we do not wish at this time to become involved in status and role theory, we must note that the broadest cross-referencing behavior in the communication system relates directly to these aspects of interaction. In the section above we discussed the function of stance shift (or change) in providing structural frames for extended stretches of kinemorphic constructions. Such cross-referencing behavior gives us data for recognizing that even on the kinemorphic level human beings do not communicate through an additive series of independent messages. In kinemorphics we were concerned with demonstrating that the system contains a variety of behavioral shapes which tie together least pieces of activity. We are now reversing our procedure to examine those cross-referencing signals which tie together the broadest possible amount of inter-actional behavior. Among such behavior, that **which** we call body base is, theoretically, sustained

throughout any interactional sequence:

Figure 4

#### Body-Base Types

Position

Sex

Age

State of Health

Body Build

Rhythm Phase

Territoriality

Mood

Toxic State\*

N-States

\* and organic confusional and deficit states.

This list of body-base types has been derived from a set of recognition behaviors some of which probably occur in all social groupings, animal or human. As we originally worked with these categories, it seemed to us that not only were these the broadest of the cross-referencing patterns, but also that they were somehow "closer" to the physiological base of the species. Certainly, with the exception of "position", which related to the order participation of a member of a group vis a vis his or her group associates, and "territoriality", which refers to systematic space occupation, all of these types seemed to have primary, physiologically constituted accompanying behavior. It seemed justifiable, therefore, to refer to these as "primitive"--somehow implying priority in an evolutionary sense. As I worked with these categories increasing confidence was gained that such states are characteristic of social groups--at least of mammalian groups--and probably of a number of fowl groupings. I had the opportunity to talk at some length with Dr. Konrad Lorenz who concurred in the tentative conclusion that these are probably requisite to sustaining the basic division of labor necessary for adaptation in the animal groupings. In the light of this, it may be suggested that if we are justified in calling these recognition states "primitive", it is with reference to the order of their appearance in social groups rather than in terms of anatomic characteristics.

The detailed description of the body-base types has been purposely avoided since these types are kinesic categories--not behaviorally specific constructs. Body-base

constitutes the basic image of other members of the social group which must be internalized by the group-member in the socialization process. Body-set constitutes behavioral derivatives from the expectancy pattern of an associated member against which are measured the body qualities or situationally variant signals basic to any interactional sequence. Body-base, then, constitutes the zero line which any communicant must have internalized in order to recognize the special cross-referencing message carried by the body-set signal-complex.

No member send or expresses any of the types as a unitary activity to the exclusion of others. Even the limited survey of films which I have attempted makes it clear that these types are neither specific (in an organic sense) nor independent from each other. In every case that we know anything about there is a complex relationship between the various types. Until further extensive cross-species research has been carried out we can only say that body-set is complexly patterned and learned. As we gain more knowledge cross-culturally, both about the patterning of these types and the predominant shapes of body-set in particular cultures and in particular individuals within the group, we shall be able to provide a more substantive base for cultural character and temperament studies.

In the discussion above, body-base was described as the patterned, learned zero-line against which body-set is measured. Body-set represents the particular cross-referencing

signals introduced in the particular interactional scene. A brief glance at Figure 5 below will make clear the relationship between the body-base zero-line and the body set which appears in the communicational sequence. It will be noted that for each of the body-base types we have derived a parallel body-set of social recognition value.

Figure 5

<u>Body-Base</u>	<u>Body-Set</u>
Position	Status
Sex	Gender
Age	Age Grade
State of Health	Health Image
Body Build	Body Image
Rhythm Phase	Rhythm Image
Territoriality	Territorial Status
Mood	Mode
Toxic State	Toxic Image
N-State	N-Status

Before discussing body-set states a word of caution must be introduced. While it is possible heuristically to abstract the ten states and to use these as frames for the collection of data, such abstracted units are never behaviorally isolated categories. Communication, intrinsic to culture, is patterned and systematic. As such it is constituted of a number of interacting sub-systems, the appearance of which is determined by the complex demands of the particular interaction situation. Since a particular (in space-time) cross-referencing system is shaped by the exigencies of a particular interaction system, it would be surprising if any specific state could be reacted to without modification by other state representations.

In the process of the establishment of actor and interactional baselines, I have found it necessary to analyze large stretches of behavior on a base-set model. In every case, at least five and at times all ten set-states categories are represented. The particular kinemorphs or kinemorphic constructions, the organization of stance shifts and postural positions, as well as the selected body motion qualifiers (to be discussed below), all combine to give us a cross-referencing statement of the quality of the interaction.

To avoid confusion, I have purposely avoided examples in the preceding discussion. Since this present chapter is not designed to provide the reader with data for the analysis

of particular sub-cultural situations but is aimed at orienting the reader to the Doris-Gregory-Billy interview, the discussion has been purposely general and theoretical. However, the reader may gain more perspective if a somewhat stereotyped example is presented at this point.

If we were to consider a situation in which a 35 year old junior Vice President talks to the 63 year old Chairman of the Board of his firm, we might find the following states manifested in the two participants in the interactional scene. These states cross-reference the discourse within the situational frame. While the problem of representation makes the diagram below appear like parallel or contiguous individual behavior, the reader is again reminded that the behavior of each is a function of reciprocation--the cross-referencing signal is a reciprocal, i.e. part of an interaction--not an individual expression.

Figure 6

(See Appendix 6 for the symbols used in the macrokinesic recording)

BODY-SET STATES

<u>Categories</u>	<u>Vice President's Behavior</u>	<u>Chairman's Behavior</u>
Status	¥ 4 K " p T p Hq or Hq -LL-alternating with L/L OO:::(intermittent)	4 ¥ b T b, AxbA Hn ... Hfbb, s00s
Age grade	+00+ Hqn Hq + N -LL- ~L/L	Set + 3
Gender or Sex Grade	¥ 4 K " [xx - xx]	4 ¥ + intermediate R 2 p
Health Status	p T p Qualifier!!! / +...+	Set + 3
Mode	+ 00 + - LL -	Hfbb, s00s AxbA
Body Image	— —	— —
Territorial Image	interaction centered	movement projection to whole room
Rhythm Phase Image	+ 00 +	s 00 s
Toxic Status	— —	— —
N-Status	— —	— —

This scene, which is purposely oversimplified by having its opening and closing phases eliminated, covers the length of a discourse, marked at its beginning by "settling in" behavior and terminated by interruption and departure behavior. We are concerned neither with the content of the scene nor with the linguistic or kinesic detail of these cross-referencing signals. The signals above are internally congruent--the young man's over-youthful, clear-eyed "sincerity", with appreciative humor, is consistent with the slit-eyed belly holding and genital scratching of the older man. The [Y4K] (knee over knee leg cross) is the congruent seating posture reciprocal for the older man's [4Y] (broken or open 4 leg cross). The seeming reversal here in which the older man uses a leg cross customarily seen in younger men is modified and tempered by his [AxbA] (bi-manual belly hold)--just as the [4YK] (knee over knee) which is formal and at the same time within the range of the feminine leg cross arc is tempered by the [+00+] (eyes with distal aspect crinkled) and the [Hq] (head cock).

These body-set cross-referencing signals may be seen as overall frames for the system of interaction occurring within their boundaries. The example above is simplified in that the interview portrayed contains no major shifts, that is, this complex of behavior extends throughout the interaction. Body-set signals are extremely important in assessing interaction topography; often the first signal of a parameter shift in an interaction is signalled by a set-shift--either in the linguistic or in the kinesic area, or in both. As our

understanding of the dynamics of interaction increases, it is clear that in body- and voice-set shifts we have one method for measuring "movement" in psychiatric as well as other interviews.

Returning to the example, it is to be noticed that under territorial image the younger man's behavior is described as "interaction centered", whereas the older is noted as "movement projection to whole room." The absence of macrokinesic recording here relates to my own uncertainty. However, at the moment it seems likely that space control has something to do with eye focus and convergence behavior modified by activity discussed below under motion qualifiers and motion markers. Of theoretical and methodological importance is the fact that while such behavior may be experimentally "located" in one rubric of the kinesic and parakinesic system, it may nonetheless be multi-functional.

Using the above as a background for recognizing the interdependence of quality behavior, the role of base and set as related to health and health image may be elaborated. Pathological conditions in the muscular, skeletal, and neurological system often directly emerge as limitations or specific underlying determinants of motor or dermal behavior. There can be little doubt that the form of the message sent or received is strongly influenced by the state of the organism qua organism. Probably the largest section of the bibliography concerned with visible body movement is related to the specific or generalized symptomatology of neurological disturbance. Theoretically, all specifically idiosyncratic body behavior lies

outside the field of kinesics, whether such behavior gains its peculiar cast from organic sources or from some special conditioning experience on the part of the actor or viewer. Yet it is essential to the methodology of kinesics as it is for linguistics, that the behavior of any participant in an interaction situation be described as idiosyncratic only after the patterned aspects of the behavior have been exhaustively described. That is, in the process of classification and testing, individuality is assigned after not before the fact of data exhaustion. Our theoretical framework provides us with an approach to the problems of allocating data to pre-kinesic or to macro-kinesic levels, but only when cross-cultural research provides us with clear indications of symptomatic activity concurrent with specific organic malfunction can we be secure in our assessment of particular pieces of behavior.

While anthropologists have long been aware of differing cultural emphases on disease or accident, the literature is exceedingly thin with regard to the specific variations in symptom presentation. Discussion of this problem with physicians whose practices are limited to the ethnic variations of an American city, has convinced me that practitioners are aware of the difficulties involved in treating symptoms expressed by various groups, as though there were a common and universal symptom structure for a given disease. This point was repeatedly stressed by M. D.'s whose practice included the range of variation provided by a Santa Fe or an Albuquerque hospital. Yet to my knowledge the data remains essentially impressionistic.

Perhaps as the World Health Organization expands its research area, specific and extensive attention will be given to the cross-cultural examination of the social structuring of symptoms. Such data as would be supplied by these studies--properly organized--should help us to be more explicit about the separation of pre-kinesic and kinesic behavior.

My own convictions in this area derive from experience gained while doing research on the social structure of two adjacent but differing sub-cultures in central Kentucky. Not only did the "BlueGrass" and "Hill" Kentuckians differ in their attitudes toward disease in general, but their choices of favorite ailments varied as systematically as did other aspects of their social organization. This research was done prior even to the preliminary systematization of kinesics, yet we were aware of the fact that there were styles of symptom presentation in both verbal and kinesic statements of illness which were sufficiently different in the two areas as to lead to misunderstanding between them. The discussion to follow is based on insights gained during this community research project, measured against the material gathered by a number of investigators in the cross-cultural sphere, and reinterpreted through the recent formalization of communication research.

Although Dry Ridge was only about fifteen miles into the hills from the Blue Grass community, Green Valley, the health set of this area is markedly different from that characteristic of the Valley. As a culture, more rigorously individualistic

and puritanical than Green Valley, signess was patterned in Dry Ridge into "non-reference to health" and "critically ill." Ideally, any variation between these two states is to be ignored or, at least, should remain a private matter. Ideally one is forced to go to a doctor, take medicine, or go to bed. The kinesic message that one is critically ill (although conscious and not yet bed-ridden) is best covered by the gestural reference, "stiff upper lip." This includes retraction of the scalp, tightening the skin of the forehead (with a significant reduction of brow markers), reduction of smiling, carrying the torso in hyper-erect, reduction of velocity in hand and arm movement, increased precision in gross movement (decreased overkick--anterior and posterior--while walking) and increased "foot-planting" (both feet--heel and ball--on floor while standing or sitting). If this does not elicit response from responsible kindred, this general quality is sporadically interrupted by "sag" behavior of about 2 to 5 seconds duration followed by "pulling together" behavior of about 2 to 4 seconds duration. The sag and pull-together should not take place very often or the quality shifts and the behavior is reacted to as malingering or as an infantile appeal. I have never, in over a year of watching this behavior, seen the sag and pull together used by males more than once in fifteen minutes except by the very young and the very old. Females, on the other hand, sag and pull together more frequently--several as often as two or three times in 5 minutes. This statement of variation is probably over-precise but there is quite obviously

a difference in expectancy here. A child, an old person, or a woman may engage in sag and pull together at greater frequency within a time span without being considered as malingering. It is perhaps unnecessary to stress the point that in Dry Ridge the full cross-referencing system is made up of "stiff upper lip" plus "sag and recover." It is perhaps of interest to note that the health image quality behavior of "stiff upper lip" differs from the mood image of anger in Dry Ridge in only two behavioral aspects that I have been able to trace. First, in eye convergence and focus--in anger the Dry Ridger avoids focussing on the eyes of others--looking to either side of other communicants, whereas, in sickness, he looks at his communicant with in and out of focus variation. Second, in aspiration presentation: in sickness he engages in intermittent pronounced chest presentation with audible aspiration (usually through the nose). Paralinguistically this is very close to a sigh. In anger, he uses deep, measured visually perceptible breathing which is usually inaudible.

In Green Valley the situation differs both linguistically and kinesically. A kith and kin community, health is used as a device for establishing interdependent interaction. Ill health is discussed and, in a manner of speaking, "enjoyed." A public affair, any manifestation of physical malaise occasions group diagnosis and comparison of symptoms. Accompanied by extensive verbalization, the kinesics of all communicants are characteristically directed with kinesic area markers. The etiquette of illness even in Green Valley (both of these

communities are, after all, American) demands that the viewer initiate verbal discussion of the actor's debility. Thus, the community member introduces a cross-referencing appeal which is sustained until it is responded to by other participants in an interactional scene. In Green Valley the kinesic illness behavior is characterized by first to third degree medial compression of the brows accompanied by first degree brow raise. The lids sag and there is tensing of the lateral aspects of the orbit plus upper cheek sag. The lips fill and the lower lip falls slightly away from the lower teeth. The neck is out-of-tonus often with a forward or forward and lateral thrust. The upper torso sags anteriorly as do the shoulders. Belly may be presented. Arms and hands may hang at the side or moved in over-slow velocity with lower arm performing any arc at greater velocity than do the hands. Feet drag while walking, or rest anteriorly on heels while sitting. There is, of course, variation in completeness or duration of this quality behavior--but it is my conviction that this variation is a function of the lack of response on the part of the other communicants rather than of the seriousness of the debility represented. This is supported by the fact that as soon as the malaise of the initiator is responded to, the body moves into tonus and a verbal recital of symptoms is accompanied by pointing--touching--rubbing--caressing of the ostensibly involved body parts. Even persons who are apparently (from doctor's diagnosis) quite ill become animated,

with eyes in focus--mouth at zero, and body at increased frequency of response during such conversations. Such activity is intermittently interrupted by "sag and recover", if the responses get "too" general in nature. I am somewhat unsure about this, but it is my feeling that malingering is suspected in this community when the "sick" person does not interrupt his or her performance with sympathy and empathy activity, when the traded symptoms are introduced by other participants in the conversation. An actor's preoccupation with his own health is a signal that his appeal is not simply a statement of illness.

These are neighboring systems and there is some intermarriage between the two groups. With this range of difference, it is easy to see that some misunderstanding arises in an intermarriage situation. It is perhaps of no consequence to this present chapter, but it is interesting to note that Dry Ridge, an economically poorer region than Green Valley, has produced four doctors since 1890 while Green Valley has produced but one.

Further discussion of body-base and body-set must await a more extensive presentation. These examples should serve, however, to illustrate the general propositions concerning the function of this aspect of the parakinesic system as a cross-referencing system. This discussion and these examples may be somewhat misleading for they do not properly underline



the point that while we are able to abstract some fairly precise movements as central indicators here, such behavior may congruently or incongruently be modified on the macro-kinesic level, which contains kinemorphic constructions, the constituent behavior of which may function on both levels of systematization. Further, our analysis must not omit what is probably the most critical (and least adequately analyzed) level of parakinesics. This area includes that behavior which I have termed the motion qualifiers, and the kinesic action and interaction modifiers. Although they in general refer to shorter stretches of behavior than do the base and set cross-referencing systems, these parakinesic qualifiers and modifiers may cover activity as limited as a kinemorph or a single kinemorphic construction or stretches of behavior of such duration as to make us feel that they may ultimately be relegated to the base-set level.

#### Motion Qualifiers

The stream of body-motion behavior has thus far been discussed as though there were a somewhat mechanical all-or-nothing quality to the production of the components of the kinesic system. The student analyst in his training tends to move from a period of concentration on the "expressive" or personality indicative, or idiosyncratic behavior to one of atomistic recording of the finite particles. It soon becomes evident that the range of variation in production of body-motion interaction is not a simple matter of idiosyncrasy or "style",

or, on the other hand, is it as highly patterned as is kinemorphic construction. Out of an extended range of production behavior three aspects of the motion qualifiers deserve special attention because their performance seems so intimately tied to the structure of the most complex arrangements of kinemorphic constructions. These include intensity behavior, durational behavior, and range behavior. For most middle-majority American movers these seem each to be distributed on a three to five-degree scale which is outlined below with the symbols I am presently employing for their notation.

Figure 7

Motion Qualifiers

<u>Intensity</u> (or degree of muscular tension or production of kine (or kinemorph)	Overtense	== ... ==
	Tense	— ... —
	N	
	Lax	o ... o
	Overlax	oo ... oo
<u>Duration</u> (or length) of kine (or kinemorph)	Stacatto	
	N	
	Allegro	
<u>Range</u> (or width) of movement in perform- ance of given kine (or kinemorph)	Narrow	‡
	Limited	‡
	N	
	Widened	‡
	Broad	‡

These motion qualifiers are roughly analogous to supra-segmental phenomena in language, that is, they may occur across or cover segments of a complex construction. They function to modify the kinesic meaning of the construction, but so far as I am able to ascertain, an increase or decrease of intensity, the rate of production, or the breadth of the

performance of any kine or kinemorph in a kinemorphic construction cannot serve as a substitute for one or more of the kines or kinemorphs in that construction. In other words, the modification function of one of the qualifiers, regardless of its extent of distribution within a bound form, seems to extend over the full kinemorphic construction. Or, to say it still differently, at least so far as our examination of American movers are concerned, there are no kinemorphs composed of variation in intensity, duration, and range.

If we try to evaluate these phenomena with relation to the present or allied research, the motion qualifiers take on special significance. While present research indicates that the 5 degrees of intensity and breadth and 3 degrees of duration have kinesic significance for all middle-majority American movers, the "distance" covered by a particular mover in the performance of the qualifiers will vary widely. This distance is of significance in the assessment of motion qualities. Further, the qualifiers seem to be especially related to that area of psychiatric symptom description called "flattened affect". Flattened affect in the kinesic behavior seems, at least in part, to be an incongruent narrowing of qualifier extent, the term incongruent, in this case, being related to the general or interactional system. Not altogether in jest we have been using another term "fattened affect" which occurs when the spread of qualifier extent becomes incongruent with the interactional sequence. This discussion of interpretation may seem somewhat out of keeping at this point in the

chapter but I think the reader must be warned with respect to a methodological point. The qualifiers may be looked at from two analytic points of view: first, as patterned modification phenomena which vary the kinesic meaning of a kinemorphic construction and, second, in their extent aspects as part of the general cross-referencing system of the full interaction. In analysis these must be kept separate since in their discovery quite different operations are involved. The particular qualifier behavior noted for a particular construction is of kinesic significance and is determined as variations of behavior within the base line of the actor. The quality aspect of the qualifier behavior is determined by comparative analysis and has interactional significance.

#### Action Signals

Since one of the purposes of this chapter is to serve as a progress report on the attempts at data exhaustion in kinesic raw materials it is perhaps justifiable to include in this already programmatic chapter a series of behavioral categories whose position and function are far from worked out. Something of a waste-basket category, the action signals include the action modifiers which are descriptive of an entire body in motion, the interaction modifiers which involve the full body behavior of two or more participants in an interactional scene, and the action markers. Perhaps the material at present handled under these headings will become data for the description of motion quality and/or

for the analysis of the base-line, but for the time being I am more comfortable in recording them under these less definitive categories.

The literature covering "expressional behavior" contains a number of sets of more or less descriptive categories of individual behavioral types. Many of these provide useful concepts based on careful observation and brilliant intuition. In the training situation, however, such borrowed concepts prove the adage that one can never get a borrowed bucket clean. Since we have attempted to make sure that each of the concepts utilized in kinesics and parakinesics relates both to a specific order of behavior and to the operations by which such behavior is abstracted, a new set of terms and categories is required. The following outline includes those modes of behavior which have been sufficiently examined to give us some confidence in their presentation. Such a systematization does little more than scratch the surface of possible categorizations. The nine modifiers listed below are what remain of forty-one paired types which I worked with in 1955. As systematic research proceeded, most of these were discarded as over-generalization of kinemorphic constructions. As it became clear that the "gesture" was a closely bound stem-like morph which signalled a constructional core, it also became evident that the classification of gesture types as indicators of cultural character tendencies must await systematic cross-cultural research. Further, the

development which followed the recognition of the cross-referencing function of the base-set activity further limited this list. I have no doubt but that this list will be lengthened and rearranged as research proceeds, but I present these categories as they now stand in the hope that other workers will find them useful. All of my testing indicates that they have some kind of communication function, but I am not at all sure how they fit into the remainder of the data.

Action Modifiers

The categories listed in outline below under the action modifiers include a series of paired types that cover the mode of behavior of the body as a whole. In all cases these are included because they elicit patterned responses from communicants and because they seem in "normal" movers to vary from situation to situation within the behavioral system of the particular member.

Action Modifiers

<u>Type</u>	<u>Behavior</u>
<u>Unilateral--Bilateral:</u>	Mover favors right or left side of body, contrasts with inclusion of both sides in performance (not just handedness).

Specific--Generalized<sup>1</sup>:

Mover tends to utilize one body area for major proportion of kinesic activities as contrasted to more extensive utilizations.

Rhythmic--Disrhythmic<sup>2</sup>:

Mover tends to adopt a definite rhythm within which he moves (often marked by kinemorphic or stance shift junctures) as contrasted to a clearly defined pattern of rhythm interruption (not just non-rhythmic).

Graceful--Awkward:

Mover tends to make major proportion of movements in a directed, minimally interrupted manner, as contrasted to a start-stop-proceed action with a series of abortive inclusions. (Grace is characterized by containing minimal "searching" behavior in contrast to awkwardness where searching is maximized.)

<sup>1</sup> There is probably a closely allied pair which covers "lost" or avoided body parts. This is not now included since cross-cultural research is needed to determine how idiosyncratic or set-quality patterned this is.

<sup>2</sup> See Interaction Modifiers, below.

Fast--Slow:

(Not to be confused with the duration qualifier.) Mover tends to high velocity of production of kinemorph and kinemorphic constructions as contrasted to a low production rate.

Integrated--Fragmented:

Integrated mover tends toward harmonic organization of various body parts (whether generalized or specific) whereas fragmented mover may divide body into non-harmonic--even apparently contradictory parts. A finger, a hand, or an eye may seem to have existence independent of remainder of body activity. May involve the full division of the body into two spheres as: above and below pelvic girdle or (in one case) right through the middle of the body leaving a right and left sphere.

Intertensive--Intratensive:

Intertensive mover tends to be highly responsive to behavior of other communicants--engages in consistent check and modification behavior as contrasted to

the intratensive mover who appears to engage in extended auto-stimulation but with minimal apparent strenuous rejection. At first this seemed to be an aspect of the encounter-interaction process but, as research continued, it became clear that this behavior continued even after an interaction was clearly in progress. As in the case of the "self-possessed/self-centered" type which follows, this type probably has special significance for clinical observation.

Self-possessed--Self-contained:

A dubious category (see discussion). These types are easy to recognize once seen but are difficult to objectify. I suspect that this is a complex category and perhaps should not be included in this list of modifiers. However,

The self-possessed mover is characterized by a reduction of qualifier width without incongruence, by the harmonic organization of the body parts, by minimal searching behavior and by what might be loosely characterized as "poise." Only the fact that self-possession seems to appear intermittently within or beyond and apparently quite

this category is so useful in the analysis of psychiatric interview material that it is included.

independent of the qualities persuades me that this is a category of another order than quality. Self-possession appears to relate to social "ease" and "confidence" in interaction (neither of which terms have more than impressionistic value in this presentation). Our description of self-containment is equally impressionistic, characterized by seeming intra-tension; the general feeling is one of restraint and "avoidance" of stimuli. Category by category the behavior is congruent, but it is best characterized as systematically resistant to any change in the interaction beyond narrowly established limits.

Only extended research can establish a clear perspective on this pair of types. The difficulty may lie in the pairing which I have utilized in the modifier assignment. Self-possession may be a special complex more adequately described under ~~the motion~~ qualities, while self-containment may be a special pathological condition paired with another poorly defined pattern that I have been calling "identity loss." Identity loss has been characterized by a high incidence of "echo" behavior or of pieces of behavior that have no apparent relevance to the interaction situation. If self-containment is characterized by exclusion or avoidance of stimuli, identity loss seems to be made up of over-reactivity to them.

#### Interaction Modifiers

While, by definition, kinesic research is only concerned with body motion behavior with a demonstrable communicative function (and this implies an interactional frame), the action modifiers were concerned with the behavior of a given actor in an interactional context. The interaction modifiers are concerned with the classification of comparable behavior which appears in a sequence involving two or more actors. In the outline below is presented a series of three paired types of interaction modifiers.

#### Interaction Modifiers

##### Mirror-Parallel:

Mirror behavior is characterized by one or more actors acting in mirror image of a central actor. Parallel behavior occurs when two or more actors move in parallel.

It is recognized that when more than two actors are involved, some by limited possibility are in parallel, others in mirror, interaction. Our very limited observation of group interaction has not revealed any particular patterning to this variation. Perhaps when kinesic observation is combined with the linguistic and studied in association with devices like Chapple<sup>1</sup>'s chronograph this material will have more consequence.

##### Rhythmic-Disrhythmic:

When the interactional behavior of two or more actors contains a clearly perceptible beat, introduced either in parallel, or in series such interaction is termed rhythmic. Disrhythmic interaction occurs when established rhythms are repeatedly interrupted.

##### Open-Closed:

An interaction is termed open when the behavior is characterized by the searching of the environment for other stimuli. To the extent that the participants are so highly interactive that they do not respond appropriately to other stimuli in the milieu, the interaction is closed.

"Searching" as used here refers to focusing the eyes or ears, or other sensory receptors on objects or people outside of the interaction area, "squirming" (non-congruent shifts in stance), foot shuffling, finger drumming, etc.

Systematic research has thus far been directed almost exclusively to the examination of two- and three-person interactions. Even within this limited universe, there are a number of other interaction modifiers which are being examined. Their behavioral limits are not yet clear, however, and discussion of them should await further analysis. Needless to say, the interaction modifiers appear both in association with speech behavior and through periods of silence.

---

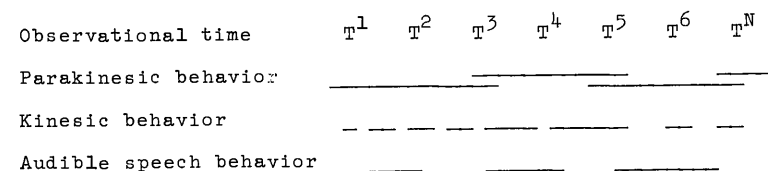
<sup>1</sup> Chapple, 19

### Motion Markers

The discussion of that aspect of body motion behavior which is classifiable only in direct association with verbal behavior has been saved until the remainder of the material had been presented. This chapter is not the place to discuss the intimate relationship between speech and movement in the communicational process. When we turn to examine the role of communication in society, these systems are dynamically interrelated with each other and with the other communicational systems as well. Up to this point, with few exceptions, body movement has been treated as a universe different from that of speech behavior. This methodological separation has been more than a convenient rendering unto Caesar. The internal consistency of language has been revealed by systematic research based on the proposition that linguistic phenomena are organized into a system which can and should be examined without reference to other social systems. This rigorous abstraction provided both a model for kinesic research and a set of clear frontiers which facilitated the abstraction of kinesic material. This entire discussion has rested upon the proposition that every interaction is based upon continuous communication carried on through the medium of patterned, discrete but interlocking and cross-referencing symbols. Looking only at the two modalities, speech and body movement,

but inspecting them from the point of view of the kinesicist, a model might be constructed to illustrate the temporal aspects of this process.

Figure 8



"Gestures" and "posture" and "facial expression" are probably the body-motion events most accessible to the American "folk"-viewer. That is, these phenomena represent public abstractions or shorthands for the much more complex behavior described in the pages above. As such they may be included in literary description, stage instructions, and even in etiquette prescriptions. As our discussion above has demonstrated, these are derived systems and are to be finally analyzed only in the complexity of the full communicational process. The motion markers, while less public in the sense that only a portion are sufficiently abstracted to be taught, seem very close to awareness in American speakers and movers. At least, an American audience seems to have relatively little difficulty in seeing them and "explaining" their function, once they are demonstrated. Yet, as with "gesture," "posture," and "facial expression" their apparent

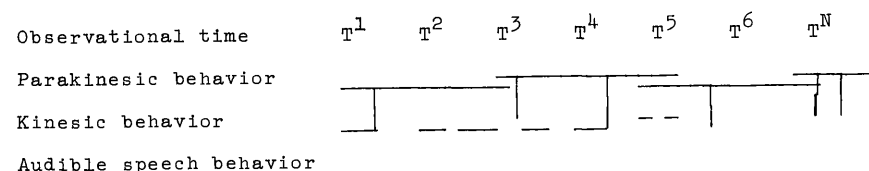


accessibility creates confusion and pseudo-understanding when we attempt to analyze them. Special attention is given here to the motion markers, because of their tremendous importance in measuring the congruity of the linguistic and kinesic systems and because, in the interview situation, they are immediately available to the observer.

The kinesicist, recording, let us say, from a muted sound film of hitherto unanalyzed material, records a stream of kines. As his analysis proceeds, he orders these into kinemorphs and kinemorphic constructions. As he enlarges his procedure to include the qualifier behavior, he develops a multi-level record which is internally consistent. Yet, as he scores this record, he can detect a particular distribution of kines as kinemorphs, and he can observe narrowly limited stretches of qualifier shift which combine to punctuate certain portions of his data sheet. Upon turning up the sound, it becomes clear that these specially marked passages are very frequently co-extant with speech phenomena, although this is not always true, for under a variety of circumstances a communicant may verbalize sub-audibly, so that his speech behavior is visible rather than audible. At this stage in the research, the kinesicist's abstraction of such phenomena will provide him with a data-series which resists systematization except in so far as it constitutes discernible patterned movement that occurs in association with speech behavior. Such body-motion behavior tends to have a different shape if the

the mover is speaking. Auditor behavior often includes the same order of punctuating events. The model shown in Figure 8 may now be expanded to:

Figure 9

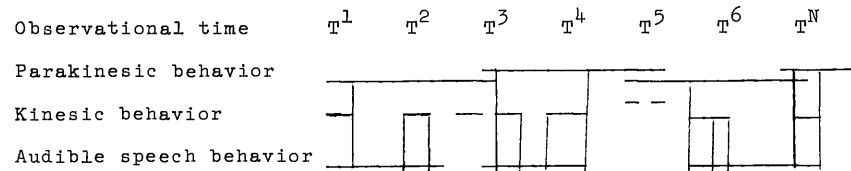


Until some of the linguistic and paralinguistic analysis is completed, however, we have no way of explaining this evident intersection of the linguistic and kinesic systems. When we turn to protocols which include both the linguistic and kinesic material, it is possible to abstract a series of linguistic situations which seem to demand a particular kinesic accompaniment. A more explicit description of some of the punctuation behavior is probably called for, therefore, before we proceed with the analysis. Utilizing our abstracted figure of kinemorphic construction, a record may read:

// [Ẋ Ẏ Ż] [N] [O Ṗ Q̇] / / [L Ṁ Ṅ] [O] [Ṗ Q̇ Ṙ] //

In this example the kine Y, in the kinemorph [XYZ], stands for brow rise which is held for one degree of over-long as compared to X and Z. Q, which may stand for lip pursing, is

Figure 9



comparably over-long in the [O P Q] kinemorph. In contrast N in the [L M N] kinemorph, here standing for head nod, is overshoot, and [P Q R], here a mid-face kinemorph, is marked by first degree of overlaxness. Since nothing never happens, this variation must be accounted for in the process of data exhaustion. The kines (as kinemorphs) [N] and [O] cannot from this record be abstracted as potential punctuation. However, when we match this record with a record of the speech events, we may well discover that either or both have a punctuational function as well as an observable bound place in the kinemorphic construction.

Theoretically, it should be possible to analyze this punctuational behavior without recourse to the linguistic or paralinguistic behavior which accompanies it. At the present time, however, our knowledge limits us to the conjecture that these events will eventually be orderable into some kind of suprasegmental form, analyzable in purely kinesic or parakinesic terms. Certainly the events do have a certain regularity of occurrence and of individual shape. Until either or both a binding or a linking kinesic principle is detected in their operation, they must be considered punctuational forms to be classed, since they are abstracted both behaviorally and functionally, as motion markers.

The motion markers, behaviorally, seem to fall into two general types: those constructed from qualifier variation, and those composed of kines-as-kinemorphs and of "gestures" as bound kinemorphs in a kinemorphic construction. Either type

may appear in the behavior of a speaker prior to and at the cessation of phonation--but not at the beginning or end of all phonation. They also appear in conjunction with special internal arrangements of complex sentences, but a speaker may very well emit certain complex sentences without punctuating them with markers. Similarly an auditor may or may not modify his speech-related behavior with motion events of the order of markers. In other words, while the punctuational behavior can be located in the speech context in certain positions, the analysis has not yet reached a point where we can posit obligatory binding between linguistic and kinesic events. With this caveat, we may list a series of derived functions that markers play in the interaction sequence. By "derived function" I mean an observable set of behaviors in a given context which can be abstracted and interpreted as related. Since my confidence in such interpretations is, at the moment, relatively low, I prefer to use "derived function" rather than some kind of "meaning."

(See Appendix A, 1961)

Figure 10

Motion Markers

<u>Types by derived function</u>	<u>Punctuation behavior</u>	<u>Motion marker</u>
I CUE:		
A- Signalling anticipations of interruption	Examples would include hand, foot, and head	(Q)
B- Signalling anticipated termination of phonation	nods, raised eyebrows, stance shifts, lid closure and duration to 2nd degree, sustained incomplete kinemorphs, palm presentation, pursed lips,	
C- Signalling anticipated initiation of phonation	visible breathing, eye focus shifts.	
D- Signalling "proceed, I'm listening."		
E- Signalling "completed phonation."		
II SELECTION:		
A- Selected item in series of items	Examples would include qualifier shift, head nod, head sweep or arc,	(S)
B- Selected connection between items in series	special lip protrusion or retroflexion, torso nod, hand nod, foot nod, digit nod, brow nod.	
C- Selection of certain items as related to other items		
III DURATION:		
A- Increase	Duration qualifier shift to staccato or allegro;	(D)
B- Decrease	lateral sweeps of hands, feet; eyeball sweep.	

IV AREA:

A- Nearby locale	Range of "gesture" includ-	(A)
B- Distant locale	ing "pointing," with head,	
C- Traversing distance	hands, feet, torso, hand	
	sweeps, head sweeps, etc.	
	(always encased in construc-	
	tion).	

V PRONOMINAL REFERENCE:

A- Speaker	Same as in IV above except	(P)
B- Auditor	that pointing is directed	
C- "We"	toward subject with suppor-	
D- "They"	tive construct.	
E- "It"		

These five markers (Q), (S), (D), (A), and (P) represent contextual appearance of a wide variety of punctuation behavior. Assignment of marker status to any particular punctuation thus represents an abstraction from context. Only extended contrast research in interaction situations can strengthen our confidence in the organization of the marker categories. For the present, the five point system represents a tentative working base which has proved useful in the examination of interview material. A sample recording of kinesic marking following extensive analysis might read:

(Q) (P) (S) (P) (A)  
I told John, Mary, and Bill to put it in the back part  
(S) (Q)  
of the big, red barn.

The reader will note that the markers here are added to the simple English orthography. As will be clear in later chapters, the position of the marker may be seen to have even more significance when the full linguistic-kinesic protocol is assembled for the assignment of symptomatic and diagnostic features.

### Interpretation

At this stage in the development of kinesics, interpretation must always rest upon the adequate measurement of the context of an occurrence. Throughout the preceding discussion I have stressed the fact that no kinesic event, whatever the size or the shape, is a carrier or invariable stimulus with its own emergent causal component. From the point of view taken within this discussion, no kinesic form is a vehicle with a constant load, no kinesic event, an encapsulator of meaning. I have tried to make it clear that the question, "What does X mean?" is non-admissible unless the system within which X operates has been subjected to sufficient analysis so that X in its multiple of transforms can be described. However, to reject the over-simple question is not to repudiate the responsibility for weighing the role of the event within the system. Perhaps a summary of certain aspects of our discussion will make this position less ambiguous.

When we have repetitively isolated the forms A, B, and C, established within the preliminary descriptive frame of the investigator, as least discriminable variations from an established zero point, we can say that for the investigator the meaning of A is not that of B is not that of C. This procedure provides us with units whose value for the subject is yet to be determined. If repetitive manipulation of the forms A, B, C demonstrates that, for the subject, they are in fact

not substitutable for each other in all frames, it is then necessary to describe them as having (for the subject) discriminational meaning.

If, on the other hand, we establish the fact that the preliminary units which the investigator distinguishes in a given position (e. g.  $A_1, A_2, a_1, a_2$ ) are substitutable for each other in that context without (for the subject) varying the function of the form, we may then say that these units are in this context in free variation and have, for the subject, the same perceptual value. That is, the kine variants, while having for the investigator discriminably different values, have for the subject identical perceptual value. They belong to the same class of events and they derive their meaning from their class membership. Thus  $A^1, A^2, a^1, a^2$  have a single perceptual meaning A. Yet we have said nothing as to the meaning of A as a kine. We can, however, discuss the structural value of A when we systematically examine the kinemorph and kinemorphic-construction bound-forms which contain A. Again we are not saying what A in and of itself means. What we are saying is that A will occur in certain kinesic contexts. While our analysis has not yet gone this far, in the future we may very well be able to list those kinds of constructions in which A does not appear. There is also the possibility that we will discover a systematic nature to kine positioning which will allow us to perform the complementary distribution analysis so characteristic of linguistic analytic procedures.

On the next level of analysis we can determine the relationship between certain groupings of kines and their complex associations under some kind of suprasegmental binding system. Through analysis we can determine that certain of these bound forms will exist in association with other bound forms under some kind of cross-referencing system which serves to distinguish one complex bound series of movements from another comparable but differing cross-referencing series.

But the question still remains, once these forms have been distinguished, ordered, and conceptualized in their complex organization, how do we then determine their significance in the interactional sequences in which they appear? Throughout the sequences discussed above, our procedure has been dominated by a series of methodological canons. (A) Establish and maintain a given level of analysis. (B) Isolate units for manipulation. (C) Establish the independent identity of these by contrast analysis. (D) Weigh the analytic value of these newly established units by the examination of the contexts in which they regularly appear or never appear.

These same canons prevail in the analysis of the social meaning of any form or series of forms. The social meaning of a form is established by the description of the shift in field or context occasioned by the presence or absence of a given complete form. However, let us reemphasize one point. This procedure cannot be accomplished before the full analysis of the form--which includes the assignment of the form to its

role within a pattern<sup>?</sup> has been carried out. We cannot simply count the forms present and derive the special meaning of the forms. Unless we carefully separate our levels of analysis, we shall be unable to deal with those patterned arrangements in which the value of a pattern is shifted by the absence of a component which is normally internally bound. In other words, no running list of kine variants will ever inform us as to the role of the kine in the interactional sequence. Only in a pattern, composed of complex bound forms, does the form enter into associations on the social interactional level.

Because we are dealing with a patterned system, our analysis, once completed, serves to make public incongruities which appear within the system at any given level. The statement that the behavior which we are analyzing at any given level contains incongruities, however, does not permit us to assume that these incongruities will introduce incongruities into the social interactional sequence. One of the most important functions of parakinesic activity is that of introducing cross-referencing signals that indicate that what appears to be an incongruity is congruous within some larger system. Such statements as "Everything to follow (or everything just said) is a joke," or "I am imitating" or "to quote so and so" or "this is play" provide us with examples which can be kinesically rendered in such a manner that apparently incongruous statements are cross-referenced into congruity. As we shall see in later discussion, this is the very area in which personally

distorted systems become mal-operative. Only systematic research with contrast analysis in multiple contexts will permit us to evaluate the particular incongruity. Within our basic assumption that "nothing never happens", the incongruity is itself a message if it remains uncompensated for within the larger system. Its interpretation, again, will rest upon its repetitive contextual appearance.

This is not the place to attempt to analyze the relationship between the two communicational systems, language and body motion. Suffice it to say that we already have considerable evidence that these systems cross-reference each other and establish full patterns of conversational performance which operate in the social interactional sequence. Man does not merely move and see movement, or talk and hear, in an interaction. Body motion and language, on this level, form a complex pattern in which they are only analytically separable. The full pattern must be assessed before we can hope to weigh the role of either within the interactional sequence.

Finally, even the most exhaustively analyzed conversational pattern does not exhaust the systems in operation in any sustaining association. That is, communication analysis as discussed here does not constitute a final analysis of culture or its component situations. The final answers to "What does X mean" can only be arrived at when all of the other social systems interacting in any situation are equally thoroughly analyzed.

#### CHAPTER 4

1968

#### Implications for Psychiatry

Sept. 30, 1968  
Henry W. Brosin, M.D.

TABLE OF CONTENTS

<u>Page</u>	
1	4.1 -- The Task of Clinical Recording and Analysis of Non-Lexical Behavior
13	4.2 -- Psychiatry's Interest in Non-Lexical Behavior
33	4.3 --Convergence on the Problem from Parallel Disciplines
42	4.4 --Need for a New Technical Language
47	4.5 --The Natural-History and the Experimental Method
63	4.6 --Advantages of More Precise Observation and Analysis and the Training Value of the New Techniques
74	4.7 --Possible Disadvantages of the New Methods of Recording, Transcription, and Analysis

"The history of science is largely a history of technique..."

Topley and Wilson, 1946

"The successful mapping of large molecules has needed courage and persistence as well as imagination, but the way to further maps will be made plainer... Yet, in spite of that, the biologist has little hope of escaping the immense variety of his material. Even when the molecular structures are mapped there will be organisms to deal with, the vast range of plant and animal form and in the end the differences between one man and another."

Lord Adrian, 1955

or--

"There are two laws discreet,

Not reconciled,--

Law for man, and law for thing;

The last builds town and fleet,



But it runs wild,

And doth the man unking."

Emerson

"For this reason no man of intelligence will venture to express his philosophical views in language, especially not in language that is unchangeable, which is true of that which is set down in written characters."

Plato

#### 4.1 The Task of Clinical Recording and Analysis of Non-Lexical Behavior

A physician along with other students of human behavior may well ask, after reading the exceedingly technical discussion of the preceding chapters, what advantages appear for him in the linguistic-kinesic analysis and synthesis of raw clinical data. Is such elegant technicality worthwhile? Is the already overburdened physician being asked to master a new area in addition to those being described in the more than 165,000 articles and 85,000 book titles published in the biomedical field last year? Our exposition in this chapter will focus on the interests of the physician-psychiatrist in order to meet space limits, although we are keenly aware of the contributions and relevance of other behavioral scientists.

Psychiatry has been defined as the study of the disorders of language and movements of the body. Freud (1901, 1905) pointed out that human communication included speech, gestures, dreams and writing; "that the interpretations made by psychoanalysis are first and foremost translations from an alien method of expression into one which is familiar to us." (1913, p. 170)

... "The language of dreams may be looked upon as the method by which unconscious mental activity expresses itself. But the unconscious speaks more than one dialect. ... While the gesture-language of hysteria agrees on the whole with the picture-language of dreams and visions, etc., the thought language of obsessional neurosis and of the paraphrenias (dementia praecox and paranoia) exhibits special idiomatic peculiarities which, in a number of instances, we have been able to understand and inter-relate. For instance, what a hysteric expresses by vomiting an obsessional will express by painstaking protective measures against infection, while a paraphrenic will be led to complaints and suspicions that he is being poisoned. These are all of them different representations of the patient's wish to become pregnant which have been repressed into the unconscious, or of his defensive reaction against that wish". (1913, p. 177-178)

Symptoms and neurotic behavior are seen to be meaningful, although disguised

communication, and that the therapist's efforts to understand and help the patient are based upon his ability to understand the ways in which forgotten or repressed feelings, conflicts, and memories are transformed in body language and behavior which are the focus of his complaint. Because the task of the therapist is the translation of these disguised patterns into conventional language, it seemed to us that stable recordings which permit re-examination by multiple judges over time would be an aid in systematic study, both for exploration, verification and refinement of models which seemed appropriate to workers in the clinic. Because there are a large number of relevant clinical observations, and also more recent laboratory investigations from workers in many disciplines to be reported which are useful to serious students, but not of primary interest in an introductory statement such as this, it was decided to place these examples, together with some historical comments into Appendix 0000. (Insert proper number for Appendix b - "Bibliographic Citations" as outlined in McQuown Table of Contents - 1968). Here we will keep to the modest goal of describing a method for improved

recording of out patient's behavior in the hope that in some organic disorders, it will help sharpen the focus of these disorders and thereby help us understand better the mechanisms present, and that in behavioral problems, whether genetically, organically or psychologically based, we can better comprehend the structure and grammar of the communication systems at work. Many, if not most, of our problems concern the human relations of the patient to the people and events surrounding him. As scientists, we have seen in all of the related fields, elaborate searching devices developed to examine apparently simple patterns. The value of these or any other techniques must ultimately be determined in terms of the usefulness of the instrument--not in terms of how complex, elaborate, or difficult to master they are. We are aware that the intensive study of human communication is of ultimate concern to all behavioral sciences, but will not try to go beyond immediate clinical psychiatric implications in this preliminary statement.

Experience teaches us that simple memory is too unreliable to permit the objective evaluation of the subtleties of behavior available in any interaction. Such study requires a clean and transmittable recording

and storage system. Recent technological advances in sound and film recording seem to provide some answers to these problems of storage and recall of social circumstances.

Existing methods for annotating, analyzing and interpreting the social interaction of any species, i. e., the transactions and the interactions between the individuals of a species, however, have remained inadequate. It seems to us that one reason for this is that a considerable proportion of the significant signals in any human interactional system have not been systematically recorded for analytic purposes. All of us can appreciate the theoretical importance of inadequate recording of signals at various levels of interaction. And even though we can legitimately expect extensive improvement in understanding with improved technique, we are also aware that recording and analysis will probably never be complete, or free from technical difficulties. We recognize that the more minute gradations of sound and movement escape even the sound-camera and the most skilled observer. We also remember that smell, color, temperature, skin resistance, pulse and respiratory rate, pupillary size changes, and

other visceral indicators are not being recorded for analysis in our studies. The recording of these modalities is also in an early stage of development (Watson and Kanter, 1956). The extensive possibilities for explaining tactile communication, even though recognized clinically in both psychotic children and adults, have not been adequately developed (Frank, 1957; Geldard, 1960; Wiener, 1966). The need for such records is evident. They provide an invaluable tool for the furtherance of investigations involving interviews. However, let us be clear. Because of the very nature of human experience it is doubtful whether even the most complete records will ever be an adequate equivalent for a direct personal experience with a patient.

There are other immediate increments from the use of these tools. Within the past few years physicians have had an opportunity to hear themselves in action. The revealing, but at times uncomfortable insights derived from this experience have been multiplied in the case of those who have had the much less common opportunity of hearing and seeing themselves on sound film. After the initial shock of recognition and

incredulity, most clinicians are surprised at the difference between their image of themselves, their transactions with the patient and that revealed by the record. It is not surprising that these distortions should occur. The physician, as a member of his society, shares its image-making system. It is not entirely reassuring to remind ourselves that the distortions evident in even the most sophisticated self-image of a clinician can be both customary and functional. It is important, however, to recognize that within the past five decades, with Freud's emphasis upon the countertransference relations, there has been a significant shift in the order of magnitude of self-awareness on the part of the physician.

The shift in the search for self-awareness presents new responsibilities. The previous chapters have laid the groundwork for the explication of this theme. We may go directly to consideration of the relationship between methods for achieving social awareness and clinical practice. The work done thus far by the use of film and sound-tape in clinical situations and the systematization of interpretations by clinicians practiced on masses of raw data are notable for pointing the way to more widespread usefulness.

But we must not be too sanguine. The high hopes of 1945 are not justified by the experimental work up to 1968, but the rate of increase in the number and quality of publications in this area since 1960 furnishes new hope.

At present, many clinics have mausoleum-like closets piled high with tapes and films which defy their possessors to abstract from them significant data which can be ordered into patterns affording interpretations which are intelligible and useful. The publications which report the work which does attempt to surmount the many difficulties inherent in the "documentary" approach merit our deepest respect and challenge us to more conscientious exploitation of such sources of clinical insight.

During 1953-1960 we might mention the work of Will and Cohen (1953), Gill, Newman, and Redlich (1954), Rogers and Dymond (1954), Wolberg (1954), Deutsch and Murphy (1955), Deutsch (1963), Deutsch (1966), Kahn and Cannell (1957), Hoch and Zubin (1958), and Pittenger, Hockett, and Danehy (1960), all dealing with the psychotherapeutic interview in one way or another.

The mass of data and the valuable hypotheses which have emerged from the study of the behavior of patients by physicians and their colleagues in the behavioral sciences in clinics, hospitals and in private practice is

impressive. Much good work continues to be done in these areas, utilizing the methods of observation and correlation, group comparison and the tracing of genetic development, with the clinician as a self-conscious participant-observer. Yet the shortcomings of these methods have been increasingly obvious to the investigators themselves. Incompletely recalled and recorded interviews provide insufficiently reliable data for intensive review by other colleagues or even by the clinician himself. Without a manageable system for transcribing significant data, based on a comprehensive theory, it is not certain how successful our analyses can ever be.

The physician is clearly subject to the errors of his own perceptions and interpretations, however skillful, talented and industrious he may be. The problems of the observer, i. e., of reviewing his subjective recollections and interpretations during and after an interview, are well recognized. This process has not been sufficiently studied to be satisfactorily broken down into its numerous component parts Renneker (1960); Lennard and Bernstein, (1960); Thorne, (1961); Rosenthal (1966); Moscovici, (1967); Berkowitz (1967); Chassen (1967); Friedman (1967). One of the rewards of the study of film-

tape interviews is the larger perspective gained on the ways in which the participant-observers, consciously or unconsciously, alter the course of the interview or the climate of communication. It has long been clear that the lexical record, i.e., the words in the typescript, is in all cases made peculiarly significant by the vocal modifiers and body motion frame in which the words occur. These have not been adequately covered in previous attempts to comprehend clinical material. The new technology provides us with stored records which hopefully can be mined with new and explicit techniques. We must say, however, at the outset, that we will not know how to use these methods most effectively until they have been given an adequate trial under favorable experimental conditions. One problem is the choice of the size and shape of a sample of film and tape, while another is the nature and magnitude (micro-, semi-micro, or gross) of the analyses which are most appropriate to the experimental task. A priori judgment based upon insufficient evidence will not demonstrate either the virtues or the faults of the methods concerned. Unfortunately, the methods for linguistic-kinesic analysis are costly of time, of energy and of talent. It will be some years before trained workers in adequate

numbers are available. For this reason, the present experiment was undertaken to test the usefulness and the limitations of such analysis in a single selected situation .

The use of the oscilloscope would seem an excellent way to study speech sounds with stable recording of patterns. Initial efforts to utilize this method have uncovered many difficulties in making these methods useful in studies of personality, social interaction and the psychotherapeutic process, but the hope remains that technical advances may give us a new tool. Ostwald (1963, 1964, 1966) has furnished excellent examples of acoustic manifestations of emotional disturbances, including the demonstration of speech disturbances as seen in a spectrographic study of a schizophrenic adolescent. (Ostwald, 1966, p. 40-49)

Our emphasis in this brief summary on problems of clinical recording and analysis of non-lexical behavior should not obscure the steady progress made by clinicians in the firsthand study of such components of communication, namely, the physiological, vocal and body motion activities which characterize the interactions between people.

Psychiatrists trained to observe these non-lexical activities will find much that is familiar in the case material presented here, but we hope that there is sufficient novelty in the data and in the concepts to encourage them to make use of these new techniques.

#### 4.2 Psychiatry's Interest in Non-Lexical Behavior

The novelty will be more evident as we review some of the relevant work in the field. To do justice to the subject, one should begin with Charles Darwin's classic study, THE EXPRESSION OF THE EMOTIONS IN MAN AND ANIMAL, originally published in 1872, before reviewing current work.<sup>1</sup> However, for the purposes of this chapter it seems justifiable to place our emphasis immediately in the more recent psychiatric context. Psychiatrists basic interest in the observation and recording of patients' thinking, feeling and acting, akin to the medical model, was enhanced with the writing of Freud and his pupils. In order to furnish easily available samples to students, these will be furnished in Appendix 000(7b) in order not to delay this narrative.

---

<sup>1</sup>For general orientation, the reader is referred to the standard textbooks and histories of psychiatry, and to a few among many special treatises as those of Zipf (1949), G. A. Miller (1951), Ruesch and Bateson (1951), Meerloo (1952), C. Morris (1955) Lamb and Mahl (1956), Ruesch and Kees (1956), Cherry (1957), Ruesch (1957), Hoch and Zubin (1958), Strupp and Luborsky (1962), Knapp (1963), Sebeok, Hays, Bateson (1964), Hymes (1964), Rioch and Weinstein (1964), Osgood and Sebeok (1965), Gottschalk and Auerbach (1966), Smith, A. G., (1966), Berkowitz (1967), Hildum (1967), Shlien (1968). The annual reviews of anthropology (linguistics) psychology, psychiatry and psychoanalysis report current work.

It is difficult to write a succinct account of the development of recording and filming devices and methods because there is relatively little publication of the early attempts even though we know from direct experience that many were made prior to World War II. Gill, Newman and Redlich (1954) have described previous recordings from 1947 to 1953. Their statement that Earl Zinn is probably the first to record therapeutic interviews is supported by his pupil, H. T. Carmichael (1956, 1966) who was with Zinn in the early 1930's. They provided a valuable discussion of the instruments and methods employed, and set an admirable example in specifying the details of equipment, and how the material was actually used. (Also see Dittman, Stein and Shakow, 1966). Many of the earlier papers fail to do so. Also, many of the earlier workers only used recordings as the basis for a transcript until later experiments by Bierer and Ström-Olsen in 1948 and by Ruesch and Prestwood in 1949 demonstrated the superiority of studying the linguistic components of the communication stream. The writings of H. H. Strupp and his associates together with the references to other active workers during the past two decades already

cited will provide the reader with descriptions of the progress in experimental technique made by himself and others. (Strupp and Luborsky, 1962, and Strupp, 1966)

Our studies began with McQuown's linguistic analysis in 1955 of the Will and Cohen (1953) recording which was published in 1957.

The earliest linguistic transcription of psychiatric material known to us is that of Stanley S. Newman and V. G. Mather (1938) who worked on the linguistic properties of patients with affective disorders and found significant patterns of constrictions. Kramer (1963) has given us an excellent review article dealing with the non-lexical properties of speech up to 1962.<sup>2</sup>

---

2

Robert E. Miller (1966) has reported dramatic experiments about the transmission of affect in monkeys. Haggard and Isaacs (1966) have carefully studied fleeting facial changes lasting one-eighth to one-fifth of a second, and observer reliability, and thereby generated several new hypotheses worth testing in a controversial field. Haggard and Isaacs (1967) is a worthwhile summary of recent and current work from this active group. Condon and Ogston (1967-68), using film of a chimpanzee found an organization of vocalization and body motion similar to that of human behavior.



Several excellent demonstration films of psychiatric patients with presentations by Heinz Lehman were made in 1938 by the Mental Health Division, Department of National Health and Welfare and the National Film Board of Canada.

Leighton and Lidz (1942) together with John P. Lambert made silent films of patients for demonstration purposes in syndromes not readily available such as paretic and manic reactions. Similar films were made by Dr. Franklin G. Ebaugh, Charles Rymer and their colleagues, between 1931-33. The recording devices before 1949 were not reliable and most workers, under the pressure of other duties and projects did not pursue this means of investigation. Herman Serota (1964) used home movies of children as correlative developmental data in the psychoanalysis of adult patients before 1940, but did not publish it until 1964. The technical and conceptual difficulties which prevented these and other workers from making progress became more obvious in the publications which followed Ruesch and Bateson (1951) and Gill, Newman and Redlich (1954), McQuown (1957), Pittenger, R. E., and Smith, H. L., Jr. (1957),

Pittenger, R. E. (1958), Eldred and Price (1958), and Pittenger, Hockett and Danehy (1960). The feasibility of making meaningful transcriptions, which had been developed by Pike (1945), Trager and Smith (1951) and Trager (1958) between 1942 to 1958, had been shown to be of interest to psychiatrists in studying the therapeutic process in ways not hitherto available and which might help improve objective recording. The difficulties are many. Linguists do not always agree upon the equivalence of the raw data, and use different conceptual schemes for recording. (Hill, 1958, 1968 and Longacre, 1964, Pike, 1967 and Stetson, 1951) It is our experience that the differences in working out transcriptions are relatively small compared to the large areas of agreement. They also vary in their opinions about interpretations which was summarized by Dittman and Wynne (1961) in their conclusion that the details of linguistic analysis are reliably describable, but "probably have little psychological relevance" (p. 203), while paralinguistic phenomena "have higher psychological relevance, but cannot be coded reliably." Trager's (1966) comments on this judgment are

worth repeating to illustrate the complexity of the phenomena involved and the need for more intensive and imaginative studies. "It is probably true that linguistic details as such have little psychological relevance, though they are basic to the sociocultural placement of the speaker in therapy; and the paralinguistic phenomena are of much higher psychological relevance in psychotherapeutic situations, though it must be understood that as behavior events they occur in a matrix of language. But I must question the conclusion that paralinguistic details cannot be coded reliably. At the time that the work was done with Dittman and Wynne, the analysis of paralanguage was still in a very preliminary stage, though good agreement was reached by the three of us in evaluating the material. Since their study, a notation system for paralanguage has been worked out, and work has been done in recording minutely the paralinguistic phenomena in English as well as in some other languages." (p. 81) The new notation in his article referred to by Trager (1966) has since been augmented by Crystal and Quirk (1964). Kramer (1963), in a review article on the

judgment of personal characteristics and emotions from nonverbal properties of speech (timbre, inflection, and stress) find evidence that some validity of judgment is possible, in spite of many methodological difficulties such as "inadequate measures as the independent criterion for the traits being judged", and no method for totally eliminating the effects of the verbal content, nor has adequate attention been given to the individual differences among listeners, or the relationship of voice to psychopathology. In our own group, the importance of individual differences was a source of frequent comment, while the diagnostic importance of the interpretations of a very small sample of Doris' vocal production, described in Chapter IX - Collation, can scarcely be over-estimated even though it cannot be reproduced in most experimental conditions. Kramer cites Sanford (1942) who noted that "common experience seems to accept the existence of connections between voice and personality, and if 'the analytic-experimental approach... reveals no relationship, we should be forced to conclude that it may be the fault of the approach' ". (Sanford, 1942, p. 838).

Markel (1965) using college students to score rating scales developed a quick method for coding pitch, loudness and tempo and found high inter-rater and test-retest reliability for these elements.

Weiner and Mehrabian (1968) discuss some of these problems, with the introduction of the concept of "immediacy" to other "channels" of communication, which must be of interest to all students of communication, particularly in connection with the concepts of "proxemic" behavior (Hall 1959, 1966, 1968), and the improved methods of the study of interpreting emotions from facial expression. (Davitz, 1964). Davitz, like Kramer (1963) and Sanford (1942) before him, comments on the wide differences in accuracy reported even though there is general agreement that facial expressions communicate feelings beyond chance expectancy. "These differences are probably a function of the stimulus used, the kinds of discriminations required in the experimental procedure, the categories of emotional meaning considered in a research, and the individual differences in ability among those who express feelings and those who are asked to identify the expressions." (Davitz, 1964, p. 14)

A further example of the different methods of obtaining data resulting in differences of opinion about another important question, namely the synchrony between body movement and speech, is the paper by Dittman and Llewellyn (1968). They found the results to be significant, but the amount of movement variance accounted for a much smaller one than that reported by Pittenger et al (1960) and Scheflen (1964). Dittman and Llewellyn believe that their method allowed direct tests on films with the conclusion that the claims of Pittenger et al (1960) and Scheflen (1964) "of very close speech movement relationship were found to be exaggerated." Dittman and Llewellyn also explore the convergence between Kendon's (1967) analysis of gaze direction at "hesitation pauses" and "phrase boundary pauses," even though direct comparison of data is not possible. The history of science is replete with examples of growth in concepts and improved methods for data gathering through the work of numerous workers comparing results. We need more active workers to test methods and concepts for recording, analyzing, and interpreting human communication more quickly and accurately so that we may be able to study the therapeutic process

and other interactions more effectively. It seems to this author that resolution of the "usefulness" of linguistic and paralinguistic data will become more apparent when the units under consideration include the kinesic (body movement) components as a single unit. We hope that more workers will find this concept useful so that the synchrony of speech and body motion found by Condon (1964) and further elaborated in Condon and Ogston (1966, 1967A, 1967B, 1968, 1969) will find corroboration and/or refinement of method.

The many perplexing problems concerning the segmentation of seemingly continuous behavior into "units" which has descriptive and experimental value is central to many other issues, and requires close examination of basic assumptions. In much of the work of Condon and his associates, the concept of "Etic" segmentation, i. e., the analysis of the physical, articulatory structure of sound emission as contrasted with the "emic" segmentation based upon classes of sounds was employed. The analysis of body motion has, thus far, been primarily etic in nature. Emic analyses will emerge as inquiry proceeds. No review of this topic

will be attempted here because many chapters of this book are relevant, and the discussions by Pike (1967, pp. 37-72) and Stetson (1951) provide adequate background.

We regret that lack of space prevents a review of many more examples from our selected bibliography of 1900 titles. The number could easily be doubled if one included additional studies in linguistics, anthropology, speech and speech pathology, neurology, infant and child development, and the linguistic philosophers. The beginning student can probably orient himself most easily by scanning Gottschalk and Auerbach (1966), A.G. Smith (1966), J.M. Schlien (1968), P.H. Knapp (1963), Hymes (1964) and Sebeok, Hays and Bateson (1964) since many of the most active workers in the field are represented in these volumes, and the bibliographies are exceptionally full.

We will not cite additional articles by Bateson, Birdwhistell, and McQuown since they are represented in their chapters. Attention is called to the work of Scheflen, a longtime associate of Birdwhistell's, whose experiments and expositions in the analysis of film and tape at

several levels have brought to public attention much new data (1965 a, b, 1967, 1968). In addition to Condon and Ogston (1966, 1967a, 1967b, 1968) already cited, the examples of coordinated regulation of lexical content, and body motion are shown by Charny (1966), Loeb (1967), Kendon (1967), and Brosin has reviewed some of the work of Condon and his associates (Brosin 1964, 1966) Condon, Ogston and Pacoe (1968). Schossberger, Condon and Ogston (1968) have closely studied the film of one case of infantile autism which shows synchrony to non-human sounds and indifference to the human voice. Pacoe (1968) is studying the effects of delayed auditory feedback showing dyssynchrony between the movements of fingers and speech, akin to aphasics, in standardized situations. In unpublished work Condon has several samples of dyssynchronies in stuttering, petit mal, aphasia and the schizophrenias. In over 50 control cases Condon has not found these dyssynchronies. If this can be corroborated by other works using similar methods, it may be possible in many cases, even if not in all of them to make significant observations about the presence or absence of various kinds

of disturbed behavior at these levels. Most observers at first are highly skeptical, but Condon has convinced outside visitors who are willing to study the film closely. Kendon (1967a, 1968), utilizing a similar approach, has tentatively confirmed some of the findings concerning self and interactional synchrony.

We will mention some of the more current psychoanalytic publications which are of interest. Carmichael (1956, also reprinted in 1966) and Bergman (1966) furnish the only two publications of a sound film of psychoanalytic therapy known to us, and both articles are worthy of study. Carl Rogers has also made a film of a psychotherapeutic process (date unknown) but cited by Carmichael (1956). Gill, et al (1968) are making audio-recordings of psychoanalyses and developing ways of studying them. They discuss the obstacles, re-evaluate objections and conclude that "recorded research analysis needs to be done, that the crucial problems may lie in the therapist as much as, if not more than, in the patient, and we believe that such an analysis can be successfully performed. If not, that demonstration in itself will constitute a valuable addition to our fund of information about psychoanalytic treatment and provide other leads for

future psychoanalytic research." (p. 242-243) In this same issue, there are excellent articles by George L. Engel (1968), and Sadow et al (1968) on psychoanalytic research. The discussion groups on language which are reported in psychoanalytic and psychiatric journals to date contain material of general, but not of special interest to our major topic. (Sternberg, Chapman and Shakow(1958) and Shakow (1960) also furnish excellent critiques of the sensitive problems surrounding the intrusions of privacy in psychotherapy research. Kiesler (1966) does a highly useful evaluation of the recent attempts at quantification and rigor, and their underlying assumptions which prematurely close off areas of inquiry. He examines the "patient uniformity assumption," the "therapist uniformity assumption," the "spontaneous remission myth," the lack of comparability of various patient groups, "the myths that present theories provide adequate research paradigms, "the distinctions between process research and outcome research; independent, dependent and Underwood's (1957) "confounding" variables in penetrating ways which support his thesis that no single one-shot method will definitely establish the value of psychotherapy.

(p. 127). The essay calls to mind the need for historical perspective described by Kuhn (1962) in *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* which emphasizes the need for long preparatory periods of "normal science."

Examples of explorations which may be of interest to our field are Schossberger (1963), K.K. Lewin (1965), V.H. Rosen (1958), (1967), Meerloo (1964) on linguistic components, Kestenbergl (1967) on movement patterns, Calogeras, R.C. (1967) reviews the psychoanalytic literature on silence. Matarazzo, et al. (1968) have presented an excellent review of their work on silence. Jaffe (1968) has numerous publications concerned with computer assessment of dyadic interaction rules from chronographic data in psychiatric interviews representing fifteen years of intensive quantitative research. While this work does not involve films, it is a dramatic example of the potential use of the computer, and some of the data and hypotheses will be of interest to linguists.

Ekman (1968) also has many publications dealing with nonverbal behavior in psychotherapy research, which will undoubtedly be of interest to linguists who have a different bias about the size, shape and duration of

units of behavior, levels of interpretation, and the kinds of messages which can be decoded from verbal and nonverbal information.

Haggard (1967) and his co-workers were cited earlier, but are mentioned again together with George A. Mahl (1968), another long-time investigator of clinical subjects because their writings cover a wide range of subjects which we cannot abstract here. His papers on gestures and body movements contain many valuable observations, and appreciation that the study of nonverbal behavior has much bearing on general psychology and that unconscious conflicts are a vital component of much behavior. Saslow and his associates (see Matarazzo, et al.) have published consistently over many years in the effort to bring more rigor into psychotherapeutic studies. The urgent recommendation of Saslow, et al. that "the time may be appropriate for a centrally administered, nationwide study of psychotherapy along the lines established for prenatal and post-natal development, cardiac disease, etc." (p. 393) may be a realistic one in the near future.

N.H. Greenberg (1967) has embarked on a large program for filming infants at closeup and distant views at any time for any reason while the infant is alone or in interaction with the mother or other caretakers. Although space does not allow an abstract of some of Greenberg's protocols, we can expect much new and detailed data from his use of "time lapse photography" and "infant stimulation" experiments.

The work of E. Hess (1968) since his earlier publications following 1960 has stimulated considerable interest because change in the size of pupils upon exposure to stimuli seems to be a measure of attitudes or change in attitudes. The measure is complicated by the fact that one component, the autonomic response, must be taken into consideration. Since potential applications are many, even if it does not become a stable index or objective measure of progress during the course of psychotherapy, it merits attention from students of the therapeutic process.

We have called attention to the interest of social psychologists to "social space" and want to mention the work of Argyle and Kendon (1967) Kendon (1967), Mehrabian (1968), Wiener, M. and Mehrabian (1968) Argyle (1969),

Sommer (1967), Rosenfeld (1967), and the popular books by the anthropologist Edward T. Hall: THE SILENT LANGUAGE (1959) and THE HIDDEN DIMENSION (1966), and a recent review article "Proxemics" (1968), because of the great importance of the larger units of social interaction which have not received adequate attention to date. The journal, EKISTICS is publishing reviews on the problems of human settlements from the Athens (Greece) Center of Ekistics by C.A. Doxiades which will be of interest to those students who wish to study the large social units.

Parenthetically, the journal, IKON, the International Review of Filmology, published since 1948, is the official journal of the Agostino Gemelli Institute for the Experimental Study of Social Problems of Visual Information in Milan, Italy. I am indebted to Dr. David Shakow of NIMH for the information that the Gemelli Institute has equipment which includes motion picture facilities, closed-circuit television, equipment for photographing audience reactions by infrared light, and neurological and psychophysiological recording apparatus. Associated with the Institute

is also the Central Register of Scientific Research on Visual Information. Their studies have included EEG response to the film situation, the use of film as stress stimulation in the study of psychophysiological response, posture changes and other overt responses of schizophrenic and normal subjects while viewing films, development of children's responses to humor in films, and several studies using a filmed projective technique developed by Professor Gilbert Cohen-Seat of the University of Paris. The quality of research reported is comparable to our own.

N.A. McQuown (1964) has recommended "particularly in view of the fact that the paralinguistic and kinesic disciplines are in their very beginnings, that we associate as closely as possible with people who are working in the animal behavior area, and in our initial training we must include films and tapes not only of a variety of human groups but also of higher primates" (p. 222). In addition to other citations to animal behavior studies by ethologists and comparative psychologists in this chapter, I would recommend the studies of R.E. Miller (1967) as



a useful introduction among the scores of exciting studies in this field.

The omnibus volumes by Rioch and Weinstein (1964), Thorpe and Zangwell (1961), Burkhardt (1967), and Altman (1967, and the articles by Sebeok (1965a, 1965b) furnish perspective on the current scene.

Occasional inquiry about the possible relevance of the well known writings of Chomsky (1965) cause me to comment that up to this time his followers have not yet undertaken to show the usefulness of his theories in the linguistic-kinesic analysis of human behavior. At least some of the reasons for this are suggested in the reviews by Lamb (1967) and Hockett (1968).

#### 4.3 Convergence on the Problem from Parallel Disciplines

Medicine is traditionally rooted in biology. The art of medicine which has always been an important aspect of the practice of clinical medicine has more recently, however, been increasingly subjected to systematic study by psychiatrists and social or behavioral scientists, as seen in the previous section. In turn, leaders in medicine have gradually opened the door to studies in experimental and social psychology, in cultural and social anthropology, in interpersonal relations, in medical economics and sociology, and in ecology and ethology. Among such pioneers is Gregg (1941, 1957). Acceptance of the need for such studies is far from general, however, since objective methods have not sufficiently demonstrated their usefulness. A few medical educators are willing to assume responsibility for helping workers in these fields develop methods and data relevant to the practice of medicine. The majority, if not indifferent because of preoccupations with their own work, prefer to wait until the fields themselves are sufficiently developed to be clinically useful.

A change in the attitude toward the non-medical disciplines, however, can be seen, for instance, in the growing recognition of the importance of genetics, especially of the value of those innovations made possible by advances in chemical and population genetics. Most psychiatrists have been much interested in the behavioral sciences because they recognize the need to study man in his social environment for the better practice of medicine.

Reviewing the development of psychiatric theory and the types of data with which psychiatrists deal, we find that they have much in common with other behavioral scientists, especially with the anthropologists, some of whom have attempted to explore the interrelations of pathology and normalcy in human behavior. Among these are Kroeber (1953 and 1956) (the latter in Thomas, 1956, pp. 292-311), Kluckhohn (1944 and 1949), LaBarre (1955 and 1964), Mead (1952, in Alexander and Ross, 1952, pp. 401-448), and Sapir (1933 and 1938, reprinted in Mandelbaum, 1949, pp. 7-32 and 569-577, respectively.)

The essays by Margaret Mead, LaBarre, Kluckhohn and Sapir especially clarify many of the links. Both disciplines have been criticized, with some justification, for the following reasons: too much theory with insufficient data or inappropriate data, lack of convincing correlations under specified conditions, lack of data which can be publicly examined and directly experienced by other observers, lack of clear definitions with operational meaning, lack of concern over the specifications of experimental conditions, lack of precision in formulations and inferences which are acceptable to workers from other fields, too great concern for raw data with insufficient generalizations of high quality, as in many case histories, and unwillingness to present both the methodology and the direct experimental data.

It seems probable that anthropology, along with other behavioral sciences, will have much to contribute to the understanding of the conditions, both internal and external, which determine human health. With better data, more precise and explicit methods of observation, and

with more carefully defined concepts, behavioral studies should make an appreciable contribution to the human clinical field.

Edward Sapir deserves mention as the pioneer linguist who clearly stated the hypothesis that "language" and "communication" in all aspects are cultural systems which can be studied and understood, even when dealing with deviant behavior and unconscious motivation. His associations with Harry Stack Sullivan enriched the insights of both men and provide us with the background to pursue clinical studies in a more rigorous way. Kenneth L. Pike (1967) is another pioneer who is credited by Hockett (1968) as having "given us some remarkably important insights: phonological hierarchical structure is one; his brief exposition of linguistic-like approaches to the discussion of other phases of culture (1967, the first few sections) is another" (p. 33). Hockett continues on with a severe criticism of Pike's terminology and style, a view also shared by others. (See N. A. McQuown, 1957). In spite of these difficulties, it may be worthwhile to study the ambitious attempts of Pike

to establish a conceptual framework of trimodal interlocking hierarchies, and building bridges between linguistic structure and the structure of society, and of nonverbal behavior. We need better hypotheses for the nature of the units to be studied from the split second eyeblink to events occurring over hours, months or years. At more basic levels we also need better understanding of the biological roots of language behavior which have been studied by Lenneberg (1964, 1967).

Clinicians cannot expect to take over ready-made either a set of hypotheses or an integrated body of fact which will immediately solve their most pressing problems. For one thing, such a body of unified fact does not yet exist, but is is in the making and it will soon be available to us, particularly if we contribute to its formation. Furthermore, there are no coherent theories of personality and culture, advanced either by anthropologists or by psychiatrists, which have been sufficiently tested to be accepted as definitive. That contributions of this order may be mutual is evidenced by the facts that the most unified theory of personality is probably provided by psychoanalysis, and that psychoanalytic theory has

been extensively used in various ways by social anthropologists.

Many anthropologists have long been familiar with psychoanalytic theories. Almost all of them are devoted to the naturalistic methods of direct experience, observation, comparison, and correlation and to the historical method, vigorously defended by clinicians. It is not surprising, therefore, that anthropologists have much in common with psychoanalysts. Kroeber (pp. 211-293 in Thomas, 1956) and Ackerknecht (pp. 117-125 in Spencer, 1954) hold that naturalistic methods have been useful since the advances in comparative anatomy in pre-Darwinian days, and that subsequent progress, using such methods, is impressive, even though less dramatic than progress in experimental science.

Although differing in disciplinary focus from the anthropologists, the work of the ethologists seems to be converging on similar problems. Darwin's great monograph *THE EXPRESSION OF THE EMOTIONS IN MAN AND ANIMALS* (1872) may be regarded as a cornerstone of behavioral studies and of comparative psychology. Such studies are only now receiving a

powerful new impetus in the direct study of behavior by the ethologists Lorenz (1952 and 1966), Tinbergen (1953), Thorpe (1956), E. Hess (1968), R. Hinde (1959), J. P. Scott (1958) and their colleagues on the one hand, and in the linguistic-kinesic analysis of film-tape recording of such behavior, on the other. However different on preliminary examination these methods may seem to the physician of today, it requires but little review of medical history to see that such methods are in the direct line of clinical medical development.

Physicians trained in the Oslerian tradition, that the history of science and of medicine is a most important source for understanding the gradual development of various experimental and clinical techniques, will recall the long chain of events which culminated in the modern disciplines of anatomy, physiology, pathology and microbiology. The complex sequence of investigations by workers in many fields which brought about the conquest of poliomyelitis is a clear example of such a chain of development. Similarly the numerous single advances which contributed to the understanding of the diabetic process have given us a

considerable control over most forms of the disease.

The life-saving qualities of insulin, on the other hand, as those of small-pox vaccine and digitalis, were successfully utilized long before the physiological and chemical complexities of their action were understood. Even in such diseases as tuberculosis, syphilis and malaria, where we have achieved reasonable comprehension of the chemical factors involved, we have not achieved similar understanding of the socio-economic factors. It is not extravagant to say that much in medicine has been built upon empirical research in the clinical setting. It would repay us in contemplating the advances of recent years to study more carefully the methods of empirical work in order to use our new technology more effectively. Klopsteg (1960) makes several good points in illustrating from the history of science that "instruments are unifying elements which help self-centered disciplines shed their isolationism," (p. 1913) and also that "the application of research to the furtherance of research is as basic as the research itself" (p. 1922), thus echoing a famous A. N. Whitehead theme. Klopsteg emphasizes the growing importance of scientists

whose research deals with instruments for research. We need to reward able men who will find this an interesting and satisfying career because our progress will be accelerated if we do so.

#### 4.4 Need for a New Technical Language

Physicians as well as other workers in the field of human relations, however, should not expect future hypotheses to be stated in classical cause-and-effect sequences. The failure of mechanically simple means of verification in the analysis of more complex systems has led the physician to seek other working models for his investigations. For example, as we become able to record more accurately slices of behavior, the sizes of which run from that of an eyewink (1/10 of a second or less), to that of a hereditary culture pattern with a duration of at least three generations (90 years), we will have available data whose ordering will require new theoretical frames.

Special attention must be given to the mechanics of the ways in which such data is obtained, so that we may understand the regularities which are introduced by the observer. At the same time it will continue to be absolutely essential to subject the clinical material to similar rigorous examination. It is evident that common-sense,

man-in-the-street hypotheses are not adequate for describing such complex systems. We may even discover that neither words nor pictures based upon mechanical models are the most economical or accurate means by which to describe and interpret a sequence of human behavior (see Dirac, 1947, pp. vii-viii, 10).

It is necessary, for example, for the clinician to find better ways to describe various states of being-in-action. Particularly are we concerned with the transformations of states such as that of an anxious-depressed-ambivalent patient who in the course of his illness develops aggressive-obsessional defenses which apparently have the option of becoming schizophrenic or manic under specified sets of conditions. Some physicians use clinical descriptive-genetic-dynamic formulations which allow for the possibility of such potentials. It seems probable that the anthropological, and particularly the linguistic and kinesic tools, which may now be utilized in clinical descriptions, may make possible a more flexible conceptual frame which permits a more systematic assessment of alternative pathways of development.

With the possibility of objectively recording and minutely describing human activities underlying multiple-level messages being sent and received between two people, we shall be in a position to make use of the more complex concepts described in the earlier chapters. The persons in a clinical setting may now be viewed as exchanging complex messages during relatively short periods of time. Such complex messages, at several levels of awareness, must be analyzed on several levels of behavior: physiological, linguistic and kinesic, and social organizational, by a set of parallel operations whose end-points are intelligible only in the total matrix of the social communication pattern (Sapir, 1931; reprinted pp. 104-109 in Mandelbaum, 1949, and Pike, 1967). We might also include as objects of analysis the large and ill-defined group of implicit processes which grow out of overt behavior and which may be rather vaguely referred to as 'social suggestion.' (Sapir, 1931, p. 105) Such multilevel messages are received by each participant, and responded to with equally complex messages. In the flow of communication there are no totally irrelevant

messages, no superfluous information, and no randomly produced data.

On the contrary, it is the obligation of the observer to assign values and meanings to all such component communicative items after studying the contexts in which they appear. We must recall that many messages which apparently relate only to the immediate present are nonetheless connected with events of the remote past and foreshadow future happenings. A good clinician can sense when the patient is bringing into the communication situation large components of the past or future and acts upon this understanding. An example is the state-of-regression which may be detected in the lexical components, or in the return to childhood vocal modifiers or body movements or even in physiological activities which signal childhood patterns. The separation of the "Here and Now" from the "There and Then" in the patient's report is one of the therapist's primary interpretive tasks.

It seems inevitable that we reconstruct our teaching programs in such a way as to give our future clinicians that familiarity with the social matrix which can only come from extended investigations carried on by them at

first hand or from the distillate of such investigations carried on by others. In fact, the degree of specialization required for effective control of such background information will probably lead to the employment of highly skilled behavioral scientists who work in the medical setting in the role of basic scientists. Such a pattern is now familiar to us in the basic science divisions of a medical school.

#### 4.5 The Natural-History and the Experimental Method

Although the conceptual frame and the method derived from it which is used in the investigations presented in this book are by no means new, perhaps a review of the salient differences between this naturalistic-field study-participant-observer method and the currently more popular "experimental" methods will help clarify some of the inherent problems. It is highly enlightening and stimulating to read now Claude Bernard's INTRODUCTION TO THE STUDY OF EXPERIMENTAL MEDICINE, written in 1865, which outlines in masterly fashion a combination of the naturalistic-descriptive with the experimental method, years before Francis Galton (1822-1911) had introduced statistical analysis, and introduced a vogue which came to be considered a cornerstone of all social scientific investigation. Today it is quite clear that much naturalistic-descriptive and experimental spade work must be done before statistical methods become appropriate and rewarding in the solution of the problems presented by our investigations (G.A. Miller 1964, 1965). We make no claim that one method is per se superior to any other, since to us it seems



quite clear that each is appropriate to a particular phase or aspect of a problem. We enter into an investigation with the assumption that all human behavior is amenable to study, but we want to apply, as far as we can, the most appropriate methods to the problem in hand as we now understand it. It appears to us that naturalistic methods are essential in the early phase of the development of a science. Until a sufficient body of data is available for the application of critical hypotheses we cannot test by crucial experiments. Naturalistic methods seem to have many advantages in the exploratory phases, since they lay the groundwork for the application of the more controlled treatment characteristic of the experimental method of verification.

Excellent methodological discussions are to be found in Pumpian-Mindlin (1952), Tinbergen (1953), Thorpe (1956), Harlow (1957), P.S. Sears (1957), Hinde (1959), G.A. Miller (1965), GAP Report No. 63 (1966), Scheflen (1966), and Schlessinger, et al (1966).

Lest this attitude toward the planning of medical research seem insufficiently critical, let us quote from an eminent physicist (Oppenheimer, THE OPEN MIND, 1955):

"You know that when a student of physics makes his first acquaintance with the theory of atomic structure and of quanta, he must come to understand the rather deep and subtle notion which has turned out to be the clue to unraveling that whole domain of physical experience. This is the notion of complementarity, which recognizes that various ways of talking about physical experience may each have validity, and may each be necessary for the adequate description of the physical world, and may yet stand in a mutually exclusive relationship to each other, so that to a situation to which one applies, there may be no consistent possibility of applying the other. Teachers very often try to find illustrations, familiar from experience, for relationships of this kind; and one of the most apt is the exclusive relationship between the practicing of an art and the description of that practice. Both are a part of civilized life. But an analysis of what we do and the doing of it--these are hard to bed in the same bed. (*ibid.*, pp. 82-83)

The analogy to the problems facing the clinician in describing his own activities, or the behavioral scientist in delineating his methods in an ongoing transaction, is obvious. As investigators grow more certain of their aims and methods, they will probably develop more tolerance for a wider range of techniques and eventually aim at the goal of finding the best tools for each job.

In order to clarify our reasons for adopting, in this preliminary investigation, the natural history approach rather than the experimental

approach, I will try to contrast some of their comparable properties:

Naturalistic Approach #1

The organism or group under study is maintained as far as possible under the customary conditions of living. We are probably dealing with a spectrum rather than an antithesis in many instances.

Experimental Approach #1

The subjects are usually approached in the more-or-less specialized artificial conditions.

Naturalistic Approach #2

Because real life situations are under observation, their problems are more likely to be comparable to practical problems in the clinic, school or business organization.

Experimental Approach #2

The specialized nature of many experiments may make the results more difficult to apply to real life situations.

Naturalistic Approach #3

The quality of the responses of the subjects is more likely to be "life-like" and in keeping with customary behavior under observed conditions.

Experimental Approach #3

The responses are more apt to be altered to suit the experimental demands of the tighter design.

Naturalistic Approach #4

The responses are more apt to appear on a wider scale such as is available to the organism under its ordinary conditions of living.

Experimental Approach #4

The responses are more apt to be constricted because of the restraints imposed by the experiment.

Naturalistic Approach #5

The time-span is apt to be longer and thus to afford more information for longitudinal studies and for the dynamics associated with larger patterns.

Experimental Approach #5

Although so-called cross-sectional experiments do utilize significant time segments for the study of behavior, and some of them may study relatively larger units of behavior, they tend to be limited to manageable units which have limited meaningfulness in clinical situations. Recording of responses is determined by the test design. These experiments have great power and are highly desirable where the experimental conditions can be sharply and meaningfully defined. It is reasonably argued that long-term studies can only be pursued profitably after a sufficient number of cross-sectional experiments have been carried out so as to make the methods of investigation worthwhile by providing essential facts about the interacting variables under experimental conditions. The fact that cross-sectional experiments are more apt to be sharply defined, amenable to relatively concrete manipulation and interpretation, take less time, energy and manpower, makes them more suitable in many research situations.

Naturalistic Approach #6

Because naturalistic methods usually permit the study of several variables, more potential latitude exists for discovering new relations between them. The naturalistic methods are essentially comparative, correlational, or genetic-developmental in their strategy. More opportunities are asked for focussing upon new problems in which controlled experiments or observational techniques provide greater precision. Clinicians have long insisted that the study of patients is the best starting point for more intensive experimental design with much evidence to support their claim that this approach will best serve our needs to understand and master clinical problems.

Experimental Approach #6

The strength of most tightly designed experiments consists in focussing upon one dependent variable which can reliably be measured in its relations to several independent variables which in themselves can be manipulated. As with the naturalistic method, this may be a shortcoming if the material manipulated is inappropriately chosen, or if the operation produces trivial results.

Naturalistic Approach #7

Ideally, although this is not uniformly true, hypotheses grow out of the data produced by the longitudinal studies of a participant-observer in the field situation. Unfortunately many clinicians laboriously accumulate much data which they cannot process in any manner productive of meaningful experiments or vitalizing hypotheses. Unfortunately, too, raw data is too often interpreted in such a manner that the generalizations derived from it are not convincing or useful in further research, or worse, cannot be put to crucial tests. Too often, also, too many investigators view natural history explanation as an end in itself and arrive at results which are essentially sterile.

Experimental Approach #7

Ordinarily, a well designed experiment, with controls, contains a more or less well formulated apriori hypothesis which is being tested. Where the question being asked is a good one and the method appropriate to the interrogation, the results may be applauded. Too often neither of these conditions is fulfilled and the results are disappointing. Again the experiment

can become an end in itself and lead only to trivial results when the experimentalist becomes a slave to his tools.

Naturalistic Approach #8

When a number of variables are being studied over a relatively long period, it is essential to make a selection of what appear to be the most significant relations between important variables among the available events. Errors in judgment and interpretation are bound to occur, but the hope remains that other investigators may examine the data for new interpretations, since such critical review and reassessment is as necessary to the natural history method as is repetition in the experimental method. In important areas it may be necessary to re-do the observations repeatedly until the significant variables are recorded with sufficient accuracy to establish confidence.

Experimental Approach #8

Although the by-products of a well designed experiment, as contrasted to the pointed questions asked, may be productive of new experiments, the artifacts inherent in the experimental design at times force inconclusive

or negative results of little significance for future work. This is a necessary cost where the aim is to simplify and cut down all data subordinate to the primary aim of the experiment. Here the attempt is to eliminate "irrelevant" factors, to place variables in vivid contrast, and, if it is not possible to "randomize" the samples, it is possible to avoid undesired effects of individual differences. The possible conclusions are, in a sense, built into the experiments.

#### Naturalistic Approach #9

In general, natural history observation requires more continuous working time and more energy spent on recording than the classic experimental method. Clinicians as a rule have not provided sufficiently complete records which may be profitably studied by colleagues who wish to gain new insights. Clinicians, furthermore, as a result of differences in training, may not produce records which are sufficiently comparable. Finally, clinicians may tend to withhold salient information, either to protect patient privacy or as a measure of self-protection.

#### Experimental Approach #9

The recording of data and its interpretations are much more concrete and obvious in planned experiments. This is one of their most attractive features.

It has been suggested that an investigator who works in a field as a participant-observer over a relatively long period, becomes devoted to the subject matter, and unduly motivated by his interest in it. In this, self-selection plays an important, but as yet indeterminable role. Such an investigator may be better suited to his role than the investigator whose primary investment is in technique. The former may be more inclined to use techniques which are appropriate to the material because the material is of primary concern. This generalization becomes vital in clinical work where the investigator also has responsibility for the welfare of patients. Unfortunately there are times when consideration for the patient has absolute priority over any investigative procedure. Perhaps some clinicians at times utilize this necessary rule either consciously or

unconsciously to save themselves embarrassments or unwanted insights. It seems highly probable that this serious barrier can be overcome with intelligent planning and hard work, but we do need to take more seriously Freud's comments about the difficulties involved in trying to maintain a stable therapeutic attitude simultaneously with an experimental attitude. (Freud, 1912)

Freud also points out in this paper that the therapeutic ambition to achieve results which will impress others is a serious affective danger which must be controlled for the welfare of the patients. There is no reason to believe that the experimentalist is any the less affected by these variables. The participant-observer pays a considerable price among his "tough-minded" colleagues for his freedom to form hypotheses in a continuing study because his records are seldom complete enough to convince skeptics, and his personal interpretations are usually vulnerable to negative criticism or subject to alternative interpretations.

Hypotheses drawn from relatively small samples of behavior often sound like unquestioning belief or self-deception to workers trained to expect quantitatively displayed, immediately demonstrable evidence. While such hypotheses may frequently be inadequate, some of them may lend themselves to experimental verification. Although such experimental verification may seem remote, freedom to formulate such hypotheses is absolutely essential to the progress of a discipline. Experimentalists will learn that for every gain in method, there may also be a loss. It is true that the undisciplined participant-observer may introduce many artifacts into the system which may not easily be detected, thus abusing at times the inalienable "right of every scientist to be wrong," but with communication we need not fear that such error will long remain undetected. All too often the critics of the participant-observer method are simply unaware of the extensive training required by the practitioner of this method. Several highly skilled individuals trained in the same methods of observation and recording provide for indispensable self-correction in

the descriptive and analytic process.

The experimentalist, in the classic quantitative tradition, has his role as an observer more or less defined by the experiment. His recording is designed to be adequate for his specific purpose and for that alone. Like the modern physicist he must be aware that as an observer he becomes a part of the system in question. He must also be alert to the possibility that he may, himself, be introducing a portion of the regularities which he perceives in the system. Like the physicist, he, too, is bound by the limitations of his observational instrument and by the restrictions inherent in the experimental design. However, these inherent drawbacks in his method tend to be much less questioned, and he may hope to be rewarded for the loss of some degree of freedom by the greater security of his results. In one gratifying way, even though it leads to more frequent disappointment, he may have an advantage over the naturalistic participant-observer in that he may be more likely to attain much sooner and with greater certainty the end point of his experiments. He can also make a definite statement as to the positive or negative results. "The

refinement of techniques for the prompt discovery of error serves as well as any other as a hallmark of what we mean by science." (Oppenheimer, 1955, pp. 93-94)

A further advantage of the experimental method is that in many experiments less training is required for the experimental technician once the experimental frame is set. During the early exploratory phases of discipline, the choice of the unit to be studied is apt to be determined by the previous experience or bias of the investigators. Only with experience do the units chosen for study become more appropriate to the task in hand, and different workers can find some equivalence in their experiments.

The units of observation of a young science, therefore, necessarily are often not so precise as one might wish. Of course, the units chosen for observation are those which seem to be most significant within the bias or conceptual scheme of the investigator. (Pike 1967). Sometimes units are chosen primarily because they are most amenable to manipulation, or the best provided by the tools

available. Such factors can be studied and improvements made as the nature of the problem in hand is better understood.

#### 4.6 Advantages of More Precise Observation and Analysis and the Training Value of the New Techniques

It was in the hope that reliable descriptive units might for the first time be made available that the demonstration described in this book was undertaken. From the outset our goals were modest, and they have remained so. A brief description of the raw material under investigation will be presented here in anticipation of a fuller presentation to follow (Chapter 5). For our examination Mr. Bateson was able to provide 1,800 feet of sound-film of the activities of the members of a family of three who had volunteered as subjects for his investigation of the etiology of schizophrenia. These sound-films, supplemented by a fuller, but in part simultaneous recording on tape, were taken in the home of the family by Mr. Bateson's technical assistants utilizing a sound camera and with sufficiently fast film that special lighting was unnecessary. The special problems involved in the filming and recording are discussed later in this chapter and elsewhere in this book. The scenes show father, mother and son eating dinner together; the family entertaining neighbors; father



bathing the son; father and Mr. Bateson talking; mother, son and Mr. Bateson talking. This last was chosen for varying degrees of analysis both of vocal and of body motion activity. Interpretations of all activity were made independently by all members of the team, each on the data of his own medium, although these interpretations were necessarily strongly influenced by the intensive work done together. In one experiment the attempt was made to categorize the nature of the activity of one person on the film in each of three modes of action, as independently as possible by one member of each discipline. These results are reported in Chapter 9. From our preliminary work we feel justified in suggesting the following advantages inherent in the new procedures:

1. Intensive repeated viewing and listening to a limited and carefully selected segment of behavior (1 second to 50 minutes), particularly after tutoring by trained workers in linguistics and in kinesics, in an enriching and gratifying experience. It is one which gives the student new perspectives on and insights into his daily clinical techniques, his strategies and maneuvers, both conscious and unconscious. One is

mindful of the shock of being confronted by auditory and visual evidence of behavior of which we were previously unaware. On occasion, one may even want to repudiate such evidence because it does not conform to one's self-image or to the conscious plan of one's therapy. Through practice, such shock is gradually overcome, just as our clinical defenses are brought under control through supervision in therapy and by personal psychoanalysis.

2. This technique forces an increased awareness of minimal cues, and permits a greater selectivity among them. We are forced to accommodate to the unusual in ourselves and in others by the growing recognition that we also are biological organizations capable of sending very diverse messages to ourselves and to others. Some of these messages may seem trivial, but many are vital to life and to well-being. A good clinician is by definition one who can lift the more relevant facts from the mass of data made available through the new techniques. The good clinician will now have the chance to sharpen his own techniques of observation through repeated and focussed viewing and listening. Not only will he be able to detect more of the cues upon which he makes

judgments, but he will be able to be more explicit about the bases of these very judgments.

3. The technique promises to objectify affects-research at the interpersonal level. Transcribed taped and filmed materials will provide the means by which his fellow clinicians may inspect the interaction and proffer focused criticism and comment. It will no longer be necessary to balance one intuition against another with no hope of arriving at a consensus. Clinical insights can now be analyzed and substantiated by specific reference to data observable by all. Such data can also be supported (or refuted) by information on visceral activity. Both externally observable and internally measurable data may be combined to show that a patient or doctor is angry, disturbed, depressed, or indifferent; that a particular system of activity represents masculinity, femininity, regression, aggression, or ambivalence; that another action or state such as politeness, passive aggressiveness, prudery, exhibitionism, or paranoid defensiveness is idiosyncratic or patently culture-bound.

4. The clinician, sensitized by the observation of linguistic-kinesic microanalysis to the patterning inherent in the clinical data, will also gain a tremendous support for his assumption that all events in a human interaction are interrelated and that events and sequences are meaningful precisely because there is mutual triggering. Even in our relatively limited experience with sound-films, we have found sequences which are startling in their strictly patterned dance-like shapes. This is especially impressive in studying the relations between a mother and her child. The closeness of their interdependence as evidenced in their vocal and body-motion activity lends additional credence to theories that each individual, whatever his biological-genetic heritage, is also a strict product of his orientation, his training, and his environment.

5. It seems likely that with the new orientation the clinician will learn to ask highly specific questions which can be answered by the presentation of linguistic-kinesic data. We may gradually learn not only to ask questions of this type, but also consciously to seek out specific con-

stellations of vocal and body-motion behavior bearing on their answers. Such questions as the following were asked (and in part answered) in the course of our investigation:

"Does this person act his age?"

"Does he speak and act in keeping with his training and present position?"

"How do his body motions suggest bisexuality?"

"Has he regressed in his identification with his father?"

"Is this behavior similar to that of adolescent revolt?"

"Is the behavior of this person incongruently mature?"

6. The clinician acquires new evidence that very small samples of interaction between persons may be the source of considerable information about the participants and the way they relate one with

another. There is evidence that smaller and smaller segments of behavior can be regarded as significant samples. Pattern-analysis shows that such small sequences, if carefully chosen, may be iconic for much larger stretches of behavior. This will hardly be news to the many skillful clinicians who make use of the device of analyzing one small segment of behavior, such as a patient's entry into the room, or his walking across a room, or the associations to one dream, as a means of writing out in detail many factors of a patient's conflicts and their defenses.

7. The purposeful viewing of selected film-tape sequences will encourage a much wider use of them as training devices and will probably raise their general utility to a higher level because they offer a relatively stable medium for such training.

For many clinicians in training, the prolonged and repeated viewing of small quantities of such film-tape materials will bring home the thesis that the price of civilization is repression. While we all know this intellectually, it is a shock to experience it viscerally when the

clinician finds himself uncomfortable because of the intimacy which such viewing and listening not only allows but forces on him. This shock phenomenon deserves careful study, inasmuch as it will be necessary to allow for it in a clinician's training; perhaps highly specific aversion patterns and defenses against them may be recognized and allowed for, thereby both accelerating self-awareness and facilitating the supervisory process.

Training in intensive film-tape observation will help free many clinicians from their need for large amounts of clinical history and interactional data before making judgments. Certainly a great deal of data must be compiled, analyzed and compared before reliance can be placed on decisions based on small samples.

8. Although we have made only a beginning toward that end, the new techniques allow us to hope that we may eventually arrive at a dynamic diagnostic formulation of an illness from an analysis of relatively small samples of behavior in typical settings. It will be possible to specify

with greater accuracy the criteria upon which a diagnostic formulation is made. Similarly, the possibility exists that we may be able to make much better prognostic evaluations with recommendations for greater specificity of therapeutic intervention than is now possible. There is no question that as our observational skills improve (with the gradual discovery of the definable indices of vital processes occurring during a mental illness and during the therapeutic intervention) more reliable ways will be found to detect and chart the phenomena which signify improvement or regression. Prognostic studies will be more meaningful as we are able to chart more tangibly (from the results of analysis of film-tape materials) those activities which are common to all patients with an active neurotic process.

9. The new techniques have obvious advantages for comparing and evaluating therapeutic procedures. With increasing experience, it is conceivable that studies can be designed to investigate more adequately those therapeutic activities which are common to the several allied types of therapy. Studies on "control" patients may then become possible in a

way which does not seem very rewarding now, and a genuine basis for comparing therapeutic techniques will become feasible. It may eventually be possible, through careful study of filmed-taped interviews of student-candidate activity with patients, and of the procedures followed by particular practicing therapists, to prescribe more economically a course of training for the individual student, and to assign more efficiently particular patients to particular therapists.

10. Trial demonstrations of materials obtained through these techniques have already shown their usefulness in case presentation.

It is likely that as clinicians become more at home with the techniques involved in efficient use of film-tape materials, such film-tape data will become increasingly popular at professional meetings when clinical problems are presented. Such evidence will have greater impact and be more convincing than is the usual written report.

We cannot expect too much too soon. It is important to remember that the use of recording devices in medicine (microscope, camera,

x-ray, EKG) required many years for their full development. It is possible that with more accurate and more objective data, improved identification, isolation and definition of various classes and categories of disease processes will emerge. The change in these processes can be followed in greater detail and thus provide valuable clues to the crucial factors which make possible significant and enduring change. It is certain that these methods in themselves will not solve all the clinical problems of diagnosis, of interviewing, or of psychopathology. It is quite likely that shortcomings will appear which are now not obvious. Viewing film is not an adequate substitute for a face-to-face interview with a live patient in a closed room. Extensive empirical experience will be required before we can fix on those problem areas where these techniques will have their most powerful application.

4.7 Possible Disadvantages of the New Methods of Recording, Transcription, and Analysis

At this early stage of exploration it is not possible to spell out with entire certainty the limitations of these methods. From our experience a number of such limitations can be described as inherent in them. Only further experience, however, will demonstrate to what extent these can be minimized so that they will not constitute insurmountable barriers to obtaining vital information. It is certain that, even though reasonable control over new methods is gained, new difficulties will appear as we deal with more (or less) complex patterns of communication.

1. It is unlikely that sound-films will ever entirely replace face-to-face interviews in clinical research, since the recording media inevitably tend to distort some aspects of the scene under observation. Even technical improvements in reproducing the scene in color will not provide the information furnished by odor, by temperature, or by other sensory cues in their native setting, for which better methods of measurement need to be devised. Even though it may be true that working

intimately with film-tape material will enable a worker to "translate" with much greater insight and accuracy, we need more studies to demonstrate the conditions under which more reliable translations may be produced. The experienced investigator must learn to develop compensatory mechanisms to counter-balance the sensory limitation inherent in the film-tape medium.

2. The privacy of the therapeutic transaction is usually cited as an absolutely essential condition for good work. This may well be true for some transactions, but we need evidence from experience to be sure that this requirement has not been overestimated for many types of psychotherapy. There are many investigators who do not share this belief, for in the types of interviewing they have studied they have not observed serious interference with the patient's production attributable to the obvious presence of recording devices. This does not preclude the possibility that more subtle transactions may be prevented from coming into being by the nature of the recording situation. (Carmichael, 1956; Bergman, 1966; Gill, et al, 1968)

It has been observed, moreover, that the therapist, however experienced, is more likely to remain acutely aware of his "public" role, in such a situation, even after considerable practice, and is apt to introduce artifacts into the transaction. This is in keeping with Freud's observations (1912) already mentioned that therapeutic ambition to impress others presents the greatest affective danger to the therapist, and that in some transactions one may find it difficult to be an experimentalist and a therapist simultaneously. Occasionally, the therapist will alter the course of the therapeutic process for "the good of the patient." It will require much sensitive cooperative work to establish the legitimacy of some of these claims which seem to be motivated by the needs of the therapist rather than by the needs of the patient. (Gill, et al, 1968)

Judicial and patient review by valued colleagues will make possible the selection of therapists who are especially fitted for this exacting experimentation. It is not unlikely that the therapist as well as the patient should be protected by as much anonymity as possible by an inbuilt system in the experimental design whereby the therapist has ample

social and professional support and consultation opportunity. With increased skills, it is probable that the system of protective devices for the therapist can be so arranged that his needs and those of his colleagues will not interfere with the external requirements of the experimental design which have been agreed upon at the beginning of the experiment. Planning of this order will eliminate some of the difficulties which might prevent proper examination, recording, and interpretation of raw data about the therapeutic process.

3. It is not known at this time how much special training may be required for the average worker to recover quickly from the "cultural shock" of intensive viewing and listening. It is necessarily a transient phenomenon which must be met by skilled tutoring, if the problem of overcoming our resistances to focused viewing and examination of another person's privacies is not to undo the defenses provided by good clinical training. We suggest that we need to develop better methods of dealing with the observer's feelings under the conditions of repeated intensive listening and viewing.

4. There is a potential danger that the mechanics of recording will introduce artifacts, distortions and omissions which will lead to misinterpretations, particularly if the recorded materials are collated too literally with physiological data or with the overt content of such materials. It was not, for example, immediately detected on film that one patient had dyed her hair, even though the shading was different in films taken several months apart. Although investigators must be alert to such artifacts and distortions, it seems unlikely that they can be of serious consequence since they were detectable upon review by several workers who had the total context in mind. Our original difficulties in detecting that mid-face motility in Doris which clinically we identify as grimacing, and in identifying on tape the voice with "flattened affect", have diminished with increased practice.

5. There is the potential danger that the very setting of micro-linguistic-kinesic examination and interpretation may lead us into distorted evaluations because we have lost proper perspective both for the elements

of the transaction and for the relations between them. Freud's valuable caution about this problem in dream analysis is contained in his quotation (Freud 1900) from Hanns Sachs:

"If we look in our consciousness at something that has been told us by a dream about a contemporary (real) situation, we ought not to be surprised to find that the monster which we saw under the magnifying glass of analysis turns out to be a tiny infusorian."

The primary safeguard is provided by the multiple listening and viewing technique leading to the establishment of patterns prior to the reevaluations provided by narrow focussing. A further safeguard provided by a multi-disciplinary team approach and by a larger number of experienced workers will facilitate detection of misuse either of concepts or of data. It is well known that preoccupation with very small samples of behavior may cause distortions, but, as long as the possibility of correction exists, this need not constitute an insurmountable barrier. We know that some patients may vary from one extreme (apathy) to another (panic) within short periods, but appropriate sampling will provide us with the proper evidence.



6. Some psychiatrists have suggested that large expenditures of time, energy, and highly specialized skills required for work on small samples may not be warranted in view of the relatively limited amount of new information obtained. They claim that a similar expenditure of effort in the more traditional recording and analysis of data would probably be equally rewarding. It is not possible to prove conclusively at this time that this claim is not justified, but the probabilities are very strong that micro-analysis will produce much new data and unforeseen hypotheses which are inaccessible to ordinary recording methods.

7. The lack of adequate comparative data (such as that provided by the social anthropologist and the social psychologist) in our own and in other societies makes accurate evaluation of our micro-linguistic-kinesic data difficult. We urgently need comparative series in the field of social behavior, not only for the evaluation of our linguistic-kinesic data, but also for the interpretation of the larger behavioral constellations in which such data are embedded.

8. Some workers are skeptical about the value of micro-analysis in view of the paucity of comprehensive psychological theories of personality which are susceptible of testing. It seems likely that the most unified theory of personality, the psychoanalytic, will become increasingly available for testing as we are able to record the data on which it rests more completely and more accurately. We may anticipate a flood of new theories released by the new data being made available by microanalysis.

9. Everyone who has worked with these recording methods knows that there is a tremendous danger of being swamped by the mass of accumulated data. The methods described in earlier chapters provide the possibility of better selection and increased mastery of significant material once the more important patterns of signal transmission between individuals are recognized. The cost in time and skill in analyzing samples may be large at this time. But this is a new method of examining data in an orderly manner with the possibility of introducing new concepts.

Although future experience may show up areas of minimal return due to redundance or relative degrees of inapplicability, it now seems probable that the gains will far outweigh the disadvantages. It does provide objective techniques in which scholars may be trained to check upon each other's findings. Until this is done we cannot finally decide the degree of parallelism, mutual support or possible contradiction between lexical, linguistic, kinesic and physiologic behavior. Only then can we finally determine the value of these developments for the physician and other students of human behavior.

## CHAPTER 5

The Actors and the Setting

Gregory Bateson

#### Conditions of Filming

The films which provide most of the data used in this book were made under the following circumstances:

A therapist, who knew of my interest in collecting film data on family interaction, told Doris (who was his patient) about my project. It so happened that Larry and Doris had attended a public lecture which I had given some months previously and, therefore, were receptive to the idea of having some part in our research. As recounted elsewhere (reference to Brosin's chapter on the case history), this couple had in the past made considerable efforts to find solutions to the problems which troubled them and their world. They had participated in dianetic sessions and had maintained an interest in communicational problems as part of their quest for answers.

Shortly after the therapist's conversation with Doris, Larry telephoned me to express their interest in cooperating with my work. I then made an appointment to visit the household with Mr. David M. Myers, our cameraman, bringing camera, tape recorder, and lights. Myers and I arrived punctually at four o'clock and found the house empty. We withdrew and telephoned 15 minutes later verifying that Doris had returned. She told us to come right over, but when we arrived we found her quite flurried. She had just returned from a session with the therapist, had picked up Billy from the house of a neighbor, and was exhibiting the expectable responses of a housewife unprepared to receive her visitors--let alone cameras and lights.

After introducing ourselves, Myers and I started to move the equipment into the house, and she hastily defined the situation by pressing upon us mugs of home brew. The exchange of social amenities was somewhat

confused by the business of setting up the recording equipment.

The "interview," of which this book is the natural history, followed immediately, while Billy intermittently played on the floor in front of us, and came and went.

A word must be said about the human problems of such photography. David Myers is singularly skillful in the art of effacing himself behind his camera most of the time, while still being able to be a participant in the group whenever this is called for. He lets children examine the camera and look through the view finder and feel that the camera is not alien to themselves. He is instructed not to shoot for any particular type of behavior but to get as complete a record as possible of what occurs. He and I rapidly made a general plan that we would try to concentrate action at one end of the living room around the couch. Beyond this, he would make no effort to have people perform for the camera or place themselves for his convenience. It was his job to record whatever happened as best he could without interfering with it.

Under these circumstances, a family rather rapidly settles down to a freedom from camera consciousness and this process is, I think, aided by me: I, too, am being photographed and I am accustomed to it. It is also worth mentioning that the film was made before I had any contact with micro-kinesic or micro-linguistic analysis. I was not aware that my smallest movement and intonation would later be examined and, therefore, I was not in a position to communicate any self-consciousness of this sort to Doris.

Doris's living room was, fortunately, rather longer than the usual so that the camera could be well separated from the action. The far end

of the room, opening to the kitchen, was provided with a dining room table at which the family usually ate. We set up camera and tripod beside this table, facing down the length of the room to the couch.

The general decor of the room was such as might be expected in the house of people with intellectual interests. There were two or three objects of Mexican art, and books, but the impression given by the room as a whole was unsettled. The family had lived there for two or three years, but I had no feeling that they liked living there or that the various objects belonged in their various positions. The living room windows opened onto a lawn where we later photographed Billy playing various games, and beyond the lawn was the railroad track. The interview, conducted in the late afternoon, was repeatedly interrupted by the noise of commutertrains, at which Doris was markedly distressed, while trying to laugh off the indignity of this interruption.

The record was also interrupted by the recurrent need to reload the camera. Doris and I simply went on talking, paying little attention to what Myers was doing so that, while the tape record of our conversation is complete, the film record has gaps in it every  $3\frac{1}{2}$  minutes when a new 100 feet of film had to be inserted.

About 5:30 p.m. Larry returned from work. He and I then had a conversation which was filmed while Doris prepared supper. Larry was mainly interested in talking to me about scientific theory and other impersonal matters, in the discussion of which he and I had the role of intellectual colleagues.

When supper was ready, the lights and camera were moved so as to film the family meal, and in this context Billy, for the first time, began

to play up to the camera. The record shows him holding up a leaf of salad, spreading it in his two hands somewhat in the manner of Saint Veronica, while he waits impishly to see how his parents are going to handle this misbehavior with the camera's eye upon them.

In subsequent visits, films were made of Billy playing in the garden and of Billy being bathed by his father. We also made a film of an informal party of friends gathered one evening at the house. This family, unlike the majority of the much more troubled families with whom we have worked, maintains an active informal social life with a number of acquaintances, neighbors and young married couples, most of them Larry's professional colleagues. Surprisingly, Doris's guests showed remarkably little tension under the stress of being photographed. This rather unusual situation arose in the following way. We had arranged to go over in the evening to photograph Billy being bathed and put to bed. When we arrived, we found the party in progress. We then filmed Billy's behavior with the company and later moved the camera to the bathroom.

The material from these subsequent filmings has all been a part of the background of the present study, though no part of it was actually used for micro-analysis.

In retrospect, and especially in the light of subsequent experience in filming other families, my main criticism of the collection of data is that it contains too little interaction between Larry and Doris. Interaction between Doris and Billy and between Larry and Billy are both well documented, and there is adequate documentation of each of the three family members in interaction with me. It is difficult, however, to get from the filmed material an adequate picture of the relationship

between the parents without which it is difficult to understand the total dynamics of the family.

In its subsequent history, the family has moved into more luxurious surroundings. They now live in the country in a house of which Doris is proud, and in the building of which Doris and Larry did a considerable part of the work. Their present living room does not have the unsettled appearance mentioned above, and they are no longer troubled by the sound of commuter trains. Billy now has a little sister who, Doris laughingly says, is "spoiled rotten."

The film record was made on Eastman TRI-X film, using a 100 ft. Auricon camera with an optical sound track. The lens used was a Pan-Cinor "4" Reflex 70 zoom lens. We also made a magnetic tape recording of the entire proceedings.

# END

## MICROFILM COLLECTION OF MANUSCRIPTS ON CULTURAL ANTHROPOLOGY

FORMERLY: MICROFILM COLLECTION OF MANUSCRIPTS ON  
AMERICAN INDIAN CULTURAL ANTHROPOLOGY  
AND  
MICROFILM COLLECTIONS OF MANUSCRIPTS  
ON THE MIDDLE AMERICAN CULTURAL ANTHROPOLOGY

Series: **XV**

No: **95**

Photographed by:

Department of Photoduplication - The Joseph Regenstein Library  
University of Chicago - Chicago, Ill. 60637

REDUCTION RATIO: **12**