



Australian
National
University

THESES SIS/LIBRARY
R.G. MENZIES LIBRARY BUILDING NO:2
THE AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA ACT 0200 AUSTRALIA

TELEPHONE: +61 2 6125 4631
FACSIMILE: +61 2 6125 4063
EMAIL: library.theses@anu.edu.au

USE OF THESES

This copy is supplied for purposes
of private study and research only.
Passages from the thesis may not be
copied or closely paraphrased without the
written consent of the author.

TEMPORAL ASPECTS OF JAPANESE
AND AUSTRALIAN CONVERSATION

Rob H. Elzinga

A thesis submitted in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy
of the Australian National University

February, 1978

This thesis describes original research carried out by the author during the tenure of an Australian National University Postgraduate Research Award in the Department of Psychology of the Australian National University from May, 1975 to May, 1978.

Rob H. Elzinga

ACKNOWLEDGEMENTS

I am very grateful to all the persons who have contributed in various ways to make the present research possible. I owe a great deal to my Japanese friends and colleagues, who initially suggested I undertake further research.

I am indebted to my supervisor, Professor Pat Pentony, whose patience and fruitful suggestions helped me through the more difficult phases of the project.

I sincerely appreciate and acknowledge the inspirational and technical guidance of Dr. Adam Kendon.

Special thanks are due to Mr. Alec McHoul, for teaching me the rudiments of transcribing conversations, and to all the other members of the interdisciplinary workshop on the analysis of natural conversation.

To all Japanese and Australian participants in the investigation, I would like to extend my appreciation, for their time and co-operation.

I also want to thank my colleague post-graduate students and academic and technical staff, in particular in the Department of Psychology who have all given valuable suggestions and assistance. I acknowledge the Australian National University for awarding me a post-graduate research scholarship.

I am grateful to Ms Norma Chin for typing the thesis and Mrs. Jill Ruse for preparing the graphs.

Finally, I thank my wife, Lynn, who provided, in more than one way, the link between Japan and Australia.

CONTENTS

ACKNOWLEDGEMENTS	iii
LIST OF TABLES AND GRAPHS	viii
ABSTRACT	x
 CHAPTER I. INTRODUCTION	 1
1.1 General Framework	1
1.1.1 Reasoning and Rhetoric	2
1.1.2 Value and Value Orientations	2
1.1.3 Audible and Visible Behavior	4
1.2 Methodological Framework	9
1.2.1 Basic Concepts	9
1.2.2 Movement	13
1.3 Approaches to the Study of Communication	15
1.3.1 The External Variable Approach	18
1.3.2 The Internal Variable Approach	20
 CHAPTER II. LITERATURE REVIEW OF SELECTED VARIABLES	 23
2.1 Introduction to Face-to-Face Interaction	23
2.1.1 Gaze-Direction in Interaction	24
2.1.2 Methodological Issues and Some Research Findings	25
2.1.3 The Functions of Gaze and Some Research Findings	30
(a) Expressive	30
(b) Cognitive	32
(c) Monitoring	33
(d) Regulatory	34
2.2 Speaker Turns in Interaction	35
2.2.1 Interaction Chronography	36
2.2.2 Ethnomethodological Approach	41
2.2.3 Signals for Turn-Taking	45
2.3 Listener Responses in Interaction	48
2.4 Gaze-Direction, Speaker Turns and Listener Responses in Japanese Interaction	51

CHAPTER III. TEMPORAL ASPECTS OF CONVERSATION	58
3.1 Introduction to Part II	58
3.1.1 Overview of the Studies	58
3.1.2 Participants	60
3.1.3 Procedure	61
3.1.4 Material	62
(a) Filmed Material	62
(b) Video Material	63
3.2.1 Gaze Direction, Speech and Listener Responses comparing one J-J with one A-A Dyad	65
3.2.2 Method	
3.3 Results	70
3.3.1 Gaze-Direction in General, when Speaking and when Listening	70
3.3.2 Speaker Turns	76
3.3.3 Gaze-Direction in Relation to Utterances	81
3.3.4 Listener Responses	90
3.4 Discussion	94
CHAPTER IV. TESTING HYPOTHESES ON TEMPORAL ASPECTS OF JAPANESE AND AUSTRALIAN CONVERSATION (VIDEO)	99
4.1 Gaze-Direction, Speech and Listener Responses for 24 Participants	99
4.2 Method	100
4.3 Results	102
4.3.1 Gaze-Direction	102
4.3.2 Speech	104
4.3.3 Listener Responses	106
4.4 Discussion	109
CHAPTER V. TEMPORAL ASPECTS OF SPEAKING STYLE IN A BILINGUAL PARTICIPANT (FILM)	113
5.1 Speaker Turns and Listener Responses in a Bilingual Participant	113
5.2 Results	114
5.2.1 Speaker Turns for Bilingual Participant	114
5.2.2 Listener Responses for Bilingual Participant	118
5.3 Discussion	122

CHAPTER VI. TEMPORAL ASPECTS OF INTERCULTURAL CONVERSATIONS (VIDEO)	131
6.1 Gaze-Direction, Speech, Frequency of Listener Responses and Gesticulation in Intercultural Interactions	131
6.2 Method	132
6.3 Results	133
6.3.1 Gaze-Direction	133
6.3.2 Speech Data	137
6.3.3 Listener Responses	141
6.3.4 Mean Time Spent Gesticulating	145
6.4 Discussion	145
CHAPTER VII. FINAL DISCUSSION AND CONCLUSIONS	153
7.1 Summary and Integration	153
7.2 Interactional Rhythms	162
7.2.1 Macro Rhythmic Levels	164
7.2.2 Micro Rhythmic Levels	165
7.2.3 Molar Rhythmic Levels	168
REFERENCES	173
APPENDIX I Symbols Used in Transcription	189
APPENDIX II Listener Responses (Pilot Study)	190
APPENDIX III Frequency of q-Directed Gaze	191
APPENDIX IVa Gaze-Direction during a 5-Minute Period (Japanese)	192
APPENDIX IVb Gaze-Direction during a 5-Minute Period (Australian)	195
APPENDIX V Frequency and Duration of Speaker-Turns	198
APPENDIX VI Data for Individual Participants of q-Directed Gaze Only	199
APPENDIX VII Data for Speech for 7 Japanese Interactions during P ₂	200
APPENDIX VIII Data for Speech for 5 Australian Interactions during P ₂	200

APPENDIX IXa	Speaker Turns (Japanese Condition)	201
APPENDIX IXb	Speaker Turns (Australian Condition)	202
APPENDIX X	Gaze-Direction in Intercultural Interaction (Random 10 Second Intervals)	203
APPENDIX XI	Gaze-Direction in Japanese Participants in J-J and A-A Interactions (N = 9)	204
APPENDIX XII	Gaze-Direction in Australian Participants in A-A and A-J Interactions (N = 7)	205

TABLES AND GRAPHS

Table I	Gaze-direction during a 5 minute period (CJA II/2 and CJA III/1)	72
Table II	Frequency and average duration of q-directed gaze and mutual gaze	75
Table III	Speaker turns (CJA II/2 and CJA III/1)	77
Table IV	Pauses, switching pauses, simultaneous speech (CJA II/2 and CJA III/1)	79
Table V	Speaker turns, frequency of words, word duration (speech analysis 2, CJA II/2, CJA III/1)	80
Table VI	Frequency of utterances in relation to speaker's gaze-direction at both beginning and end of the utterance	82
Table VII	Latency between speaker utterances in relation to speaker's gaze-direction	86
Table VIII	Relationship between duration of speaker utterances and q-directed listener gaze	89
Table IX	Frequency of listener responses for CJA II/2 for a 5 minute period (Japanese)	92
Table X	Frequency of listener responses for CJA III/1 for a 5 minute period (Australian)	93
Table XI	Gaze direction for random 10 second periods based upon 24 participants	103
Table XII	Analysis of variance for q-directed gaze for 24 participants	104
Table XIII	Frequencies and average duration of q-directed gaze for 24 participants	105
Table XIV	Means for cumulative speech and frequency of speech based upon 24 participants	106
Table XV	Frequency of listener responses in 7 Japanese dyads over a 5-minute period	107
Table XVI	Frequency of listener responses in 5 Australian dyads over a 5-minute period	108
Table XVII	Speaker turns (CJA III/2 and CJA III/3 — bilingual participant)	115
Table XVIII	Average duration of speaker turns and switching pauses (bilingual participant)	118

Table XIX	Speaker turns, frequency of words, word duration (speech analysis 2, bilingual)	119
Table XX	Frequency of listener responses for CJA III/3 (Japanese condition)	120
Table XXI	Frequency of listener responses for CJA III/2 (Australian condition)	121
Table XXII	Gaze-direction in intercultural interactions	134
Table XXIII	Gaze-direction in Japanese participants in J-J and J-A interactions (N = 9)	135
Table XXIV	Gaze-direction in Australian participants in A-A and A-J interactions (N = 7)	136
Table XXV	Data for speech in intercultural interactions	139
Table XXVI	Data for speech in Japanese participants in J-J and J-A interactions (N = 9)	140
Table XXVII	Data for speech in Australian participants in A-A and A-J interactions (N = 7)	140
Table XXVIII	Frequency of listener responses in intercultural interactions	142
Table XXIX	Frequency of listener responses for Japanese in intra- and intercultural interactions	143
Table XXX	Frequency of listener responses for Australians in intra- and intercultural interactions	144
Table XXXI	Gesticulation in intra- and intercultural interactions	146
Graph 1	Frequency distribution of q-directed gaze (in percentages)	73
Graph 2	Frequency distribution of speaker turns (in percentages)	78
Graph 3	Frequency of q-directed speaker's gaze at beginning and end of utterances	84
Graph 4	Frequency distribution of speaker turns for bilingual participants (in percentages)	117

ABSTRACT

Recent years have seen a proliferation of cross-cultural studies in many areas, but the cross-cultural study of actual face-to-face interactions has been paid lip service only.

Within the study of face-to-face interaction, conversations have received most attention. Although no proper framework is generally accepted, some features of conversations are by now quite well documented, at least for English-speaking persons. In any conversation, irrespective of its purpose, contents, length, outcome or language used, participants typically alternate as speaker and listener. This provides a natural orderliness to the behavior of the participants that can be described and measured.

It remains an important task to include speakers of languages other than English into our investigations of conversational behavior. The present investigation is concerned with the general question of how the use of language employed for conversation affects the organization of it. The thesis is organized in three main parts.

The first part consists of two chapters. In the first chapter a general framework and the methodological approaches in the study of face-to-face interactions are outlined. Broad areas of special importance in comparing communication cross-culturally are discussed. Because *what actually happens during an encounter* is crucial, our focus is on observable behavior. Following the lead of interaction theory, we adopt a framework to describe what occurs in interaction in universal operations, namely temporal patterning. Concepts such as communication, interaction and movement as well as methodological approaches, traditionally employed in the area are discussed.

In chapter II, literature on selected variables - gaze-direction, speaker turns and listener responses - found to be important in conversation, is reviewed. This literature mainly deals with conversations between English-speaking persons, but some literature relevant to Japanese interactions is also discussed. The present

study investigates the temporal patterning of conversational behavior as related to whether the conversations are conducted in Japanese or in English.

The second part consists of four chapters, reporting the results of four studies of two types (film and video). As an introduction to part II an overview of the four studies and the methods and procedures used is given. The actual data part of the study commences in chapter III where a detailed study, comparing one Japanese-Japanese and one Australian-Australian conversation, is reported. In this investigation we explore whether there are any differences in the temporal patterning of conversational behavior according to whether the conversations are conducted in Japanese or in English. It is found that the Japanese conversation has a different temporal organization from that of the Australian-English conversation. All the behavioral variables investigated are shown to be closely related. Turns at talk in Japanese are briefer, other-directed glances last shorter and in general the Japanese conversationalists maintain a higher rate of response than is the case in the English conversation. It is suggested that the regulatory function of gaze-direction, in particular at the end of utterances, is possibly of varying significance for different cultural groups.

Based upon the initial isolation of this pattern, a number of hypotheses, testing aspects of it, are investigated in the second study, reported in chapter IV. Following an external variable approach, results, based on twelve conversations recorded on video, reveal shorter cycles for speech and gaze-direction and a higher number of listener responses in conversations where the participants speak in Japanese than in conversations where English is spoken.

The third study (on film), reported in chapter V, examines in detail the relationship of temporal aspects of speaking style in a bilingual participant. In the Japanese condition, interacting with another Japanese, the bilingual participant speaks in turns of shorter duration, with a higher frequency of listener responses, than in the Australian condition, interacting with another Australian. Rather than explain the results in terms of code-switching and two clearly distinct speaking styles, we found evidence that a fusion of

different aspects of speaking style in the two conditions described our results better. Some factors relevant to differences in the structure between the Japanese and the English language are briefly discussed, and further research in this area suggested.

In chapter VI, a study of nine intercultural conversations on video, in which Japanese participants converse with Australian interactants in English, is reported. This investigation explores what happens to the temporal organization of a conversation in such a situation. Following an external variable approach, the important implications of the temporal organization and its change in this condition, are discussed. The findings are explained in terms of the different functions of gaze-direction and the relationship between the variables. One interesting result is the increased use of gesticulation for Japanese participants speaking English.

The third part of the thesis consists of chapter VII. Here the main conclusions are integrated and some further lines of inquiry indicated. We restate that the temporal structure of a conversation is different according to the language that is used and point out the importance of temporal organization of conversation for any theoretical and methodological framework in the area of face-to-face interaction. A suggestion for a possible link between different child rearing practices in the two cultures and taking turns in conversation is made. Finally, on a more speculative note, to attempt to substantiate the notion of temporal or rhythmical organization, the way in which various researchers have investigated rhythms is reviewed, and further investigation on the molar rhythmical, or turn, level suggested.

CHAPTER I

INTRODUCTION

1.1 GENERAL FRAMEWORK

When we are comparing the way in which persons from different cultures interact, there are certain considerations that arise out of cultural differences.

Perhaps it is fairer to say that we take some things for granted when discussing communication within one culture which require special attention when comparing communication cross-culturally.

These considerations arising out of cultural differences are most clearly manifest in different "rules" persons from different cultures employ in social interaction. Each participant in an encounter is backed up by accounts for his actions in terms of a unified set of rules and plans and a coherent system of meanings [Harré and Secord, 1972].

Naturally individuals too differ in their set of rules, but when they are of the same cultural background these differences are likely to be less discrepant than when they are from different backgrounds.

"Rules" are not the only category in which differences occur. Basically we can distinguish three broad areas which deserve special consideration when we compare communication cross-culturally. In descending order of difficulty to make them amenable to study, they are: (1) "reasoning", "rhetoric", thinking patterns, cognition or some other term for the way in which we organize our thoughts and resolve problems; (2) value and value orientations; and (3) audible and visible behavior [Condon and Saito, 1974].

Before elaborating on what might be included in each of these broad areas it has to be realized that they are closely related and

exert mutual influence on each other.

1.1.1 Reasoning and Rhetoric

In this broad area we are referring to the way in which we organize our thoughts and resolve problems in order to structure the world in which we live. "Ethnorhetoric", cognition and perception are terms used to describe some of these.

Although this area is largely a blend of values, audible and visible behavior, it is often treated in the literature as separate. One reason for treating it separately is the assumption that while values and audible and visible behavior differ at least to some extent from culture to culture, what is "reasonable" does not or at least should not. In the rhetorical tradition — from before Aristotle continuing down to the present time — there is almost no concern with cultural differences. What is logical in one place should be logical in another and yet this does not seem to be quite the case. The saying about someone from a different culture — "a foreigner can never come to understand what is going on in the minds of these people" — illustrates this point.

There have been attempts especially in the past few years to describe culturally influenced ways of posing and solving problems, ways of asking the "right questions" and giving culturally appropriate answers to them [Moore, 1967; Nakamura, 1964, 1967; Cole and Scribner, 1974].

In its totality, however, we still know very little about the relationship between culture and ways of thinking. It seems highly likely that research in this area will expand considerably.

1.1.2 Value and Value Orientations

Values or value orientations are perceived as the organization of data, attempts to grasp a whole pattern, and also abstractions from

discrete bits of evidence that gave shape to that pattern [Condon and Yousef, 1975, p.50]. They are abstractions, concepts, or ways of organizing a large amount of apparently unrelated behaviors.

A value system represents what is expected or hoped for, required or forbidden. It is not a report of actual conduct but is the system of criteria by which conduct is judged and sanctions applied [Albert, 1968]. Kluckhohn [1951] pointed out that about the only defining point about values which is generally agreed upon is that they have to do with normative as opposed to existential propositions. Values have to do with what is judged good or bad, right or wrong. Statements based on values describe the ideal, the standards by which behavior is evaluated; they do not necessarily describe the actual behavior.

A culture's values give it a sense of unity, consistency and, above all, helps its members to make choices. Traditionally, researchers have attempted to describe human behavior in terms of biological, psychological and sociological universals. Individual, social and cultural behaviors are interpreted within this framework, providing a common denominator for all societies. Some such systems are hierarchical, requiring that certain needs be satisfied before others, while other systems are simpler listings of "basics", which can be symbolized in many forms. Food, love, recognition, territorial rights, and so on, are basics which have been assumed to be universal and thus are a basis for comparing societies (cf. Condon and Yousef, 1975].

These attempts to compare the value systems of different cultures have not been very successful. An important reason for this is in our opinion the ambitiousness of such an undertaking. Only in recent years improved measurement techniques and improved operationalization of constructs included for investigation have met with more success (for an example see Triandis *et al.* [1972]).

One development which might contribute to a renewed interest in studying values cross-culturally, albeit at a lower level of abstraction, is the emergence of "rules" as a theoretical construct. Following the success obtained by linguists and sociologists,

psychologists are now also trying to incorporate "rules" into their theories [Harré, 1974; Segal and Webb Stacy, 1975]. It is not difficult to see that there exists a close relation between values and rules. Values are a higher order set of rules (a subset of rules related to certain behaviors is a value). The difference between a rule and a value lies in that a rule prescribes an action and is far less abstract than a value [Harré and Secord, 1972].

From our point of view, the cross-cultural study of the rules followed in specified situations could contribute significantly to an increased understanding of interaction. Such a study could also be directly related to observable behavior, the topic of the present study.

1.1.3 Audible and Visible Behavior

As far as the two areas we have briefly discussed are concerned, one can say that they can only be studied by inference. Both values and thinking patterns are concerned with abstract concepts which, so to speak, "exist within a person's head". Studies dealing with these issues typically make use of questionnaires, interviews, surveys and the like. Behavior pertinent to the broad area referred to here, however, can be observed directly, without inference. One can hear and see people interact. Of course cognitive behaviors may mediate the processes of interaction and participation, but only the visible and audible behaviors are the media or code of communication [Schefflen, 1973]. This is not only so for the investigator but also for the participants themselves, those "doing the interaction".

The term interaction is often not clearly specified (see also next section) with the result that the concept remains somewhat vague and elusive. In a review article on intercultural communication, Brein and David [1971] point out that not the social interaction or contact experience *per se* or the way people are influenced by it, but the processes that occur during an encounter are of crucial importance. As yet very few researchers have taken up the challenge to investigate *what actually happens during an encounter cross-culturally*. This is,

to say the least, somewhat surprising because the relevant methodology to do just that has been available for some time. As early as 1940, Chapple and Arensberg put forward a series of operations to describe what occurs in interaction in such a way that a universal application of their methods was possible. Rather than asking questions such as what? why? where? and how?, which are not universal and often inferential, these authors proposed to ask who? when? in what order? and how often?, questions which are universal and observable without prior knowledge of the events observed [Arensberg, 1973]. Admittedly, in the days of Chapple, observation of interaction was somewhat selective in practice.

The availability nowadays of sound-, video- and film-recordings has made it much more feasible to study what exactly happens during an encounter. Using these types of permanent records allows one to analyze the same data over and over, making possible much more precise analysis than would otherwise be the case. When comparing interactions cross-culturally, audible and visible behavior include the most clear-cut differences between members of different cultures. Language, one important kind of audible behavior, is probably the most obvious barrier to understanding someone from another country. Language is a system which can be described and studied and this has been done for most languages in some detail. A large part of the curriculum in secondary education is spent in learning another language. The value of this education lies presumably in the fact that it allows one to come into direct contact with a different culture and this is believed to be beneficial to the learner. One important aspect of language is that, when one does not understand a language, not only does one not know the words, but one also knows that one is not understanding [Condon and Yousef, 1975]. For the other part of observable behavior, what we have referred to as visible behavior, but which is more popularly called "nonverbal" behavior, this is not necessarily the case. One can perceive another person smiling, nodding his head and so on, and think to understand its meaning in a particular context, but actually misinterpret it. Learning another language, difficult and painstaking though it may be, is less of a problem than being sensitive to the range of "nonverbal"

behaviors that differ [Bell, 1973]. The popular terminology "non-verbal" behavior is really a misnomer, for it suggests more clearly what is not included than what is. Moreover the distinction between pure language and kinesic, or bodily behavior, is a rather arbitrary one.

We chose to talk about audible and visible behavior because the two go together and are intrinsically related. All the variables traditionally called "nonverbal" such as head-nods, gestures, posture shifts, intonation, gaze-direction and so on, more often than not, accompany speech rather than substitute for it. One of the greatest contributions of pioneers like Birdwhistell [1970] was to show that kinesic movements are not random but closely related to speech. Unlike language, it is only recently that these behaviors have received the attention of researchers, that is to say in any systematic fashion. From time immemorial people probably had some understanding that persons from different cultures not only spoke a different language but also behaved in a different fashion, in speaking that language.

Efron [1941], another pioneer in investigating the influence of culture on gesticulation, compared the way migrants from European Jew and Southern Italian background living in New York gesticulated while speaking. Birdwhistell [1970, p.102] again writes how he had access to newsreel films just after the second world war on which a popular politician spoke Italian, Yiddish and American English: "To me, at that time, the astonishing thing was that even with the sound removed, any observer who knew the three cultures could immediately detect whether he was speaking English, Yiddish or Italian." What Birdwhistell's anecdote and Efron's more systematic study showed is that, without actually hearing what was said, there is something characteristic about the behavior of the person which is typical of the language he is speaking. If this is so, it follows that learning another language not only involves learning the grammar and the proper pronunciation and intonation, but also a certain behavioral style typical to that language. Or as Lomax [1977] puts it more broadly and elegantly: "Culture and communication are both rooted in co-identification — in common ways of handling space, force, timing,

dominance, sequence, sex, or organizing and synchronizing activity and interaction. Such shared patterns of identification enable culture members to calibrate and synchronize their behavior, thus ensuring the ongoingness of work, social life, and, of course, conversation. They form the warp of communication through which the web of particular relationships, acts, and interchanges is woven. Cultural style thus shapes the river-beds along which the streams of human dialogue flow." While a few researchers have dealt with the study of interaction cross-culturally in a broad sense, even fewer have concerned themselves with conversations cross-culturally.

And yet conversation is the sub-form of interaction we engage in for a large part of everyday life. Within interaction in a single culture, it is the conversation which has received most attention from researchers, in particular in English-speaking countries.

The studies to be reported in this thesis are concerned with the study of conversation cross-culturally. In particular we will focus our attention on the relationship of language/culture and behavioral style.

Our earlier remarks about available methodology to make such a comparison possible, led us to adopt a framework closely related to interaction theory [Collins and Collins, 1973]. For this reason we chose to concentrate on temporal aspects of conversations. To do so allowed us to quantify our audible and visible variables in a way which is universal, avoiding as much as possible problems of inference. As Chapple *et al.* [1970] have pointed out, rhythm, or rather various rhythms, act as the organizing principles for activities and interaction.

In order to study the temporal aspects of conversation cross-culturally, the way in which participants from different cultural backgrounds converse is compared. Since as we have pointed out we know that the form of behavior and relations are shaped by sex, social class and so on we try to hold the different contexts of the participants relatively constant while observing variations depending upon their cultural background [Schefflen, 1973]. In particular these

backgrounds in the present study are Japanese and Australian. Australians are studied because the study is carried out in Australia. The reason for choosing Japanese participants is threefold:

- Firstly the author has spent over three years in Japan, prior to commencing his research in Australia and therefore has some knowledge of the Japanese language and some understanding of Japan.
- Secondly Japan has been isolated from Western influences for a considerable period of time and has therefore presumably developed more or less independently from Western influences until quite recently (1868).
- Thirdly the way Japanese people interact with one another appeared to the author intrinsically different from the way either Dutch- or English-speaking people do. In this respect an observation by McDermott and Wertz [1976] is interesting. According to these authors "culture and personality theorists have recognized that a Japanese reality, for example, is different from an American reality; but they have seldom given attention to how Japanese realities get done in everyday life, how the Japanese self has, to borrow a phrase from Hallowell, a behavioral environment that allows the self to maintain a Japanese reality" (p.170). What can be more typical of everyday life than conversation?

The author's interest in interaction in a broad sense, and more specifically in relation to clinical populations, started some 10 years ago [Elzinga, 1972]. Three years of research carried out in Japan increased my awareness of the lack of knowledge about interaction-processes in general. At the same time it became more obvious how different "realities" can be accomplished in various ways. With this frame in mind, the topic for the present study emerged. Because of my background in clinical psychology the approach taken here is somewhat akin to a clinical case study approach in which the Japanese and Australian conversation constitute the case-study. A similar method could be applied to other cases.

Most of the studies to be reported compare the way two Japanese converse with the way two Australians do so — this is to say a cross-cultural comparison. Another study deals with the way a bilingual

participant converses while speaking Japanese and while speaking Australian. The final study investigates the way Japanese participants converse with Australian participants — i.e. an intercultural comparison. A central question in all studies reported when we analyze the data is the extent to which there is a mutual relationship in the temporal aspects between certain visible and audible elements.

Summarizing we have discussed three broad areas of special importance in comparing communication cross-culturally. We briefly pointed out what is to be included in the areas — patterns of thinking, values and visible and audible behavior — pointing out what is of crucial importance in interaction is what actually happens during an encounter. Therefore we limit our focus to the study of observable behavior only — thus audible and visible behavior. Following the lead given by interaction theory we decided to focus upon universal operations to describe what occurs during conversation — the most commonly engaged in subset of interaction.

In this first section we have used terms such as communication, interaction in a broad sense, without actually defining what we understand these terms to mean. Before discussing the specific variables investigated, we will first define these terms more precisely and provide a theoretical and methodological framework, traditionally employed for the study of interaction in general.

1.2 METHODOLOGICAL FRAMEWORK

1.2.1 Basic Concepts

According to a dictionary of psychology [Chaplin, 1967] the term communication refers to: (1) the transmission of energy changes from one place to another as in nervous transmission; (2) the process of transmitting and receiving signals or messages; (3) a message or signal in itself; and (4) information given by a client to a psychotherapist. The word "communication" is derived from the Latin "communis" which means "common". As Kendon [1974] points out, if we

study how behavior functions in communication, the question of interest will be different, according to which sense of the term "communication" we follow.

When using the term communication it is important to specify to which meaning of the term one is referring. For the present purposes two of the meanings: "it refers to a message or signal, or rather the ideas or information conveyed by it" and secondly "it refers to a process or system by which these messages are conveyed" are important.

In its "message" meaning we focus our attention on the individual person as an entity who has some news to transmit. We study his behavior in order to find out what that news might be.

In its "process" meaning we are not so much interested in the message, but in the behavioral relationship between individuals and our attention will be focussed on the description of this behavioral relationship and its function [Kendon, 1974, p.526].

Historically psychologists in particular have been largely concerned with the behavioral systems of individuals and the message-contents to the relative neglect of the process by which individuals interact [Danziger, 1976]. The behavioral relationship between the activities of two or more individuals in a social system during interaction is of relatively recent interest. In the next section it will be mentioned that the technical innovation of cinematography as well as the development of an appropriate theoretical framework were necessary for this shift in emphasis. Only in the last twenty years have a number of researchers investigated the basic processes in interaction rather than being solely concerned with the outcome of the interaction. Because of this shift in emphasis the term communication is surrounded with confusion. Even nowadays many psychologists using the term communication insist on the inclusion of message-content intended by the person communicating.

Ekman and Friesen [1969] define communicative behavior as "those acts which are clearly and consciously intended by the sender to transmit a specifiabile message to the receiver" (pp.55-56).

von Cranach and Vine [1973] define communication as "the conventional exchange of messages between individuals in interaction. It depends on the existence of a system of rules, the code, that is, at least in part, common to both communicating partners and relates specific signs or sign-systems to specific meanings" (p.17).

von Cranach and Vine's definition is wider than the one provided by Ekman and Friesen, in which message-content and intention are prerequisites. Birdwhistell [1970] and Scheflen [1966], however, argue that since the behavior of one interactant influences the behavior of the other interactant the term communicative has to be taken broadly.

In their sense of the term, any aspect of behavior, not just verbal behavior or not even just intended behavior, may be communicative.

Any action of one person may "communicate", whether or not it communicates what was intended, and even whether or not the person was aware of having communicated anything at all. Moreover any action is always multiply communicative. Kendon [1977a] explains how when one person says something to another, then not only do the words uttered convey something, but the very occurrence of the utterance as opposed to its non-occurrence is informative. And the tone of voice in which the utterance is spoken, with a certain body motion, posture and so on, all reveal something about the utterance. So the context in which any communicative act occurs is considered to be of crucial importance.

For researchers as Birdwhistell, Scheflen and Kendon, communication is synonymous with the organization of the behavioral relationship of interacting persons. For these authors, communication and interaction are terms which are interchangeable, because for them it is the process between individuals which is important.

For Ekman and Friesen [1969], however, the term interaction has a much more limited meaning: "They are acts by one person in an interaction which clearly modify or influence the interactive behavior of the other person(s)" (p.56). Because Ekman and Friesen make intention a prerequisite for any behavioral action to be communicative,

they distinguish informative acts that influence the interaction (and are therefore interactive if there was no intent to communicate). For Schefflen *et al.* [1966] any behavioral action may be communicative because one *has to* see it in relation to what the other person in the interaction does.

Even attempts to avoid interaction or any reactions to each other's presence and appearance "communicate". For Schefflen, Birdwhistell and Kendon, one cannot not communicate. Whenever two people are in each other's presence communication takes place.

An interesting point about the term interactive is made by von Cranach and Vine [1973] who point out that all behavior may be treated as interactive if it embodies some reference to a partner, even in the absence of direct face-to-face interaction.

Throughout this thesis communication and interaction will be used interchangeably. Because of the misunderstanding the use of the word communication creates, the term interaction is perhaps preferable. What it refers to is what actually occurs in the interaction, that is to say its process meaning and the way in which persons mutually influence another. Perhaps less disagreement exists about the term "informative". Informative behavior conveys information about the state of the individual to an observer.

Goffman [1963] made the distinction between "giving information" and "giving off information": "An individual may give information through the linguistic means formally established in society for this purpose, namely, speech or recognized speech substitutes such as writing and pictorial signs or gestures. One speaks here of an individual sending messages to someone who receives them. But the individual may also give information expressively, through the incidental symptomatic significance of events associated with him. In this case one might say that he emits, exudes, or gives off information to someone who gleans it" (p.13).

Any behavior, for example when I adjust my hair, is informative. Many informative behaviors are concerned with the smooth operation of one's own behavioral performance. Many informative behaviors, which

can only be detected after detailed analysis, are self-informative, which can be detected by a scientific observer but not necessarily by a fellow interactant [von Cranach and Vine, 1973]. For example, many parts of the body may be synchronized to a person's own speech or to the speech of the other [Condon and Ogston, 1966]. Detailed analysis of a film might reveal this, but the interactants themselves might not detect it.

1.2.2 Movement

In order to study and quantify systematic behavior a crucial aspect of it is that it occurs in *spatial and temporal* contexts. These contexts involve both history and present environment that together define the interaction situation [von Cranach and Vine, 1973, p.15].

In the studies to be reported our main focus will be on the temporal dimension and organization of interaction in its present context. The spatial dimension in its present context will only be used in so far as certain variables investigated such as head-nods and gestures require it. The historical environment, incorporating both biological inheritance as well as cultural tradition, must be regarded as one in which human biology and culture interact in a mutually dependent process [see Williams, 1975]. Behavior systems of interacting individuals can be quantified to a certain extent independently, that is to say without taking the behavior of the co-participant into account.

Most studies investigate some aspect of movement, the spatio-temporal change both of the whole body in relation to the environment, and of parts of the body in relation to each other [von Cranach and Vine, 1973, p.15]. The choice of certain spatial and temporal dimensions depends largely on theoretical presuppositions and the development and availability of technical equipment used to measure these dimensions.

Usually movement is studied directly, as in research on gestures

and gaze-direction. Sometimes, however, movement is studied by analyzing some of its immediate consequences as in speech, i.e. the acoustic consequences of movements of the vocal organs [von Cranach and Vine, 1973].

In the studies which follow, movement is investigated in gaze-direction, i.e. whether a person is looking at the facial region of his co-interactant or away from it; in head-nods, which is moving the head from one position to another; in speech, indirectly as the acoustic consequences of movements of the vocal organs; in gestures, as movements of the hand or of both hands while speaking. These movements are investigated to a large extent in relative isolation, but attempts are made to relate variables to one another. The temporal aspect of these movements is the cornerstone in each of the investigations.

Von Cranach and Vine [1973] contrast movement with appearance, which includes static features and a limited number of dynamic features (such as blushing). Movement aspects or dynamic aspects are contrasted with static aspects. When something is static there is no spatial temporal change. Some parts of the body can be static in relation to the body as a whole while some other parts are in motion. Or, the body may be static in relation to the environment, while some of its parts may be moving in relation to each other. For example, a person can be observed in a certain posture, while he is speaking and waving his hands about in order to gesticulate. von Cranach and Vine point out that what counts as movement and what as static is, in a sense, arbitrary; it depends upon the spatio-temporal span of our conceptual or measurement focus.

With presently available high speed cameras the span can be considerably expanded. In an example of micro-analysis, Condon [1973] shows how a number of movements involving the head, the eyes, the mouth, four fingers and one shoulder all sustain their relationship to each other during $3/48$ ths of a second, which it takes the speaker to utter the first phone of the four involved in the word "ask".

In practice, most investigators of movement do not go into such

detail. As a matter of fact, most investigations of movement use successive static positions for research rather than movements *per se*. In relation to this point the footnote on page 63 takes on special significance (on the difference between frame-by-frame and movement phrase boundary analysis).

The term appearance includes things like dress, hair-style, skin-colour and so on. A lot of research on facial expressions also investigates static features, since most often photographs of particular facial expressions are used as stimuli.

In the following studies we do not concern ourselves directly with appearance aspects of the participants involved, although it must be clear to everybody that Japanese and Australian participants are different in appearance.

Summing up we have discussed the different ways in which terms such as communication and interaction are used in the literature. We will use both terms throughout as referring to the behavioral relationship between individuals, with a preference for the term interaction. In order to quantify these behavioral relationships a crucial aspect is that both the behavior system of individuals as well as the behavioral relationship between individuals occur in spatial and temporal contexts. Movement, which is the spatio-temporal change of parts of the body or the whole body in relation to the environment, is contrasted with appearance. In the studies to be reported the main focus will be on the temporal organization of movement in the present environment, while we pointed out that the historical environment must be regarded as one in which biological factors and cultural factors interact in a mutually dependent process. Before discussing the specific variables included for investigation we will discuss the different methodological approaches commonly used to study communication.

1.3 APPROACHES TO THE STUDY OF COMMUNICATION

In this section we will discuss briefly the major research

orientations in the study of communication. It has to be realized, however, that, while distinguishing a number of approaches for orientation, it is often difficult to slot any one researcher into a particular tradition. Within each of the major approaches, researchers differ from each other in that they emphasize certain aspects of the approach more than other aspects.

Initially Duncan [1969] distinguished the external variable approach from the structural approach, while emphasizing that the two approaches are complementary. Some years later von Cranach and Vine [1973] in their introductory chapter distinguished four approaches: (1) the assessment of human movement dimensions; (2) the external variable approach; (3) the ethological approach; and (4) the structural approach.

The emphasis in this section is upon methodology, although certain theoretical considerations have to be taken into account to make methodological prescriptions meaningful. The approaches outlined by von Cranach and Vine can be ordered to the extent that theoretical considerations influence methodology. The assessment of human movement approach does not depart from any theoretical viewpoint. In this approach quantification of movement based on instrumental measurements or systematic observation methods is the primary aim. The external variable approach's contributions are mainly in terms of prescriptions for designing an experiment to test a specific hypothesis and using statistical procedures to analyze results. Toward the other end of the scale the ethological approach is deeply rooted in evolution theory. In this approach, researchers typically apply the theoretical and methodological principles of animal ethology to human communication. The structural approach is most specific about its theoretical commitments. In this approach communication is studied as a tightly organized and self-contained social system, which operates according to a definitive set of rules. The method followed by most researchers in this tradition is known as context analysis. In his detailed exposition of the method of context analysis, Schefflen [1973] points out that the birthdate of context analysis can be considered to be a multidisciplinary collaborative study-group studying in Stanford in 1956. This group — Bateson, Birdwhistell, Brosin, Fromm-Reichmann,

Hockett and McQuown — studying a psychiatric interview from many different perspectives has been extremely influential for the study of communication. Many of the group's original ideas have been elaborated upon further by Scheflen [1966, 1968, 1973], Kendon [1972, 1977b] and others.

Scheflen [1973] describes five major operations involved in context analysis as: (1) obtaining an audiovisual record of a transaction; (2) mapping the behavioral events; (3) delineating the units of communicative behavior; (4) determining the units of transaction or communication; and (5) the contextual analysis of the behavioral processes. Scheflen acknowledges that his account of the method of context analysis is somewhat idealized and that the criteria may be unnecessarily stringent for many researchers' purposes: "Many researchers do not care to use all of these operations. Some, for instance, may not be interested in cross-cultural sampling, while others may be interested only in comparing certain forms of behavior cross-culturally. Neither of these may wish to study the total structure of any given transaction" (p.312).

The five operations described do, however, cross the boundaries of older disciplines and Scheflen offers his account as a broader framework than the compartmentalized ones of the past.

Besides the influence of theoretical approaches and the development of an appropriate theoretical viewpoint, the technical innovation of cinematography has also played a part. Kendon [1977a] traces the historical use of film recordings in psychology and ethnography as well as the emergence of a theoretical viewpoint in which the *behavioral relationship* between interactants is the primary focus of attention. Kendon points out that although film has been used for more than fifty years, the appropriate theoretical viewpoint for the study of interaction has only been available for the last twenty years.

Without going into detail about their method of studying conversations [Sacks, Schegloff and Jefferson, 1974] here — see the section on speaker turns — there exists an interesting parallel between conversational analysis and context analysis.

As McDermott and Wertz [1976] observe, both approaches are concerned with: "providing behavioristic accounts of how the social world is made up of people with both material and identity investments in particular ways of making sense of each other, for accounts not only of what people know, but also of how people help to generate the world in terms of whatever it is that they know helps them to get by, no matter how painfully" (p.172).

After these brief remarks about the approaches we do not employ in the present study, we will now turn our attention to the research methods followed here. Two methods will be discussed in greater detail, namely, the external variable approach and what we will call "the internal variable approach".

1.3.1 The External Variable Approach

The name for this approach was introduced by Duncan [1969] to set the more traditional social-psychological experimental method apart from the structural method. In these studies, one is seeking relationships between the (nonverbal) behaviors and other variables, such as personality characteristics, situation and observers' judgment [Duncan, 1969, p.119]. External variable studies typically place heavy emphasis on hypothesis testing usually through controlled experiments. The "externality" of the variables investigated, e.g. personality measures correlated with foot movement patterns, is characteristic for this approach. Ekman and Friesen [1968, 1969] discuss the methodology of the external variable approach most fully, but Duncan [1969], Argyle [1969, 1976] and von Granach and Vine [1973] also discuss the approach. Ekman and Friesen [1968] distinguish between indicative and communicative studies. The former measure the statistical relationship between designated nonverbal behaviors with other variables of interest.

The latter are concerned with the meanings ascribed to designated nonverbal behaviors by observers. It must be pointed out here that the meanings observers ascribe to certain behaviors it not necessarily equivalent to the meanings a person actually participating in the

interaction gives to it. Another point of interest is the reliance upon the use of statistics. Von Cranach and Vine [1973] point out that the external variable approach is mainly followed by social psychologists who stress, besides observational procedures and content analysis, the usefulness of experimental design and statistics, in particular multivariate experimental techniques. The emphasis on design and statistics led to a proliferation of rigorously designed experiments, which test hypotheses. Usually studies in which only one or two variables, preferably distinguished as dependent and independent are studied in strictly controlled laboratory conditions.

Summarizing an external variable approach study takes a more traditional psychological approach, typically seeking individual and group differences in the use of certain behaviors by relating the rates and durations of these behaviors to other external variables, such as characteristics of the interaction situation, scores on personality inventories, the reactions of judges to the interaction, or very rarely other behaviors [Duncan and Fiske, 1977, p.12].

The external variable approach has yielded a large quantity of isolated findings, which are difficult to integrate into a general theory. Ekman and Friesen [1969] do not concern themselves very much with the organization of behavioral acts into a larger framework. The approach, however, can be very useful to test hypotheses once they emerge from more naturalistic exploratory investigations.

Especially in the last few years there have been attempts to modify the strict adherence to hypotheses testing studies, rigorously designed and the use of sophisticated statistical techniques, in a move toward more naturalistic exploratory studies, using descriptive measures.

Argyle [1969] writes about a "New Look" in research on social behavior; and Duncan and Fiske [1977] advocate a move away from the strict external variable approach in favour of the study of action sequences.

To take account of some of the criticism on the external variable approach as well as to describe the methodology used by a number of

studies, which do not neatly fit into the traditional approaches, we propose the term "internal variable approach" [Kendon, 1977 — personal communication].

1.3.2 The Internal Variable Approach

This term is used to describe the methodology followed by a number of studies which fall somewhere between the structural approach and the external variable approach.

Studies adhering to this approach concern themselves with the relationship of behavioral variables within or internal to an interaction-system. In these studies, the main focus is on how one behavioral variable is related to another both within the same person as well as in relationship to another co-interactant. In these studies, it is of prime importance when and where, in the ongoing interaction, designated behavior occurs and also how this behavior relates to or is influenced by the behavior of the co-interactant participating in the interaction under investigation. Because of this emphasis, micro-analytic data analysis is often characteristic. Therefore the use of permanent records of the interactions studied, in the form of film or video-tape, is of great importance.

The approach differs from the external variable approach in a number of aspects. Firstly it starts out exploratory as opposed to hypotheses testing. Secondly, it concerns itself solely with the relationship between communicative acts rather than studying the relationship between communicative and non-communicative variables. Therefore it usually is concerned with more than just one variable in isolation. Thirdly, it is less concerned with research design and statistics.

Naturally occurring interactions are preferred as data, eliminating a great deal of experimenter influence or control. Because it is realized that the behavior of one interactant influences the behavior of the co-participant and vice versa, the use of statistical procedures is limited. The main statistics used are of a

descriptive nature. The internal variable approach differs from the structural approach in its more modest goal. Usually only a limited number of variables are studied as opposed to the structuralists's all-encompassing hierarchical structure of behavioral organization.

The internal variable approach can also be considered a sub-method of the structural — and the external variable — approach. An example of the approach is Kendon's [1967] study on the relationship between speech and gaze-direction.

Summarizing, we have briefly discussed the major research approaches used in the study of communication. The approaches differ in their theoretical starting points, resulting in a variety of methodological prescriptions. In this section, the emphasis was on methodology. For the research to be reported, two approaches are of special importance and were discussed in more detail. Initially, the internal variable approach is used to analyze in detail one Japanese-Japanese conversation and one Australian-Australian conversation. In this analysis, the emphasis is on the dyad as an integrated system. Although the variables investigated — gaze-direction, speaker turns and listener responses — are reported for each participant separately, particular attention is paid to the relationship between the variables, both within an individual and between them.

Once the micro-analytic study has revealed the relationship between variables, the external variable approach is used to test hypotheses relating to the pattern. For each of the variables a hypothesis is put forward and tested in a larger sample. Data for each interactant are entered as scores and statistical analysis used.

A study on a bilingual participant using the internal variable approach is reported. In that study the bilingual participant interacts with a Japanese in one encounter, and with an Australian in the other.

In the final part, the same variables — gaze direction, speaker turns and listener responses — in addition to gesticulation are investigated for the same individuals, in one interaction with someone of the same culture and in the other with someone of the other culture.

This study is exploratory in nature.

In the next chapter we will discuss the literature of the variables selected for study.

CHAPTER II

LITERATURE REVIEW OF SELECTED VARIABLES

2.1 INTRODUCTION TO FACE-TO-FACE INTERACTION

Interaction has been defined as the behavioral relationship between individuals. In order for people to interact with one another it is necessary that the behavioral systems of individuals are co-ordinated and integrated with each other for the interaction to be successful.

When I want to interact with you I have first to make it clear that I want to do so. I can walk up to you, I can ring you, write you a letter, make an appointment and tell you my intention. On the other hand you will have to agree to my intention or you can avoid me. In short, there are many conditions to be met, which all involve interaction to a certain extent, before any verbal exchange can take place. All these aspects are topics for research in itself. In the studies reported here however, we do not deal with these aspects [Kendon, 1973].

Our starting point is after all these preliminary steps, once two persons have entered a certain space and have a certain purpose to engage in interaction. We also concern ourselves with dyadic interactions as opposed to larger group interactions. The topic is conversations in face-to-face interactions. Conversations are probably the most well-studied episodes in interactions. A large number of variables pertinent to conversations have been documented and researched intensively.

One of the most important aspects of any conversation is that participants alternate in speaking and listening. This alternation of action sets a conversation apart from a monologue. For research purposes, the alternation of action between participants is a natural

unit of attention. The process of turn-taking involves changes in other behaviors such as gaze-direction, gestures, and so on. These changes are also related to a certain extent on who is having the turn. In order to compare conversations cross-culturally, the most logical focus of attention in our point of view is this phenomenon of alternation of actions. In this section gaze-direction, speaker turns and listener responses — variables that have received considerable attention from researchers — are discussed in some detail.

2.1.1 Gaze-Direction in Interaction

In this section we will discuss some of the research dealing with questions such as: to gaze at someone or not to gaze; when and where to look; who looks at whom and how much; how does one look; what does such and such a look mean; etc. The interest in gaze is an ancient preoccupation of man. In earlier days the mythology surrounding "the evil eye" is widely documented in the literature. For a review see Tomkins [1963].

Almost a century ago the effect and function of gaze made Magnus aware that people communicate by means of a "language of the eyes" ("die Sprache der Augen" [Magnus, 1885]).

In a phenomenological essay, Heron [1970] describes gaze, along with touch, as the most intimate mode of interpersonal encounter.

A similar point had been made earlier by Simmel [1921] who described mutual gaze as "a wholly new and unique union between two people ..., which represents the most perfect reciprocity in the entire field of human relationships" (p.358).

An existentialist such as Sartre [1943] goes even further by stating that it is through our experience of the Look (sic) that we most directly apprehend another person as being with consciousness and intention of his own. More recently, results of experimental studies led Scaife and Bruner [1975] to the suggestion that the early appearance of mutual orientation and shared attentiveness between

mother and child, marks the beginning (for the child) of "intersubjectivity", i.e. awareness of the thoughts or intentions of another. For a similar point see also Trevarthen [1975].

Only in the last decade has gaze become the object of considerable systematic research attention. Numerous studies on gaze have been carried out following various methodological approaches, using different terminologies. Some authors have suggested the importance of certain physiological processes in looking, while others omit to mention these. A number of functions of gaze have been put forward. Most studies dealing with gaze have been carried out with American- or English-speaking subjects.

2.1.2 Methodological Issues and Some Research Findings

Looking behavior has been investigated in various ways. Most of the physiological studies have followed what von Cranach and Vine [1973] call the assessment of physical movement dimensions approach. This approach aims at quantification in which the change of the eyes or certain parts of them, in relation to the physical environment, and the speed and variability of this process, are the objects of investigation. Considerable research interest was generated, following Hess [1965], in the investigation of pupil dilation and constriction under various conditions.

These conditions can be stimuli, internal to the organism, such as problem-solving or memory tasks. Both Hess [1965] and Kahneman and Beatty [1966] showed that the pupil of the eye dilates during problem-solving or memory tasks. Bernick and Oberlander [1968], however, found that talking, as opposed to thinking, produced dilation.

Or these conditions can be external stimuli, that is to say external to the organism, such as pupil constriction when subjects looked at unpleasant pictures, as found by Hess [1965].

Stass and Willis [1967] undertook an interesting study of pupil size with people actually interacting. Subjects came to take part in

an experiment and were asked privately to choose between two people as their partners. One of these people had pupils which were dilated with a drug, the other one had normal pupils. When choosing between either two males or two females, both sexes preferred the one with enlarged pupils. Relative attractiveness or unattractiveness was controlled for, so it was concluded that it was really pupil size which accounted for the participant's choices.

Hindmarch [1973] has shown that a mother's pupil dilates to the sight of her child and suggests the child, in turn, could be aware of the dilation of the mother's pupil and this could therefore be important in early dyadic communication.

The importance of these studies is to point out that a variety of stimuli cause physiological change in the size of the pupil, which in turn might be perceived by the other person with whom one is interacting.

Besides physiological movement studies, there are many more investigations in which looking, variously termed — looking behavior, gaze-direction, eye contact, mutual glances, visual interaction, line of regard, and so on — has been studied. It is with this type of study where, broadly speaking, *looking behavior* is investigated, that we are primarily concerned here. Since the term looking behavior is also used in e.g. studies of exploratory behavior, we will be more specific and describe the prime focus of the present study as gaze-direction at the face of the other or away from the other's face, as it occurs within the context of an ongoing dyadic interaction.

The direction of gaze has also been studied by many disciplines, resulting in a variety of methodological approaches and use of terms and definitions. Developmental studies of gaze-direction between mother and infant have mainly been carried out within the ethological framework. In this approach, working on the basic assumption that human behavior depends to some degree on phylogenetically inherited, adaptive predispositions, the function and structure of gaze-direction has been studied by analogy with related species. The ontogeny of facial-visual signalling is considered especially important [Vine, 1973].

There is substantial evidence of innate contributions to the widespread use of gaze-direction as a social signal [Diebold, 1968; Vine, 1973; Rosenfeld, 1977]. The onset of gazing at another organism's facial region, as a signal, indicates a readiness and intention to engage in an interaction and has been considered a "releaser" of social behavior [Ahrens, 1954; Chance, 1962]. Goffman [1963] makes a similar point, although he does not use the term "releaser". Once attention is drawn to the mother's face, the varied stimulus properties associated with it (movements and later configurations) become increasingly attractive to the developing child [Ahrens, 1954]. Vine [1970] has further argued that the evolution of the head and face area as a dense source of a variety of nonverbal signals may be associated with the fact that the human visual apparatus can only focus upon a limited area of the visual field.

Developmental studies [Stern, 1971 and 1974] have shown that the gazing patterns between a mother and her infant constitute the first dyadic system over which both have equal control. It is this early control of gaze-direction which allows the infant to regulate from instant to instant the amount of social visual contact [Robson, 1967]. Termination of gazing at another person's face reduces the intensity of an interaction and has been compared to social "cut-off" behaviors in other animal species [Chance, 1962]. There is evidence that gaze avoidance is particularly directed by the infant to his own mother [Richards, 1971]. For our purpose, the most important contribution of the developmental studies is to show the importance of innate mechanisms contributing to the use of gaze-direction in interaction.

Most studies on gaze-direction, however, employed what Duncan has termed the "external variable approach" [Duncan, 1969]. In this approach, the function or meaning of gaze-direction is inferred from its connection with non-communicative variables [von Cranach and Vine, 1973]. An example of this approach is a study by Kleinke, Meeker and La Fong [1974] in which actors portray the role of "engaged" couples in a videotaped interview. In half of the tapes the "engaged" couples gazed at each other fairly often during the interview. In the other half the couples never looked at each other. College students who viewed the videotapes and rated the couples were led to believe that

the interviews were genuine and that the couples were actually engaged. Couples who gazed at each other were rated as significantly more genuine, relaxed, co-operative, intelligent, and attentive toward each other than couples who did not gaze. In addition, couples who gazed at each other were seen as liking each other more, as having better potential for a successful marriage, and were rated as better liked by the viewers than couples who did not gaze [Kleinke, 1975]. It will be obvious that calling the subjects "engaged" bears no direct relationship to the communicative behavior (amount of gaze) investigated. This approach is mainly followed by social psychologists who stress, as was pointed out earlier, besides observational procedures and content analysis, the usefulness of experimental design and statistics.

In particular, within the so-called external variable approach, another methodological distinction is of some importance. Ellsworth and Ludwig [1972], in their review article on visual interaction, distinguish between whether gaze-direction is methodologically studied as a measured dependent or as a manipulated independent variable. A group of experiments in which gaze is studied as a manipulated independent variable are those studies where visibility is reduced experimentally and its effect measured [e.g. Argyle *et al.*, 1968]. These studies have indicated the importance of one of the functions of gaze, namely the monitoring function during conversation [Rosenfeld, 1977].

As an independent variable, in methodological terms, whether one gazes at one's co-interactant has been shown to influence emotional responses and cognitive attributions. Adults increase in autonomic arousal in reaction to being stared at [Nichols and Champness, 1971]. In another study [Kleinke and Pohlen, 1971], college men played a competitive game with a male opponent who either gazed at them constantly or not at all. The men who received constant gaze from their opponent had significantly higher heart rates than the men receiving no gaze.

Gaze can also serve as an aversive stimulus to which people react with "flight" responses. In a field study carried out by Ellsworth, Carlsmith and Henson [1972] drivers waiting at a red traffic light

were confronted by a person who stared at them until the light turned green. Sometimes the staring person was standing on a corner and sometimes he was on a scooter. Drivers who were stared at crossed the intersection with significantly higher speed than a control group, who were not stared at. This difference occurred regardless whether the driver was male or female and regardless of the sex of the staring person.

The contribution to the understanding of gaze and its functions made by studies following the external variable approach, are difficult to sum up. Many factors related to gaze-direction have been experimentally investigated. It is the external variable approach which is most frequently employed to investigate the expressive function of gaze.

A relatively small number of studies have employed what was earlier described as the "internal variable" approach. It is within this approach that the different functions of gaze have been studied in actual occurring interactions. Especially the close relationship between gaze-direction and other communicative acts, in particular the relationship between utterances and gaze-direction, have been studied in detail. Kendon's [1967] study is probably the best example of this approach. Other examples are Nielsen [1962] and more recently, albeit from a conversational analysis viewpoint, Goodwin [1975] (see also the section under speaker turns, p.46). These studies have shown how important it is to relate gaze-direction to when and where in the ongoing interaction, and when and where in relation to speech, gazing at the other's face or gazing away, occurs. It is within the micro-analytic "internal variable" approach that the present study must be placed. Rather than examining when and where in the interaction gaze at one's co-interactant occurs, most of the studies on gaze have measured its frequency, or the duration of gaze, or both as main features.

Before going on to discuss the functions of gaze, it must be pointed out that although we have outlined five main approaches, it is not always easy to classify studies mentioned in any of these approaches. Some studies combine two or more approaches in one study.

Some of the methodological issues mentioned are useful distinctions in didactic terms, but not always meaningful when we try to understand gaze behavior. For a review of the terminology used in the research on gaze see von Cranach and Ellgring [1973, pp.421-422].

2.1.3 The Functions of Gaze and Some Research Findings

The different functions of gaze can be summed up under the headings of (a) expressive, (b) cognitive, (c) monitoring, and (d) regulatory (see Kendon [1967] and Knapp [1972]). Before describing each of these functions, one must realize that *the same gaze can have any or all of these functions simultaneously*. It has taken quite some time before the functions of gaze, as outlined here, were described. Many researchers are still mainly concerned with the relationship of gaze to other characteristics of the person gazing or some external stimulus. Kendon [1967] points out how it was Goffman [1963] who described gaze-direction as important not only in initiating encounters, but also as allowing two interactants to monitor each other. Nielsen described the functions of gaze-direction mainly from the viewpoint of the speaker [Nielsen, 1962]. In Kendon's study [1967], the function of gaze-direction is examined "both as an act of perception by which one interactant can monitor the behaviour of the other, and as an expressive sign and regulatory signal by which he may influence the behaviour of the other" (p.24).

(a) *Expressive*: The expressive function of gaze direction deals with the degree of involvement or arousal of the subject for whom the behavior is measured, or as La France and Mayo [1976] put it, "with the information, which is available through gaze about another's psychological state or trait" (p.547).

The bulk of the studies on gaze direction are concerned with its expressive functions. The research method most frequently employed is the external variable approach. The focus in these studies is usually on the individual, sometimes on groups, and on the personality and social-cultural characteristics that underly interpersonal visual behavior. Within the studies on the expressive function of gaze-

direction, all traditional psychological variables have been put to the test — sex-differences, individual and personality differences, social class differences, cultural differences, etc. — and proved to be of some importance.

Some of the findings, on each of the traditional psychological variables, are:

Women engage in more "eye contact" than men [Exline, Gray and Schuette, 1965], later qualified by Kendon and Cook [1969], who found only a slight tendency towards the same result, with a great deal of variability.

Subjects look more at higher status persons according to Mehrabian [1969], while Fugita [1970] added, that this was true only if the higher status person also approved verbally, smiled and nodded. Fugita's and Mehrabian's study used American subjects. In Watson [1970] it is stated that Nigerians told him "that an elder or *person of higher status* must not be looked directly in the eye in a conversation, as that conveys an impression of aggressiveness, which is impolite" (p.104).

Earlier we briefly mentioned individual differences in gaze-direction. Nielsen [1962] reports differences from 8% to 73% for looking at each other's face, while Kendon [1967] finds 28% and 70%. Gaze *patterns*, however, are consistent across interactions with different partners, except on the measures of length and frequency of gazing while listening and, to a lesser degree, while speaking [Kendon and Cook, 1969].

A number of studies have shown that autistic and schizophrenic persons look significantly less at the face of the other than average people do. This avoidance of eye contact has been interpreted as an attempt to withdraw from interaction with other people, which to the autistic and schizophrenic person is presumably emotionally arousing and uncomfortable [Hutt and Ounstead, 1966; Lefcourt *et al.*, 1967; Rutter and Stephenson, 1972].

Hore [1970] found that high social economic mother-child pairs

exchanged more mutual glances than low social economic pairs during verbal interaction [see also Schmidt and Hore, 1970].

There have been a number of studies relating gaze-direction to cultural variables. Thus Collett [1971] found that Arabs, who have a high level of gaze, preferred those Englishmen who, among other behaviors, looked more during interaction. Ingham [1972], in a comparative study of Swedes and Englishmen, found Swedes to look less often, with much longer glances; there was more looking while listening than when speaking, and a greater percentage of gaze that was mutual among Swedes than among Englishmen. Ingham relates his findings to results obtained with an introversion-extraversion questionnaire.

These findings are only some of the large number of studies dealing with the expressive function of gaze. For comprehensive reviews of the literature, see Ellsworth and Ludwig [1972] and Argyle and Cook [1976]. We hope that the foregoing sample of studies illustrates the sorts of questions posed, and how often a particular finding is immediately qualified or contradicted by another finding. It is important, however, to realize that, even when studying gaze-direction in an ongoing social interaction, the variables mentioned under the expressive function of gaze might have some influence on the way the interaction develops.

(b) *Cognitive*: Gazing at the other interactant's face or gazing away from it is also influenced or affected by cognitive functioning. Whether a person directs his gaze at the face of the other or away from it is partially governed by the operation of such processes as the planning of a new activity, finding a difficult to retrieve fact, etc. In general, speakers tend to look away to aid speech encoding of their thought during longer utterances.

Exline and Winters [1965] found discussion of simple topics to produce twice as much gaze as discussion of topics that were cognitively difficult. The suggestion for the influence of cognitive functioning upon gaze has been most clearly spelled out by Kendon [1967], and although the phenomenon is mentioned in other studies, the

cognitive function is never regarded as a separate function. The reason for this is probably that the same gaze can have different functions simultaneously. When the speaker in an interaction directs his gaze away, when planning a complex utterance, his away-gaze can of course also function to signal the other interactant not to interrupt or take over at this point. Different authors therefore include the influence of cognitive processes on gaze-direction under its regulatory function.

Where a person directs his gaze as related to cognitive processes is presumably physiologically based and can therefore be assumed to be universal. The implication of this is that the human being is capable of dealing with only limited amounts of information at once, and paying attention to a listener and planning what to say are in some cases incompatible activities [Kendon, 1967]. The capacity of the speaker to send and receive messages, virtually simultaneously, through visual orientation is apparently limited. For these reasons, it seems warranted to single out the cognitive function of gaze-direction as being separate from the other functions.

(c) *Monitoring*: The monitoring function of gaze-direction refers to seeking feedback, primarily by the speaker, while he is engaging in his turn at speech. Gaze directed at one's co-interactant's face is used in this instance by the speaker to check if the listener is still "with him". According to Kendon [1967], subjects may look at their interactants to indicate the conclusion of a meaningful unit -- which refers to the regulatory function of gaze -- and to check their interactants' attentiveness and reaction -- its monitoring function.

At least on the part of the listener, the monitoring function of gaze-direction may be partially accounted for by an innate physiological mechanism. Diebold (1968) states that the acoustic orienting reflex involved in "sound localization" is an innate characteristic of humans and other species (for details see Diebold [1968, p.551ff]). "When not directly facing a speaker, a sound shadow at the far ear can block or alter many high-tone inputs. Binaural balance is achieved by means of a reflexive 'noncompensatory nystagmus'

which moves *the eyes* toward the locus of the perceived sound source" [Diebold, 1968, p.555]. The orientation reflex also has autonomic correlates, some of which are important for communicative behavior, since their evocation in one member of a dyad is in turn liable to be perceived by the other [Diebold, 1968].

Finally, some other evidence that the monitoring function of gaze is also important for the listener comes from linguists. Linguists interested in speech production and perception have long been aware that "seeing the speaker" noticeably affects the hearer's perception and comprehension of the former's speech signals.

(d) *Regulatory*: The regulatory function of gaze-direction concerns the control of the communication channel. Gaze-direction initiates, maintains or terminates the interactional flow and influences the other participant. We have already pointed out that gazing at one's co-interactant, or gazing away, be it to express lack of interest or because one is trying to formulate a difficult idea, influences the other in an interaction. More than that, it determines also, at least in part, how the interaction will progress. The regulatory function of gaze-direction has lately been investigated for early mother-infant interaction.

Significant "coupling" between the gaze of infant and mother appears early in the infant's first year [Jaffe *et al.*, 1973] and so does the connection of gaze-direction with the turn-taking pattern of vocalization [Bateson, 1971]. These mathematical regularities in the gross temporal pattern of mother-infant gaze are identical to those found in adult conversations. This suggests that such regularities describe a property of interaction, that predates the onset of speech and is first seen in an attention-regulating behavior [Jaffe, Stern and Peery, 1973; Stern, 1974; Peery and Stern, 1976].

Tronick, Als and Brazelton [1977], in a study on three months old infants and their mothers investigated gaze-direction as well as head and body position, vocalization, etc. They conclude that the infant can sequence expressive displays in an appropriate serial order while allowing for the turn-taking necessary to the reciprocal exchange of

messages. They also show that the infant is able to communicate his or her intent to respond to the expressed intent of the mother. Thus, long before language, the infant is a skilful communicator. From the foregoing, it is clear that visual-movement and vocal-auditory systems become linked very early in the developmental process.

Kendon [1967] showed that responses may be demanded or suppressed by gaze-direction. Whether one looks at one's interlocutor is related to whether one speaks or listens [Exline, 1963], and this in turn is related to the speaker having to plan his utterances. Listeners spend more time looking at speakers than do speakers at listeners [Exline *et al.*, 1965; Kendon, 1967]. At least this is the case for Caucasian speakers of American and British English. A recent study, however, has challenged this finding. La France and Mayo [1976] claim a reversal of the conversational role for American blacks. They point out that it is in the regulatory function of gaze-direction that cultural differences are most likely to emerge.

2.2 SPEAKER-TURNS IN INTERACTION

The fact that participants in a conversation take turns in speaking and listening has been observed and discussed by many researchers. The American college dictionary defines a turn as the time for action or proceeding which comes in due rotation or order to each of a number of persons.

According to Duncan [1972] the term "turn taking" has been suggested independently by both Yngve [1970] and Goffman [1970] (cited in Duncan [1972]). Goffman [1963] however already uses the term: "... The second step deals with focused interaction, that is, the kind of interaction that occurs when persons gather close together and openly co-operate to sustain a single focus of attention, typically by taking turns at talking" (p.24).

For Yngve [1970], turn-taking is nearly the most obvious aspect of conversation (p.568), while Miller [1963] had suggested some years earlier that it is a language universal: "But consider, for example,

the remarkable fact that the conversational partners alternate between talking and listening. This reciprocity, which I assume is universal, ..." (p.418). The universality of turn-taking has been disputed during a recent conference where some argued that in certain cultures conversations were carried out without any turn-taking at all [Kendon, 1975]. Before we can elaborate upon this argument, it is necessary to go into what is meant by a "turn" in conversation.

Nearly every author recognizes, in one way or another, the importance of the alternation of performance between individuals in interaction. The definition as to what is regarded as a turn in conversation depends largely upon the theoretical assumptions and research methods adhered to by researchers in this field.

Basically, three main approaches can be distinguished: firstly, Chapple's study of "action" units, although not using the term "turn", certainly included the study of turns. This approach is the oldest and probably takes more contextual variables into account than any of the other approaches. Secondly, we have those others who define a "turn" in a functional sense, asking the question: "how does the next speaker know that it is his turn?" The more narrow scope functional approach, as opposed to Chapple's, is of more recent origin and is manifest most clearly in "ethnomethodology" and in an approach advocated by Duncan [1972]. Thirdly, we have those who define a "turn" operationally in terms of some specified behavior. The best example of this approach is probably Jaffe and Feldstein's 1970 study.

2.2.1 Interaction Chronography

The alternation of action and inaction by each individual is the most basic fact of any conversation interaction. For a long time this rotation of performance has been recognized by many authors. Chapple [1940a] however was the first to develop methods to measure these actions and silences during interaction.

During the 1930s and 1940s, Chapple began to develop a technique for the objective measurement and description of "personality" based

on an analysis in time units of the interaction between an interviewer and an interviewee. Chapple's idea was that if we want to predict how people will act, the way to do it is to watch how they *do act* and not to infer their behavior from what they say without any means of observational check [Chapple and Donald, 1946, p.199].

Chapple stated further that the content of the interview, what the person says, is a much more difficult variable to quantify than is how he acts while saying it. The latter is an observable event that can be recorded by an observer and the question of "what the person meant by what he said", i.e. the enduring problem of low interjudge reliability in interpretation of levels of explanation of the content, is avoided [Matarazzo, 1962].

The core of Chapple's interaction theory is the realization that the patterning of action in time and the interrelation of actions are dimensions of social interaction of major significance [Matarazzo, 1962; Argyle and Kendon, 1967]. Chapple [1940a] arrived at his conclusion that time was an important variable for describing human relations after he and his early collaborator, Arensberg, found that their field work as anthropologists was unduly hampered by the lack of precision and communicability of the various "subjective" variables which anthropologists and other behavior scientists were then using to describe human relations in the family, tribe, interview situation, etc. Through comparing frequency distributions by length of actions and silences for several individuals in several conversations, Chapple [1940b] found that these distributions and the way they change over different conversations are characteristic of the individuals concerned.

In his early study of interaction patterns during interviews, Chapple placed little restriction upon the interviewer other than that he should use a nondirective interview of the type described by Rogers [1942]. Chapple soon discovered, however, that every interviewer was different not only in the way he behaved but also in the results that he obtained from the same subject [Chapple, 1949, p.296; Chapple, 1953, p.23; Chapple and Donald, 1946, p.203; Matarazzo, 1962, p.475]. Chapple's finding that the degree and manner of co-ordination of

actions in an encounter may have important consequences for its outcome [Argyle and Kendon, 1967] led him to develop a standard interview in which the length and timing of the interviewer's actions and silences are governed by fixed rules [Chapple, 1953]. Chapple's initial breakthrough led to a proliferation of research in the area [Chapple, 1940, 1949; Goldman-Eisler, 1951; Saslow and Matarazzo, 1959; Guze and Mensch, 1959; Kendon, 1963].

A device often used in these studies was Chapple's Interaction Chronograph, which is activated by an observer from the other side of a one-way mirror. This device permits the continuous and simultaneous measurement and recording of such interviewee and interviewer interaction variables as: the number of units of each person's interaction; duration of each of these communicative actions; the duration of each silence; the frequency with which the interviewer purposely reacts with silence by failing to respond to the interviewee's last utterance; the frequency and duration of each participant's interruptions; the resulting dominances and submissions; and several other variables [Matarazzo, 1962].

Matarazzo [1962] reviewed the work of Chapple and that of his own research group on the reliability and validity of non-content aspects of interview interaction behavior and reports one of his own studies in which the method of interview observation and measurement was the Interaction Chronograph. In common with Chapple, Matarazzo [1962] also found that interview action and silence are usually reliable and stable interviewee characteristics.

Not all studies used an interview situation. Goldman-Eisler [1951] demonstrated that, in free conversation, each of her 7 subjects had interaction patterns (short and long silences and short and long actions) which were rather stable for any particular individual. She also found that the silence variable was a somewhat more stable personality characteristic than was action. Kendon [1963] also confirmed that individuals were consistent in their patterning of action from one conversation to the next, and that their behavior in the conversations was related to their behavior in Chapple's standard interview.

It must be realized that originally not only speech or its absence, but also other body movements, were taken into account in the extensive body of research that used the Interaction Chronograph. In the instruction to the observer it was stated: "all facial (and other) muscle activity must have come to an end for action to be scored as ceased". In principle, the chronograph could be applied to a conversation which used sign language. The purpose of the instrument was the study of "action" units rather than speech *per se*. These units of action were subsequently used as the basis for "action length" and "inaction" scores. The realization that activities other than vocalization only contribute to whether someone has his turn, is an important one, although it later became a largely neglected one.

The main focus in the forementioned studies, with the possible exception of Kendon's [1963], though upon behavior, is upon behavior as a means of measuring "personality". For a more extensive review of this early research see the reviews of Matarazzo [1962] and Argyle and Kendon [1967]. The conclusion of the latter on this research is that a properly developed theoretical framework is still much needed. In particular, a definition of Chapple's units is needed which specifies their relationship to other possible units of behavior into which the social performance may be analyzed. Most of the studies used interview and therapeutic situations as opposed to naturally occurring conversations. The realization that the duration of time an individual spends in active speech or gesture is a major factor in interaction remains important. Chapple's findings raised high hopes, but when hoped for contributions did not immediately eventuate, interest in this approach waned [van Doorn and Lammers, 1962].

As was pointed out in the foregoing, Chapple's aim was the study of action units rather than speech *per se*, but gradually studies in this tradition focussed on speech only.

It was the desire to automate the procedure that led to the focussing on speech only. Matarazzo and Wiens [1972, p.9] justify their exclusion of behaviors other than speech: "... the occurrence of smiling, head nodding, and other forms of nonverbal communicative gestures in isolation (without concurrent speech) rarely constitute

more than a very small fraction of any sample of conversation. Consequently, these nonverbal gestures can be omitted from the definition of a unit of utterance with little or no loss in fidelity. This latter is what we have done for approximately the past ten years."

It must be realized that this is far removed from Chapple's original idea. Verzeano [1950 and 1951] was the first to develop an automatic analyzer of verbal interview behavior. The method of automatically recorded speech data, thus doing away with the human observer, was also used by Hargreaves and Starkweather [1959] and Hargreaves [1960]. Their speech data consisted of a single person delivering a lecture; a two-person role playing situation; two-person interviews; a husband and wife alone and in interaction with another couple, etc.

With the waning interest in interaction chronography, the method of automatically analyzing speech data increased in popularity, culminating when Jaffe and Feldstein [1970] published their monograph entitled "rhythms of dialogue". Their initial interest was the study of dyadic interaction in the framework of psychotherapy research. According to Jaffe and Feldstein, their approach essentially combined two earlier approaches; one was Chapple's [Chapple, 1939] and the other the application of probability theory to the time sequence of telephonic communication [Norwine and Murphy, 1938]. It has to be emphasized that automatic measurement of speech is different from Chapple's, who took contextual variables into account to a large extent.

Jaffe and Feldstein [1970, p.2] explain their application as follows: "The application within the framework of psychotherapy dictated a focus upon certain interpersonal or system features of conversation that might be relevant to the communication of 'mood', to the phenomenon of 'empathy' and to the breakdown of effective dialogue. Investigations of such problems are confronted with a mass of clinical interview material which, in all its richness, is largely unmanageable. *Expediency*, therefore, led us to concentrate on the on-off patterns of vocal signals in face-to-face conversation. This restriction

simplified the transduction of signals via the use of voice-actuated relays, which treat any sound above threshold as equivalent, and made large scale automated data processing possible."

Summarizing their findings of 600 conversations, Jaffe and Feldstein [1970, p.11] write: "that the pacing of conversational interaction is

- (a) characteristic of the speakers involved
- (b) stable within any particular conversation and
- (c) consistent from one conversation to the next for the same two speakers.

At the same time, the temporal pattern of an individual is capable of modifying, and of being modified by, that of his conversational partner to such an extent that it need not remain stable from conversation to conversation when the successive conversations involve different partners and different topics." Their findings lend support to Chapple's notion [Chapple, 1939] that the timing of interpersonal actions — among which Chapple included nonverbal actions — could provide the behavioral data needed for a reliable assessment of personality [Jaffe and Feldstein, 1970, p.117].

Again it is clear that the emphasis in Jaffe and Feldstein's work is upon behavior — and more precisely only on speech behavior which is measured automatically — as a means of ultimately measuring "personality". From Verzeano onwards, for all those researchers who measured speech automatically, a turn at speech is defined operationally in terms of a continuous vocalization measured by such and such device (e.g. voice-actuated relays). It will be apparent that this is far removed from Chapple's approach.

2.2.2 Ethnomethodological Approach

A second functional approach, albeit more narrow in scope than Chapple's, dealing with among other aspects turn-taking in conversations, is represented by a group originating from among those

for whom the name ethnomethodologists has been coined [Garfinkel, 1968]. This group started during the sixties, when a group of sociologists rediscovered everyday conversations as a legitimate object of study. According to Turner [1974, p.11] their aim is: "the description and analysis of members' resources (i.e. a rather vague term for what a person talks about and how he treats a certain topic) for finding what they find and doing what others will find them to have done." In more recent work, instead of examining conversation in order to describe the speaker-hearer practices that accomplished the conversation, the internal structural arrangement of conversation itself is examined. This shift in focus raises questions about the place of this type of analysis in ethnomethodology [Mehan and Wood, 1975]. It might therefore be better to speak of conversational analysis.

As far as conversations are concerned, conversational analysts basically study the rules governing the sequential relationship between utterances [e.g. Schegloff, 1968]. The methodic practices of people in their everyday conversations are the topic of concern. As Schegloff and Sacks [1973, p.290] put it: "insofar as the materials we worked with exhibited orderliness, they did so not only for us, indeed not in the first place for us, but for the coparticipants who had produced them. If the materials were orderly, they were so because they had been methodically produced by members of the society for one another, and it was a feature of the conversations that we treated as data that they were produced so as to allow the display by the coparticipants to each other of their orderliness, and to allow the participants to display to each other their analysis, appreciation, and use of that orderliness." The emphasis in the study of internal organization of conversation is upon the speaker-hearer as practical analyst, as practical reasoner. In studying conversational materials the orientation of conversational analysts is to the speaker-hearer's production and recognition of interactionally consequential features of talk [Turner, 1974, p.11]. Turner [1970] finds Austin's distinction of performatives [Austin, 1962] important because Austin explicitly ties talk (or utterances) to activities. Instead of treating talk as indexing underlying values, attitudes, or beliefs,

Austin suggests that talk can be analyzed as the doing of activities.

Kendon [1975] describes the approach by ethnomethodologists as follows: "... suppose that participants in a conversation collaboratively develop a solution to the problem of time-sharing which inevitably arises whenever we are dealing with a situation in which simultaneous talking cannot occur. Features of the sequential relationship between utterances are then examined in actual examples of conversation, and rules which the participants appear to be following are then postulated. The rules that are thus put forward are regarded as the solution that was arrived at. Such rules are seen to be applicable, however, in all conversations, and the substantial differences that may appear in the surface structure of different conversations are nevertheless to be seen as but the outcome of different modes of application of these rules. The reason that these rules are quite general, it would appear, is not because they are rules that exist 'in society' in some sense, and which people learn to follow, but the inherent properties that conversationalists have because they are human beings" [Kendon, 1975, p.10].

Sacks, Schegloff and Jefferson [1974] proposed a number of these "facts" and then offered a system of rules to account for these "facts", which they believe to hold true for "conversation". Among those features are the following:

- speaker change recurs, or, at least, occurs (1)
- overwhelmingly, one party talks at a time (2)
- occurrences of more than one speaker at a time are common, but brief (3)
- transition from one turn to a next with no gap and no overlap between them are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions ... (4)
- turn allocation techniques are obviously used. A current speaker may select a next speaker (as when a current speaker addresses a question to another party); parties may self-select, in starting to talk (12)

- various turn-constructural units are employed for the production of the talk that occupies a turn. Turns can be projectively "one word long" obviously, or, e.g., they can be sentential in length ... (13)

[Sacks, Schegloff and Jefferson, 1974, pp.700,701].

Philips [1976] remarks that no explicit claim of universality has been made for these features, while at the same time, no effort has been made to identify the specific social groups or situations for which they might hold true.

As mentioned earlier, Kendon [1975] observed that during a discussion on turn-taking it was maintained that the very different patterns of conversational organization that are said to occur in different cultures could not be accounted for by some universal set of rules of conversation, and that each culture had elaborated its own rules and its own means of structuring events that might be called conversations. The controversy about this issue was apparently not resolved. Furthermore, the discussion of it did not last very long. Kendon suggests that the most important reason is that we are so lacking in examples of conversational organization from other than English-speaking societies that everyone realized the futility of further argument in the absence of observations to back up his position [Kendon, 1975, p.11].

Other points of criticism of the ethnomethodological approach are that the data employed for study dealt until recently exclusively with data on the verbal structuring of interaction in the form of transcriptions of tape recordings. The information they offer is presented in terms of what one speaker after another does verbally that contributes to the regulation of talk [Philips, 1976]. But of course for every speaker at a given point in time, there is also assumed to be at least one listener. Tape recordings only capture part of the listener's contribution to the regulation of interaction (the verbal part); and a speaker-by-speaker sequence analysis does not allow room, analytically, for consideration of the listener [Philips, 1976]. In the earlier mentioned study by Goodwin [1975], however, he shows how a speaker may re-structure his turn while

producing it, depending upon whether the listener is looking at him.

Furthermore most of the papers by ethnomethodologists incorporate a certain amount of jargon which make them difficult to assess for someone unfamiliar with the area. As Meltzer *et al.* [1975] note, any attempt to grasp the nature of ethnomethodology must come to grips with convoluted, opaque prose. Additionally, one must acquire a degree of facility with a large array of esoteric concepts, such as "bracketing", "deep rules", "documentation", "et cetera clause", "glossing", "idealization", "reflexivity", "second order conceptions", "tag questions", "typification", etc.

Ethnomethodologists are concerned with the study of speech only, in particular how speakers order their utterances sequentially. In general they do not take any interest in the timing of utterances or other actions in conversations, and only recently have [Schenkein, 1972; Jefferson, 1973; Goodwin, 1975] they shown interest in the contribution of the listener. In this sense their approach to the study of interaction is of a much more narrow scope than Chapple's. Chapple's approach can be applied to any type of interaction. The most important contribution they have made is a set of descriptions of methods and procedures for the analysis of verbal interactions, while they also focus more on how interactions are collaboratively produced as opposed to studying the individuals, who produce them.

2.2.3 Signals for Turn-Taking

Another functional approach emerged at about the same time as conversational analysis. Its approach to the phenomenon of turn-taking is to propose signals for its mediation. The phenomenon of transition by which one interactant talks, stops talking and another starts in a more or less smooth, synchronized manner is the focus of studies by Duncan *et al.* [Duncan, 1972 and 1974; Wiemann and Knapp, 1975; Duncan and Fiske, 1977].

The central question in this approach is how participants in a conversation can avoid continually bumping into each other in a verbal

sense. Duncan [1972] proposes that the turn-taking mechanism is mediated through signals composed of clear-cut behavioral cues, considered to be perceived as discrete. Kendon [1975] describes the essence of the signals-approach as follows: "Participants in a conversation are supposed to inform one another whether they are ending a turn, and so providing a turn for someone else; whether they are beginning a turn, and so preempting conversational space for themselves; whether they are continuing; or whether they are continuing as an auditor rather than as a speaker. Audiovisual records of conversations have been transcribed and examined for this purpose. Behavioral features that occur at points where turns are requested, yielded, continued or exchanged, and that are found in an apparently regular fashion are then postulated as the set of turn-taking signals" [Kendon, 1975, p.10]. "Whether or not such features are regularly associated with turn exchanges is then tested statistically. In this way it is hoped to demonstrate what the turn-taking signals are, in terms of particular kinds of behavior, and what the rules are that govern their use [Kendon, 1975, p.10].

In his 1972 article, Duncan describes the signals, cues and rules of turn-taking as they occurred during the first 19 minutes of two dyadic interactions [Duncan, 1972]. Taking an inductive approach, Duncan examined these interactions, and then described the behavior that co-occurred with speaking role exchanges. Duncan has put forward three rules, which are supposed to operate during successful exchanges of speaking turns: (1) turn-yielding cues, which are used by speakers to signal that they are concluding their remarks and that the auditor can take the floor; (2) suppression of speaking-turn claims, which are exhibited by speakers in conjunction with turn-yielding cues, when they intend to maintain their speaking turn; and (3) back channel cues, which are displayed by the auditors to indicate that they do not want the floor, even though the speaker is displaying yielding cues.

Wiemann [1975] elaborated upon Duncan's work by proposing a fourth turn-requesting rule. Auditors employ such a rule, according to Wiemann, to let the speaker know that they want the floor without violating the respect due to the speaker because he or she is not the speaker. Taking a more deductive approach than Duncan, Wiemann

describes the behavioral cues co-occurring with both the yielding and requesting aspects of turn-taking [Wiemann and Knapp, 1975]. When the turn-taking mechanism is functioning correctly, the speaker relinquishes his speaking role (and turn) upon the completion of the meaningful unit being communicated at the time the request is made.

It will be obvious that Wiemann allows a greater time-delay between a certain behavioral cue and its supposed consequence in terms of turn exchange. Precisely this — the question of the time-delay between a certain behavioral cue and turn exchange — seems to be one of the problems with the signal approach to turn-taking. For example, Duncan's proposed back-channel cues supposedly cover the category where the listener does not want the floor, but merely gives feedback to the speaker. Wiemann's turn-requesting cues introduce another category, for which the consequent turn exchange is the criterion of definition.

When, however, a speaker starts a sentence by: "either" ... and so on, the listener already knows that it will be followed by two classes to be compared. Therefore the listener can anticipate the end of the speaker's meaningful unit. In this and many similar cases, other behavioral cues may not be necessary for successful turn exchange. And even if behavioral cues are found to co-occur at turn exchanges, they do not necessarily cause successful exchange. One could go on indefinitely by proposing other rules which cover certain classes of conversational events. Another point is that the same behavioral cue can be used to signal a variety of turn-negotiations, while other behavioral cues, such as pointing a finger, which is not included in any of the cue-systems mentioned, can be and are used to obtain a turn. Duncan and Wiemann have proposed a number of regularities which are found to occur at strategic points in interaction to cover all aspects of turn-taking, and this seems, to say the least, an oversimplification.

It must be pointed out, however, that the "signals-approach" to turn-taking takes the listener analytically more into account than either interaction chronography or ethnomethodology. Taking the listener into account and focussing attention on a number of

regularities occurring at strategic points in interaction, which might be related and have a function in interaction, is its main contribution. By doing so, the signals-approach has also moved away from the more narrow scope which tended to concentrate on speech data only.

2.3 LISTENER RESPONSES IN INTERACTION

In this section the behavior of the listener, while his co-interactant is engaged in speech, will be discussed. As Yngve [1970] has pointed out, when someone has a turn he engages primarily in listening activities. The distinction between having the turn or not is not the same as the traditional distinction between speaker and listener. Both the person who has the turn and his partner are simultaneously engaged in both speaking and listening. Yngve [1970] therefore proposed the term back-channel, over which the person who has the turn receives short messages such as "yes" and "uh-huh", etc., without relinquishing the turn. The partner, of course, is not only listening, but speaking occasionally as he sends the short messages in the back channel. These messages do not consist of only verbal behavior but also of other body movements, such as head-nods.

Yngve was not the first researcher to focus attention on this phenomenon. Fries [1952] had noted earlier many variants of verbal listener responses. According to Kendon [1967], he did not attempt any analysis of the differences in function of various of these listener responses, and even implied that there are no such differences.

Birdwhistell [1962] has shown that the listener influences the speaker's behavior through his head-nods, placed in relation to aspects of the structure of the speaker's speech.

A substantial body of research on verbal conditioning [Greenspoon, 1962] has shown how a speaker's behavior can be influenced to a remarkable extent through the use of such utterances as "Mmh", "huh-uh", etc., and head-nods by the listener, placed in strategic points

in relation to the speaker's speech. The research on verbal conditioning came forth out of Skinner's [1948] behavioristic approach toward verbal behavior. Certain verbal behaviors by the speaker and what we call here listener responses were conceptualized as stimulus-response contingencies, whereby the former increased in frequency as a result of the latter, which acted as a reinforcer. The large body of research (for reviews see Krasner [1958]; Salzinger [1959]; Greenspoon [1962]) was largely carried out within clinical psychology. The phenomenon was investigated for many different populations, using a variety of speaker verbal behavior such as nouns, adjectives, personal pronouns, etc., but also sentences or certain topics. Many different experimental settings were used; interviews, psychotherapy sessions, etc. Some of the more well-known studies are those by Kanfer [1958], Sarason [1958] and Verplanck [1955].

As an example of this approach one study will be mentioned. In a study on a geriatric population, Dignam [1960] used three types of listener responses, "mmhmm", a head-nod and a smile, singly or in combinations. Stimulus words were written on cards and the subjects had to form sentences. All of the three listener responses either used singly or in combinations increased the verbal behavior of the subjects.

To the author's knowledge nobody has ever re-interpreted the results of the verbal conditioning experiments individually in terms of listener responses. As mentioned, the person doing so would have a difficult task because of the differences in methodology, stimulus words used, reinforcement schedules used, etc. Naturally occurring conversations do not lend themselves, by definition, to that type of manipulation.

At least, Kendon [1967] incorporated the verbal conditioning literature as being of significance for any classification of the wide variety of forms which accompaniment signals (Kendon's term for listener responses) take. He describes accompaniment signals as the short utterances that the listener produces as an accompaniment to a speaker, when a speaker is speaking at length. His tentative classification of listener responses was the first in the literature

and is based upon an analysis of where, in relation to the speaker's behavior, the listener places his attention signals. Kendon [1967] distinguishes two main classes of accompaniment signals: the attention signal proper, in which the speaker appears to do no more than signal to the listener that he is attending, and following what is being said, and what Kendon refers to as "point granting" or "assenting" signal (mhm, mhm). It occurs when the speaker is developing an argument by presenting a series of points for which he asks the listener's agreement. Examples are: "yes quite" or "surely", "I see", etc. Kendon [1967], however, also mentions exclamations and laughter, attempted interruptions and short questions and he points out that further variants may be conveyed by variations in the stress and intonation patterns within the utterance, by the use of shorter or longer syllables, and by the way in which he places his utterances in relation to what the other is saying.

Another researcher who has investigated listener responses is Dittmann. Dittmann and Llewellynn [1967] have argued that spoken language is decoded by the listener in word groups called phonemic clauses and that listener responses typically occur at the ends of the speakers' phonemic clause rather than within them, regardless of whether or not there were hesitations. In another study [Dittmann and Llewellynn, 1968] it was shown that head-nods and verbal listener responses often co-occur, again almost exclusively at the ends of rhythmical units in the speaker's talk (phonemic clause). When head-nods and verbal listener responses did occur together the head-nods were found to precede the vocal response slightly. It is Dittmann who coined the term "listener responses".

In 1972 Duncan dealt with listener responses in his paper on turn-taking, and points out that Yngve [1970], Kendon [1967] and Dittmann and Llewellynn [1967 and 1968] all roughly have the same phenomena in mind. Following Yngve [1970] he uses the term auditor "backchannel communication" [Duncan, 1972, 1974; and Duncan and Fiske, 1977]. Duncan [1972] points out that a backchannel communication does not constitute a turn or a claim for a turn and that they comprise a large and complex set of signals. He describes four classes of verbal listener responses; (a) Mm-hmm, etc.; (b) sentence completions;

(c) brief requests for clarification; (d) brief restatement, and one nonverbal one, head-nods and headshakes, and documents their occurrence in two dyadic interactions.

2.4 GAZE-DIRECTION, SPEAKER TURNS AND LISTENER RESPONSES IN JAPANESE INTERACTIONS

After summing up the argument presented thus far, we will review in this section some of the literature relevant to Japanese interactions.

We started by pointing out that especially when comparing communication cross-culturally, it is important to focus attention on what actually happens during the interaction. Therefore we choose to investigate the area of language — that is to say some aspects of it — and its associated behavioral style. We discussed and rephrased our understanding of the term communication as referring to the behavioral relationship between individuals. Since the term interaction is less "loaded" than the term communication, we prefer to use this term.

We have discussed how behavioral systems of individuals as well as the relationship between individuals is quantifiable in spatial and temporal contexts. Our focus of attention is on the temporal organization of movement in its present environment. Movement is defined as the spatio-temporal change of parts of the body in relation to each other or of the body in relation to the environment. We discussed five approaches used in the study of communication of which two approaches — the internal and external variable approach — are employed for the research carried out.

Whenever two or more persons relate to one another, we speak of interaction. The investigations to be reported here, however, deal with only one episode of interaction, namely with conversations. To limit ourselves even further we only study dyadic face-to-face conversations between male participants. The reason for choosing males is both pragmatic — they were more easily available at the time of the study as far as Japanese participants was concerned — as well

as methodological. It excludes any sex variation which has been found in this area [Zimmerman and West, 1975].

Conversations have generated a good deal of the attention of researchers and some of its aspects are well documented. In reviewing the literature, gaze-direction especially has been extensively studied, but its role in conversations has received far less attention. We discussed a number of methodological and theoretical issues relating to the study of gaze, and briefly described four functions of gaze-direction in interaction: expressive, cognitive, monitoring and regulatory. The importance of physiological processes underlying gaze-direction was pointed out. Cultural variability in the use of gaze has been found to occur, especially in its regulatory function. One of the cultures for which this variability in gaze behavior is suggested, is the Japanese.

Bond and Shiraishi [1974] in an experimental study on body lean and status found that the number of glances (frequencies) is a more relevant characteristic for Japanese subjects than the average length of looking. In two other studies [Bond and Iwata, 1976; and Bond and Hwa Young Ho, 1974] it was also found that in (experimental) interview situations, Japanese males gave eye contact only about 20% of the total time to a steadily gazing male interviewer. Gaze in this study is methodologically studied as a manipulated variable on the part of the interviewer, but the result reported here concerns the measured dependent gazing behavior of the subjects. In all three studies nothing about the relationship between gaze-direction and utterances is mentioned.

According to Morsbach [1973] people in Japan do not look each other in the eye much, but are taught to look at the neck — which would bring the eyes and face into peripheral vision, except at close distances. Morsbach bases his statements on subjective impressions. Argyle and Cook [1976] state: "that in Japan little use is made of the visual-facial channel, either in encoding or decoding (p.33), but they do not offer any experimental or naturalistic data to warrant such a conclusion. None of the studies mentioned so far has taken the relationship of gaze-direction to speaking and listening into account,

but all studies seem to suggest that gaze-behavior among Japanese is somehow different from that among English or American speakers.

As mentioned earlier for a different cultural group — American blacks — La France and Mayo [1976] claim a reversal of the conversational role. In their study, blacks look more at the listener while speaking, and less while listening. Even if the co-participant is white, the same pattern emerges for blacks. As Ellsworth and Ludwig [1972] have pointed out, it is not simply a matter of a member of one culture looking "too much" or "too little" for someone of another culture, but more likely a matter of a conflict between two highly complex, totally overlearned (sic) patterns of behavior, which creates cross-cultural communication difficulties. La France and Mayo [1976] add to this that it is especially the regulatory function, which relies on a shared system of cues, that is likely to be a natural arena for the expression of cultural differences. Although they distinguish looking during speaking and listening in their study, they do not investigate those points at which gaze-direction occurs in relation to utterances. While blacks might well look more during speaking and less while listening, the question is at which "strategic points" in relation to utterances is their gaze directed toward the face of the other.

The question where a speaker directs his gaze in relation to utterances has been investigated in detail by Kendon [1967] who found that for Oxford undergraduates, speaker glances tend to be of shorter duration than hearer glances (p.37). He also found a relationship between frequency of gaze-direction toward co-participants at the beginnings and endings of turns to talk (p.33): "... The speaker tends to look away as he begins a long utterance (more than 5 seconds), and in many cases somewhat in advance of it, ... he looks up at his interlocutor as the end of the long utterance approaches, usually during the last phrase, and continues to look thereafter." When the speaker does not direct his gaze at the listener as he finishes speaking, the frequency of delayed response and no response is much greater (p.37).

Although his findings are widely quoted, his study has never been

replicated let alone systematically investigated for different cultural groups. And yet, some studies do suggest that there is some cultural variability in gaze-patterns in relation to speaking and listening. Philips' [1976] anecdotal descriptions on American Indians suggest that gaze-direction at the other's face is used less overall: "... Both speaker's and hearer's gaze rest on their individual co-interactants for shorter periods of time ..." "Both speakers and hearers spend more time looking away from their co-interactants. Speakers look directly at hearers less of the time and hearers look directly at speakers less of the time" (pp.89,90). An anthropological field-study by Moerman [1969] in a Thai-Lue village suggests that the Lue often do not look at another while they talk and he offers various explanations for this "eye-separation". Both of these studies show that participants from some cultures gaze less at one another's face during conversations, than the researchers expected from their English speaking backgrounds.

For the Japanese, Taylor [1975] writes in a study based on subjective impressions, focussing on American-Japanese interactions: "... the American, who expects considerable eye contact when conversing, notices the Japanese constantly breaking his eye-contact as soon as it is made -- the Japanese quickly looks down or off into the distance" (p.120).

It is apparent from forementioned studies that a certain amount of cultural variability in the use of gazing at the other's face in general, and in its regulatory function in particular, occurs. Most of the investigations however are based upon generalizations from the literature or upon participant observation. Permanent data of conversations in different cultural groups, which can and have been analyzed and re-analyzed, are seriously lacking. The question, already posed by Diebold [1968] several years ago, as to what extent the regulatory function of gaze is cross-culturally invariant, has still not received much attention from researchers. Earlier on, we have pointed out that animal and physiological studies have shown that looking somewhere is partly controlled by innate physiological processes, a fact which is neglected by many studies on gaze-direction. These physiological processes will account to a large extent for

similarity in patterns of gazing at someone's face during conversations, even when comparing gaze-direction cross-culturally. One would also expect differences in gaze-direction to occur at those strategic points in an ongoing interaction, where an interactive rather than a physiological process determines the resulting behavior. In particular, where gaze-direction has an interactive regulating function only, differences are likely to occur. One of the cultures for which this cultural variability is suggested by some investigations is the Japanese, but no systematic study of gaze-direction in general, and the relationship between gaze-direction and utterances in particular or gaze-direction in its regulatory function, in Japanese conversations has been carried out. It is to the investigation of these issues that the first part of our study, a comparison of gaze-direction in Japanese-Japanese and Australian-Australian dyads, addresses itself.

In order to investigate the relationship of gaze-direction and utterances, it is a prerequisite to have detailed information on the speech occurring during the conversations. We have already pointed out that the alternation of action and inaction by each individual is the most basic fact of any conversational interchange. The turn-taking phenomenon, as it has now become known, is most clearly related to who is speaking and who is listening at any particular point in time. Rather than focussing on the content of conversations and its difficult measurement problems, we choose to focus our attention on the observable and quantifiable temporal aspect of each turn at speech. The reason for doing so is based on the ideas of Chapple, namely that the patterning of action in time and the interrelation of actions are major dimensions in interactions.

It will be recalled that Chapple and Arensberg [1940] originally started research on temporal aspects in human relations in order to provide more objective criteria to study and describe relations in the family, tribe, etc. As anthropologists, they found their field work unduly hampered by the lack of precision and communicability of the "subjective" variables en vogue just before the second world war. Arensberg [1973] summed up their operations, as proposed in their 1940 monograph as a paradigm-shift from asking questions, which were non-

universal and inferential (what?, why?, where? and how?) to asking questions which were universal and opened the door to empirical comparison and measurement (who?, when?, in what order? and how often?). Although an anthropological textbook [Chapple and Coon, 1942] appeared, the application of interaction theory was never seriously applied by them to interactions in different cultures. The initial formulation of the theory however was based upon Chapple's fieldwork in Yankee City and Arensberg's work in Country Clare in Ireland. One reason for departing from further cross-cultural comparisons is probably that Chapple got involved in the application and refining of his ideas in organizational settings with considerable success and therefore until quite recently he [Chapple, 1970] did not pursue his early anthropological application. His recent book, however, makes it clear that systematic and quantitative observation of social interaction can lead to valuable insights. Arensberg also continued his interest resulting in a project on folk song style and culture [Lomax *et al.*, 1968].

In Collins and Collins [1973] the development of interaction theory and its offshoots are traced historically. In a sense our investigation comparing temporal aspects of conversation cross-culturally can be seen in the light of the early application of interaction theory which, as we have pointed out, has never been followed up. In these early stages, interaction theory was applied at a micro-level by studying specific situations. The more recent publications are much more concerned with a macro-level either in a theoretical sense as Chapple's [1970] recent book or on an empirical level [Lomax, 1968]. One development which makes a micro-level study of conversations cross-culturally much easier is the availability of audio-visual records.

We have also discussed the way in which conversational analysts investigate conversations. We pointed out that their descriptions of methods and procedures to arrive at transcriptions are an important contribution. The emphasis on the general organizing principles of conversations and its sequential development in conversational analysis make their method of transcription a very suitable one for our purposes. By obtaining speech records in this way, it is possible

to investigate some other aspects of conversations, mainly relating to the behavior of the listener. One important advantage of our approach is that we have, besides detailed transcriptions, a permanent visual record available. This visual record makes precision timing a far easier task than it was in Chapple's early days, when the interaction could only be observed once, as it occurred, through participant observation.

In the literature relevant to Japanese interactions, we have not been able to discover anything which deals with temporal aspects of Japanese conversations. The only indirectly relevant observations are on paralanguage across cultures. In an article by that name, Cammack and Van Buren [1967] make some comparisons between Japanese and English. On pauses and affirmation they write that while there is a considerable range of individual differences in these areas, conversational pause, particularly in formal situations appears to be more acceptable in Japanese than in English. Pauses of more than a second or two embarrass English speakers, and provide a common problem for Japanese speaking English [Cammack and Van Buren, 1967, p.10]. On listener responses further on in the same article they write that for the English speaker, Japanese insert too many "yes's" and nods of affirmation, thus lending the impression that the listener will agree to anything and cannot therefore be trusted (p.47). Again their observations are based on subjective impressions and no supporting data of any sort are offered. In order to investigate Cammack and Van Buren's investigations, listener responses are compared. For this purpose an adaptation of what Duncan [1972] described as backchannel communication signals is used.

From our discussion it will be clear that the variables investigated are all closely related and cannot be meaningfully studied, independently from each other.

CHAPTER III

TEMPORAL ASPECTS OF CONVERSATION

3.1 INTRODUCTION TO PART II

3.1.1 Overview of the studies

In chapter II we discussed at some length various aspects of behavior in conversation almost entirely carried out on speakers of English. As we have shown it remains an important task to expand studies of this sort, using established methods of assessment, to include subjects of different language and cultural backgrounds.

The studies to be reported consist of an investigation of temporal patterning of utterances and glances as they occur in natural conversations between dyads. The investigation addresses the general question of how the use of language employed for conversation affects the temporal organization of it. More specifically in the present study we investigate whether the temporal patterning of behavior varies according to whether the conversations are conducted in Japanese or in English.

Accordingly, four studies are reported. In the first two studies comparisons are made between samples of Japanese and Australian conversations. In the second two studies, a comparison is made of conversational behavior when the same individual changes from speaking in Japanese to speaking in English.

In order to provide the participants with a joint interaction focus, a version of de Bono's L-game was used. The purpose of using it was solely to provide a focus for subsequent conversation. In all cases audiovisual records were made of the conversations to be studied, and the analysis of utterance and gaze patterns were done from these records. Of the four studies reported, in two of these (reported in the present chapter and in chapter V, respectively) the recording was made with 16 mm film with synchronome sound recording. The remaining two studies (reported in chapters IV and VI) were recorded on video

tape. The film recordings made possible intensive detailed analysis which were undertaken by means of transcription from the film by the author. The video recordings provided a large amount of data which was used to check the patterns found in the detailed film studies. Measurements of utterance patterns and gazing behavior was done by using pairs of raters who, watching the video tapes in real time, operated switches to indicate when they observed the relevant behavior variables to occur.

The first study, of which an account will follow in this chapter, reports a detailed examination of only one Japanese-Japanese conversation compared with one Australian-Australian conversation. In both filmed conditions, participants speak in their own native language. In this investigation we examine whether there are any differences in the temporal pattern of utterances and glances, according to whether the participants speak Japanese or whether they speak English. This study is exploratory in nature and the emphasis is upon how the duration of certain behaviors of one participant are related to each other and to those of one's co-participant.

After the initial isolation of a different pattern of temporal organization in a Japanese - and in an English language condition it was decided to test our findings for a larger sample of participants. In chapter IV this second study will be reported. In this investigation, video-taped material, involving 24 participants, is used for data analysis. Again each of the participants conversed with a fellow-countryman in his own native language. Fourteen Japanese participants formed seven Japanese pairs and ten Australians formed five Australian pairs. It is predicted that in Japanese turns at talk are briefer, other-directed glances of shorter duration and in general that the rate of response a Japanese conversationalist maintains is higher than is the case in English conversations.

The third study, reported in chapter V, concerns itself with a detailed filmed study of the relationship between temporal aspects of speaking style and language use in a bilingual participant. In this investigation, a white Caucasian Australian of British descent is first filmed interacting with another Australian, conversing in

English, and subsequently with another Japanese, conversing in Japanese. It is predicted that in Japanese the conversation is co-produced to a greater extent than when speaking in English.

The fourth and final study, to be reported in chapter VI, addresses the question of what happens to the temporal organization of a conversation, when Japanese participants converse with Australian participants using the English language. For this purpose nine video-taped interactions were used for data-analysis. Nine of the Japanese and nine of the Australians, all of whom participated in the second study, together formed nine intercultural pairs. The data obtained for individual participants in the intercultural condition are compared with data for the same participants in their own native language conditions. This investigation is also an exploratory one, with some expectation, however, that the temporal organization of their conversation will be more like the pattern found when conversationalists speak in English.

3.1.2 Participants

The Japanese participants were solicited through a Japanese acquaintance of the author. The Australian participants were either acquaintances of the author or persons who were recruited by these acquaintances to participate. In this way 27 pairs of male interactants, who had never talked at length with one another before, participated. Nine pairs consisted of Japanese males, ranging in age from 20 to 34, with an average of $24\frac{1}{2}$ years. All of the Japanese had come to Australia for study and were presently living in Canberra. All but one had been in Australia for less than two years. Six pairs consisted of Australian males, ranging in age from 18 to 40, with an average of 25 years. All of the Australians were living in Canberra and, with the exception of the bilingual participant, who was in employment, studying here.

Ten of the Japanese and ten of the Australians together formed ten intercultural pairs, and thus participated twice. Two additional interactions consisted of a bilingual participant interacting with an Australian in one interaction, and with a Japanese in another. There were no differences in social class and for the purpose of the

interactions, pairs were matched as much as possible for social status, (both Ph.D. students, both undergraduates, or both part-time students).

Some of the participants knew each other, but only casually, and care was taken to match, as interactants, persons who had not previously talked at length with one another. The reason for this is that it has been found that friends establish different patterns of interaction than do participants who are initially unacquainted (Chapple, 1940b; Guze and Mensch, 1959; Kendon, 1963).

3.1.3 Procedure

On the days of recording, the participants were met at an appointed time and, after introducing them to each other and to the experimenter, the order of their appearance in pairs on film or video was decided in collaboration. It had been explained that their interactions would be recorded on either film or video. No one raised any objections to this.

Participants were fully informed of and shown the setting beforehand. It was hoped that, by explaining everything beforehand, the participants would be put at their ease. None of the participants had any difficulty with the situation and questioning afterwards revealed that, after the usual initial awkwardness, their interactions went quite smoothly without any evidence of self-consciousness about being filmed. Watching the records gave the experimenter the same impression.

In each session the experimenter explained briefly to the participants that he was doing research in interactive and cognitive processes and was particularly interested in crosscultural differences. For that purpose he wanted them to play a game together, which was then explained to the participants.*

* L-Game: The L-game was chosen because its rules are fairly simple and to disguise partially a focus upon the interaction itself. After completing the interactions, all participants were told the experimenter was only interested in the interactions themselves, and none of them had had any suspicions. The rules, which were explained literally to the participants, are as follows: Each player has an
(continued over)

He added that after finishing the game the participants could talk with one another about whatever they liked. After approximately 20 minutes the experimenter would come back and talk with them about it.

For those participants who participated twice two sessions were recorded in succession. In one session the participant conversed with someone of his own country (intra-culture), in the other Japanese-Australian conversations (inter-culture) were recorded.

3.1.4 Material

The material thus obtained consists then of two parts. The first part contains 6 filmed dyadic interactions recorded on 16 mm sound film, while the second part contains 21 dyadic interactions, recorded on 1-inch video-tape.

(a) *Filmed material*: The filmed material was gathered on two separate occasions. The films will be referred to as CJA II and CJA III respectively. CJA II contains one intercultural dyad (Japanese-Australian) and two intracultural dyads (both Japanese-Japanese). CJA III contains one intracultural dyad (Australian-Australian) and two dyadic interactions where an Australian bilingual, in the English and Japanese language, interacts with another Australian in one interaction and with a Japanese in the other.

L-shaped piece which covers four squares on a 16 square board. In addition there are two small neutral pieces, each covering one square, which do not belong to the players but can be moved by either of them. Taking it in turns each player must move his L-piece to a new position. The piece may be picked up, turned around, turned right over, etc., and then placed back on the board in any new position whatsoever. A position is deemed to be new if only one of the squares covered by the piece has been changed. The piece may be placed anywhere on the board so long as it covers an exact arrangement of squares and does not overlap another piece. After the L-piece has been moved, a player may - if he so wishes - move either neutral piece to any unoccupied square on the board. The object of the game is to force the opponent into a position from which he cannot move. The game is won when the opponent cannot change the position of his L-piece. The L-piece must always be moved before a neutral piece is touched. The game originated from de Bono (1967).

During the intercultural and one Japanese-Japanese interaction proportionally too much time was spent on playing the game, so that it did not yield sufficient data on speech and gaze for our criteria (5 minutes), although other interesting information on co-ordination of actions by the participants emerged.

For each filmed interaction the two participants were seated in comfortable armchairs, placed in one corner of the room, with the camera placed in the corner diagonally opposite. Only during the change between two dyads did the experimenter enter the room to adjust the focus of the camera and to film the change-over of two pairs. At all other times during the interactions, the participants were left to themselves and the camera unattended.

The filmed interactions were recorded in a large room in such a way that both participants in the interaction were seated at a distance of 2 m at about a 145 degree angle towards each other and a 60 degree angle towards the camera, and fully visible from head to foot on the film at all times. A small table was positioned between and slightly in front of the participants.

The film camera used was a CP-18 equipped with a 12-120 zoom lens and a 1200 ft magazine, which made continuous filming for half an hour possible. On both occasions three interactions were filmed in succession, without any interruptions. Sound was recorded using a NAGRA IV-2 tape recorder. The film was exposed at a rate of 24 frames per second. After initial viewing of the original copies, it was decided to get workprints made, superimposed with a so-called frame-count B-roll, which numbers each frame consecutively, yielding 24 frame-numbers per second (see van Vlack, 1966).

Transformation of the permanent film records into manageable scores, based upon frame-numbers for both speech and gaze-direction and upon frequencies for listener responses, was carried out by the author. Details of the procedures followed are given in the method sections of chapters III and V. No formal reliability studies on the scores thus obtained of the filmed material were carried out.

(b) *Video material*: The video-material was obtained over a period of about five months between November 1975 and March 1976.

Altogether six video-tapes, each of them lasting for approximately 70 minutes, were used. These video-tapes contain 21 dyadic interactions and a number of pilot studies.

The setting for the video-taped interactions was a small room, where the participants were seated in comfortable armchairs 2 m apart at a 145 degree angle between them and at a 60 degree angle towards the camera. The position of the camera was such that both participants were fully visible from head to foot at all times. A small table was positioned between and slightly in front of the participants. The camera was focussed before the participants entered the room and left unattended during the interactions.

To record the interactions, 2 Sony 1-inch video recorders model number EU-310 CE and EU-320 CE, with 2 Sony integrated amplifiers model number TA 1055 were used. The cameras were 2 Sony video cameras with electronic viewfinders, equipped with Cosmocar lenses 15.5 mm 1:1.9 and 25 mm 1:1.9. A microphone AKG D 200 C was placed between the two subjects. For viewing the material, 2 Shibaden monitors were used and for analysis a bigger Philips monitor. The video-recording tape was 1-inch V-12-60 750 meters. Initially two cameras were used but after one session it was decided to use one camera only, equipped with Sony TV zoom lens 1:1.8, 12.5 mm - 75 mm (positions 1.8 - 12.5 mm - 6.5 meters). In a later stage of data analysis a Data System Design video frame number generator model number 440-2 was used to number each frame consecutively. The video-tape scans at a rate of 50 rasters per second, so each second of tape yielded 50 numbered rasters. Transformation of the material from the video tapes into manageable scores was achieved by having raters attend to the material. The tapes were played at normal speed and the raters pressed a switch, whenever an action they were instructed to attend to, occurred. In total 6 raters, all undergraduate students, were employed. Two of the raters attended to speech, two to gaze-direction and two to gestures. The responses of the raters were stored in a computer and the data reported, with the exception of those for gaze-direction, are based upon the responses of each of these raters. The data on gaze-direction are based upon the author's rating in a fashion analogue to the one used for the filmed material. For details of the procedures followed see the method sections of chapters IV and VI.

Reliability studies were carried out on the responses of the raters. For this purpose part of the material was scored a second time by the same rater, attending to the same action of the same participant he was attending to (test-retest reliability) or by having two raters attend to the same action of one participant (inter-rater reliability). The results of these reliability studies are reported in the relevant sections of the chapters to follow.

3.2.1 GAZE-DIRECTION, SPEECH AND LISTENER RESPONSES COMPARING ONE J-J WITH ONE A-A DYAD

In the first study a detailed examination of one Japanese-Japanese conversation is compared with one Australian-Australian conversation. The filmed material on which data analysis was performed consisted of two 5-minute segments. They were taken from the later parts of CJA II/2 (Japanese-Japanese) and CJA III/1 (Australian-Australian), both generated "natural conversations" which lasted approximately 10 minutes each. It was decided to take the later parts so that the participants would be better adapted to the new environment and any awkward first moments would be solved.

In both conditions, participants converse in their own native language. Our aim in this investigation is to examine whether there are any differences in the temporal patterning of utterances and glances and their relationship, according to whether the participants are Japanese and converse in Japanese or whether they are Australian and speak in English. Particular emphasis is given upon how the duration of certain conversational behaviors within one participant are related to each other and to the conversational behavior of the co-participant.

3.2.2 METHOD

Segmentation of speech and gaze-direction was coded independently. Coding of gaze-direction was carried out according to whether gaze was directed toward the face of the other or directed away from the other's face. Inference from head position and orientation as well as the position of the pupil were the main criteria used for judgements. Following Kendon's [1967] notation, gaze directed at the face will be referred to as q-directed gaze, while gazing away from the other's face will be referred to as a-directed gaze. The beginning and end of each look at the other's face was synchronized with frame-numbers,

after which q-directed gaze was marked on graph paper* thus resulting in a flow-chart, where gazing at the other's face and its absence could be easily seen. The results reported are based on these flow-charts.

Segmentation of gaze-direction in the film was accomplished by means of a Bell & Howell 16 mm film projector model number 8309 which has been modified so that, in addition to regular automatic forward and backward speeds, it is equipped with a hand-crank to permit both frame-by-frame and frame-movement** (2, 3, or 4 frames) manual examination.

Data on speech occurring during the two interactions were examined. Transcriptions of the conversations were made twice separately. For the first analysis, transcriptions of the conversations were used to segment speech into frame-numbers at a rate of 24 frames per second. The original audio-tracks were transcribed in such a way that all pauses, verbal slips, other speech disfluencies and inaudible or unclear sections were clearly marked. The Japanese transcriptions were done by a native Japanese, after consultation on the methods to be followed and segments which were unclear discussed till consensus was reached. A literal translation of Japanese into English was also made.

The Australian transcriptions were done by the author. In some cases the participants in the interactions themselves were asked what exactly was said. These cases all referred to names of persons mentioned during the interaction.

The transcriptions thus obtained were typed and the beginning and

* Graph paper used had a scale of 10ths, $\frac{1}{2}$ and 1 inch, where each 1/10th of an inch represents 1 frame of the film.

** It is important to make a distinction between frame-by-frame and frame-movement analysis. In the former static pictures are compared one by one. With the help of a hand-crank it is possible for a skilled operator to sweep the handle showing movement (2, 3 or 4 frames at a time) at close to real time and to do so repeatedly, considerably facilitating the segmentation of movement boundaries. For a more detailed discussion of these and other aspects see Kendon [1977, appendix, pp.225-240].

ending of each participant's turn at speech synchronized with frame-numbers. Speech and non-speech were then marked on graph paper, resulting in a flow-chart for each participant in the interaction.

A speaker turn was basically defined operationally as beginning when one speaker starts solo talking and ending when another speaker starts solo talking [Jaffe and Feldstein, 1970; Markel, 1975]. However one functional distinction was also incorporated. Following Yngve's [1970] distinction, as mentioned earlier, all verbal back-channel communication, in which the person who has the turn receives short messages such as "yes", "yeah", "uh-huh", "mmh", etc., without relinquishing the turn, were not considered speaker turns. In some instances, however, a speaker would have completed some meaningful speech unit, then pause for e.g. $1\frac{1}{2}$ second, followed by the listener responding with typical short "back-channel" communication. In those instances back-channel communication was taken as a turn at speech. The frequency of turns is the number of times this solo talking occurred during the 5-minute period. The average duration of speaker turns was obtained by dividing the total number of frames spoken by the frequency of turns.

Switching pause relates to one speaker stopping solo talking and the other or the same speaker (depending on the context) continuing. Longer pauses however, reflecting a lull in the conversation, were excluded from this category and appear separately.

The average duration per switching pause was obtained by dividing the total number of frames switching pause by the frequency of turns. Simultaneous talking was defined as both speakers talking simultaneously.

For the second analysis, tape-recordings of the original conversations covering the same conversations were once more transcribed, independently from the first analysis, following a procedure as advocated and widely used by ethnomethodologists. The transcriptions thus obtained were used to determine the number of units occurring in both interactions. The transcription method [Jefferson, 1973; Sacks, Schegloff and Jefferson, 1974] in which each

turn of solo speaking by one person is numbered consecutively, used the following symbols (see appendix I for symbols used in transcription). The same transcriptions were also used to count the number of words uttered for each consecutive turn. If the same person continued speaking after a short pause, the number of words for two or three consecutive numbers were added up, until the other interactant started speaking. Laughs, audible breathing, coughs, etc., other than verbalizations were excluded from the count. Following this method, listener responses, to be dealt with separately, were also included, accounting for a greater number of utterances than in the first analysis and a high number of one or two word utterances. The pauses or gaps between utterances were timed by means of a digital counter timer, equipped with two leads. As soon as a person started talking, the gate of the timer was opened by touching the end of one lead; when the person stopped talking, touching the other lead closed the gate. This procedure was repeated two or three times and found to be accurate at a 0.2 second margin.

The second analysis of the speech data together with the film was also used to examine the occurrence of listener responses. After a pilot study, it was decided to score the data according to a slightly modified version of Duncan's [1972, 1974] classification of what he calls "auditor back-channel signals" [Yngve, 1970; Duncan, 1974]. The pilot study consisted of explaining the rules of the L-game to a number of participants, both Japanese and Australian. After scoring the data according to Duncan's categories, it was found that all of the categories used occurred in virtually the same proportions as in Duncan's original data [Duncan, 1972] (see appendix II). Listener responses [Dittmann and Llewellyn, 1968] are thus similar more or less to "back-channel signals" [Duncan, 1972] or "auditor back-channel signals" [Duncan, 1974] and also roughly equivalent to what Kendon first called "accompaniment behavior" [Kendon, 1967].

Listener responses consist of short verbal utterances such as "mm-hmm", "yeah", "yes", etc., as well as head nods, both from the listener. They are regarded as not constituting a turn at speech, but are used by the listener to provide the speaker with feedback on the

progression of his turn. The various categories of verbal signals that were distinguished, are the following:

reinforcers: all those short verbalizations such as "m-hm", "uh-huh", "yeah", "right", and so on. Included in this group are also Kendon's [1967] examples of "yes quite", "surely", "I see", and "that's true". Most of these reinforcers may be used only once or in repetition as in "yeah, yeah".

sentence completions: sometimes the listener completes a sentence, which the speaker had begun, and does not continue beyond the brief completion, after which the speaker would continue his turn. Sentence completions have also been reported by Yngve [1970]. Another term for this category is helpful completions, a term used by ethno-methodologists.

request for clarification: in some instances the listener will briefly insert a few words or a phrase to request the speaker to clarify what he is referring to, during his turn at speech. Again the speaker will continue his turn as if uninterrupted. These requests do occur sometimes, but in our comparison of the two interactions under investigation they did not.

brief restatement: sometimes the listener restates in a few words an immediately preceding thought expressed by the speaker, after which the speaker again continues as if uninterrupted.

Finally, although not a verbal utterance, a listener might respond to something said by the speaker with a smile or laughter of short duration. This behavior can serve a similar function as any of the foregoing and essentially provides the speaker with feedback about the progression of his turn.

head-nods and shakes: contrasting with the verbal utterances mentioned so far, the listener may nod or shake his head. These nods and shakes can occur by themselves or accompany any of the verbalized listener responses. However, head-nods can also occur when the person is speaking. Head-nods may vary in duration from a single nod to a rather protracted, continuous series of nods. In our data analysis,

we distinguish

- whether head-nods are accompanied with a verbal statement or without
- whether the head-nod is emitted while speaking or while listening
- whether the head-nod is multiple (more than 3 in succession), a triple nod, a double one, a single one. Sometimes one and a half nods can be observed, where the head is moved up and down and then up again, after which the head stays in that position. Similarly, half head-nods are sometimes observed, in which the head either moves from its rest position down (head-dip) or up and down and then remains in that position.

And finally, it was noted whether the head-nod was a head-shake, where the head moves horizontally rather than vertically as in a head-nod. Most of the head-shakes occur as negations of something previously said. In the tables, head-nods are symbolized with the letter N.

3.3 RESULTS

3.3.1 Gaze-Direction in General, When Speaking and When Listening

After transcription of speech and gaze-direction was completed, the 7200 frames (= equivalent to five minutes) in each filmed interaction, were arbitrarily divided into 72 parts of 100 frames (= 4.17 seconds) each and the number of frames in each of the following categories was written down: S (speech) and q-directed gaze; S (speech) and a-directed gaze; NS (non-speech) and q-directed gaze; and finally NS (non-speech) and a-directed gaze. The data in table I show the totals for the 7200 frames and the averages in each of the categories per 100 frames, for each participant in the interaction separately.

The Japanese participants glance at the face of their co-interactant (q-directed gaze) approximately 40% of the time, whereas the Australians do so for approximately 72% of the total time.

For the Japanese participants the gaze-patterns are very similar: participant I gazes at the other's face 13% while speaking, participant T 15%. While listening interactant I gazes at his co-interactant 27% of the total time, while T does so for 26%.

For the Australian participants the pattern shows some differences between the two, mainly due to participant V engaging more time in speaking, but regularities are apparent: participant R gazes at the face of the other 27% of the total time, while speaking, participant V 28%.

When listening, interactant R engages in q-directed gaze for 49% of the total time, whereas participant V does so for 41%.

The ratios given at the bottom of table I express the percentage that one participant engages in q-directed gaze while he is not speaking himself but his co-participant is. For the Japanese participants these percentages are 69% and 75% respectively, whereas for the Australian participants the percentages are 97% for participant R, with participant V engaging in more q-directed gaze than his co-interactant R is actually engaged in speaking.

From table I it is apparent that both of the Japanese participants look less at each other's face than do both of the Australians, and this is so both while speaking and while listening.

Although both dyads look more at the other's face when listening than when speaking and their glances are of longer duration when listening as compared to while speaking, looking away (a-directed gaze) is predominant for both Japanese, whereas looking at the other's face (q-directed gaze) is more characteristic for the two Australians.

The arbitrary division of the data into 72 parts, however, does not show the actual average duration per q-directed gaze. For Japanese participant I and Australian participant R, frequency distributions were plotted (see graph 1). Both participants had approximately the same number of glances at their co-interactants - I 68 and R 71. For participant I, the average duration per q-directed gaze is 42.22 frames (or 1.76 seconds) with a standard error of the

Table I
Gaze-Direction during a 5-minute Period
JAPANESE

	I				T			
	S		NS		S		NS	
	q-gaze	a-gaze	q-g	a-g	q-g	a-g	q-g	a-g
Total Speech	2450 34%		4750 66%		2805 39%		4395 61%	
Total Frames	937 (13%)	1513 (21%)	1934 (27%)	2816 (39%)	1084 (15%)	1721 (24%)	1840 (26%)	2555 (35%)
Per 100 Frames	13.01	21.01	26.86	39.11	15.06	23.90	25.55	35.49
In Seconds	0.54	0.88	1.12	1.63	0.63	1.00	1.06	1.48

$$\frac{\text{Gazes List I}}{\text{Total Speech T}} = \frac{1934}{2805} = \underline{0.69}$$

$$\frac{\text{Gazes List T}}{\text{Total Speech I}} = \frac{1840}{2450} = \underline{0.75}$$

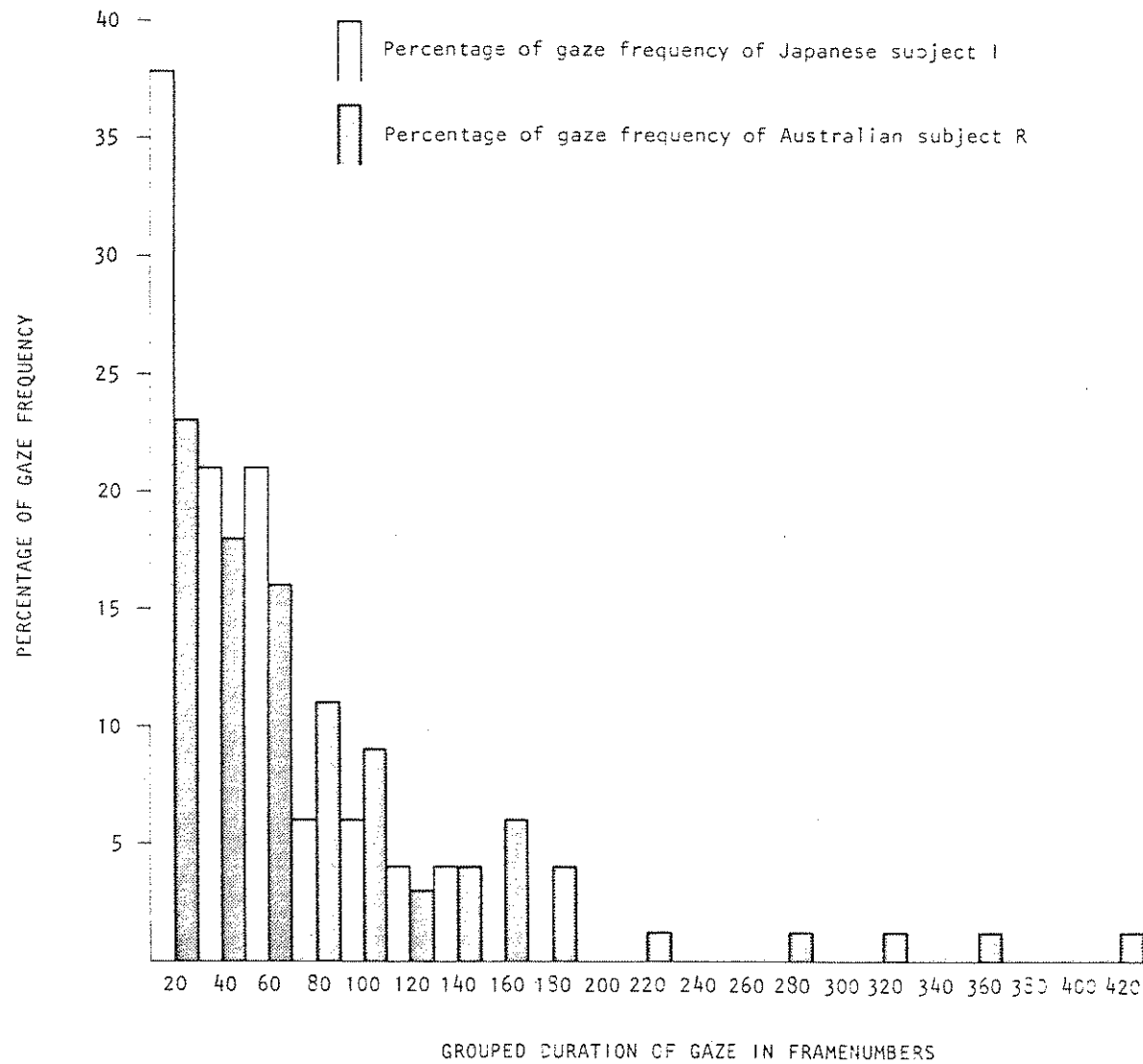
Gaze-Direction during a 5-minute Period
AUSTRALIAN

	R				V			
	S		NS		S		NS	
	q-gaze	a-gaze	q-g	a-g	q-g	a-g	q-g	a-g
Total Speech	2739 38%		4461 62%		3626 50%		3574 50%	
Total Frames	1936 (27%)	803 (11%)	3519 (49%)	942 (13%)	2025 (28%)	1601 (22%)	2967 (41%)	607 (9%)
Per 100 Frames	26.89	11.15	48.88	13.08	28.13	22.24	41.21	8.43
In Seconds	1.12	0.46	2.04	0.55	1.17	0.93	1.72	0.35

$$\frac{\text{Gazes List R}}{\text{Total Speech V}} = \frac{3519}{3626} = \underline{0.97}$$

$$\frac{\text{Gazes List V}}{\text{Total Speech R}} = \frac{2967}{2739} = \underline{1.08}$$

Graph 1. FREQUENCY DISTRIBUTION OF GAZE AT THE OTHER (IN PERCENTAGES)



mean of 4.14 frames. For participant R, the average duration per q-directed gaze is 76.83 frames (or 3.20 seconds), with a standard error of the mean of 9.72 frames. The median for interactant I is 32 frames, and for R 53 frames. Graph 1 illustrates that the q-directed gazes of the Japanese participant tend to be of shorter duration in general, with a high number of glances of very short duration (about 30% of them last less than 1 second). The glances of the Australian participant are more evenly distributed, with some glances of very long duration.

The average duration of q-directed gaze for the other two subjects is 4.20 seconds for interactant T, and 5.20 seconds for participant V. In table II q-directed gaze only is shown. Frequency of q-directed gaze for each participant was determined, irrespective of whether the person is speaking or listening. In the same table, data on mutual gaze, where gaze-direction of the two participants is toward each other's face, are given. The Japanese participants gaze both simultaneously at each other's face for about 23% of the total time that they are engaged in interaction as compared with the Australians who do so for approximately 52% of the time. During the 5 minute period mutual q-directed gaze occurs 47 times for the Japanese dyad and 82 times for the Australian pair. The average duration of a mutual gaze is shorter for the Japanese dyad — 1.36 seconds — than for the Australian dyad — 1.89 seconds. Mutual glances in proportion to each participant's total time spent looking, however, differs only slightly between the two dyads, 53% and 52% for the Japanese participants, and 68% and 75% for the Australian interactants.

Summarizing the results presented so far: participants in the Japanese-Japanese interaction look less toward each other's faces than participants in the Australian-Australian interaction. Participants in both dyads look more at each other, and in general their glances are of longer duration, when listening rather than when speaking.

The two Australian participants spend more time looking at each other simultaneously and their average mutual look also lasts longer. Looking away is more characteristic for the two Japanese, whereas looking at each other's face is more predominant for the two

Table II
Frequencies and Average Duration of Q-Directed Gaze and Mutual Gaze

Japanese		
	I	T
Total number of frames q-directed gaze	2871	2924
q-directed gaze as a percentage of total time	40%	41%
Frequency of q-directed gaze	68	29
Average number of frames per q-gaze	42.22	100.83
Average duration per q-gaze in seconds	1.76	4.20
Total number of frames of mutual gaze	1532	
Percentage of total time	23%	
Frequency of mutual gazes	47	
Average number of frames per mutual gaze	32.60	
Average duration of mutual gaze in seconds	1.26	
Mutual gaze in proportion to q-directed gaze of I and T	53%	52%

Australians		
	R	V
Total number of frames q-directed gaze	5455	4992
q-directed gaze as a percentage of total time	76%	69%
Frequency of q-directed gaze	71	40
Average number of frames per q-gaze	76.83	124.80
Average duration per q-gaze in seconds	3.20	5.20
Total number of frames of mutual gaze	3729	
Percentage of total time	52%	
Frequency of mutual gazes	82	
Average number of frames per mutual gaze	45.48	
Average duration of mutual gaze in seconds	1.89	
Mutual gaze in proportion to q-directed gaze for R and V	68%	75%

Australians. Finally, the gaze-direction patterns found within each dyad for each of the two interactants resemble each other more than they do between the dyads. Before we can examine the relationship between gaze-direction and speech in more detail, we will first turn our attention to the speech data itself.

3.3.2 Speaker Turns

The two Japanese interactants and one of the Australian participants all are engaged in speaking for approximately the same total time, 34% and 39% and 38% respectively of the total 5 minute period analyzed. The other Australian participant speaks slightly more, about 50% of the total time (see table I).

For the first analysis of the speech data, the Japanese dyad has almost twice as many speaker turns as the Australian dyad. Participant I has 49 turns at speech and his co-interactant T 48. The Australian participant R has 25 turns at speech and his co-interactant 28^{*} (see table III). The average duration per speaking turn is for Japanese participant I 50 frames (or 2.08 seconds) with a standard error of the mean of 8.53 frames. His counterpart in the interaction T has average speaker turns of 58.44 frames (or 2.44 seconds) with a standard error of the mean of 7.51 frames. The median of speaker turns is for participant I 27½ frames and for participant T 40 frames (table III).

For the Australian participants, the average speaker turn duration is 109.56 frames for interactant R, which is equivalent to 4.57 seconds, with a standard error of the mean of 24.29 frames. For his counterpart V the average speaker turn duration is 129.50 frames (or 5.40 seconds) with a standard error of the mean of 34.34 frames.

* For a period of 10 seconds during the 5 minute period selected for analysis, the camera was focussed eliminating the Australian participants from full view on the film. For the purpose of data-analysis, this segment was excluded from analysis. It so happened that on both occasions, when interrupting analysis and when resuming, participant V was engaged in speaking. This added 2 speaker turns to his total, although neither turn was included in the analyzed material from beginning to end.

Table III
Speaker-Turns

	Japanese		Australian	
	I	T	R	V
Frequency of speaker turns	49	48	25	28
Average duration per speaker turn in frames	50.00	58.44	109.56	129.50
Standard error of the mean (in frames)	8.53	7.51	24.29	34.34
Average duration per speaker turn in seconds	2.08	2.44	4.57	5.40
Number of speaker turns above group-mean	9	12	10	11
- in percentages	18%	25%	40%	39%
Number of speaker turns below group-mean	40	36	15	17
- in percentages	82%	75%	60%	61%

The median speaker turn for the Australian participants is $65\frac{1}{2}$ frames and 45 frames respectively (table III).

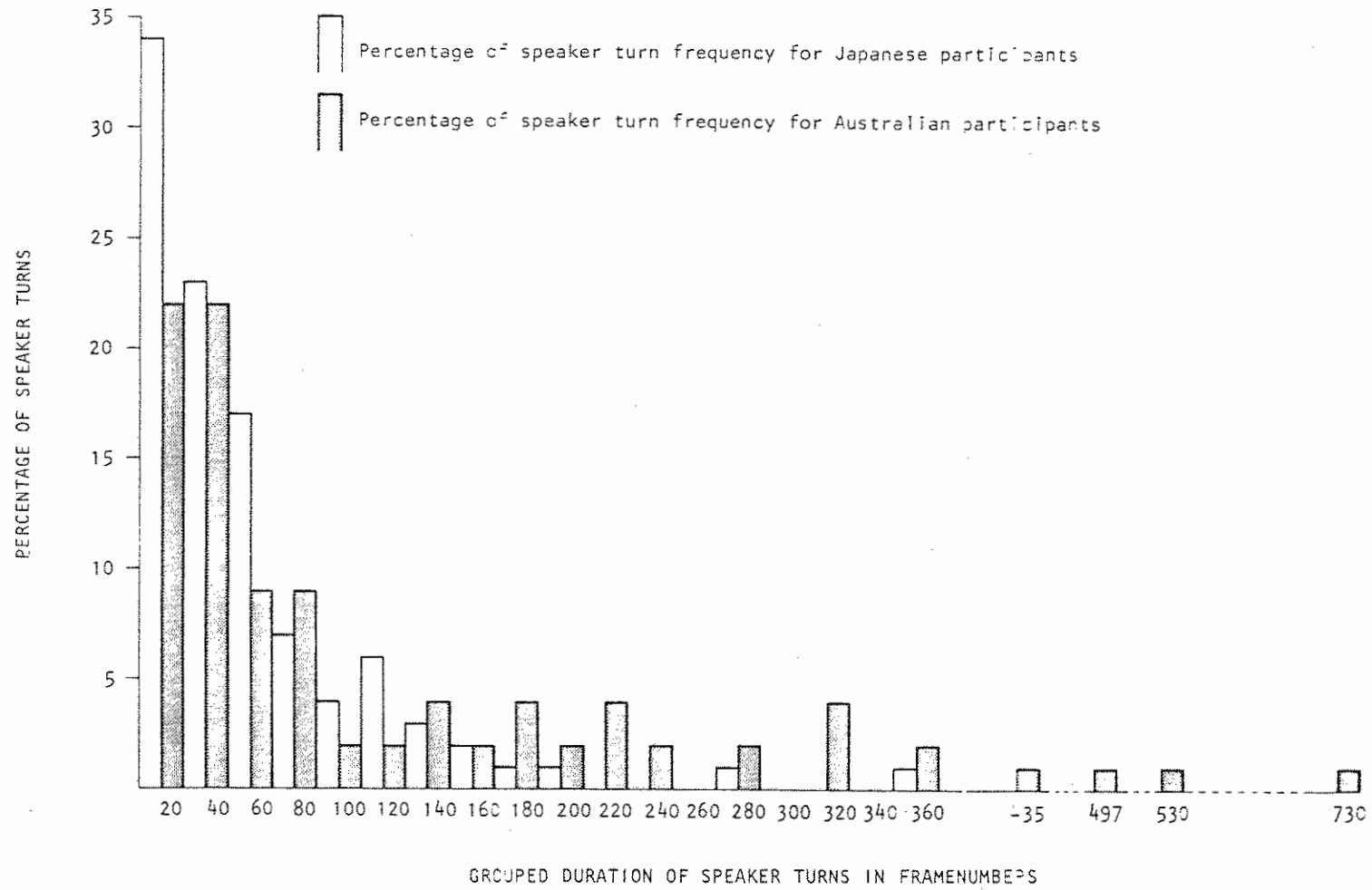
In table III the number of speaker turns falling above and below the group mean — i.e. the mean of all speaker turns in both interactions — is shown. It can be readily seen that proportionally more speaker turns of the two Japanese interactants fall below the mean than for the Australian interactants.

Graph 2 illustrates how the number of speaker turns are distributed for various durations. It is apparent that for both dyads bimodal curves appear. The two peaks for the Japanese participants fall at the shorter end of the scale as compared with the peaks for the Australian interactants. It must be pointed out that graph 2 includes the data for both participants in each of the dyads and that the frequency of speaker turns are converted to percentages to make comparisons possible.

From the data discussed so far it is clear that the two Japanese

Graph 2.

FREQUENCY DISTRIBUTION OF SPEAKER TURNS (IN PERCENTAGES)



interactants co-produce conversation to a greater extent than the Australian participants do.

Table IV summarizes data on pauses, switching pauses and simultaneous speech. Switching pauses for the Japanese participants are 0.25 (I) and 0.43 (T), whereas for the Australian interactants the average duration is 0.56 (participant R) and 0.51 (participant V).

Table IV
Pauses, Switching Pauses and Simultaneous Speech

	Japanese		Australian	
	I	T	R	V
Total number of frames spoken	2450	2805	2739	3626
Frequency of speaker turns	49	48	25	28
Total number of frames switching pause	288	496	338	345
Average duration per switching pause in seconds	0.25	0.43	0.56	0.51
Total number of frames pause	1206		348	
Total duration of pause in seconds	50.25		14.50	
Frequency of pauses	7		8	
Total number of frames simultaneous speech	45		196	
Frequency of simultaneous speech	4		8	

In the Japanese interaction rather lengthy pauses occur, a total of 50.25 seconds. During the Australian interaction only 14.50 seconds is spent in pauses. The Japanese dyad engage for 45 frames in simultaneous speech, whereas the Australians do so for 196 frames.

Table V summarizes data based on the second transcription-method of the speech data. According to this way of analyzing the data, the average duration of speaker turns for the Japanese participants is for I 38.28 frames (1.60 seconds) and for T 44.52 frames (or 1.86 seconds).

Table V
Speaker Turns, Frequency of Words per Utterance and Word Duration

Number of Words	Japanese				Australian			
	I		T		T		V	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
20	3	5%	1	1%	3	7%	7	16%
19 - 20	1	2%	-	-	1	2%	3	7%
17 - 18	2	3%	1	2%	-	-	5	11%
15 - 16	2	3%	-	-	3	6%	1	2%
13 - 14	-	-	3	5%	1	2%	1	2%
11 - 12	3	5%	5	8%	4	8%	1	2%
9 - 10	2	3%	4	6%	1	2%	1	2%
7 - 8	8	12%	7	11%	8	17%	3	7%
5 - 6	8	12%	15	24%	7	15%	3	7%
3 - 4	13	21%	15	24%	7	15%	3	7%
1 - 2	22	34%	12	19%	12	26%	17	37%
	64	100%	63	100%	47	100%	46	100%

391	383	total number of words	366	516
6.11	6.08	average word length per utterance	7.79	11.28
6.27	7.32	average duration per words in frames	7.48	7.03
0.26	0.31	average duration per word in seconds	0.31	0.29
38.28	44.52	average duration per turn in Jefferson-method transcription (in frames)	58.28	78.83
1.60	1.86	average duration per turn in Jefferson-method transcription (in seconds)	2.43	3.28

For the Australian interactants 58.28 frames (or 2.43 seconds) for participant R and 78.83 frames (or 3.28 seconds) for V. Again we find that the Japanese co-produce their conversations more than do the Australians in the study.

The word count showed for the Japanese dyad 391 words for I and

383 words for T, while the Australian dyad produced 366 words (R) and 516 words (V) respectively. The number of utterances in which these words occurred are 64 for participant I and 63 for participant T, while the two Australians grouped their words in 47 (R) and 46 (V) utterances respectively. The average number of words per utterance is slightly less for the two Japanese participants — 6.11 and 6.08 words respectively — than for the two Australians — R an average of 7.79 words and V 11.28 words. The average duration per word, however, did not differ; I 0.26 seconds with T 0.31 seconds, and for the Australian dyad 0.31 and for R and 0.29 for V.

In summary then, it is found that the two Japanese interactants co-produce conversation to a greater extent than the two Australian participants. The number of turns at speech is greater for the Japanese interactants, while the average duration of each turn is shorter. This result emerged from two independent analysis of the same speaker turns separately, we will examine how these variables are related.

3.3.3 Gaze-Direction in Relation to Utterances

In this section gaze-direction as related to utterances is examined. For this purpose the number of speaker turns found in the first analysis of the speech data in each interaction and for each participant is used. In table VI the number of speaker-turns is related to whether the speaker was engaged in q-directed gaze at both the beginning and end of each turn at speech. From the data it is clear that both Japanese participants engage less often in q-directed gaze at either the beginning or the end of each turn, than do the two Australian participants. Again not looking at the face of one's co-interactant is more characteristic for the Japanese participants, whereas looking at the other's face is more dominant between the Australian participants.

In Kendon's [1967] study, he observed "a very clear and quite consistent pattern" for the relationship between gaze-direction and the production of utterances, namely; "that p (= the speaker) tends

Table VI
Frequency of Utterances in Relation to Speaker's Gaze-Direction
at Both Beginning and End of the Utterance

Japanese	I		T	
Total number of utterances	49		48	
	Frequency	%	Frequency	%
Number of utterances <i>beginning</i> with q-gaze	10	20%	19	40%
Number of utterances <i>beginning</i> with a-gaze	39	80%	29	60%
Number of utterances <i>ending</i> with q-gaze	24	49%	25	52%
Number of utterances <i>ending</i> with a-gaze	25	51%	23	48%
Australians	R		V	
Total number of utterances	25		26*	
	Frequency	%	Frequency	%
Number of utterances <i>beginning</i> with q-gaze	13	52%	15	58%
Number of utterances <i>beginning</i> with a-gaze	12	48%	11	42%
Number of utterances <i>ending</i> with q-gaze	22	88%	20	77%
Number of utterances <i>ending</i> with a-gaze	3	12%	6	23%

* Although participant V has 28 utterances over the whole period, only 26 of them can be used in relation to gaze-direction. Due to focussing of the camera during the fragment analyzed, it had to be interrupted and continued again when both participants were fully visible. At both these points V was speaking. Both utterances were incomplete as far as data on gaze-direction are concerned, and were therefore excluded from analysis.

to look away as he begins a long utterance, and in many cases somewhat in advance of it; and he looks up at his interlocutor as the end of the long utterance approaches, usually during the last phrase, and he continues to look thereafter" (p.33). The data reported in table VI

refer to all utterances and not only to long utterances, operationally defined by Kendon as lasting 5 seconds or more (p.31). The data in table VI also refer only to exactly the point where a speaker begins and ends his utterance, which is in our material 1/24th of a second.

To take account of Kendon's observation that the exact beginning or end of an utterance is not always important, since speakers tend to look away at the beginning of an utterance but in many cases somewhat in advance, while they look up at the end, usually during the last phrase, graph 3 was constructed. Since for the Japanese, utterances last on the average far less than 5 seconds, and even for the Australians long lasting utterances are not common in our data, it was decided to include utterances lasting a minimum of 60 frames (= $2\frac{1}{2}$ seconds). For the Japanese participants this criterion resulted in 33 utterances, for the Australian interactants in 27 utterances. To normalize this difference the frequency of q-directed gaze at the beginning and end of these utterances is converted to percentages. The zero point on the horizontal axis refers to the exact point of either beginning or end of the utterance. For each 0.25 second interval before and after the beginning and end of the utterance, it was determined whether the speaker was engaged in q-directed gaze. The open squares and circles refer to the Australian utterances, the closed ones to the Japanese.

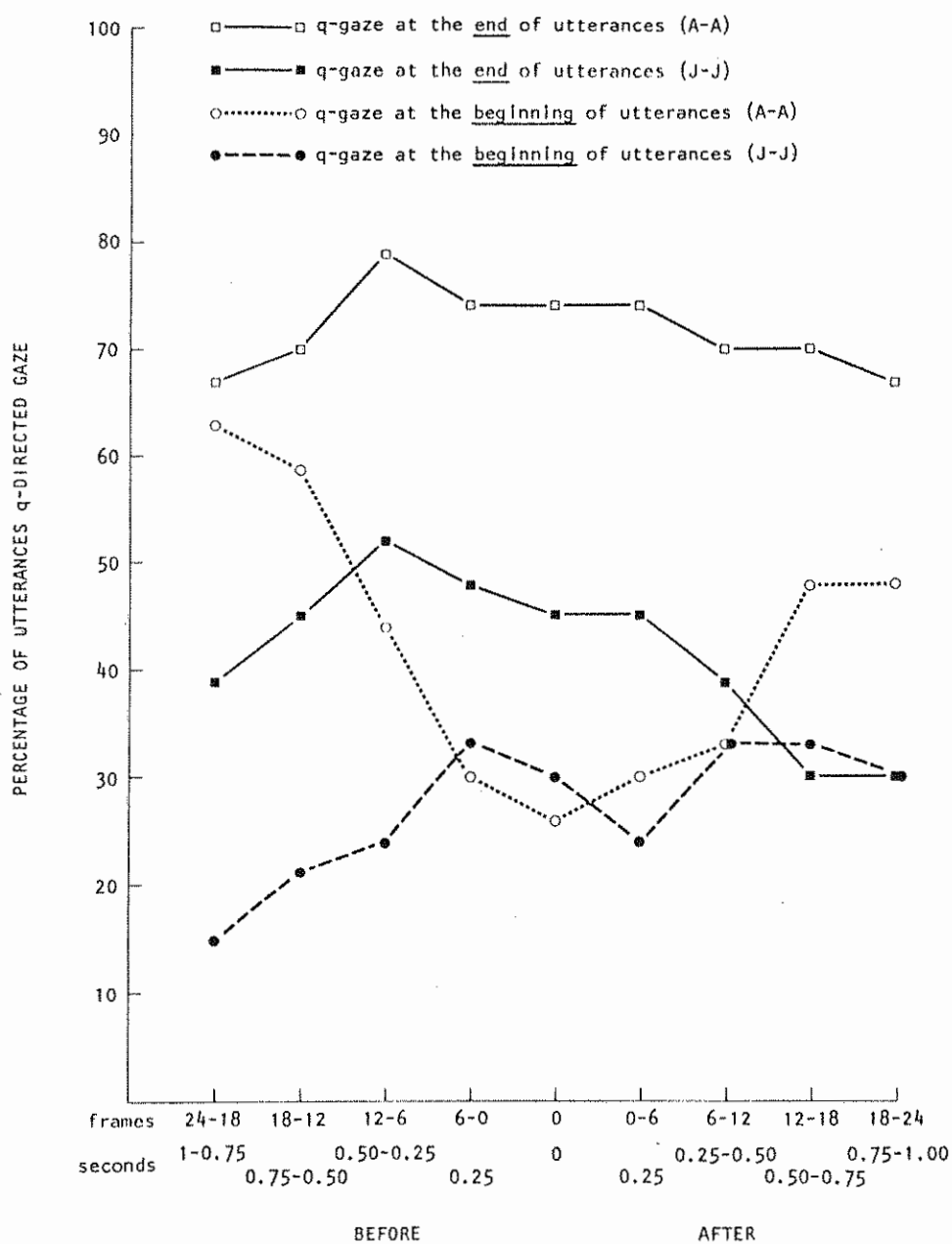
At the beginning of the utterances, the frequency of q-directed gaze is very similar at the zero point, but the curve for the Japanese participants is flatter throughout. For the end of utterances, the shapes of the curves is identical, but the Japanese participants engage considerably less in q-directed gaze. Kendon [1967] points out that the relationship between gaze-direction and utterances may be understood from two viewpoints. Firstly, in looking away at the beginning, the speaker regulates one source of input, namely the listener, and it may be that he is actually withdrawing his attention from him, to concentrate on planning his utterance. When he approaches the end of his utterance, he approaches a strategic point in the interaction. At this point any subsequent action he might take will depend largely upon how his co-interactant is behaving. So we can expect the speaker to seek information at the end of utterances

Graph 3.

q-directed gaze at regular intervals before and after the beginning and the end of utterances. (each utterance duration is 60 frames (=2½ seconds) or more).

AUSTRALIAN DATA BASED ON 27 UTTERANCES

JAPANESE DATA BASED ON 33 UTTERANCES



[Kendon, 1967, pp.34,35].

Secondly, in so far as looking away at the beginning and looking back at the end are regular occurrences, these gaze-direction changes can come to function as signals for the listener, marking strategic points in the interaction, which may be used by the speaker to regulate the listener's behavior [Kendon, 1967, p.35].

Thus, through looking away at the beginning of an utterance the speaker may be seen by the listener to be about to be engaged in an action, and the speaker may look away to forestall any response from the listener. In the same vein, through looking up at the end of his utterance the speaker gives his attention to the listener, who can perceive that the speaker now expects some response from him [Kendon, 1967, p.36]. Kendon tested the hypothesis that looking up at the ends of utterances has a signal function for the listener. When the speaker finished his utterance, Kendon noted, whether he also looked up or away. He then determined whether the listener responded immediately or whether there was a pause or no response at all. When the speaker looked away at utterance endings, the frequency of delayed or no response on the part of the listener was greater, thus supporting the hypothesis [Kendon, 1967, p.36]. With our present data we will also test this hypothesis. We will compare the actual duration of delays, rather than comparing the frequency of utterances responded to, with or without delay.

The data relevant to testing this hypothesis are shown in table VII. For the same utterances as plotted in graph 3, all lasting a minimum of 60 frames ($2\frac{1}{2}$ seconds) or more, the delay between the end of the utterance and the beginning of the next utterance by the other person in the interaction was timed. The utterances were divided according to whether the speaker looked at the listener's face (q-directed gaze) at the end of his utterance or whether he was looking away (a-directed gaze). For the Japanese interactants two rather long pauses occurred during the interaction and for the Australian participants one, and these were excluded from analysis.

For the Australian participants looking away at the end of an

Table VII
Latency between Speaker Utterances
in Relation to Speaker's Gaze Direction

Japanese-Japanese (total number of utterances = 33)				Australian-Australian (total number of utterances = 27)			
q-directed gaze		a-directed gaze		q-directed gaze		a-directed gaze	
1.	6	1.	25	1.	41	1.	30
2.	10	2.	15	2.	10	2.	15
3.	-	3.	-	3.	-	3.	30
4.	10	4.	pause of 300	4.	4	4.	pause of 60
5.	10	5.	10	5.	8	5.	54
6.	-	6.	-	6.	8	6.	73
7.	-	7.	10	7.	43	7.	45
8.	20	8.	15	8.	5		
9.	-	9.	45	9.	10		
10.	21	10.	40	10.	8		
11.	-	11.	15	11.	10		
12.	4	12.	-	12.	5		
13.	pause of 130	13.	-	13.	22		
14.	20	14.	18	14.	-		
15.	30	15.	25	15.	-		
		16.	-	16.	30		
		17.	36	17.	25		
		18.	-	18.	30		
				19.	19		
				20.	14		
131		254		292		247	
8.73	average number of frames	14.11		14.6	average number of frames	41.17	
0.36	in seconds	0.59		0.61	in seconds	1.72	

utterance delays the response of the other commencing to speak considerably — when looking up the average latency is 0.61 seconds, contrasting with when looking away when the average latency is 1.72 seconds. The Japanese interactants, however, respond rapidly no

matter if the speaker looks up at the end of his utterance or if he looks away — average latency when looking up 0.36 seconds, when looking away 0.59 seconds. We can conclude therefore that looking up at the end of an utterance is less important and has less signal-function for the Japanese dyad than for the Australian dyad.

Another analysis carried out on the present data deals with mutual gaze. For all occurrences of mutual gaze it was determined who looked away first, the speaker or the listener. For the Japanese participants, for whom mutual gazing occurred 47 times, the speaker looked away 40% of the time (19 times), while the listener changed his gaze-direction first 60% of the time (28 times). For the Australian interactants however, where mutual gazes occurred on 82 occasions, it was the speaker who changed his gaze-direction first 72% of the time (59 times), while the listener did so only 28% of the time (23 instances).

Strongman and Champness [1968] have pointed out that, if each participant in a dyadic encounter looks at the face of his co-interactant from time to time, and if each looks at the other's face independently of the other, then a predictable duration of mutual looking or eye-contact can be expected to occur simply by chance. If A is the duration of one participant's q-directed gaze, B is the duration of the other participant's q-directed gaze, and T is the duration of the encounter, the duration of chance eye-contact will be AB/T .

Rutter *et al.* [1977] recently re-examined six studies to test this prediction. They found that 95% or more of the variance of the data on mutual looking in these studies can be accounted for by chance, when applying the Strongman and Champness formula. Given that in many dyadic interactions the duration each participant is speaking is more or less equal, and assuming that their gaze-patterns are closely related with the duration of their speech, this result is not surprising.

Nevertheless, when we apply the same formula to our data, dividing the actual duration of mutual looking with the predicted chance-duration, we find for the Australian dyad a ratio-correlation

of 0.99, while for the Japanese dyad this ratio is 0.76.

For the Australian pair mutual gaze is possibly a chance product of each participants' duration of q-directed gaze. For the Japanese pair however the ratio between actual- and predicted-mutual gaze is substantially lower than either our Australian data or any of the six studies on English speaking subjects reported by Rutter *et al.* [1977].

This finding suggests that the two Japanese participants in our study avoid q-directed gaze, in particular when one's co-participant is looking. It has to be emphasized, however, that the simple formula used here to predict mutual looking takes only indirect account of the relationship between duration of q-directed gaze in relation to duration of speech for each of the interactants.

Finally, data on the length of utterances and the length of q-directed gazes are examined. From the foregoing, it is clear that for the Japanese participants in the interaction the duration of speaker utterances and the duration of q-directed gazes are on the average shorter than for the Australians in the interaction. These findings suggest that the duration of speaker utterances and listener q-directed gazes might be related. To test this hypothesis scatter diagrams were plotted for the data as earlier summarized in table I. (The 5 minute period was arbitrarily divided into 72 periods.)

On the vertical axis, speaker utterances were set out and on the horizontal axis the listener's q-directed gazes were plotted. After plotting the 72 points, each point reflecting the relationship between speaker utterance and listener's q-directed gaze over a 4.17 second period, the scatter diagram was divided into quadrants. The number of points occurring in each quadrant as well as points with zeros on either of the axes — those points where a speaker's utterance is accompanied by looking away of the listener or where the listener is engaged in q-directed gaze but the speaker silent — were counted. Table VIII shows the frequencies with which the duration of speaker utterances and the duration of listener's q-directed gaze occur in any of the quadrants. The data for the two participants in each interaction are pooled, showing 144 points for both the Japanese and the Australian interactants.

Table VIII
Relationship between Duration of Speaker Utterances
and q-Directed Listener Gaze

Number of Cases for Japanese Participants (N = 144)

	No q-gaze (= a-gaze)	Duration of q-gaze < 50 frames	Duration of q-gaze = 50-100 frames	Duration of q-gaze = 100 frames
Duration of speech = 100 frames	2	2	2	4
Duration of speech = 50-100 frames	4	14	16	-
Duration of speech < 50 frames	15	42	9	-
No speech (= silence)	30	4	-	-

Number of Cases for Australian Participants (N = 144)

	No q-gaze (= a-gaze)	Duration of q-gaze < 50 frames	Duration of q-gaze = 50-100 frames	Duration of q-gaze = 100 frames
Duration of speech = 100 frames	-	1	13	22
Duration of speech = 50-100 frames	-	6	17	3
Duration of speech < 50 frames	3	21	11	1
No speech (= silence)	41	5	-	1

It can be readily seen that for the Japanese participants more speaker utterances and q-directed listener's gazes of shorter duration co-occur. For the Australian interactants speaker utterances and q-directed listener gazes of longer duration are more common.

For the Japanese interactants, 21 cases of speaker utterance

occurred without any q-directed listener gaze; for the Australian participants only 3 such cases were found to occur. Speaker utterances and q-directed listener gaze with a duration between 2.08 and 4.17 seconds occur 22 times for the Japanese dyad, contrasting with 55 times for the Australian dyad. In view of the findings of shorter speaker utterances and shorter q-directed gazes for the Japanese participants as opposed to longer speaker utterances and longer q-directed gazes for the Australian interactants, we can conclude that the duration of speaker utterances and the duration of q-directed gaze are related. It also implies that the study of gaze-direction cannot be considered separately from the ongoing interaction, in particular from the utterance structure.

To summarize the results on gaze-direction in relation to utterances: the Japanese participants look less at the face of their co-interactant, at either the beginning or the end of each turn, than do the Australian participants. Whether the speaker looks up or whether he looks away, his co-participant responds rapidly in the Japanese interaction. In the Australian dyad, however, looking away at the end of an utterance delays the response of the co-interactant considerably. Looking up at the end of an utterance has less turn-taking function for the Japanese participants.

Analysis of data on mutual gaze revealed that Japanese participants tend to avoid looking at each other, possibly because they are more aware of looking at each other. For the Japanese interactants it was more often the listener who discontinued his gaze at the other first, whereas for the Australian interactants it was more often the speaker who did so. Finally, it was shown that the speaker turn length and the duration of looking at the other's face are related. For the Japanese participants the duration of both utterances and looks at the other's face are shorter than for the Australian interactants.

3.3.4 Listener Responses

After analysis of the data on speech and gaze-direction was

completed, the second transcription — Jefferson method — was also used to score the verbal listener responses occurring during the 5 minute conversation for each of the participants in both interactions. After that head-nods were scored directly from the films, marking them on the typed transcripts.

Tables IX and X show the frequency of listener responses occurring during the interactions for the Japanese and Australian participants respectively. It is clear that the Japanese participants display more listener responses than do the Australian interactants. This occurs despite the fact that the Australian participants engage more time in speaking. Participant I engaged in speaking for 34% of the total time and his Japanese co-interactant for 39%, whereas the Australian participants engaged in speaking for 38% (participant R) and 50% (participant V).

For the Japanese participants, I gives 21 head-nods of which 20 are accompanied with a verbal utterance and 1 without, while interactant T nods his head on 31 instances of which 26 are together with a verbal utterance and 5 without. The same head-nods can be divided in another way, namely if displayed when speaking or when listening. Participant I nods his head 4 times while speaking and 17 times while listening. For his co-interactant 12 head-nods occur while speaking and 19 while listening. Distinguishing the head-nods according to the different types: participant I displays 3 multiple head-nods, 1 triple nod, 1 double, 12 single nods, 1 head-nod in which the head is moved $1\frac{1}{2}$ times, two head-nods go only half up and on one occasions he shakes his head. Participant T gives one triple head-nod, five double head-nods, on 21 occasions he nods once and in four instances he shakes his head. As far as the verbal listener responses are concerned; interactant I gives 31 verbal responses as a listener of which 29 are reinforcers and 2 requests for clarification. Participant T makes 21 verbal listener responses of which 17 are reinforcers, one is a sentence completion and 3 are requests for clarification. The Japanese reinforcers occurring most often in the data are: "Un", "Hūn", "Uh", "so desu ne", "so ne", "Aa", "Ee", "Sō", "Mhm", "Sō ka", "A sō", "dakedo", "sodaro".

Table IX
Frequency of Listener Responses for CJA II/2
for a 5 Minute Period (Japanese)

	I	T
head-nods (total number)	21	31
(a) with verbal	20	26
without verbal	1	5
Total	21	31
(b) while speaking	4	12
while listening	17	19
Total	21	31
(c) multiple N	3	-
triple NNN	1	1
double NN	1	5
single N	12	21
1½ N	1	-
½ N ↑	2	-
½ N ↓	-	-
head-shake	1	4
Total	21	31
verbal listener responses (total number)	31	21
reinforcers	29	17
sentence completions	-	1
request for clarification	2	3
laughter, etc.	..	..
Total	31	21
Over 187 units		

For the Australian participants, person R nods his head on 13 occasions, 12 times accompanied with a verbal response and once without. Six of those nods are displayed when speaking and seven when listening. On one occasion it is a triple head nod, once a double one, six times it is a single nod, in two instances he moves his head half up and three times he shakes his head. Participant V gives 25 head-nods of which 20 are accompanied by a verbal response and 5 are not.

Table X
Frequency of Listener Responses for CJA III/1
for a 5 Minute Period (Australian)

	R	V
head-nods (total number)	13	25
(a) with verbal	12	20
without verbal	1	5
Total	13	25
(b) while speaking	6	3
while listening	7	22
Total	13	25
(c) multiple N	-	1
triple NNN	1	4
double NN	1	2
single N	6	13
1½ N	-	2
½ N ↑	2	-
½ N ↓	-	-
head-shake	3	3
Total	13	25
verbal listener responses (total number)	14	25
reinforcers	11	23
sentence completions	2	1
request for clarification	-	-
brief statement	1	1
laughter, etc.	..	..
Total	14	25
Over 176 units		

He nods his head 3 times while speaking and 22 times while listening. Once the type of nod is a multiple one, four times a triple head-nod, twice a double one, on 13 occasions he gives single nods, twice he moves his head 1½ times, and in three instances he shakes his head. The Australian participants give, 14 for person R and 25 for V, verbal listener responses. For interactant R the responses are reinforcers

In 11 cases, twice they are sentence completions and once a brief restatement. Participant V gives 23 reinforcers, one sentence completion and 1 brief restatement. For the Australian participants the most common reinforcers are: "Yeh", "Mmh", "Uhm", in that order.

Despite some individual differences there appears to be a remarkable consistency in the extent that different listener responses occur. The finding that the Japanese have a higher frequency of listener responses indirectly supports the finding that speaker turns in Japanese tend to be of shorter average duration. Again we find some evidence that the two Japanese participants co-produce conversation to a greater extent.

3.4 DISCUSSION

Following an "internal variable" approach we isolated a number of patterns after a detailed micro-analytic comparison of only one Japanese and one Australian dyad. The variables selected for investigation were gaze-direction, speaker turns and listener responses. Participants in both dyads look more at each other's face, and in general their glances last longer, when listening rather than when speaking. This finding is similar to the results reported by other investigators (Exline *et al.*, 1965; Kendon, 1967]. These studies were carried out with English speaking subjects, while Ingham [1972] obtained a similar pattern in a study comparing English and Swedish speaking subjects.

In general, participants in the Japanese interaction look less at each other than participants in the Australian interaction. The Australian interactants also spend more time looking at each other's faces simultaneously and their average mutual look also lasts longer.

When analyzing the speech data, we found that the average duration per turn is shorter and the number of turns over a specified period greater in the Japanese interaction. Supporting the greater number of speaker turns in the Japanese interaction was the higher frequency of listener responses in the Japanese conversation compared

with the Australian one.

Relating the findings on gaze-direction, speaker turns and listener responses to one another, we can summarize them by concluding that the temporal cycle of each interactant's contribution to the conversation is of shorter duration for the Japanese participants, whereas it lasts longer for the Australian interactants. Or, to put it in different words, the Japanese participants co-produce conversations to a greater extent than do the Australian interactants.

Some caution in interpreting the results however is warranted. The pattern found here is based on only one pair of Japanese and one pair of Australians. To test whether this pattern is an idiosyncratic one, related to the participants in the interactions, or a more general one, we will investigate the same variables in a larger sample of Japanese and Australians in the next chapter.

Detailed investigation of the relationship between gaze-direction and utterances revealed a proportional relationship. In general, longer utterances are associated with longer other-directed gazes and shorter utterances are associated with shorter other-directed gazes.

It must be pointed out however that deviations from this general rule occur. Sometimes an interactant speaks, but does not gaze at the other at all. In general, however, shorter utterances and shorter other-directed gazes are more common in the Japanese dyad, whereas longer utterances and longer other-directed gazes are more common in the Australian dyad. The relationship between gaze-direction and utterances also revealed that the Japanese interactants look less at the face of their co-interactant, at either the beginning or the end of an utterance, than do the Australian participants.

It was suggested [Kendon, 1967] that planning what to say and paying attention to a listener are incompatible activities. A physiological process underlies this phenomenon and therefore one would not expect to find any differences between the two interactions when comparing their looking behavior at the beginning of utterances. In both interactions participants look away at the beginning of an utterance or somewhat in advance of it. This looking away at the

beginning of an utterance may also be of importance in its regulatory function. When looking away at the commencement of an utterance, a speaker may convey to the listener that he intends to "keep the floor". This dual function, cognitive and regulatory, is absent at the end of utterances. When finishing an utterance, the speaker can signal to the listener that he has completed a speech-unit and check whether the listener is still "with him". Gaze-direction at the other at the end of an utterance takes on a more interactive quality — as part of a negotiating process with the other — and no physiological process compels the speaker to look up.

In the present study, it was shown that the Japanese participants may or may not look up at the end of an utterance, while the Australians more often than not do look up. Looking up at the end of an utterance seems to have less turn-taking function for the Japanese participants.

This hypothesis was tested by timing the delay of response when speaker change occurred. For the Japanese looking up or away does not make much difference at the end of an utterance. The next speaker responded rapidly without much of a delay. This delay was clearly present for Australians, when the speaker finishing his utterance did not look up. Our findings for Australian participants are similar to those obtained for the English speaking subjects in Kendon's [1967] study, whereas for the Japanese participants looking up at the end of an utterance has less turn-taking function.

Although looking up is less important in its regulatory function, we have presented some evidence that the Japanese participants are more aware of looking at each other. We showed that the Japanese participants avoid eye contact and that the listener (politely?) looked away first after mutual looking. In contrast for the Australians it was usually the speaker who did so. One would expect the speaker to look away first because of the incompatibility of planning one's utterances and looking at one's co-participant (the cognitive function of gaze-direction). The question as to why the Japanese avoid mutual gaze and why they look less in general, as found in the present study, deals more with the expressive function of gaze

and will be taken up in the next chapter.

As pointed out earlier, physiological processes underlying looking behavior account for a large part of invariance encountered when comparing participants of different cultural backgrounds. While actually analyzing the data, it soon became obvious that the functions of gaze-direction in conversations are to a large extent similar for both dyads. One Japanese participant typically looked up when asked a question to which he couldn't reply immediately. The Japanese participants also monitor each other during the interaction, albeit usually for very short periods. Both groups look away at the beginning of utterances, the Japanese participants slightly less. At the end of utterances the Japanese participants look up considerably less than do the Australian participants. We suggested that gaze-direction in its regulatory function differs for the participants, depending upon whether they are Japanese or Australian. Only through comparing participants of two cultural groups is it possible to investigate the importance of the different functions of looking.

It is therefore not surprising that differences occur in the regulatory function. It is this function which has most importance for the way people negotiate speaker-listener switches in interaction. At least as far as looking behavior is concerned, its importance may simply depend upon the culture of the interactants. Gaze is of course not the only "signal" relevant for turn-taking. This leaves us with the interesting question of how the Japanese signal the ending of a speech-unit to each other. The probability is that, at certain points in an interaction, looking at the other's face can be replaced by other behaviors. The greater frequency of listener responses found for the Japanese participants might not only be related to a greater number of speaker turns, but also replace the absence of gaze at certain points.

The implication of the results presented is that gaze-direction should be investigated in relation to speech, and in particular to utterances. Studies which suggest a relationship between patterns of gaze-direction and variables other than speech are dubious, unless the looking-speech relationship is also examined. It will be clear that

the speech-gaze-direction relationship has also implications for the methodology to be followed in studies of this type and that an "internal variable" approach offers advantages, especially initially where the intention is to uncover a pattern. Studies of larger samples and/or field studies can then be carried out to check the findings. This will be done in the next chapter.

CHAPTER IV

TESTING HYPOTHESES ON TEMPORAL ASPECTS OF JAPANESE
AND AUSTRALIAN CONVERSATION (VIDEO)4.1 GAZE-DIRECTION, SPEECH AND LISTENER-
RESPONSES FOR 24 PARTICIPANTS

After detailed examination of gaze-direction, speech and listener-responses in only one Japanese-Japanese and one Australian-Australian dyadic interaction had revealed differences in the temporal organization of these interactions, it was decided to test our findings for a larger sample. The differences could after all be due to some factor(s) related to the participants in the interactions examined. Therefore a larger sample of participants from both cultures were video-taped. The reason for using video-tape instead of film was that film is expensive and that detailed analysis of film is time and labour consuming. It was therefore decided to switch to video; although the image is less clear and reliable analysis more difficult.

The hypotheses to be tested are the following:

- (1) Participants interacting in the Japanese language will proportionally engage in less q-directed gaze, both when speaking and when listening, than will Australian participants. The average duration of q-directed gazes will be shorter for the Japanese than for the Australians.
- (2) Japanese participants will take more frequent turns at speech than will Australian interactants.
- (3) For the Japanese interactants the frequency of listener-responses will be higher than for the Australian participants over the same time period.

4.2 METHOD

The material on which data analysis was carried out, consisted of 5 minute periods of the later parts of the interactions. It was decided to take the later parts so that the interactions would have stabilized somewhat [Guze and Mensch, 1959]. For about half the interactions, the conversational periods came first, for the other half the game. In both cases some awkward moments occurred either when interacting with a stranger or when shifting from game to conversation. In all 24 participants, 14 of them Japanese and 10 of them Australian participated. They formed 7 Japanese pairs and 5 Australian. Segmentation of speech and gaze-direction was coded independently, while data on speech and listener responses for the whole 5 minute periods were also scored independently of one another.

Initially raters were used to monitor both gaze-direction and speech. Although data on gaze-direction were reliable — both on test-re-test and on inter-rater measures — they did not always correspond to what actually occurred during the interaction. Raters responded to body-orientation, shoulder- and head-movements. They did so consistently even when the same conversation was rated a second time. Careful micro-analytic inspection of parts of the data revealed that the cues the raters appeared to be using were not always indicative of the actual gaze-direction of the participants. Short a- or short q-gazes were often not noticed because the video was run at normal speed during rating. Therefore it was decided to do a detailed analysis on small segments of the video material to obtain data on gaze-direction.

From the interactions on video-tape 10 second segments of each of 7 Japanese-Japanese dyads and 5 Australian-Australian dyads, were selected. From each of the interactions exactly 10 seconds (= 500 frames) was selected by playing the video-tape at normal speed, till a rounded frame-number, either ending with ... 500 or ... 000 was encountered. When a segment thus obtained was not obviously a lull in the conversation, it was included, and from the starting point on exactly 500 frames were segmented first for speech, and after all the speech data, for gaze-direction.

Although not strictly speaking random, it will be clear that this

method of selection is largely unrelated to what occurs in the interaction. It must be pointed out here that the video-recorder used for analysis — a Sony 1-inch model number 310 EV — permitted through its still picture and slow-motion facility only frame-by-frame analysis (see footnote, p.63).

The twelve 5 minute conversations were rated by 2 raters. The raters were trained to monitor speech and instructed to press a switch each time the interlocutor they were attending to, spoke, and to maintain pressure until that interlocutor stopped speaking.

Any phonation except sneezing, coughing and laughter was counted as speech, and any perceptible pause was counted as a stop in speech. To carry out this task the raters were given headphones, but they also had the video-screen to rely upon. In this way one rater attended to the speech of one interactant and the other rater attended to the speech of the other interactant. The responses of the raters were recorded by means of a computer program and stored in a PDP-8 computer. The procedure for every 5 minute interaction was as follows: The raters were seated in a room with a 48 cm television monitor, in the room adjacent to where the PDP-8 was located. The raters were separated from each other by a wooden panel, placed between the chairs. The television monitor was placed at a distance of 4 meters centrally in front of the raters. On top of the monitor two lights were mounted, one red and one green.

When the raters were ready, the experimenter would put the light in the raters' room out and start the video-tape. As soon as the picture had stabilized itself, the computer was started. This automatically triggered off the green light on top of the monitor in the raters' room. For the raters this signalled that their responses were being recorded. In the adjacent room the author could monitor the interaction as well as the responses of the raters through a 18.5 cm monitor. Every time a rater would press a switch, the corresponding light, mounted next to the 18.5 cm monitor on the side of each interactant on the screen, would light up and stay on until the rater released pressure. In this way the raters' performance could be checked. After each 5 minute interaction the author would

ask the raters if they were satisfied with their rating or give them feedback based on his own observations. If any of the raters or the experimenter were not satisfied the interaction would be rated again. Whenever either of the raters would become tired a break was introduced. Listener responses were tallied for each of the 24 participants separately for the same 5 minute period as the data on speech. The method followed was essentially the same as described in Ch.III. The video-tapes were run at normal speed and with the help of a frequency counter each of the listener response categories was scored in turn. The author carried out the scoring himself, but for the purpose of determining reliability another person scored part of the material a second time. Because video was used it was impossible to differentiate the different types of head-nods, so all head-nods are lumped into one single category. As far as the verbal listener responses are concerned, the same categories as discussed in Chapter III were used: reinforcers, sentence completions, request for clarification, brief restatements, laughter, etc.

In some instances it was difficult to determine what was going on exactly, when for example a very rapid succession of listener responses occurred. In that case the video-tape was slowed down and viewed repeatedly until the author was satisfied that all listener responses had been observed and counted.

4.3 RESULTS

4.3.1 Gaze-Direction

After transcription and segmentation of speech and gaze-direction was completed, the number of frames in which the participants engaged in: Speech and q-directed gaze, speech and a-directed gaze, non-speech and q-directed gaze and finally non-speech and a-directed gaze, was written down. Table XI shows the data for the totals of each of the forementioned categories and can be compared to table I. The Japanese sample engage in speaking for 39% of the total time, the Australians for 47%. In spite of this difference, which is unrelated to the selection of the segments, the Japanese participants engage

Table XI
Gaze-Direction for Random 10 Second Periods
Based upon 24 Participants

Japanese (N = 14)				
Speech		total number of frames = 7000	Non-Speech	
q-gaze	a-gaze		q-gaze	a-gaze
2768			4232	
1359	1409	total number of frames	2002	2230
97.07	100.64	average number of frames per subject	143	159.28
1.94	2.01	average duration in seconds	2.86	3.19
(19%)	(20%)	in percentage	(29%)	(32%)

Australians (N = 10)				
Speech		total number of frames = 5000	Non-Speech	
q-gaze	a-gaze		q-gaze	a-gaze
2327			2673	
1993	334	total number of frames	2170	503
199.30	33.40	average number of frames per subject	217	50.30
3.99	0.67	average duration in seconds	4.34	1.00
(40%)	(7%)	in percentage	(43%)	(10%)

less in q-directed gaze than do the Australian interactants.

This difference is significant at the 0.5 level (table XII — $F = 5.15$, $p < 0.05$). The Japanese interactants gaze less at the other's face both while speaking and while listening. The difference in q-directed gaze is not significantly different while speaking as compared to while listening, and there is no interaction effect between gaze-direction of the two groups and speaking or listening.

The patterns of gaze-direction are remarkably similar to those shown in table I, where data on only two interactions are shown.

Table XII
Analysis of Variance for q-Directed Gaze

Source	SS	df	ms	F	p
Total	34206.313	47	-	-	-
q-directed gaze culture	3593.363	1	3593.363	5.153	0.05
q-directed gaze speech/ non-speech	540.02	1	540.02	0.77	ns
interaction of q-directed gaze culture and speech/non-speech	88.23	1	88.23	0.13	ns
Error	29984.70	43	697.32		

While speaking, the Japanese participants engage in q-directed gaze for 19% of the total time, and when listening for 29%. For the Australian participants these percentages are 40% and 43% respectively.

It must be pointed out however, that the Australian participants engage proportionally more of the time in speech, which could conceivably increase the percentage in both categories. Even so, it is clear that again looking away (a-gaze) is more characteristic for the Japanese interactants and looking at the other's face (q-directed gaze) for the Australian participants. The data for the individual participants are shown in appendix VI. Table XIII shows the frequencies with which q-directed gazes occurred in the 10 second segments analyzed, as well as the average duration per q-directed gaze. For the Japanese this is 2.32 seconds, for the Australian interactants 4.16 seconds. We can conclude that our hypothesis is supported for gaze-direction.

4.3.2 Speech

Before discussing the results of comparing the speech data, a number of reliability measurements were carried out on the conversational interactions in the study. Interactions which were used for this purpose were chosen at random and 33% of the conversational interactions were rated twice. Test-retest reliability,

Table XIII
Frequencies and Average Duration
of q-Directed Gaze (Video-Material)

Japanese (N = 14)	
Total number of frames q-directed gaze	3361
q-directed gaze as a percentage of total time	48%
Frequency of q-directed gaze	29
Average number of frames per q-gaze	115.90
Average duration per q-gaze in seconds	2.32
Australians (N = 10)	
Total number of frames q-directed gaze	4163
q-directed gaze as a percentage of total time	83%
Frequency of q-directed gaze	20
Average number of frames per q-gaze	208.15
Average duration per q-gaze in seconds	4.16

as determined by dividing the shortest by the longest times for cumulative speech and the frequency of speech was 0.92. Inter-rater reliability was 0.90.

As for the speech data itself; the responses given by the raters were stored in a computer every 300 milliseconds. Thus for each 5 minute period the rater's response was monitored 1000 times. A two-way analysis of variance was carried out for cumulative time and frequency of speaking. The total scores for each 5 minute period for these two measures were entered as data, for both the game-period and the conversational period. Each participant's scores were entered in turn. The between factor was the language spoken or cultural group of the interactants, with the type of interaction as the within variable. Data as relevant for the game period will be fully discussed in Chapter VI. Suffice it to remark here that interaction during the game-period was severely constrained by the rules and objectives of the game, minimizing talking. In table XIV the means for cumulative speech and the frequency of speech are shown. Data for the individual scores in the conversational interactions are shown in appendices VII and VIII.

Table XIV
Means for Cumulative Speech and Frequency of Speech
Based upon 24 Participants

	Cumulative Speech (in seconds)		Frequency of Speech	
	P1	P2	P1	P2
Japanese	35.61	98.76	35.64	75.43
Australian	12.42	92.10	10.20	62.20

The Japanese and Australian participants do not differ significantly for total time spent speaking (cumulative speech), but the Japanese have more frequent turns at speech than do the Australians ($F = 13.73$, $p < 0.01$).

Therefore we can conclude that our second hypothesis relating to the frequency of turns at speech is supported.

4.3.3 Listener Responses

Reliability measurements for listener responses carried out on one-third of the conversational interactions, yielded a test-retest reliability of 0.94 and an interrater-reliability of 0.90. Interrater-reliability was tested by getting a native Japanese to rate the interactions involved. Tables XV and XVI show the frequencies with which the different categories of listener responses occur during the 5 minute conversational periods for the Japanese and Australian participants respectively. From the data, it will be apparent that the Japanese participants exhibit more listener responses than do the Australian interactants. For the frequencies of head-nods and the frequencies with which reinforcers occurred, the Mann-Whitney U test was used to test whether there was a difference between the Japanese and Australian participants. In both cases the differences are significant at $p < 0.01$ (one-tailed). Because the other categories of verbal listener responses did not occur sufficiently often enough and some categories did not occur at all in any particular interaction, further tests on these categories was not warranted. Even so, it can

Table XV
Frequency of Listener Responses in 7 Japanese Interactions over a 5-Minute Period

	J ₁		J ₂		J ₃		J ₄		J ₅		J ₆		J ₇	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
Head-nods	34	27	8	8	14	36	16	21	33	37	32	34	35	17
Reinforcers	33	13	13	31	15	71	21	13	22	31	35	23	30	21
Sentence completions	1	2	1	2	1	-	1	2	-	-	1	-	-	2
Request for clarification	-	-	1	6	3	-	2	2	-	1	2	1	-	1
Brief restatement	-	3	2	1	-	4	4	2	1	2	7	8	4	4
Laughter, etc.	12	3	2	4	-	1	1	2	2	1	-	1	3	1

Table XVI
Frequency of Listener Responses in 5 Australian Interactions over a 5-Minute Period

	A ₁		A ₂		A ₃		A ₄		A ₅	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
Head-nods	10	6	3	10	9	12	3	22	19	22
Reinforcers	10	4	18	11	7	12	10	21	30	25
Sentence completions	-	-	3	2	1	-	-	-	1	-
Request for clarification	2	1	-	-	1	-	1	-	-	-
Brief statement	-	1	-	-	2	1	1	2	3	2
Laughter, etc.	7	-	4	-	-	-	-	-	1	1

be seen in the tables that requests for clarification and brief restatements and even laughter, etc., tend to occur more often between Japanese participants.

Again we can conclude that our third hypothesis — that a higher frequency of listener responses will occur between Japanese participants than between Australians — is supported.

4.4 DISCUSSION

Comparing one Japanese-Japanese and one Australian-Australian conversation on variables important in conversational behavior suggested differences in the temporal organization of conversation between the two dyads. One possible explanation for the findings reported in chapter III is that the differences found are related to idiosyncracies of the participants involved. In order to provide further evidence that the pattern uncovered is a more general one, relating to the cultural background of the participants, a number of hypotheses are tested in a larger sample.

Hypotheses for each of the variables — gaze-direction, speaker turns and listener responses — predicted shorter temporal cycles for the Japanese participants as opposed to longer lasting ones for the Australian interactants. Following an external variable approach we analyzed twelve conversations, recorded on video. Seven of these conversations involved 14 Japanese participants and in five conversations 10 Australian participants interacted with each other.

Results on gaze-direction revealed that Japanese participants look each other less in the face, both when speaking and when listening, than do the Australian participants. Furthermore the average duration of a glance is of shorter duration among Japanese than among Australians. For the Japanese conversations, looking away from the other's face is more characteristic, while for the Australian conversations, looking at the other is more typical. Admittedly the samples analyzed were short in duration, but even so a similar pattern emerged as for the detailed analysis carried out on two filmed dyads.

Results for speech showed that the Japanese co-produced speech to a greater extent than did the Australians. In the Japanese conversations, a higher number of turns at speech was found than in the Australian conversations. The findings on listener responses showed a higher frequency of listener responses among the Japanese participants than between Australian interactants. This pattern occurred when the total time the participants engaged in conversing was equated. Therefore we can conclude that, in general, conversations between Japanese participants are produced by shorter temporal behavior cycles, whereas between Australian interactants these cycles last longer.

Even though we have now established that the pattern of differences in the temporal organization is related to the interactants being Japanese or Australian, one has to be cautious in interpreting the results. The Japanese participants were all living in Australia and participants from both groups all belonged to what can be described as the higher-middle to upper socio-economic classes. The results reported here cannot make claims for all classes, nor for all Japanese or Australians. Further evidence in the form of field studies carried out in natural Japanese settings would provide support for our results. Some further evidence will be presented in chapter V, where we will present data obtained from a bilingual Australian participant interacting in one instance with another Japanese and speaking in Japanese and in the other with an Australian, conversing in English.

Earlier on, literature relevant to Japanese interactions was discussed. Most of the literature dealt with gaze-direction among Japanese, while one reference pertinent to listener responses was mentioned. All of the authors wrote about gaze in general and, in particular, about its expressive function. Only Taylor [1975] mentions gaze-direction in conversation.

All these authors mentioned based their observations on subjective impressions and none of them offered any data on which it is possible to check their findings. Over-generalizations such as the observation that in Japan people do not look each other in the eye

much [Morsbach, 1973], Japanese constantly break eye-contact [Taylor, 1975] or in Japan little use is made of the visual channel [Argyle and Cook, 1976] are the result.

None of these authors have considered the possibility that their observations on gaze might be related to the way conversations are organized in Japan. From our data we would suggest that their observations are at least partly influenced by differences in the temporal organization of conversations. Again we want to emphasize the necessity for including speech data for analysis when one investigates gaze-direction as a variable.

The authors writing about gaze in Japanese interactions invariably relate the diminished use of gazing at the other's face to Japanese cultural rules. The Japanese are described as being "reserved", "formal", "cautious" and so on. Since gaze can play an important part in taking initiative, dominance and the expression of intimacy their reasoning is that Japanese avoid gaze to a certain extent in order not to appear dominant or too intimate.

There is some evidence to suggest that the expression of intimacy among Japanese is subject to more stringent rules and therefore to a certain extent avoided. Research carried out by the author [Elzinga, 1975] showed that Japanese touch each other less than do Americans or Australians. Presently research is in progress to find out more about the meaning of being touched. Japanese and Australians of both sexes are asked what a certain way of touching on certain parts of the body conveys to them [Elzinga and Iwawaki, 1977]. As far as topics in conversation are concerned, Japanese also disclose less about themselves to others [Elzinga and Iwawaki, 1977].

In so far as looking at the other is important as an expression of intimacy, looking at the other's face might be more salient and hence avoided. The lesser degree of mutual looking found for Japanese participants in this study as compared with Australian participants would support this notion. The fact that it was the listener who generally looked away first after mutual looking in the Japanese dyad (see chapter III) and the speaker in the Australian interaction,

suggest a greater degree of deference on the part of the Japanese participants. (The speaker will look away because of the incompatibility to plan an utterance and to look at the listener.) In this chapter we do not concern ourselves directly with the relationship between gaze-direction and utterances. There is some evidence that in an experimental no-vision condition subjects speak in shorter and more frequent utterances than in an experimental vision condition [Cook and Lalljee, 1972]. It could be that, for Japanese participants, not looking the other in the face, which is predominant in their conversations, and more frequent and shorter utterances are related.

Another possible explanation for the more frequent speaker turns found in the Japanese conversations, related to the structure of the Japanese language, will be discussed in the next chapter.

The observation by Cammack and Van Buren [1967] that Japanese insert many more "yes"'s and nods of affirmation, is supported by our findings. We have already pointed out that the higher number of listener responses is related to the higher frequency of speaker turns in Japanese conversations, but also might partly replace the absence of other-directed gaze in some instances. Again we want to stress the importance of taking the relationship between variables into account, when making any generalizations about conversations. In the next chapter we will provide some further evidence for a different temporal organization when speaking Japanese than when speaking Australian-English.

CHAPTER V

TEMPORAL ASPECTS OF SPEAKING STYLE
IN A BILINGUAL PARTICIPANT (FILM)5.1 SPEAKER TURNS AND LISTENER RESPONSES
IN A BILINGUAL PARTICIPANT

In this part some of our findings, indicating that when participants speak Japanese they tend to co-produce conversations to a greater extent than when participants speak Australian-English, are put to test in a bilingual participant.

While gathering data for comparison of temporal aspects in Japanese and Australian interactions, the author came into contact with a person who supposedly spoke both the Japanese and Australian-English language fluently. In actual fact the person — from here on referred to as RW — was born in Australia and is not bilingual in the true sense of the word. However, having studied the Japanese language at university for four years, followed by a three year study in Japan, he was regarded by the Japanese as one of the few persons fluent in their language. Some of the Japanese participants, who knew him, remarked that when talking to RW "they forgot he was not a native Japanese". Impressions of the author on observing RW interacting with two other Japanese participants in a pilot study, confirmed this.

It was then decided to arrange a session in which RW would interact with an Australian participant and separately with a Japanese. Two persons who did not know RW were requested to participate for this purpose. At the time of filming, RW was working for the Japanese embassy and through his work he was interacting with Japanese throughout the day. The material analyzed consisted of two interactions — CJA III/2, the Australian condition and CJA III/3, the Japanese condition.

For the analysis of data on speech, the first 5 minutes of both

interactions was examined. For the scoring of listener responses a further 3 minutes and 50 seconds was added to both interactions, bringing the duration for this purpose to 8 minutes 50 seconds. The hypotheses tested are:

- When speaking Japanese RW will take more frequent turns at speech than when speaking Australian-English.
- When interacting in the Japanese condition, RW will show a higher frequency of listener responses than when he interacts in the Australian condition.

5.2 RESULTS

5.2.1 Speaker-Turns for Bilingual Participant

The method followed for transcription of the conversations and the segmentation of speech was identical to that described in Chapter III. Thus again, speech data were transcribed twice. Listener responses were also examined in the same manner as described in Chapter III.

In the Japanese condition RW speaks for 44% of the total time, while his Japanese counterpart engages for 31% in speaking. In the Australian condition RW is talking for 46% of the time, whereas his Australian co-interactant engages 42% of the total time in speech.

Even though RW engages slightly more time in speech, when speaking Australian-English, he has only 30 turns at speech as compared with 54 turns when speaking Japanese (table XVII).

The first analysis of the speech data shown in table XVII shows that his Australian counterpart has 25 turns at speech and his Japanese co-interactant 55. The average duration per speaking turn is for RW 110.56 frames (or 4.61 seconds) in the Australian condition, with a standard error of the mean of 32.20 frames. His Australian co-interactant has average speaker turn durations of 119.64 frames (or 4.99 seconds) with a standard error of the mean of 30.89 frames. The median duration per speaker turn is 40 for RW and 50 for participant R (table XVII — in frames). In the Japanese condition however the

Table XVII
Speaker-Turns (CJA III/2 - CJA III/3)

	Japanese Condition	
	H	RW
Frequency of speaker turns	55	54
Average duration per speaker turn in frames	40.47	58.48
Standard error of the mean (in frames)	6.80	13.13
Average duration per speaker turn in seconds	1.69	2.44
Median duration per speaker turn in frames	30	30
Number of speaker turns above group-mean	6	13
- in percentage	11%	24%
Number of speaker turns below group-mean	49	41
- in percentage	89%	76%
	Australian Condition	
	R	RW
Frequency of speaker turns	25	30
Average duration per speaker turn in frames	119.64	110.56
Standard error of the mean (in frames)	30.89	32.20
Average duration per speaker turn in seconds	4.99	4.61
Median duration per speaker turn in frames	50	40
Number of speaker turns above group-mean	10	12
- in percentage	40%	40%
Number of speaker turns below group-mean	15	18
- in percentage	60%	60%

average duration per speaker turn is 58.48 frames (or 2.44 seconds), with a standard error of the mean of 13.13 frames for RW. The other Japanese participant has average speaker turn durations of 40.47 frames (or 1.69 seconds) with a standard error of the mean of 6.80 frames. For both participants in the Japanese condition the median duration of speaker turns is 30.

In table XVII the number of speaker turns falling above and below

the group mean — which is the mean for all speaker turns in both conditions — is shown. It is apparent that proportionally more speaker turns in the Japanese condition fall below the mean than do speaker turns in the Australian condition. ($\chi^2 = 8.76$, $p < 0.01$)

Graph 4 illustrates how the frequency of speaker turns are distributed for various durations. Again we can observe that the two peaks of the bimodal distributions fall at the shorter end of the horizontal axis (duration) for the Japanese condition as compared with the Australian condition. To facilitate comparisons graph 4 has been constructed for RW scores only, and the frequency of his speaker turns in both conditions has been converted to percentages.

Table XVIII shows data on switching pauses and simultaneous speech. In the Japanese condition, the average duration of switching pauses is 0.58 seconds for RW and 0.76 for his co-interactant, whereas in the Australian condition it is 0.87 seconds for RW and 1.02 for his Australian counterpart. It must be pointed out, however, that the category of switching pauses in this case also includes short pauses, which were lumped into this category because no clear lulls occurred in the conversations. The actual duration of switching pauses could therefore be somewhat lower. In the Japanese condition, simultaneous speech occurs for 132 frames, in the Australian condition for 160 frames.

Table XIX summarizes data based upon the second transcription method of the speech data. Analyzing the data in this way, the average duration of speaker turns in the Japanese condition is 1.88 seconds for RW and 1.34 for his co-interactant (or 45.11 and 32.26 frames respectively). In the Australian condition average turn duration is for RW 52.65 frames (or 2.19 seconds) and for his counterpart 48.24 frames (or 2.01 seconds). The word count yielded 451 words for RW and 306 words for H in the Japanese condition, and 536 words for RW and 390 words for R during the Australian condition. The average number of words per utterance was in the Japanese condition 6.44 (RW) and 4.43 (H) and in the Australian condition 8.51 (RW) and 6.29 for R. The average duration per word, however, did not differ very much, Japanese condition 0.29 seconds (RW) and 0.30 (H) versus

Graph 4.

FREQUENCY DISTRIBUTION OF SPEAKER TURNS FOR BILINGUAL PARTICIPANT
(IN PERCENTAGES)

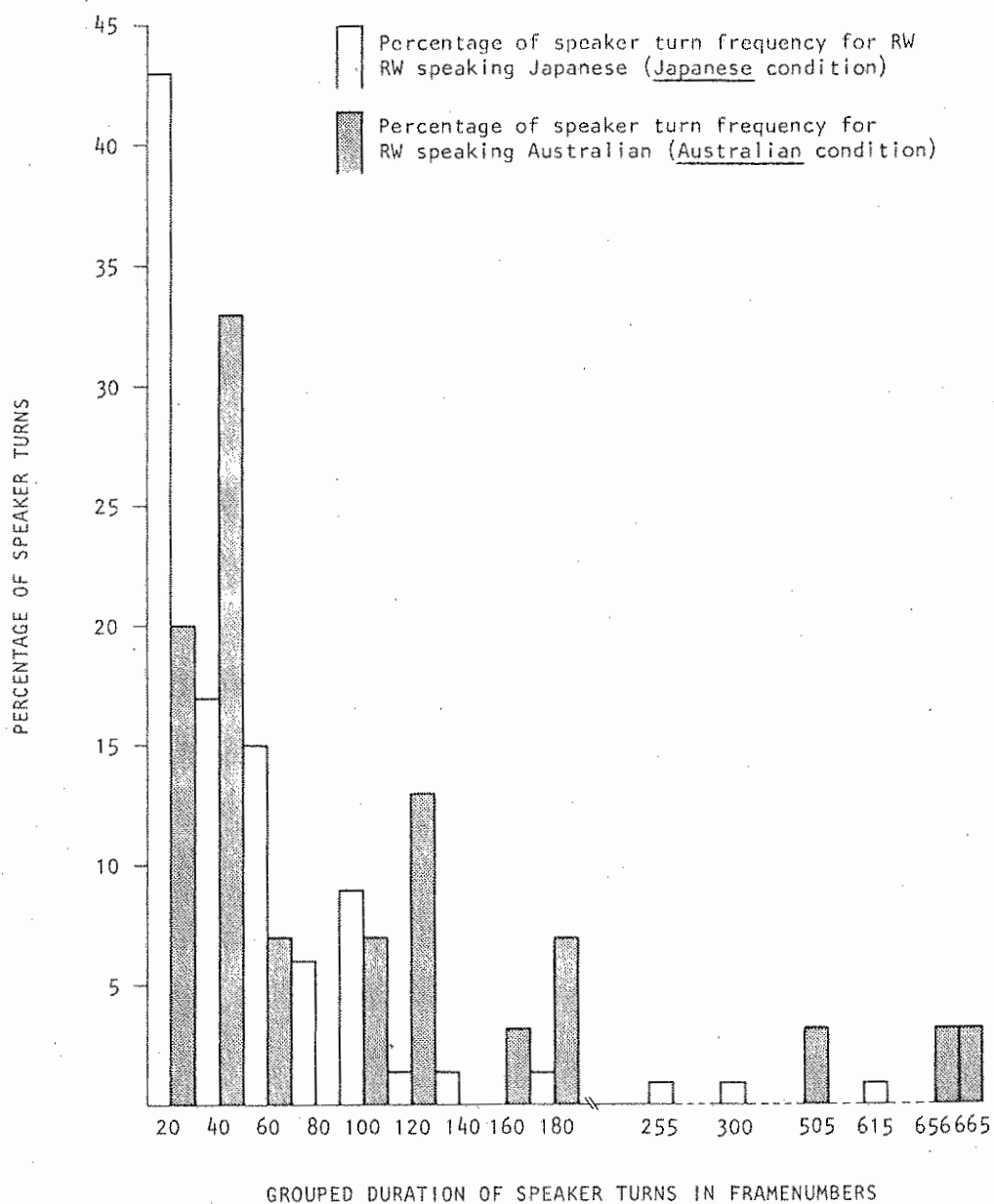


Table XVIII
Average Duration of Speaking Turns and Switching Pauses
(*Australian Condition*) over a 5-Minute Period

	R	RW
Total number of frames spoken	2991	3317
Frequency of speaking turns	25	30
Average number of frames per turn	119.64	110.57
Average duration per turn in seconds	4.99	4.61
Average number of frames per switching pause	24.52	20.80
Average duration per switching pause in seconds	1.02	0.87
Total number of frames simultaneous speech		160
Frequency of simultaneous speech		6

Average Duration of Speaking Turns and Switching Pauses
(*Japanese Condition*) over a 5-Minute Period

	H	RW
Total number of frames spoken	2226	3158
Frequency of speaking turns	55	54
Average number of frames per turn	40.47	58.48
Average duration per turn in seconds	1.69	2.44
Average number of frames per switching pause	18.33	13.98
Average duration per switching pause in seconds	0.76	0.58
Total number of frames simultaneous speech		132
Frequency of simultaneous speech		11

the Australian condition 0.26 (RW) and 0.32 (R).

We can conclude that even the bilingual participant takes more frequent turns at speech when speaking Japanese than when speaking Australian-English. In the Japanese condition conversation is co-produced to a greater extent than in the Australian condition.

5.2.2 Listener Responses for Bilingual Participant

After the analysis of the speech data was completed, the

Table XIX
Speaker Turns, Frequency of Words, Word Duration
for Bilingual RW-Speech Analysis 2

Number of Words	Japanese Condition				Australian Condition			
	H		RW		R		RW	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
20	-	-	3	4%	4	6%	4	6%
19-20	-	-	1	1½%	3	5%	-	-
17-18	-	-	1	1½%	-	-	4	6%
15-16	-	-	1	1½%	-	-	3	5%
13-14	2	3%	2	3%	1	2%	5	8%
11-12	1	2%	6	9%	2	3%	2	3%
9-10	5	7%	4	6%	5	8%	12	19%
7- 8	10	14%	10	14%	4	6%	5	8%
5- 6	12	17%	9	13%	9	15%	2	3%
3- 4	13	19%	15	21½%	12	19%	11	18%
1- 2	26	38%	18	26%	22	36%	15	24%
	69	100%	70	100%	62	100%	63	100%
306	451	total number of words				390	536	
4.43	6.44	average word length per utterance				6.29	8.51	
7.27	7.00	average duration per word in frames				7.67	6.18	
0.30	0.29	average duration per word in seconds				0.32	0.26	
32.26	45.11	average duration per turn in Jefferson-method transcription (in frames)				48.24	52.65	
1.34	1.88	average duration per turn in Jefferson-method transcription (in seconds)				2.01	2.19	

occurrence of listener responses were tallied for a period of 8 minutes and 50 seconds in both the Japanese and the Australian condition (tables XX and XXI).

During the Japanese condition the bilingual participant RW nodded his head on 86 occasions. 65 of these nods were accompanied with a

Table XX
 Listener Responses CJA III/3 (Bilingual Participant --
 Japanese Condition) (8 min. 50 sec.)

	H	RW
head-nods (total number)	76	86
(a) with verbal	48	65
without verbal	28	21
Total	76	86
(b) while speaking	20	23
while listening	56	63
Total	76	86
(c) multiple N	17	12
triple NNN	7	9
double NN	8	14
single N	33	44
1½ N	-	1
½ N ↑	1	1
½ N ↓	4	3
head-shake	6	2
Total	76	86
verbal responses (total number)	51	66
reinforcers	46	53
sentence completions	-	2
request for clarification	1	4
brief restatement	1	7
laughter, etc.	..	..
Total	51	66
Over 370 units		

verbal utterance and 21 of them without verbalizations. Of the same 86 nods, 23 were performed while speaking and 63 while listening. Breaking the nods down into their different types: 12 were multiple nods, 9 were triple nods, on 14 occasions RW nodded twice in succession, 44 were single nods, once his head moved 1½ times, once the nod was half up and on 3 occasions did RW dip his head, while he

Table XXI
 Listener Responses CJA III/2 (Bilingual Participant —
 Australian Condition) (8 min. 50 sec.)

	R	RW
head-nods (total number)	20	60
(a) with verbal	16	39
without verbal	4	21
Total	20	60
(b) while speaking	4	11
while listening	16	49
Total	20	60
(c) multiple N	-	5
triple NNN	1	8
double NN	4	17
single N	13	25
1½ N	-	-
½ N ↑	1	-
½ N ↓	-	1
head-shake	1	4
Total	20	60
verbal responses (total number)	37	45
reinforcers	30	37
sentence completions	1	1
request for clarification	2	1
brief restatement	4	6
laughter, etc.	..	..
Total	37	45
Over 351 units		

shakes his head twice.

His Japanese counterpart displayed 76 head-nods, of which 48 were accompanied with verbalizations and 28 without; 20 nods were performed while speaking and 56 while listening. According to type of nod, the 76 nods can be broken down into: 17 multiple nods, 7 triple

nods, 8 double nods, 33 single nods, one half nod up, 4 half head-dips down and 6 head-shakes.

As far as verbal listener responses are concerned, RW gave 66 in total of which 53 were reinforcers, 2 sentence completions, 4 requests for clarification and 7 were brief restatements. His co-interactant uttered 51 verbal listener responses of which 46 were reinforcers, one a request for clarification and 4 brief restatements.

In the Australian condition the number of listener responses occurring was less, although higher than expected. Why this could be so will be treated in the discussion. RW displayed 60 head-nods, 39 of them with a verbalization and 21 without. These same nods occurred 11 times while speaking and 49 times while listening. Broken down into their respective types: 5 multiple nods, 8 triple head-nods, 17 double and 25 single nods, one half head-dip down and 4 head-shakes. His Australian co-participant displayed only 20 nods, of which 16 in relation to verbalizations and 4 without. Of the same nods 4 were displayed while speaking and 16 while listening. According to their type: one was a triple head-nod, 4 double nods, 13 single head-nods, once his head moved half up and on one occasion he shook his head. RW displayed 45 verbal listener responses; 37 reinforcers, one sentence completion, one request for clarification and 6 brief restatements. His Australian co-interactant exhibited 37 verbal listener responses of which 30 were reinforcers, one a sentence completion, two requests for clarification and 4 brief restatements.

We can conclude that the bilingual participant shows a higher frequency of listener responses when he interacts in the Japanese language than when he does so in the Australian-English language. Our hypothesis finds support.

5.3 DISCUSSION

In this chapter, data on the temporal organization of conversation were reported for one bilingual participant. Following an "internal variable" approach speaker-turns and listener responses

were examined in detail in a Japanese and an Australian condition. Based upon our investigations earlier it was predicted that shorter temporal cycles and a higher frequency of listener responses would occur when the bilingual participant spoke in Japanese as contrasted with longer temporal cycles and less listener responses when he would speak Australian-English. Results provide support for the hypothesis, while also revealing some other interesting aspects.

When conversing in Japanese the bilingual participant speaks in shorter lasting turns than when speaking in Australian-English. The average duration in both conditions is very similar to those obtained from other participants reported earlier. An interesting finding is that, in the Japanese condition, the bilingual participant has some rather long lasting speaking turns, resulting in a high standard-error (13.13). This deviation from the average duration is substantially higher than that of his co-participant or those of the two Japanese participants reported in chapter III. It was already pointed out that RW was not a "true" bilingual in the strict definition of the term -- a person speaking two languages, neither of which dominates the other [Lambert, 1955]. Some long lasting speaker turns suggests that something of a typical Australian temporal style is still operating when he is speaking in Japanese. Rather than two distinct speaking styles, a fusion of two styles seems present in both conditions.

In the Japanese condition, the bilingual participant gives more listener responses than in the Australian condition. The number of head-nods in the Australian condition is higher than expected in comparison with the other Australian participants. It is suggested that this higher number of head-nods is more typical of a Japanese style of conversing. On asking the bilingual participant about this afterwards he suggested that because he was interacting with Japanese speaking in Japanese in the course of his work, it was likely that he used behavioral elements of both languages when speaking either of them. Rather than code-switching from distinctive style to another, a fusion of the two codes seems to explain our findings better.

Another interesting observation occurring in the Japanese condition deals with posture. Whenever his co-interactant would ask a

question, the bilingual participant would typically shift himself backward, bringing his head further away and in some instances rest his hands behind his head. This posture-shift did not occur in the Australian condition. Because both participants in the Japanese condition were wearing spectacles and the bilingual participant in the most difficult position (due to the camera-angle) it was impossible to obtain reliable data on gaze-direction. This is regrettable because it would have been interesting to test our findings obtained in the preceding chapters.

Again our findings support differences in temporal aspects of speaking style related to the effect of language spoken.

One explanation for these differences would be one based on the differences in the structure of the Japanese and Australian-English language. To pursue this hypothesis would require an extensive investigation in itself and lead far beyond the scope of the present study. Here we will only briefly discuss some of the issues involved in such an investigation. It is popularly believed that the Chinese and the Japanese language have a lot in common. This is in fact incorrect, except for shared vocabulary, due to borrowing. True, the early influence of Chinese civilization including the Japanese writing system, an adaptation of the Chinese writing system, have had a profound influence on Japanese culture.

Whereas Chinese is a branch of the Sino-Tibetan language family, Japanese is distantly related to the Altaic language family. As such the Japanese language might have more in common with Turkish, Finnish, Hungarian and also with Korean, than it does with Chinese.

Chinese is an isolating language while Japanese is agglutinating. That is, in Chinese, each word tends to have an immutable form, but in Japanese, a word can have a number of shapes and be made of various morphemes that combine to represent different meanings.

Within a word in Japanese the morpheme divisions are *not* marked. When you hear the word: siki there is no obvious way to tell whether you are confronted with a one-morpheme word such as that for ceremony (written with a single Chinese character) or a two-morpheme word such

as that for the four seasons of the year (written with two Chinese characters) [Martin, 1975].

Not only did the Japanese borrow vocabulary from Chinese, it also includes words borrowed from Portuguese, Dutch and American-English. Throughout Japanese history new words were introduced whenever a culture had a predominant influence on Japan. Even nowadays the Japanese create their own Japanese-English expressions.

After this historical digression, let us now consider some of the structural differences between Japanese and English. One of the most obvious formal properties of utterances is their bracketing into constituents of various types, i.e. the "tree structure" associated with them [Chomsky, 1965]. Among such constructions are self-embedded constructions, multiple-branching constructions and left- and right-branching constructions. All these constructions as well as combinations of constructions are found in all languages, but some languages prefer one type of construction in preference to other constructions.

To the extent that a particular language prefers certain constructions as opposed to others and it would take more or less time to use one construction in contrast with others, some effect on the duration of utterances could occur. To give an example^{*} of one of the differences between Japanese and English language, English uses the mechanism referred to as "right-branching", e.g. a relative clause which modifies the head-noun comes to the right of it. (This is the house ... is expanded to its right by: This is the house that Jack built.) In Japanese, branching goes in the opposite direction from English, i.e. to the left of the head-noun. (Kore wa uchi desu ... will be expanded into: Kore wa Jakku ga tateta uchi desu.)

So we can see that the way in which sentences are "embedded" differs between Japanese and English.

Another structural difference dealing with the word order in the

* Examples given here are derived from Inoue [1977]. In each case the Japanese example is equivalent to the English example.

two languages is concerned with what is called "gapping" or "deleting". In English there is a process of deleting all but the first one of a series of identical verbs or verb phrases in co-ordinate sentences, i.e. "forward gapping". For example: John ate an apple, Jack ate a pear, and Bill ate a peach ... is simplified to: John ate an apple, Jack a pear, and Bill a peach. In Japanese, gapping works in the opposite direction and is called "backward gapping". All but the last of the identical verbs is deleted. The English example translated in Japanese would be: Jon ga ringo o tabemashita, Jakku ga nashi o tabemashita, soshite, Biru ga momo o tabemashita ... is simplified to: Jon ga ringo o, Jakku ga nashi o, soshite, Biru ga momo o tabemashita.

So far we have illustrated two types of constructions in which the structure of a Japanese sentence differs from the structure of an English sentence. First of all, these constructions referred to are theoretical constructions. Linguists derived them from breaking up sentences into various parts and describing the rules apparently followed to construct them. From an inspection of the two examples it will be apparent that to arrive at these constructions might have some impact on memory or processes occurring within a person, they do not seem to have any direct relationship to the duration of natural spoken language. Secondly, in so far as one language prefers one type of construction over another, the speakers of that language will make this construction as natural as speakers of other languages would make other constructions.

There are a number of other differences between Japanese and English. Japanese is a verb-final language. The basic word order of transitive sentences in Japanese is that of subject-object-verb, with a very rigid constraint that the verb appear at the end of the sentence. There is, however, considerable freedom as for the positions of subject and object in a sentence. For an example in English: John writes a letter, the Japanese equivalent would be: Jon ga tegami o kakimasu (or: tegami o Jon ga kakimasu). In the Japanese example the subject is marked by *ga* and the object by *o*. Verb final languages have a strong tendency to have "postpositions" instead of "prepositions" and in Japanese there are about seventy of them. *Ga* and *o* are two examples of these so-called "postpositions".

When there is an understanding between the speaker and listener as to what they refer to, noun phrases are often omitted in Japanese. So, if it is clear from the context in the earlier example that we talk about John, one might just say in Japanese: *tegami o kakimasu*. Since the postpositioned *o* marks that *tegami* is the object of the sentence, it is clear that somebody is doing the action. So an English translation of: *tegami o kakimasu* would be — (he) writes letters. Another important difference is related to the use of intonation in English and its non-use in Japanese. What we are referring to here is that the formation of, for example, a question in Japanese is done by an interrogative particle at the end of the sentence, and not like in English by changing the word order of the sentence and using an upward intonation at the end.

The way in which topics are marked in English is also done by intonation contours. By using a different intonation contour the meaning of the sentence can be altered. In Japanese, this distinction is not made by intonation, but by using different syntactic markers which are placed behind the word they refer to as, for example, when you are supplying information about the points of contrast between grammatically parallel adjuncts in two sentences, e.g. between two subjects or two objects. In English you will normally use stress (= "accentual prominence") to call attention to the two *items* in contrast, but in Japanese both the items are subdued (with *wa*) in order to play up their points of *difference*: *This* is big, but *that* is little, and its Japanese equivalent: *kore wa ookii ga, sore wa tiisai* [cf. Martin, 1975].

There are many other differences between the structure of the Japanese and the English language, such as the use of deferential language which is much more complex in Japanese, or the different forms used by either males or females which is much more prominent in Japanese [Inoue, 1977].

It is rather complicated to relate these structural differences directly to temporal aspects of speaking style, that is to say without further detailed investigation. Without clear evidence in which the effect of these and possible other differences has been investigated,

it is difficult to say whether such differences would account for all of the temporal differences found in spoken language between Japanese and Australian-English. Further investigation into the relationship of the structure of a language and the temporal aspects of spoken language warrants attention.

Another possible explanation of differences in the temporal aspect of spoken language is related to the use of consonants and vowels. Japanese is a monosyllabic language in which virtually all sounds are made up of consonant-vowel combinations, whereas in English there are many instances in which two or three consonants follow each other. (See also the difference between stress-timed and syllable-timed languages [Martin, 1972].)

Japanese utterances are analyzed in terms of moras. Each mora consists of either consonant plus vowel, or vowel alone, or mora nasal, or mora obstruent, and it functions as the unit of length in the language. By pronouncing each mora evenly it receives approximately the same duration. It is possible that to produce these mora in Japanese, which usually consist of consonant-vowel clusters, is a smoother process and takes less time than it would to produce two or three successive consonants, which is often the case in English. It has to be realized that what we are referring to here is different from the linguistic constructions mentioned earlier. Again we are not aware of any research which would enlighten us in the matter. Experimental studies comparing such differences would not satisfactorily answer all these questions. Without further evidence we cannot dismiss this explanation either. However, overall similarities in our speech rate values — average duration of a word in seconds — contradict differential production rates depending upon the language spoken.

It seems to us that at this stage linguistic aspects relating to the structure of Japanese versus English are insufficiently explored to explain the differences in temporal aspects of spoken language. Until counter-evidence proves us wrong, a more plausible explanation at present is an interactive one as opposed to one based on purely theoretical technical linguistic constructions.

The problem of time-sharing with which two participants in a conversation are faced may be handled differently when the participants speak different languages. When a person speaks in English there is an appropriate way of expressing oneself and the duration with which any speaker holds the floor might be part of this. In English one is allowed to take relatively long turns. In Japanese, however, the optimal rhythm is in general of shorter duration. Naturally, all sorts of situational variables can alter the nature of this process. It might be different when the participants are of different status, whether the occasion is a formal one, and so on. In the studies we have presented here the situation is uniform for everyone and status-differences or sex-variables are controlled.

Earlier on, studies in which "the Japanese" were described as being more formal, reserved, polite, and so on were mentioned. It is possible that certain rules about interactions and conversations relating to being more formal and careful in expressing oneself have some bearing upon the actual duration that a person holds the floor. On the basis of the present material, it is very difficult to make any generalizations about this. It might, however, provide a link to the numerous studies on Japanese personality and culture carried out just after the second world war. Presumably some of the observations made by the authors of those studies were based upon the way Japanese conversed with one another.

The data presented here are derived from only one participant, mainly because true bilinguals in the Japanese and Australian-English language are not very commonly found. It is therefore difficult to generalize from the findings presented here. In a study by Gallois and Markel [1975] we find some support for the effect of language spoken and differences in temporal aspects of speaking style. The comparison in their study is between Spanish and English, spoken by male Cuban bilinguals. Participants spoke longer and more frequently per turn in English than in Spanish, which Gallois and Markel [1975] interpreted as reflecting their identification with two different social groups.

Our results show differences in temporal aspects of speaking

style between Japanese and Australian-English, with Japanese being more like Spanish in having shorter and more frequent turns.

An interesting question is in what way Japanese and Spanish conversations are more alike than are American-English and Australian-English conversations. An answer to this question might provide some additional information to explain temporal differences in speaking style as related to language spoken.

Having established that there is a temporal difference in speaking style we will now turn our attention in the next chapter to what happens when Japanese participants interact with Australian participants speaking in English.

CHAPTER VI

TEMPORAL ASPECTS OF INTERCULTURAL
CONVERSATIONS (VIDEO)6.1 GAZE-DIRECTION, SPEECH, FREQUENCY OF
LISTENER RESPONSES AND GESTICULATIONS
IN INTERCULTURAL INTERACTIONS

In this part gaze direction, speech, the occurrence of listener responses and in addition the time spent gesticulating are examined in 9 intercultural (Japanese-Australian) interactions. Earlier Japanese-Japanese and Australian-Australian interactions examined were found to differ in their temporal organization. An investigation on one Australian "bilingual" participant in a Japanese and an Australian condition revealed this difference in temporal organization for speech and the occurrence of listener responses related to it, to be related to the use of either the Japanese or the Australian language.

Here the behavior of 9 Japanese interacting with 9 Australian participants in the English language is examined. Ideally the reverse condition of Australians interacting with Japanese using the Japanese language should also have been included. Apart from the earlier mentioned "bilingual" participant, however, no other Australian participants with sufficient fluency in Japanese could be found.

The investigation reported here is exploratory, so no specific hypotheses are put forward. Based upon results obtained from the bilingual participant and common sense, one could expect the Japanese participants to behave like Australians if they have "mastered" the English language fully. The Japanese interactants participating in the study have all been taught what can be described as "American English" for a period of six years or more. It must be pointed out, however, that this teaching covers traditionally grammar and vocabulary mainly. Only in more recent times conversational "American

"English" is becoming more and more fashionable. Even now, however, many Japanese do not have sufficient chance to interact with "English" speaking foreigners. The Japanese participating in this study can be described as having more exposure to "English" speaking people than most of their fellow countrymen. It will be recalled (see page 97) that even they are, with one exception, relatively speaking recent arrivals in Australia — they have been in Australia for less than two years, and about half of them as little as two to three months.

6.2 METHOD

The method used for data analysis is identical to the one described in chapter IV. The only addition is the rating of time spent gesticulating during the interaction. For this purpose the 21 conversational interactions were rated again, in a similar way as described earlier (see page 100). The same raters, who had monitored speech, did so again, while two new raters were trained to attend to gesticulation. The training consisted of obtaining agreement about what was considered a gesture and what not. The training period did not take a long time, because the raters knew intuitively how to distinguish gestures from other movements. According to Kendon [1975, p.353], gesticulation is to be distinguished from movements such as those engaged in when the individual scratches his head or adjusts his clothing. Gesticulation occurs only when someone is actively engaged in an utterance, usually within a focussed interactional event.

Again, every time the interlocutor, to whom they were attending, gesticulated they pressed a switch and maintained switch pressure as long as the interlocutor gesticulated. The ratings of gestures were used in a similar fashion as mentioned for the speech data earlier.

Reliability measurements carried out on the rating of gesticulation yielded 0.88 for test-retest reliability and 0.90 for inter-rater reliability.

The data reported in this part compare 9 Japanese participants in intra- (Japanese-Japanese) and inter-cultural (Japanese-Australian)

exchanges and 7 Australian participants in similar conditions. Each time the behavior of an individual participant is compared over these two conditions, taking into account the behavior of the other interactants in both conditions. Most of the comparisons will be made for the conversational interactions (P2), but data on the game-period (P1) will also be briefly discussed.

6.3 RESULTS

6.3.1 Gaze-Direction

After transcription and segmentation of speech and gaze-direction was carried out, the number of frames in which the participants engaged in speech and q-directed gaze, speech and a-directed gaze, non-speech and g-directed gaze and non-speech and a-directed gaze, was written down. The tables refer to the totals of each of the aforementioned categories over the period analyzed. Table XXII shows gaze-direction in intercultural interactions. The participants engage here approximately 44% of the total time in speaking. While speaking, the participants engage in about the same percentage of time in looking at the other's face (q-directed gaze) as in looking away from his face (a-directed gaze).

While non-speaking (listening), they engage three times more in q-directed gaze (42%) than in a-directed gaze (14%). When we examine however the data for individual participants in intra-and intercultural interactions in tables XXIII (for the Japanese participants) and XXIV (for the Australian participants), it is apparent that the patterns for gaze-direction are different for the two groups.

It has to be pointed out, however, that for the segments examined the Japanese participants increase their time spent speaking from intra- to inter-cultural interactions, while the reverse is true for the Australian participants. It will be shown later that this is not typical for the total 5 minute period for which data on speech are examined.

Table XXII
Gaze-Direction in Intercultural Interaction
(random 10 second periods)

Speech			Non-Speech	
q-gaze	a-gaze	total number of frames = 9000	q-gaze	a-gaze
3981			5019	
1954	2027	number of frames	3747	1272
108.56	112.61	average number of frames per subject	208.17	70.67
2.17	2.25	average duration per subject in seconds	4.16	1.41
(21%)	(23%)	in percentages	(42%)	(14%)

In intercultural interactions, the Japanese participants show a similar pattern of gaze-direction as they do in intra-cultural interactions. They engage less in q-directed gaze, both while speaking (22%) and while listening (16%), than they do in a-directed gaze (39% and 23%). The Australian participants show in intercultural interactions a pattern of gaze-direction which is largely the same as the one they show in intra-cultural interactions. Examining the data in greater detail, there is a tendency for Japanese participants to spend proportionally more time in looking away from the other's face in intercultural interactions than in intra-cultural ones, in particular while speaking. The Australian participants, on the other hand, tend to engage more time in gazing at the other's face in intercultural interactions as contrasted with intra-cultural ones, in particular while listening.

Again we find that looking away (a-gaze) is more dominant for Japanese participants and looking at one's co-interactant more characteristic for Australian participants, even in intercultural interactions. One can go one step further and claim that in intercultural interactions the patterns of gaze-direction described as characteristic for the two groups becomes even more salient for both.

Table XXIII
Gaze-Direction in Japanese Participants in J-J and J-A Interactions

Intra-Cultural					Inter-Cultural				
Speech		Non-Speech		total number of frames = 4500	Speech		Non-Speech		
q-gaze	a-gaze	q-gaze	a-gaze		q-gaze	a-gaze	q-gaze	a-gaze	
1860		2640			2727		1773		4500
852	1008	1073	1567	number of frames	998	1729	725	1048	
94.67	112	119.22	1741.11	average number of frames per subject	110.89	192.11	80.56	116.44	
1.89	2.24	2.38	3.48	average duration per subject in seconds	2.22	3.84	1.61	2.32	
(19%)	(22%)	(24%)	(35%)	in percentages	(22%)	(39%)	(16%)	(23%)	

Table XXIV
Gaze-Direction in Australian Participants in A-A and A-J Interactions

Intra-Cultural				total number of frames = 3500	Inter-Cultural			
Speech		Non-Speech			Speech		Non-Speech	
q-gaze	a-gaze	q-gaze	a-gaze		q-gaze	a-gaze	q-gaze	a-gaze
1467		2033			934		2566	3500
1167	300	1593	440	number of frames	752	182	2504	62
166.71	42.86	227.57	62.86	average number of frames per subject	107.43	26	257.71	8.86
3.33	0.86	4.55	1.26	average duration per subject in seconds	2.15	0.52	7.15	0.18
(33%)	(9%)	(46%)	(12%)	in percentages	(22%)	(5%)	(71%)	(2%)

6.3.2 Speech Data

For the purpose of comparing the cumulative time spent in speaking and the frequency of speech units, a two-way analysis of variance was carried out in which the three groups (Japanese-Japanese, Australian-Australian and Japanese-Australian) was the between factor and the type of interaction (P1 = game and P2 is conversation) the within factor.

	P1	P2	P1	P2
	Cumulative Speech		Frequency of Speech	
J-J	35.61	98.76	35.64	75.43
A-A	12.42	92.10	10.20	62.20
J-A	12.53	103.30	12.83	69.33

Above, the means for each of the groups in each of the conditions is shown. It can be readily seen that for all three groups a difference between the game situation and conversation occurred. Interaction during the same period is naturally severely constrained by the rules and objectives of the game minimizing speech ($p < 0.01$).

The three groups do not differ on total time engaged in speaking ($F = 1.69$, ns). They do however differ from each other in the way this speech is co-produced ($F = 8.58$, $p < 0.01$). The Japanese talked most frequently, the Australian interactants least frequently, and in intercultural interactions the frequency of talking is somewhere in between the two other groups.

Overall, the three groups reacted in a similar way to the two different situations ($F = 1.59$, ns). That is to say, the Japanese participants engaged in more speech during the game than did the other two groups of participants. For the same data, speech equilibrium-ratios (ratio of the participation between interactants in each dyad) were calculated to give an indication of the reciprocity in the interaction. The most interesting result revealed by this method is that the amount of participation is less equal during the game period for all three groups of participants. As reflected by the lower equilibrium ratios in the intercultural (Japanese-Australian) interactions during the game period, exchange here is most unbalanced.

But even during conversation, intercultural interactions show less equal participation than intra-cultural ones.

Equilibrium-Ratio* Speech									
Intra-Cultural Interactions									
Japanese-Japanese					Australian-Australian				
P1		P2			P1		P2		
pairs	cum.	freq.	cum.	freq.	pairs	cum.	freq.	cum.	freq.
1	0.69	0.87	0.52	0.91	8	0.63	0.55	0.42	0.31
2	0.28	0.78	0.73	0.95	10	0.21	0.23	0.86	0.92
7	0.97	0.64	0.34	0.57	11	0.75	0.50	0.79	0.92
9	0.50	0.85	0.75	0.80	15	1.00	0.78	0.97	0.68
13	0.48	0.88	0.39	0.66	21	<u>0.37</u>	<u>0.38</u>	<u>0.73</u>	<u>0.69</u>
17	0.99	0.93	0.88	0.90		0.59	0.49	0.72	0.70
18	<u>0.70</u>	<u>0.66</u>	<u>0.88</u>	<u>0.84</u>					
	0.66	0.80	0.64	0.80					

Inter-Cultural Interactions

Japanese-Australian					
P1			P2		
pairs	cum.	freq.	cum.	freq.	
3	0.21	0.60	0.75	0.93	
4	0.33	0.67	0.50	0.51	
5	0.03	0.11	0.39	0.68	
6	0.32	0.19	0.53	0.88	
12	0.21	0.45	0.23	0.34	
14	1.00	0.50	0.82	0.82	
16	0.25	0.31	0.43	0.68	
19	0.39	0.52	0.19	0.29	
20	<u>0.77</u>	<u>0.84</u>	<u>0.73</u>	<u>0.69</u>	
	0.39	0.47	0.51	0.65	

* Equilibrium-ratios were obtained by dividing the shortest time for cumulative speech with the longest, likewise for frequency of speech.

The data on speech mentioned so far do not show the contributions of the Japanese and Australian participants in intercultural interactions. Table XXV shows the individual data for each of the participants in intercultural interactions, while tables XXVI and XXVII show data for speech comparing intra- with inter-cultural conversations for Japanese and Australians respectively. The overall time spent speaking decreases slightly for Australian participants and remains virtually unchanged for Japanese participants. The frequency of speech is similar in intra- and inter-cultural interactions for Australian interactants, while the Japanese tend to decrease the frequency in intercultural as compared to intra-cultural interactions. It must be pointed out, however, that these observations are made on the basis of the averages for the groups. Individual participants change in different directions depending upon the contribution of the co-participants, their skill in using the English language, the topic of conversation, and so on. These factors will be treated more fully in the discussion.

Table XXV

Data for Speech for 9 Japanese-Australian Interactions during P₂

	Left				Right			
	cum.	freq.	aver.	perc.	cum.	freq.	aver.	perc.
M ₁	100.5	79	1.27	33.50	134.7	85	1.58	44.90
M ₂	71.1	44	1.62	23.70	141.3	87	1.62	47.10
M ₃	45.3	40	1.13	15.10	119.1	59	2.02	39.70
M ₄	155.7	89	1.75	51.90	82.8	78	1.06	27.60
M ₅	35.4	37	0.96	11.80	153.6	108	1.42	51.20
M ₆	99.9	68	1.47	33.30	81.9	56	1.46	27.30
M ₇	177.3	81	2.19	59.10	77.1	55	1.40	25.70
M ₈	32.1	30	1.07	10.70	172.2	102	1.69	57.40
M ₉	75.6	61	1.24	25.20	103.8	89	1.17	34.60

Table XXVI
Data for Speech in 9 Japanese Participants
in J-J and J-A Interactions during P₂

	Intra-Cultural				Inter-Cultural			
	Cum.	Freq.	Aver.	%	Cum.	Freq.	Aver.	%
J ₁	144.9	67	2.16	48.3%	100.5	79	1.27	33.5%
J ₂	74.1	62	1.20	24.7%	141.3	87	1.62	47.1%
J ₃	168.6	123	1.37	56.2%	155.7	89	1.75	51.9%
J ₄	106.5	90	1.18	35.5%	177.3	81	2.19	59.1%
J ₅	80.4	72	1.12	26.8%	99.9	68	1.47	33.3%
J ₆	142.2	85	1.67	47.4%	153.6	108	1.42	51.2%
J ₇	85.5	91	0.94	28.5%	75.6	61	1.24	25.2%
J ₈	96.3	82	1.17	32.1%	45.3	40	1.13	15.1%
J ₉	91.2	63	1.45	30.4%	32.1	30	1.07	10.7%
		82		36.65%		71		36.34%

Table XXVII
Data for Speech in 7 Australian Participants
in J-J and J-A Interactions during P₂

	Intra-Cultural				Inter-Cultural			
	Cum.	Freq.	Aver.	%	Cum.	Freq.	Aver.	%
A ₁	147.9	81	1.83	49.3%	82.8	78	1.06	27.6%
A ₂	104.1	66	1.58	34.7%	119.1	59	2.02	39.7%
A ₃	68.4	55	1.24	22.8%	71.1	44	1.62	23.7%
A ₄	103.2	50	2.06	34.4%	35.4	37	0.96	11.8%
A ₅	106.2	74	1.44	35.4%	81.9	56	1.46	27.3%
A ₆	75.6	61	1.24	25.2%	103.8	89	1.17	34.6%
A ₇	103.8	89	1.17	34.6%	172.2	102	1.69	57.4%
		68		33.77%		66		31.73%

6.3.3 Listener Responses

The frequency of listener responses in intercultural interactions was also examined. Table XXVIII shows the frequency of listener responses in intercultural interactions. In tables XXIX and XXX the occurrence of listener responses for Japanese and Australian participants respectively in intra- and inter-cultural interactions are compared. Again one has to take into account the contribution made by the other interactant in each dyad. When one is faced with someone who talks very little and takes very few turns, one will naturally display less listener responses than when faced with someone who keeps talking all the time.

For the Japanese participants it looks as though about half the interactants increase the frequency of listener responses in intercultural interactions as compared to intra-cultural ones, while the other half decrease the frequency of listener responses.

Upon a closer inspection of the data, in relation to the frequency with which the other interactant contributes to speech, it can be seen that with participants J5, J6, and J8, their counterparts in intercultural interactions all speak far less than their counterparts in intra-cultural interactions. When conversing with another Japanese, J5's counterpart contributes 90 turns; while in intercultural interaction the co-interactant of J5 contributes only 56 turns. For participant J6 these figures are from 56 to 37 and for J8 from 91 to 59. Only participant J7 produces far less listener responses in an intercultural situation than in an intra-cultural one, despite his co-interactant in both cases contributing a similar frequency of turns — his Japanese counterpart 85, his Australian co-interactant 89. The Australian participants all increase the frequency of listener responses considerably in intercultural interactions as contrasted with Australian-Australian interactions. The only participant who does not do so (A7) is faced with a co-participant who only contributes 30 turns to speech in intercultural interaction, "forcing" A7 to do virtually all of the talking, while his Australian counterpart contributes 89 turns.

Table XXVIII
Frequency of Listener-Responses in 9 Japanese-Australian
Interactions over a 5-Minute Period

M ₁		M ₂		M ₃		M ₄		M ₅		M ₆		M ₇		M ₈		M ₉	
Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
33	40	16	16	16	14	25	26	11	12	15	36	42	48	48	6	13	20
25	43	9	34	10	12	21	43	19	16	19	33	21	45	26	11	27	25
1	2	1	-	-	-	-	-	-	-	-	-	-	1	1	-	2	3
1	-	-	-	-	-	-	-	-	1	-	1	-	2	2	-	-	2
6	4	-	2	4	1	5	3	1	4	5	-	-	1	1	-	5	3
-	-	3	3	1	1	1	2	13	2	3	1	1	1	6	1	2	1

Table XXIX
Frequency of Intra- vs Intercultural Listener Responses
for 9 Japanese over a 5-Minute Period

	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter
Head-nods	27	33	8	16	14	25	16	42	21	15	37	12	32	13	34	16	35	48
Reinforcers	13	25	31	34	15	21	21	21	13	19	31	16	35	27	23	10	30	26
Sentence completions	2	1	2	-	1	-	1	-	2	-	-	-	1	2	-	-	-	1
Request for clarification	-	1	6	-	3	-	2	-	2	-	1	1	2	-	1	-	-	2
Brief restatement	3	6	1	2	-	5	4	-	2	5	2	4	7	5	8	4	4	1
Laughter, etc.	3	-	4	3	-	1	1	1	2	3	1	2	-	2	1	1	3	6

Table XXX

Frequency of Intra- vs Intercultural Listener-Responses for 7 Australians over a 5-Minute Period

	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter
Head-nods	6	26	3	14	9	16	3	11	22	36	19	20	22	6
Reinforcers	4	43	18	12	7	9	10	19	21	33	30	25	25	11
Sentence completions	-	-	3	-	1	1	-	-	-	-	1	3	-	-
Request for clarification	1	-	-	-	1	-	1	-	-	1	-	2	-	-
Brief restatement	1	3	-	1	2	-	1	1	2	-	3	3	2	-
Laughter, etc.	-	2	4	1	-	3	-	13	-	1	1	1	1	1

6.3.4 Mean Time Spent Gesticulating

The final analysis deals with the mean time spent gesticulating for each participant in both intra- and intercultural conversations. Since only 7 Australians participated in both intra- and in intercultural interactions, an analysis of variance was carried out on the data of 14 participants. The between factor is Japanese or Australian participant, the within factors intra- versus interculture and the period in the conversational interaction (for this purpose the interaction was divided in five 1-minute periods). This last variable was included after the observation that some participants seemed to concentrate the use of gestures at later stages during the conversations. By splitting up the period of conversation it was hoped to find if any trend was apparent.

The data in table XXXI show the time spent gesticulating for each of the five periods (means and totals in seconds).

There is no difference between Japanese and Australian participants. The difference between talking to someone of your own or to someone of another culture (Self/other) or between periods also was not significant. However for culture $\times$ Self/other an interaction-effect emerged ($F = 5.61$, $p < 0.03$). The Japanese participants increased the use of gestures by approximately 300% in intercultural interactions as opposed to intra-cultural ones. This occurred throughout the interaction and no trends emerged for any differences in any of the stages of the conversation. The Australian participants tended to decrease the use of gestures slightly, but this finding is partially accounted for by a slight decrease in the time they spent speaking in intercultural interactions.

6.4 DISCUSSION

In this chapter data on intercultural interactions, in which Japanese participants converse with Australians in English, are reported. Following an external variable approach the conversations are examined for gaze-direction, speaker turns, listener responses and

Table XXXI

Mean Time Spent Gesticulating over Five 1-Minute Periods

	Japanese		Australian	
	Self	Other	Self	Other
1.	2.14	7.84	3.52	1.71
2.	2.91	8.74	2.66	1.07
3.	2.10	12.39	8.61	5.53
4.	5.44	10.50	1.54	2.74
5.	2.70	7.54	0.39	2.87
	3.06	9.40	3.34	2.79

Analysis of Variance

Source	df	Ms	F	P	
(a) Between subjects	13				
Culture (C)	1	351.078	3.673	0.076	ns
Error	12	95.573			
(b) Within subjects					
Self/other (SO)	1	292.902	2.944	0.067	ns
C × SO	1	416.587	5.610	0.03	0.05
Error	12	74.261			
Period (P)	4	66.075	2.129	0.09	ns
C × P	4	30.487	0.982	0.57	ns
Error	48	31.039			
SO × P	4	4.611	0.181	0.94	ns
C × SO × P	4	21.675	1.242	0.30	ns
Error	48	25.500			

gesticulation. This study is exploratory in nature and no specific hypotheses are put forward. In so far as the Japanese participants would be fluent speakers of English, results from the bilingual participant would lead us to expect their temporal aspects of speaking style to be similar to those found for Australian-English conversations. Since some of the Japanese participants have arrived in Australia only recently difficulties with the English language were to be expected, and this in turn could influence the temporal aspects

of speaking style.

In earlier chapters, patterns for gaze-direction characteristic for Australian - and Japanese interactants emerged. In intra-cultural conversations, looking away from the other's face was found to be more typical in Japanese conversations and looking at one's co-participant was found to be more characteristic in Australian conversations.

These patterns of gaze-direction become even more salient in intercultural conversations. That is to say, other-directed gaze for Japanese participants decreases in intercultural conversations as compared with intra-cultural ones, in particular when speaking. In contrast Australian participants look at their Japanese co-interactants virtually continuously when listening. Compared to intra-cultural interactions, the Australian interactants increase their other-directed gaze, in particular when listening. One of the reasons that Japanese interactants look away even more when speaking English is because looking away from the other's face can enhance cognitive processes. To some extent looking at one's co-participant and planning an utterance are incompatible activities under normal conditions, i.e. when speaking one's native language. Here the Japanese participants speak a language they have not fully mastered. Looking less at one's co-participant could facilitate the cognitive processes involved in planning and producing utterances in a language which is not their native one.

The Australian participants may increase their other-directed gaze when listening in order to facilitate speech perception, since the Japanese interactants speak English less fluently and with an accent.

Another possibility is that the different physiognomy of Japanese increases other-directed gaze among Australian participants. From the author's own experience, it took him about two or three weeks, when in Japan for the first time, before he could distinguish one Japanese from another.

Whether it is because of the different physiognomy, or whether it is because speech-perception is facilitated that Australian

participants increase their other-directed gaze in intercultural conversations could be tested. One would set up conversations between, for example, Eastern-European participants who would speak English with an accent and less fluently but who would otherwise look like Caucasians, and Australians. If the same pattern of other-directed gaze as found here would emerge, the speech-perception hypothesis would be supported. If the Australian participants do not increase the time spent in other-directed gaze, the different physiognomy hypothesis would be supported. The most likely explanation is that both hypotheses are necessary to explain the virtual continuous looking at the other which Australian participants engage in when conversing with Japanese.

A third possibility is that through other-directed gaze the Australian participants show attention and provide the speaker with encouragement to continue speaking. The plausibility of this hypothesis is increased if one accepts that something akin to a "host"-effect is operating in intercultural conversations. After all, the Australian participants are on their homeground and speak their own language. Consequently they are more responsible for getting the other to feel at ease, and one way of doing that is by encouraging him, by paying attention, and so on. Other-directed gaze would be used for that purpose. Supporting the fact that Australian participants encourage their Japanese co-participants is the higher number of listener responses the Australian participants give in intercultural as compared with intra-cultural conversations.

Ironically the increased other-directed gaze on the part of the Australian "host" might be perceived by the Japanese participant to be too imposing and embarrass him. The Japanese participant might want to avoid the imposition and one way of doing just that is through looking away. What we are sketching here is the cumulative effect opposing patterns of gaze-direction might create when these patterns function in such an approach-avoidance regulatory fashion. Maybe it would explain why the patterns of gaze-direction in intra-cultural conversations become even more salient in an intercultural one.

Our data on gaze-direction illustrate the point made by Ellsworth

and Ludwig [1972] that cross-cultural communication difficulties arise because of a conflict between two highly complex, totally overlearned patterns of behavior.

Unlike the data on gaze-direction, the data on temporal aspects of speaking style are much more difficult to interpret. As we mentioned earlier some of the Japanese participants were more fluent in English than others, and some had been in Australia for longer periods than others. This is probably the main reason why the Japanese participants differ so much in the extent that temporal aspects of speaking style change from intra-cultural to intercultural conversations. Besides these individual differences, differences are also related to the speaking style of their co-participants, while varying topics of conversation might also contribute to the variance. The majority of Japanese participants, however, tend to decrease the number of speaker turns from intra-cultural to intercultural conversations. These participants therefore adopt a style of speaking which is in its temporal aspect more typical for Australian-English conversations. The minority of participants who do not follow this pattern would either have a lot of difficulty with the English conversation or the contrast between intra- and intercultural conversation was great indeed. J1 for example converses in Japanese with a co-participant, who lets him do most of the talking. In an intercultural conversation J1, who has arrived only recently and speaks English rather poorly, has a lot of difficulty to keep up the same pace as when speaking Japanese. The resulting difference in total time spent speaking as well as frequency of turns is great and opposite to that expected.

Although the time spent talking in intra-versus intercultural conversations does not differ significantly, there is a tendency to speak more in intercultural conversations. Participants in intercultural conversations pay more attention to the other's right for his turn and silences are avoided as much as possible. Periods of silence or long pauses are comparatively rare in intercultural conversations, while simultaneous talking also occurs less in intercultural than in intra-cultural conversations. Our data dealing with the reciprocity of the two participants in the conversation (equilibrium-ratios) show

that the equal participation is less likely in intercultural interaction than in intra-cultural interaction, in particular when a task is involved, but also in conversations. In some conversations it is the Japanese who does more of the talking, in others it is the Australian. This finding might have some implications for conferences, negotiations, etc., between persons of different cultures, in particular when some sort of task is involved.

In the game period, the Japanese participants talked more than either the Australian participants or the intercultural pairs. In intercultural interactions, the Japanese participants did most of the talking. The reason for this is probably that a "smooth" relationship between the interactants when a competitive game is involved is more important for Japanese. The importance of the contribution of the co-participant is particularly important when considering the data on listener responses. The frequency of listener responses is directly related to the number of speaker turns one's co-participant makes. Having considered the contributions of co-participants in the result section, the majority of the Japanese participants, and all but one of the Australian participants, increase the number of listener responses from intra- to intercultural conversations.

For the Australian participants this increase is possibly related to what we referred to as a "host"-effect. They are on homeground and speak their own language and consequently act as hosts to the Japanese visitor. From the polite questions "where do you come from", "how long are you here", "do you have this in Japan", etc., they often took it upon themselves to get the conversation going. Once underway they would encourage their Japanese co-participants, in particular when they could not find a word or had difficulty in expressing themselves. Listener responses, together with other-directed gaze, were employed in this pattern of encouragement. The only Australian participant who did not do so conversed with a partner who contributed relatively little to the conversation. Actually in this conversation the roles were reversed. Because the Australian participant had just returned from a holiday in Japan, he started describing his experience there to his Japanese co-participant. He in turn took the role of "host" through helping the Australian to retrieve information about places he

had visited and asking him questions about it.

As far as the Japanese participants increased the number of listener responses, two explanations seem plausible. First of all, through listener responses it is possible to encourage the other to keep on continuing to speak, in that way not having to do too much work oneself. In particular, when one is having difficulty with the language used, this would seem only a sensible strategy. Secondly, through displaying listener responses at the appropriate times one can give the speaker the impression that one follows what he is saying, even when one in fact does not. Anybody, who has ever been in contact with people whose language one does not speak or whose language one only understands to a limited extent, will know the importance of smiles, nodding, etc., in order to make interaction possible.

An interesting observation on listener responses deals with a sentence completion. One of the Australian participants spoke about policemen in Tokyo wearing gasmasks as a safeguard against pollution. While saying his utterance, the Japanese participant simultaneously signed gasmask in order to complete the sentence. This is one illustration of a sentence completion carried out without using any words. Many other examples probably occur and it would seem that a classification for listener responses should be wide enough to incorporate such instances. A classification based upon verbal listener responses and head-nods only seems too narrow.

Finally we reported data on the time participants spent gesticulating in the different conversations, and found that Japanese and Australian participants talking to their fellow countrymen did not differ from each other. In intercultural conversations, however, the Japanese participants increased gesticulation by 300%, while there was no marked difference for Australian participants. It was pointed out that Japanese participants were having difficulty in finding the right verbal forms for what they had to say, because they spoke in a language which they did not fully master. In an article on gesticulation and language origins, Kendon [1975] discusses his observations and points out that gesticulating is an integral part of the whole act of uttering. He refers to de Laguna [1927] and Werner

and Kaplan [1963] who suggested that gesticulation would increase when the speaker was having difficulty expressing himself. Kendon [1975] also gives another occasion when gesticulation increases, namely where feedback from the listener suggests difficulty in comprehension. The high level of other-directed gaze, as well as the increased number of listener responses for Australian participants in intercultural conversations, makes this suggestion, in our case, unlikely.

CHAPTER VII

FINAL DISCUSSION AND CONCLUSIONS

7.1 SUMMARY AND INTEGRATION

The final chapter consists of two parts. The first part will briefly restate the arguments of the present study and suggest that the temporal organization of conversation is an important aspect of it. In the second part we will elaborate on this notion of temporal and rhythmical organization, discuss the way various researchers have investigated rhythms, and suggest further research in this area.

We discussed several levels of the study of communication. We argued that, especially when studying communication cross-culturally, it is more fruitful to study observable behavior rather than making inferences about that behavior. We then went on to discuss the general theoretical framework and the different methodological approaches which have developed in the area of face-to-face interaction.

The problem is that the area of face-to-face interaction at large suffers from a critical lack of an adequate theoretical framework. As Yngve [1975, p.313] observed: "research is impeded because we lack the understanding and insight that a satisfactory theory would give and the guidance it would provide for observational and experimental programs. Lacking an appropriate theoretical structure, we find it difficult to state results and more difficult to relate them with one another. There is a general consensus both for the intrinsic interest and importance of the phenomena studied, and the desperate need for an appropriate framework".

The only area within face-to-face interaction which has received a considerable amount of the attention of researchers is the study of conversation. The vast literature in this area has dealt, however,

almost exclusively with what goes on in conversations between English speakers. Even in the study of conversation, theory and methodology is inadequate.

Kendon [1975] states the need to develop theory and methodology and a well articulated approach before attempting to deal with the problem of the variety of ways in which communicative behavior is organized in interaction in different cultures. One of the basic problems is that data on conversations in different cultures are scarce. If we cannot investigate the organization of conversation cross-culturally until the appropriate methodological and theoretical framework is available, we cannot discuss the whole issue of cross-cultural comparisons of communicative behavior, both in its theoretical and practical implications.

The viewpoint taken in the present investigation is that until we do investigate the organization of conversation cross-culturally, even in the absence of an adequate framework, whatever theoretical and methodological framework we will ultimately come up with is sure to be deficient in that it will not be able to take cross-cultural variation into account. This assumes that our framework can be based upon our investigations of English speakers only. The fallacy of this assumption can be readily shown if we consider for a moment what has happened historically in the study of intelligence, perception, personality and so on.

In all of these areas, theories and methodological practices became established doctrines only for it to emerge later that their validity was limited to a particular culture or group of cultures. Unless cross-cultural comparisons are included from the early stages of any field in the behavioral sciences, a similar road lies ahead in the study of face-to-face interaction. It would seem to be far more cumbersome to have to undo theories than to allow for and even include variations from the outset.

While we agree that no appropriate theoretical framework for the study of conversation exists, it is arguable that an appropriate methodological framework is also unavailable. (The arbitrary division

between theory and method is difficult to maintain.) We have argued that interaction theory uses operational procedures and definitions invented for the study of human social and cultural behaviors and it does this through the discrimination, measurement and sequencing of interpretational interactions [Arensberg, 1973]. Its methods yield a way of handling the record of observation capable of empirical quantification and generalization, yet leading to deduction and prediction capable of replication and verification. To use again Arensberg's [1973] terminology, it abandons the anthropomorphism of inferred meanings and the ethnocentrism of received definitions. One of Chapple's greatest contributions was the realization that the temporal organization of interaction is one of its major dimensions. The methodology to measure these temporal aspects has been available for quite some time [Chapple and Arensberg, 1940].

In our discussion of the major variables which have received attention from researchers in the study of conversation, we paid particular attention to their temporal patterning. Although the development of the new hardware of film and video allowed us to study the variables selected in greater detail and with more accuracy than was possible in Chapple's early days, we were still faced with the need to summarize, order, compare and generalize our data into something infinitely shorter and less cumbersome than our original records [cf. Arensberg, 1973].

Focussing on temporal aspects of certain selected variables in conversation allowed for a descriptive summary in universal, comparable and measurable terms and definitions. The variables selected for measurement — gaze-direction, speaker turns and listener responses — have emerged in the study of conversation as being the most pertinent. Other variables however also play some role.

The problem investigated in the studies here suggested itself as a result of personal experience of the author. In that sense the initial phase consisted of participant observation, with the author as a participant. Ideally it is probably best to start one's research in this area along the lines of context analysis. This allows the researcher to immerse himself in a certain problem and thus to

consider the greatest possible range of variables involved. Such an initial investigation can be carried out as a participant or when permanent data are available even on a restricted sample. Here what we called "the internal variable approach" was used as a second phase to discover a pattern, or the way in which the different variables investigated were related. This approach has the advantage of limiting the number of variables involved, while still focussing on the behavioral relationship between the interactants and the relationship between the variables. Under different circumstances even the topic of conversation would be far too complex to allow us to single out variables for immediate study. Given that the variables selected in the present study were documented quite well, at least for English speakers, a more rapid progress was possible.

The problem in any research endeavour is a possible trade-off between a qualitative but vague richness, dependent upon the perceptive skills and intuitive hunches of the individual analyst, and a more precise and more replicable, but more narrow identification of patterns. This is probably the reason why the different approaches, as discussed in chapter I, developed in the first place. This trade-off is also evident in the present study. Basically the line followed here is one of gradual progression. After defining an area for research and selecting specific variables, material for analysis was gathered. In the first study reported in chapter III we searched for a pattern.

In this detailed investigation, it was found that the Japanese conversation had a different temporal organization from that of the Australian conversation. All the variables investigated were shown to be closely related. The Japanese interactants had more frequent and shorter lasting other-directed gazes than did the Australian participants. It was suggested that the regulatory function of gaze-direction, in particular at the end of utterances, is possibly less important for different cultural groups. Closely associated with the pattern of gaze-direction were the findings that Japanese participants co-produce conversation to a greater extent than do Australian participants. A greater frequency of listener responses was also found for the Japanese interactants as compared with the Australian

participants. The most important general conclusion about the first study was the different temporal organization of the Japanese conversation as compared with the Australian conversation. It was the temporal organization which related the different variables and from this finding we drew the conclusion that variables such as gaze-direction and speech related variables have to be considered in relation to one another.

In chapter IV we set out to test a number of hypotheses based upon the differences in the temporal organization of Japanese and Australian conversations. Basically we predicted shorter cycles for speech and gaze-direction and a higher number of listener responses in Japanese conversations than in Australian conversations. Once the pattern was isolated, it then became possible to test hypotheses about particular aspects of the patterns. The approach followed was the external variable approach. For each of the variables a comparison between Japanese and Australian interactants was made. As spelled out in chapter I, the measurements here were more crude in nature than in the earlier detailed investigation. Our hypotheses about shorter temporal cycles in the Japanese conversations were confirmed and basically the pattern found earlier replicated. In the study reported in chapter IV the sample was larger than in the earlier detailed study and instead of using 16 mm film, video was used. The advantages of using film for more precise measurements and the relatively low cost of video were reasons for this pattern of use. Whereas in the first study only descriptive statistics were used, in the second study more sophisticated statistical techniques were used.

Chapter V describes the detailed investigation of one bilingual participant in a Japanese — and an Australian — condition. The method followed in this investigation was similar to the one used for our search for a pattern. The major findings showed the relationship between temporal aspects of speaking style and language. In the Japanese condition the bilingual participant spoke in turns of shorter duration than in the Australian condition. Rather than talking about two distinct speaking styles and clear-cut code-switching, we argued, based upon our results that a fusion of the different aspects of speaking style described our results better. Again, however, we found

further evidence that conversation in Japanese is co-produced to a greater extent than in Australian-English. We briefly pointed out some factors reflecting differences in the language structure between Japanese and Australian-English. At this stage we cannot substantiate the relationship between language structure and temporal aspects of speaking style adequately. We suggested, however, that further research in this area should be undertaken.

In the final investigation reported in chapter VI we focussed our attention on the temporal aspects of gaze-direction, speaker turns and listener responses and gesticulation in intercultural conversations. In these conversations, where Japanese participants talked with Australians in English, we discussed the important implications of the temporal organization and its changes, when a speaker is using a different language, especially one in which he is not fluent. Because of the large number of subjects involved, video was used and the method was the external variable approach. The findings were explained in terms of the different functions of gaze-direction and the relationship between the variables. One interesting result was the increased use of gesticulation, when the Japanese participants were speaking in English.

Based on the findings presented, we would like to stress that the temporal organization of conversation is an important aspect of it and should therefore feature in any theory of conversation. We have shown that focussing on the temporal aspects of conversation leads to making cross-cultural comparisons possible. Not only is cross-cultural comparison possible, but we have found differences in temporal aspects between persons of different cultures, speaking different languages. We would like to go one step further and argue that the temporal organization of interaction is a crucial aspect of any interaction.

One implication of the present study is to suggest the temporal organization as a key concept for inclusion in any theoretical and methodological framework of conversation in particular, but even more generally in face-to-face interaction. In doing so, we imply that studying the organization of conversations cross-culturally rather than being premature can actually result in important contributions toward establishing such a desired framework.

It seems to us that one interesting feature of the focus on temporal aspects is that all disciplines studying conversations, or more generally interaction, can possibly agree that social, cultural and communicative data are natural events recurrent in time and space. In that sense a focus on temporal aspects circumvents theoretical and definitional problems which mark the disagreement between disciplines.

The realization that the temporal organization of interaction is of crucial importance dates back, as pointed out earlier, to Chapple [1940]. He studied interaction processes as they occurred and his units of action consisted of behavior that could be observed. An action is to be regarded as any overt muscular activity of an organism which involves a change from its immediate preceding activity, while a unit of action is that period of activity recorded from the initial change in muscular state until a second change brings the activity to an end [Chapple, 1940a, p.24]. Fundamentally our operationalizations of gaze-direction, speaker turns, listener responses and gesticulation can be seen as part of the integral production of what Chapple would call actions, and units of action. Analysis of our data showed that the duration of these units of action in Japanese conversation are shorter than in Australian conversation. Since these units of action recurred over and over again in different conversations and for different variables, we concluded that the rhythmical organization of Japanese conversation is different from Australian conversation. Ordinary conversation processes are to some degree endowed with rhythmic properties. According to Chapple [1970] the beat is built up from the length of each action and the interval of silence or inaction in between, the summation of which establishes the tempo (or pace) of the individual's pattern of behavior. Chapple found these patterns to be consistent for individuals. Our data suggest consistency for persons belonging to the same cultural group. These rhythms of interactional behavior include the total pattern of action. It does not merely include speech or the voice, but it is further elaborated by voice intensity, pitch, timbre, gaze-direction, posture, gesture and so on, and it compels responses from the listener. One way to obtain further confirmation of our findings would be, to measure these rhythms of interactional behavior in large samples of randomly

selected Japanese and Australian participants. We would expect that the frequency distributions of these rhythms would reveal differences related to the culture of the participants. Given that individuals differ in their rhythms, some overlap would probably occur.

Since our data show differences in the temporal organization of conversation on the turn level between Japanese and Australians, one can speculate how these differences come about. We suggest that these differences at the turn level are initially established through mother-infant interaction, and subsequently reinforced. Recent research suggests that in dealing with her infant a mother will establish a certain rhythmic pattern.

Condon and Sander [1974] have shown that, on a micro-rhythmic level (see next section), an infant's movements are synchronized with adult speech as early as the first day of life. Their study makes it plausible that time sense and rhythm are innate. One can even imagine that an unborn child responds to the rhythms of its mother, when still in the womb. Research in early mother-infant interaction suggests that the infant actively makes movements, initially in a relatively unco-ordinated way. The mother scans these movement patterns and selectively responds to certain patterns. In chapter II we discussed that significant "coupling" between the gaze of infant and mother appears early in the infant's first year [Jaffe *et al.*, 1973] and so does the connection of gaze-direction with the turn-taking pattern of vocalization [Bateson, 1971]. Tronick, Als and Brazelton [1977] conclude that the infant can sequence expressive displays in an appropriate serial order while allowing for the turn-taking necessary to the reciprocal exchange of messages. They also show that the infant is able to communicate his or her intent to respond to the expressed intent of the mother. Thus, long before language, the infant is a skilful communicator. In chapter II we pointed out that these developmental studies have shown that the temporal pattern of mother-infant gaze is identical to those found in adult conversations. This suggests that such regularities describe a property of interaction, that predates the onset of speech and is first seen in an attention-regulating behavior [Jaffe, Stern and Peery, 1973; Stern, 1974]. Summarizing these developmental studies, it is clear that

movement, in particular visual-movement, and vocal-auditory systems become linked very early and that the temporal patterns of these relations are identical to those found in adult conversations. These studies suggest that given a certain adult rhythmic pattern the infant will learn its properties very early. This then suggests that speech, seen as both linguistic and kinesic, is a basic form of human relationship and that differences, if they occur, may be reinforced from very early in life.

Some suggestive evidence for a different rhythmical pattern is available. Studies by Caudill *et al.* [1969, 1972, 1973] suggest differences in patterns of child-rearing between American and Japanese mother-infant pairs. He found that American mothers do more looking at, positioning of the body, and chatting to their infants. Japanese mothers do more carrying, rocking and lulling of their infants. American infants show greater amounts of gross bodily activity (with toys, hands and other objects) and happy vocalization; in contrast, Japanese infants seem passive, having only a greater amount of unhappy vocalization. Infants in both cultures become habituated by 3 or 4 months of age to respond appropriately to these different styles of caretaking.

The studies by Caudill *et al.* were carried out through participant observation on a time-sampling basis. The observers went into the homes of both Japanese - and American mother-infant pairs. The observations appear to be based upon molar level operationalizations of interactive behavior (see next section). The American mother's style of caretaking serves to increase her infant's happy vocalization and more broadly, to emphasize the importance of vocal communication. The Japanese mother has less vocal interaction with her infant during caretaking. For Japanese mother-infant pairs there exists a lesser reliance on vocal communication, but the importance and communicative value of physical contact is emphasized [Caudill and Schooler, 1973].

The response of the infants are strikingly in line with the general stereotypical expectations for behavior in the two cultures; in America, the individual should be physically and verbally assertive;

and in Japan, he should be physically and verbally restrained [Caudill and Schooler, 1973]. To substantiate our hypothesis a comparison of mother-infant interaction, in which the emphasis would be on turn-taking, in Japan and in Australia would have to be carried out. We would expect Japanese mothers to take generally shorter turns than would Australian mothers. The rhythm of speech (see also next section) would be an important aspect. Any *reasons* for the occurrence of these differences in the temporal organization of conversation between Japanese and Australians, have to be sought primarily in the possible differences in structure between the Japanese and English language (see the discussion in chapter V), and secondarily in the different rules, values and patterns of thinking between the two cultures.

7.2 INTERACTIONAL RHYTHMS

The second part of this chapter attempts to develop the notion of rhythm, points out some ways in which researchers have used it and suggests some further lines of inquiry. This section will be of a more general and speculative nature.

Chapple [1940] was the first researcher to realize the importance of the temporal organization in interaction, and subsequently the idea of temporal organization or "rhythmical structure" has featured in the work of several researchers. The concept of "rhythm", however, is an extremely complex one and several authors who have dealt with it have used different operationalizations to measure rhythm. The development of a fruitful approach to the study of rhythm has been hampered by a failure to distinguish clearly among the several aspects of temporal organization itself. Therefore clear distinctions and unequivocal terminology are necessary. In the field of music the distinctions have been spelled out most clearly [Cooper and Meyer, 1960]. The exposition, which follows, is largely based on the definitions and principles distinguished in the rhythmic structure of music.

There are three basic modes of temporal organization which can be differentiated. They are pulse, meter, and rhythm. Tempo, though it

qualifies and modifies these, is not itself a mode of organization. A pulse is one of a series of regularly recurring, precisely equivalent stimuli. Meter is the measurement of the number of pulses between more or less regularly recurring accents. Therefore, in order for meter to exist, some of the pulses in a series must be accented -- which is marked for consciousness -- relative to others. When pulses are thus counted within a metric context, they are referred to as beats. Rhythm may be defined as the way in which one or more unaccented beats are grouped in relation to an accented one.

Accent appears to be a product of a number of variables whose interaction is not precisely known. These variables are duration, intensity, regularity and so on. Accent must not be confused with stress, which means the dynamic intensification of a beat, whether accented or unaccented.

Rhythmic grouping is a mental fact, not a physical one. There are no hard and fast rules for calculating what, in any particular instance, the grouping is. Grouping on all architectonic levels is a product of similarity and difference, proximity and separation of the movements perceived by the senses and organized by the mind. Therefore, researchers may differ with respect to rhythmic grouping, depending on their sensitivity, perceptual abilities, and training. Rhythm is more than a mere sequence of durational proportions. To experience rhythm is to group separate movements into structured patterns. Such grouping is the result of the interaction among various aspects of the data [Cooper and Meyer, 1960]. A crucial aspect of rhythm, to which we will come back later, is that the perception of initial elements in a pattern allows later elements to be anticipated in time.

After this exposition of the definitions and principles of the several aspects of temporal organization, we can now turn our attention to interactional rhythms.

Since movement is architectonic in its organization, it is possible to distinguish different rhythmic levels. We shall describe three levels of rhythm, which appear to be commonly used by

researchers, namely: macro rhythmic levels, molar rhythmic levels and micro rhythmic levels. These rhythmic levels differ from each other primarily in their temporal time span and therefore secondarily in the way behavioral data are rhythmically grouped together. By saying that movement is architectonic in its organization, we mean that, just as phone types are combined into syllables, syllables into words, words into utterances and so on, body movement can be described in terms of changes of several bodies in relation to each other and/or the environment, or in terms of changes of the whole body in the environment, or in terms of changes of body parts in relation to each other — the trunk, the head, the left upper-arm, left fore-arm, the hand, the fingers, and so on. When one observes behavior, its rhythmic structure is perceived not as a series of discrete independent units strung together in a mechanical, additive way like beads, but as an organic process in which smaller rhythmic motives, while possessing a shape and structure of their own, also function as integral parts of a larger rhythmic organization.

Rhythm both organizes, and is itself organized by, all the elements which create and shape interactional processes. These processes, that are rhythmic, possess a coherent internal structure.

7.2.1 Macro Rhythmic Levels

On the macro rhythmic level, man's relationship to his environment or even the universe at large is related to various biological, social and cultural cycles. Chapple [1970] in his later work discusses at length the various rhythms which exist, their properties and the ways in which they can be said, by hypothesis, to act as the organizing principles for both human and animal activities and interaction.

In a rather complex way the studies of Lomax *et al.* [1968] are also partially concerned with a macro-level rhythmical structure. In their *Folk Song Style and Culture* (Lomax *et al.*, 1968] an account is given of the discovery of panhuman, worldwide correlations between culture characterizing styles of folk music, dance, expressive arts of

every human culture, on the one hand, and highly specific cultural variations in the social relationship of each community's ordering of its preferred and customary male-female, familial and public-decision-making behaviors on the other hand. More recently, Lomax *et al.* [1977] have turned their attention also to speaking styles as related to social structure and cultural traditions. The rhythmic organization of music, song, dance and speaking styles figures prominently in their operationalization of cantometrics, coreometrics and parlametrics.

On the macro level there are rhythms whose recurrent patterns occur in reference to high level units of behavior and which are realized in extended stretches of time. On this level there is a strong suggestion that the "rhythm of a culture" becomes reified in its song, dance, music and speaking styles.

7.2.2 Micro Rhythmic Levels

On a much lower level of organization there are rhythms which are organized over very short stretches of time, so-called micro rhythmic levels. On this level most of the research deals with speech. The rhythms of speech are a result of pulse-like movements of the speaker's lung muscles controlling the output of air. Unfilled pulsations, or silence, are called "silent stress-pulses" [Abercrombie, 1965].

Rhythm, then, is not merely a regulation of the sounds of speech, but also the use of muscular movements of the speaker. These muscular movements place certain constraints on the organization of human movement and this in turn has important implications for an understanding of the production of rhythms.

The constraints on speech sounds, or on any other real-time sequence of behavioral elements, that is directly implied by the concept of rhythm is *relative timing*, which means that the locus of each element along the time dimension is determined relative to the locus of *all* other elements in the sequence, adjacent *and*

nonadjacent. This is to say that sequences of sounds, speech or otherwise, that are rhythmic will possess hierarchical organization, that is, a coherent internal structure, at the sound level [Martin, 1972].

Martin [1972] in his article on rhythmic structure in speech and other behavior has developed the hierarchical notion to some extent. He observes that speaking and listening are dynamically coupled rhythmic activities, such that linguistic information is encoded rhythmically into the signal by the speaker and decoded out of it on that basis by the listener. He points out that there is extensive information available in speech at the rhythmic level that can be used by the listener. Some languages may make more informative use of rhythmic constraints on the sound patterning of speech than others. He also observes that patterned speech sounds could be redundant with respect to linguistic message elements, and that the perception of initial elements in a pattern allows later elements to be anticipated in real time. Therefore, one might say not that the listener "follows" the speaker, but rather that the listener given initial cues, actively enters into the speaker's tempo. Our discussion so far has dealt with speech. The area of speech production and perception could contribute much to enhance our understanding of micro rhythmic levels in interaction.

Movement of the body, however, is organized in relation to speech, not only within a single individual, but also between individuals [Condon, 1966]. "Intensive analysis revealed harmonious or synchronous organizations of change between body motion and speech in *both* intra-individual and interactional behavior. Thus the body of the speaker dances in time with his speech. Further, the body of the listener dances in rhythm with that of the speaker ..." [Condon and Ogston, 1966, p.338]. "From the analysis of the sound film, we see that as a person talks ... his body moves in a series of configurations of change which are precisely correlated with that serial transformation of 'phone into syllable into word of speech'. Kinesic segmentation in general seems to coincide with the etic segmentation of speech. Further analysis, focusing on interactional behavior, revealed a startling and relatively continuous harmony

between the body motion configurational change patterns of speaker and listener" [Condon and Ogston, 1966, p.339].

In a later article, Condon [1973] discusses the rhythmic properties and organization of speaker behavior. Human beings appear to have a very basic rhythm pattern which, tentatively, may be characteristic for the species. Thus a speaker's behavior when speaking appears to be composed of a continuously emerging series of one second, linguistic-kinesic forms, which are themselves composed of organized sub-forms [Condon, 1973]. One of these is documented in Condon's article, namely half second cycles. These half second cycles are usually composed of two or three words or primary syllables strung together. There tend to be one or two unstressed elements followed by a stressed element. The body motion pattern which demarcates the inner structure of the half second cycle is created by rapid kinesic acceleration or change of direction on the terminal word or syllable which is then seen as a "stressed" feature in contrast to the preceding aspects which are, contrastingly, unstressed [Condon, 1973, pp.11,12]. These half second durational forms are sometimes found alone. Most of the time, however, they seem to form an integral part of a wider 1 second rhythm form. This 1 second rhythm feature appears to be a *very basic* dimension in the on-going emergence of linguistic-kinesic behavior ... Beyond this one second interval behavior appears to enter the sphere of posturology, which is, however, intimately related to movement [Condon, 1973, p.12]. The one second rhythm dimension is also accompanied by motion in various ways, common ones being a head or arm ensemble movement(s) which sustains a given direction of movement across the emission of the, usually four to five word, phrase.

According to Condon, different cultures put these one second rhythm forms together in differing ways, indicating possible sources of cross-cultural communicational difficulty [Condon, 1973, p.23]. He does not spell out how these basic rhythm forms are put together in different ways, but observes that American blacks appear to have a different micro rhythm pattern during communication than whites [p.20].

On the micro level there are rhythms whose recurrent patterns

occur in reference to low level units of behavior that are realized over short periods of time. On this level, there is evidence that a relationship occurs between the rhythmical organization of the speaker's behavior and the rhythmical organization of the listener's behavior. Conversationists tend to synchronize their movements, entering into a continuous and jointly sustained "dance" [see also Kendon, 1970, 1973].

There is a suggestion that different cultures organize these micro rhythmical "waves" in differing ways, but in what way they do is not spelled out.

7.2.3 Molar Rhythmic Levels

To investigate macro rhythmical structures requires substantial resources over a considerable period of time. Rather sophisticated ways of analyzing a diverse range of data then reveals a structure. The data are grouped in a way, which is beyond the normal everyday realization of an individual.

Astrology is an extreme example to illustrate this point. When a horoscope is drawn up, the time of birth as well as the place of birth is related to the positions of the stars in order to predict future developments. The relationship between all these variables, however, can only be grasped after extensive analysis, and is not normally within awareness.

The micro rhythmical level of investigation progresses very slowly and is extremely laborious. The rhythmical structure which is uncovered deals with such subtle relationships between variables that they might be perceived by an observer, but probably not by the participants.

Both lines of inquiry are highly suggestive, but are concerned with architectonic levels which are beyond the level of deliberate intentional production. We want to advocate further investigation of rhythm on a molar level. The term molar is used here to distinguish

the turn, as a unit of behavior, from the sub-molar (or micro) properties of movements. Our data have shown that the behavioral acts within a turn such as speech, gaze-direction and listener responses are closely related to each other.

We have shown that shorter utterances, shorter other-directed gazes and a higher frequency of listener responses are commonly found in Japanese conversations. The Australian conversations were characterized by longer utterances, longer other-directed gazes and a lower frequency of listener responses.

This relationship between variables emerged as an organizing pattern over a conversation or over a number of conversations, but is not highly probable in any particular instance. It does, however, imply that, for such a relationship to emerge, the variables investigated must be temporally related. The level on which they are is the molar level, or the level of turn-taking.

In our opinion, the investigation of the temporal organization on the turn level, would increase our understanding of how turn-taking is brought about. We have pointed out that one crucial aspect of rhythm is that the perception of initial elements in a pattern allows later elements to be anticipated in time. Rhythmical organization on the molar level, or turn level, would allow participants to anticipate turn-endings. The research on turn-taking discussed in chapter II does not take account of the organization of a turn — or rather of the behavior within a turn — in such a way that its termination can be anticipated.

A turn in a conversation is a unit of behavior, which is intentionally produced by conversationalists. Therefore, the behavior within a turn must be organized in a meaningful way. In chapter II we discussed the various approaches to turn-taking. The oldest approach is Chapple's investigation of action units, which certainly included the study of turns. Chapple took more contextual variables into account than any of the subsequent approaches. Kendon's conclusion [Argyle and Kendon, 1967] on Chapple's approach is that a properly developed theoretical framework is still much needed. In particular,

a definition of Chapple's units is needed which specifies their relationship to other possible units of behavior into which the social performance may be analyzed. We take this conclusion to refer to the internal structure of turns.

After Chapple a number of researchers continued the study of turns focussing on speech only. For them a turn was defined in terms of speech duration, disregarding any other behavior involved in turn-taking. Of the more recent functional approaches conversational analysts are concerned with the sequential relationship between utterances. Their orientation is a return to how interactants structure conversations.

The other functional approach, what we referred to as signals approach, considers turn-taking to be mediated through signals. These signals are composed of clear-cut behavioral cues, considered to be perceived as discrete [Duncan, 1972]. This notion is of course incompatible with a rhythmical structure to turn-taking. We pointed out that one problem with the signals approach is the question of time delay between a certain behavioral cue and turn exchange.

An integration of these different approaches to turn-taking seems possible if we accept a rhythmical organization on the level of turn-taking. As we discussed on p.164, when one observes behavior, its rhythmic structure is perceived not as a series of discrete independent units strung together in a mechanical, additive way like beads, but as an organic process in which smaller rhythmic units, while possessing a shape and structure of their own, also function as integral parts of a larger rhythmic organization. Particular signals, as identified by the signals approach, may function in turn-taking, but as a part of a larger rhythmic organization on the molar level.

In order to investigate the rhythmic organization at the molar level, it might be necessary to investigate first of all the ways in which the evidence on the sub-molar level can be composed to form a higher order at the turn level. Kendon [1970] observed that when one person initiated movement another person came to follow. He also observed that these movements appeared to be related to the variations

in pitch level of the first person's speech. He suggests that these movements serve to give advance warning that the second person wants to speak when the first is finished, and further, it may be that in overtly "beating time" to the first person's speech, he may thereby facilitate the precise timing of his own entry as a speaker. This example is concerned with a turn exchange.

We are presently carrying out a preliminary investigation, comparing two 10-second stretches of material, taken from the conversations of the bilingual participant in both the Japanese — and the Australian condition. We plan to analyze these stretches of material, relating speech and body movement. Phonetic transcripts of speech will be segmented into smaller units such as, utterances, words, syllables and phone types and these in turn will be related to frame-numbers. Particular attention will be paid to voice pitch and intensity. The flow chart of body movement related to frame-numbers will consist of changes in overall picture, gaze-direction, head, trunk, legs and arms. We expect that points of change in speech and separate body parts will coincide [cf. Condon, 1973; Kendon, 1970].

Suggested differences in the use of gaze-direction (see chapter III) and the use of listener responses, in particular head-nods (see chapter V) leads us also to expect that the organization of movement on the level of a turn, will reveal differences.

If it is possible to relate micro rhythmical levels of organization to molar rhythmical levels, our understanding of how turn-taking comes about will be enhanced considerably. As it is now, research on the micro rhythmical level of organization can account for how it is that conversationalists can alternate as speaker and listener without any gap or time-delay between turn exchanges *in particular instances*, or, the phenomenon of synchrony *in particular instances*. The stretches of time these phenomena span are of very short duration.

What we are suggesting then is that micro level rhythmical structure can be organized into molar rhythmical structures. This organization is made up from a combination of variations in speech,

posture, gaze-direction, gesture and so on over intermediate stretches of time (spanning the duration of a turn). We also suggest that the grouping of these combinations might differ for different cultural groups.

REFERENCES

- Abercrombie, D.
1965 *Studies in phonetics and linguistics*. Oxford: Oxford University Press.
- Ahrens, R.
1954 *Beitrage zur Entwicklung des Physiognomie und Mimikerkennens. Zeitschrift für Experimentelle und Angewandte Psychologie*, 2, 412-454; 599-633.
- Albert, E.
1968 "Value systems" in: *The International Encyclopedia of the Social Sciences*, New York: Macmillan, volume 16, 287-290.
- Arensberg, C.M.
1973 Introduction, in Collins O. and Collins, J.M. *Interaction and social structure*. The Hague: Mouton, 9-20.
- Argyle, M. and Kendon, A.
1967 The experimental analysis of social performance. In Berkowitz, L. (Ed.) *Advances in experimental social psychology*, vol. III, London: Academic Press, 55-98.
- Argyle, M., Lalljee, M. and Cook, M.
1968 The effect of visibility on interaction in a dyad. *Human Relations*, 21, 3-17.
- Argyle, M.
1969 *Social interaction*. London: Methuen.
- Argyle, M.
1975 *Bodily communication*. London: Methuen.
- Argyle, M. and Cook, M.
1976 *Gaze and mutual gaze*. London: Cambridge University Press.
- Austin, J.L.
1962 *How to do things with words*. Oxford: Clarendon Press.
- Bachrach, A.J. (Ed.)
1962 *Experimental foundations of clinical psychology*. New York: Basic Books.
- Bateson, M.C.
1971 The epigenesis of conversational interactions. Paper presented to the Society for Research in Child Development. Minneapolis.

- Bell, R.
1973 *The Japan experience*. Tokyo: Weatherhill.
- Berkowitz, L. (Ed.)
1967 *Advances in experimental social psychology*, volume III. London: Academic Press.
- Bernick, N. and Oberlander, N.
1968 Effects of verbalization and two modes of experiencing on pupil size. *Perception and Psychophysics*, 3, 327-330.
- Birdwhistell, R.L.
1962 Critical moments in the psychiatric interview. In: Turlentes, T.T. (Ed.) *Research approaches to a psychiatric problem*. New York: Grune & Stratton, 179-188.
- Birdwhistell, R.L.
1970 *Kinesics and context*. Philadelphia: University of Pennsylvania.
- Brein, M. and David, K.H.
1971 Intercultural communication and the adjustment of the sojourner. *Psychological Bulletin*, 76, no. 3, 215-230.
- Bond, M.H. and Shiraishi, D.
1974 The effect of body lean and status of an interviewer on the nonverbal behavior of Japanese interviewees. *International Journal of Psychology*, 9, no. 2, 117-128.
- Bond, M.H. and Hwa Young Ho
1975 The effect of interviewer status and sex composition of the dyad on the cognitive responses and nonverbal behavior of Japanese interviewees. Unpublished manuscript, Kwansei Gakuin University.
- Bond, M.H. and Iwata, Y.
1976 Proxemics and observation anxiety in Japan: non-verbal and cognitive responses. *Psychologia*, 19, no. 3, 119-126.
- de Bono, E.
1967 *The Five Day Course in Thinking*. New York: Basic Books.
- Cammack, F.M. and Van Buren, H.
1967 Paralanguage across cultures: some comparisons between Japanese and English. *English Language Education Council - (ELEC) Bulletin*, 22, 7-10, continued 47.
- Caudill, W.A. and Weinstein, H.
1969 Maternal care and infant behavior in Japan and America. *Psychiatry*, 32, 12-43.
- Caudill, W.A.
1972 Tiny dramas: Vocal communication between mother and infant in Japanese and American families. In: Lebra, W.P. (Ed.) *Transcultural Research in Mental Health*, volume

2 of Mental Health Research in Asia and the Pacific.
Honolulu, Hawaii: University of Hawaii Press, 25-48.

- Caudill, W.A. and Schooler, C.
1973 Child behavior and child rearing in Japan and the United States: an interim report. *Journal of Nervous and Mental Disease*, 157, 323-338.
- Chance, M.R.A.
1962 The interpretation of some agonistic postures: the role of "cut-off" acts and postures. *Zoological Society of London, Symposia*, 8, 71-89.
- Chapple, E.D.
1939 Quantitative analysis of the interaction of individuals. *National Academy of Science, Proceedings, Washington*, 25, 58-67.
- Chapple, E.D. and Arensberg, C.M.
1940a Measuring human relations: An introduction to the study of the interaction of individuals. *Genetic Psychology Monographs*, 22, 3-147.
- Chapple, E.D.
1940b "Personality" differences as described by invariant properties of individuals in interaction. *National Academy of Science, Proceedings, Washington*, 26, 10-16.
- Chapple, E.D. and Coon, C.S.
1942 *Principles of anthropology*. New York: Holt, Rinehart & Winston.
- Chapple, E.D. and Donald, G., Jr.
1946 A method for evaluating supervisory personnel. *Harvard Business Review*, 46, 197-214.
- Chapple, E.D.
1949 The interaction chronograph: its evolution and present application. *Personnel*, 25, 295-307.
- Chapple, E.D.
1953 The standard experimental (stress) interview as used in Interaction Chronograph investigations. *Human Organization*, 12, 23-32.
- Chapple, E.D.
1970 *Culture and biological man: explorations in behavioral anthropology*. New York: Holt, Rinehart & Winston.
- Chaplin, J.P.
1967 *Dictionary of psychology*. New York: Dell.
- Chomsky, N.
1965 *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.

- Cole, M. and Scribner, S.
1974 *Culture and thought: a psychological introduction*. New York: Wiley.
- Collett, P.
1971 On training Englishmen in the non-verbal behaviour of Arabs: an experiment in intercultural communication. *International Journal of Psychology*, 6, 209-215.
- Collins, O. and Collins, J.M.
1973 *Interaction and social structure*. The Hague: Mouton.
- Condon, J.C. and Saito, M. (Eds.)
1974 *Intercultural encounters with Japan*. Tokyo: Simul Press.
- Condon, J.C. and Yousef, F.Y.
1975 *An introduction to intercultural communication*. Indianapolis: Bobbs-Merrill.
- Condon, W.S. and Ogston, W.D.
1966 Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease*, 143, 338-347.
- Condon, W.S.
1973 Communication and order: the micro "rhythm hierarchy" of speaker behavior. Paper presented at School Psychologists meeting, March 15, New York.
- Condon, W.S. and Sander, L.W.
1974 Synchrony demonstrated between movements of the Neonate and Adult Speech. *Child Development*, 45, 456-462.
- Cooper, G.W. and Meyer, L.B.
1960 *The Rhythmic Structure of Music*. Chicago: University of Chicago Press.
- Crpok, J.H. (Ed.)
1970 *Social behaviour in birds and mammals*. London: Academic Press.
- Danziger, K.
1976 *Interpersonal communication*. New York: Pergamon Press.
- De Laguna, G.
1927 *Speech: Its function and development*. New Haven: Yale University Press.
- Diebold, A.R., Jr.
1968 Anthropology of the comparative psychology of communicative behavior. In: Sebeok, Th.A. (Ed.) *Animal communication*. Bloomington: Indiana University Press, 525-571.

- Dignam, P.J.
1960 The effect of verbal and nonverbal reinforcement on the verbal behavior of geriatric subjects. Unpublished Ph.D. dissertation, Florida State University.
- Dittmann, A.T. and Llewellynn, L.G.
1967 The phonemic clause as a unit of speech decoding. *Journal of Personality and Social Psychology*, 6, no. 3, 341-349.
- Dittman, A.T. and Llewellynn, L.G.
1968 Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology*, 9, no. 1, 79-84.
- van Doorn, J.A.A. and Lammers, C.J.
1959 *Moderne sociologie*. Utrecht: Het Spectrum.
- Douglas, J. (Ed.)
1970 *Understanding everyday life*. Chicago: Aldine.
- Duncan, S.D.
1969 Nonverbal communication. *Psychological Bulletin*, 72, 118-137.
- Duncan, S.D.
1972 Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23, 283-292.
- Duncan, S.D.
1974 On the structure of speaker-auditor interaction during speaking turns. *Language in Society*, 2, 161-180.
- Duncan, S.D. and Fiske, D.W.
1977 *Face-to-face interaction: research, methods and theory*. Hillsdale, N.J.: Lawrence Erlbaum.
- Efron, D.
1941 *Gesture, race and culture*. New York: Kings Crown Press. Reprinted: Sebeok, Th.A. (Ed.) *Approaches to semiotics*, no. 9, The Hague: Mouton, 1972.
- Ekman, P. and Friesen, W.V.
1968 Nonverbal behavior in psychotherapy research. In Schlien, J. (Ed.) *Research in Psychotherapy*, volume 3, Washington: A.P.A., 179-216.
- Ekman, P. and Friesen, W.V.
1969 The repertoire of nonverbal behavior: categories, origins, usage and coding. *Semiotica*, 1, 49-68.
- Ellsworth, P.C., Carlsmith, J.M. and Henson, A.
1972 The stare as a stimulus flight in human subjects: A series of field experiments. *Journal of Personality and Social Psychology*, 21, 302-311.

- Ellsworth, P.C. and Ludwig, L.M.
1972 Visual behaviour in social interaction. *Journal of Communication*, 22, 375-403.
- Elzinga, R.H.
1972 Some aspects of phobias in Japan. Unpublished doctoraal scriptie. University of Amsterdam.
- Elzinga, R.H.
1975 Nonverbal communication: Body accessibility among Japanese. *Psychologia*, 18, no. 4, 205-211.
- Elzinga, R.H. and Iwawaki, S.
1977 Intimacy signals: a crosscultural study of Japanese and Australians. Unpublished manuscript.
- Exline, R.V.
1963 Exploration in the process of person perception: visual interaction in relation to competition, sex and need for affiliation. *Journal of Personality*, 31, 1-20.
- Exline, R.V. and Winters, L.C.
1965 Affective relations and mutual glances in dyads. In: Tomkins, S.S. and Izard, C.E. (Eds.) *Affect, cognition and personality*. New York: Springer, 319-351.
- Exline, R.V., Gray, D. and Schuette, D.
1965 Visual behavior in a dyad as affected by interview content and sex of respondent. *Journal of Personality and Social Psychology*, 1, 201-209.
- Fries, C.C.
1952 *The structure of English*. New York: Harcourt Brace & Co.
- Fugita, S.S.
1970 The effects of anxiety and approval on visual interaction. Unpublished Ph.D thesis, Riverside: University of California.
- Gallois, C. and Markel, N.N.
1975 Turntaking: social personality and conversational style. *Journal of Personality and Social Psychology*, 31, no. 6, 1134-1140.
- Garfinkel, H.
1974 The origins of the term "ethnomethodology". In: Turner, R. (Ed.) *Ethnomethodology*, Harmondsworth: Penguin Books.
- Goldman-Eisler, F.
1951 The measurement of the time sequences in conversational behavior. *British Journal of Psychology*, 42, 355-362.
- Goffman, E.
1963. *Behavior in public places*, Glencoe: The Free Press.

- Goodwin, Ch.
 1975 The interactive construction of the sentence within the turn at talk in natural conversation. Paper presented at the 74th annual meeting of the American Anthropological Association, San Francisco, December 2-6.
- Gottschalk, L.A. and Auerbach, A.H. (Eds.)
 1966 *Methods of research in psychotherapy*. New York: Appleton Century Crofts.
- Greenspoon, J.
 1962 Verbal conditioning and clinical psychology. In: Bachrach, A.J. (Ed.) *Experimental foundations of clinical psychology*. New York: Basic Books.
- Guze, S.B. and Mensch, I.N.
 1959 An analysis of some features of the interview with the Interaction Chronograph. *Journal of Abnormal and Social Psychology*, 58, 269-271.
- Harré, R. and Secord, P.F.
 1972 *The explanation of social behavior*. Oxford: Blackwell.
- Harré, R.
 1974 Some remarks on "rule" as a scientific concept. In: Mischel, Th. (Ed.) *Understanding other persons*. Oxford: Blackwell, 143-184.
- Hargreaves, W.A. and Starkweather, J.A.
 1959 Collection of temporal data with the Duration Tabulator. *Journal of the Experimental Analysis of Behavior*, 2, 179-183.
- Hargreaves, W.A.
 1960 A model for speech unit duration. *Language and speech*, 3, 164-173.
- Heron, J.
 1970 The phenomenology of social encounters: the gaze. *Philosophy and Phenomenological Research*, 3, 243-264.
- Hess, E.H.
 1965 Attitude and pupil size. *Scientific American*, 202, 46-54.
- Hindmarch, I.
 1973 Eyes, eye-spots and pupil dilation in nonverbal communication. In: von Cranach, M. and Vine, I. (Eds.) *Social communication and movement*. London: Academic Press, 299-321.
- Hore, T.
 1970 Social class differences in some aspects of the nonverbal communication between mother and preschool child. *Australian Journal of Psychology*, 22, no. 1, 21-27.

- Hutt, C. and Ounstead, C.
1966 The biological significance of gaze aversion with particular reference to the syndrome of infantile autism. *Behavioral Science*, 11, 346-356.
- Ingham, R.J.
1972 Cross-cultural differences in social behaviour. Unpublished D.Phil. thesis, Oxford University.
- Inoue, K.
1977 Japanese: a story of language and people. In: Schopen, T. (ed.) *Languages and their speakers*. Cambridge: Winthrop, ch. 7 (in press).
- Jaffe, J. and Feldstein, S.
1970 *Rhythms of dialogue*. New York: Academic Press.
- Jaffe, J., Stern, D.N. and Peery, J.C.
1973 "Conversational" coupling of gaze behavior in pre-linguistic human development. *Journal of Psycholinguistic Research*, 2, 321-329.
- Jefferson, G.
1973 A case of precision timing in ordinary conversation: overlapped tag-positioned address term in closing sequences. *Semiotica*, 9, 47-96.
- Kahneman, D. and Beatty, J.
1966 Pupil diameter and load on memory. *Science*, 54, 1583-1585.
- Kanfer, F.
1959 Verbal conditioning: reinforcement schedules and experimenter influence. *Psychological Reports*, 5, 343-353.
- Kendon, A.
1963 Temporal aspects of the social performance in two-person encounters. Unpublished D.Phil. thesis. University of Oxford.
- Kendon, A.
1967 Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22-63.
- Kendon, A. and Cook, M.
1969 The consistency of gaze patterns in social interaction. *British Journal of Psychology*, 60, 481-494.
- Kendon, A.
1970 Movement coordination in social interaction, some examples described. *Acta Psychologica*, 32, 100-125.
- Kendon, A.
1973 The role of visible behaviour in the organization of social interaction. In: van Cranach, M. and Vine, I. (Eds.) *Social communication and movement*. London: Academic Press, 29-74.

- Kendon, A.
1975 Gesticulation, speech and the gesture theory of language origins. *Sign Language Studies*, 9, 349-373.
- Kendon, A., Harris, R.M. and Key, M.R. (Eds.)
1975 *Organization of behavior in face-to-face interaction*. The Hague: Mouton.
- Kendon, A.
1977a Some theoretical and methodological aspects of the use of film in the study of social interaction. In: Ginsburg, G.P. (Ed.) *Emerging strategies in social psychological research*. New York: Wiley (in press).
- Kendon, A.
1977b *Studies in the behavior of social interaction*. Vol. 6 of studies in semiotics (Sebeok, Th. (Ed.)). Bloomington: Indiana University Press and Lisse; Peter de Ridder Press.
- Kleinke, Ch.L., Meeker, F.B. and La Fong, C.
1974 Effects of gaze, touch and use of name on evaluation of "engaged" couples. *Journal of Research in Personality*, 7, 368-383.
- Kleinke, Ch.L. and Pohlen, P.D.
1971 Affective and emotional responses as a function of other person's gaze and cooperativeness in a two-person game. *Journal of Personality and Social Psychology*, 17, 308-313.
- Kleinke, Ch.L.
1975 *First impressions: the psychology of encountering others*. Englewood Cliffs, New Jersey: Prentice Hall.
- Kluckhohn, C.
1951 Values and value orientations in the theory of action: An exploration in definition and classification. In: Parsons, T. and Shils, E. (Eds.) *Toward a general theory of action*. Cambridge: Harvard University Press, 390-406.
- Knapp, M.L.
1972 *Nonverbal communication in human interaction*. New York: Holt, Rinehard & Winston.
- Krasner, Ch.L.
1958 Studies of the conditioning of verbal behavior. *Psychological Bulletin*, 15, 148-171.
- La France, M. and Mayo, C.
1976 Racial differences in gaze behavior during conversations: two systematic observational studies. *Journal of Personality and Social Psychology*, 33, no. 5, 547-552.
- Lambert, W.E.
1955 Measurement of the linguistic dominance of bilinguals. *Journal of Abnormal and Social Psychology*, 50, 197-200.

- Lebra, W.P. (Ed.)
 1972 *Transcultural research in Mental Health*. Volume 2 of Mental Health Research in Asia and the Pacific. Honolulu, Hawaii: University of Hawaii Press.
- Lefcourt, H.M., Rotenberg, F., Buckspan, R. and Steffy, R.
 1967 Visual interaction and performance of process and reactive schizophrenics as a function of examiner's sex. *Journal of Personality* 35, 535-546.
- Lewis, M. and Rosenblum, L.A. (Eds.)
 1974 *The origins of behaviour*. Volume 1: The effect of the infant on its caregiver. New York: Wiley.
- Lomax, A. et al.
 1968 *Folk Song Style and Culture*. Washington: American Association for the Advancement of Science, No. 88.
- Lomax, A. et al.
 1977 A stylistic analysis of speaking. *Language in society*, 6, 15-36.
- McDermott, R.P. and Wertz, M.
 1976 Doing the social order: some ethnographic advances from communicational analysis and ethnomethodology. *Reviews in Anthropology*, 3, no. 2, 160-174.
- Magnus, H.
 1885. *Die Sprache der Augen*. Wiesbaden: Würzburg.
- Markel, N.N.
 1975 Coverbal behavior associated with conversation turns. In: Kendon, A., Harris, R.M. and Key, M.R. (Eds.) *Organization of behavior in face-to-face interaction*. The Hague: Mouton, 189-197.
- Martin, J.G.
 1972 Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79, 487-509.
- Martin, S.E.
 1975 *A reference grammar of Japanese*. New Haven: Yale University Press.
- Matarazzo, J.D.
 1962 Prescribed behavior therapy: suggestions from interview research. In: Bachrach, A.J. (Ed.) *Experimental foundations of clinical psychology*. New York: Basic Books, ch. 14, 471-509.
- Matarazzo, J.D. and Wiens, A.N.
 1972 *The interview: research on its anatomy and structure*. Chicago: Aldine Atherton.

- Mehrabian, A.
1969 Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin*, 71, 359-372.
- Mehan, H. and Wood, H.
1975 *The reality of ethnomethodology*. New York: Wiley.
- Meltzer, B.M., Petras, J.W. and Reynolds, L.T.
1975 *Symbolic interactionism: genesis, varieties and criticism*. London, Boston: Routledge & Kegan.
- Miller, G.A.
1963 Review of J.H. Greenberg (Ed.) *Universals of language*. *Contemporary Psychology*, 8, 417-418.
- Moerman, M.
1969 A little knowledge. In: Tyler, S.A. (Ed.) *Cognitive anthropology*. New York: Holt, Rinehart & Winston, 449-469.
- Moore, G.A.
1967 *The Japanese Mind*. Honolulu, Hawaii: East-West Center Press.
- Morsbach, H.
1973 Aspects of nonverbal communication in Japan. *Journal of Nervous and Mental Disease*, 157, no. 4, 262-277.
- Nakamura, H.
1964 *Ways of thinking of Eastern Peoples*. Honolulu, Hawaii: East-West Center Press.
- Nakamura, H.
1967 *A history of the development of Japanese thought*. 2 volumes. Tokyo: Kokusai Bunka Shinkokai.
- Nichols, K.A. and Champness, B.G.
1971 Eye gaze and the GSR. *Journal of Experimental Social Psychology*, 7, 623-626.
- Nielsen, G.
1962 *Studies in self-confrontation*. Copenhagen: Munksgaard.
- Norwine, A.C. and Murphy, O.J.
1938 Characteristic time intervals in telephonic conversation. *Bell System Technical Journal*, 17, 281-291.
- Park, R.E. and Burgess, E.W. (Eds.)
1921 *Introduction to the science of Sociology*. Chicago: University of Chicago Press.
- Parsons, T. and Shils, E. (Eds.)
1951 *Toward a general theory of action*. Cambridge: Harvard University Press.

- Peery, J.C. and Stern, D.N.
1976 Gaze duration frequency distribution during mother-infant interaction. *Journal of Genetic Psychology*, 129, 45-55.
- Philips, S. Urmston
1976 Some sources of cultural variability in the regulation of talk. *Language in Society*, 5, no. 1, 81-95.
- Richards, M.P.M.
1971 Social interaction in the first weeks of human life. *Psychiatria Neurologia Neurochirurgia*, 74, 35-42.
- Robson, K.S.
1967 The role of eye-to-eye contact in maternal-infant attachment. *Journal of Child Psychology and Psychiatry and Applied Disciplines*, 8, 13-25.
- Rogers, C.R.
1942 *Counseling and psychotherapy*. Boston: Houghton Mifflin.
- Rosenfeld, H.M.
1977 Conversational control functions of nonverbal behavior — an evaluative review. In: Siegman, A.W. and Feldstein, S. (Eds.) *Nonverbal behavior and communication*. New York: Halsted Press.
- Rubenstein, E.A. and Parloff, M.B. (Eds.)
1959 *Research in psychotherapy*. Washington: American Psychological Association.
- Rutter, D.R. and Stephenson, G.M.
1972 Visual interaction in a group of schizophrenic and depressive patients. *British Journal of Social and Clinical Psychology*, 11, 57-65.
- Rutter, D.R., Stephenson, G.M., Lazzerini, A.J., Ayling, K. and White, P.A.
1977 Eye contact: a chance product of individual looking? *British Journal of Social and Clinical Psychology*, 16, 191-192.
- Sacks, H., Schegloff, E.A. and Jefferson, G.
1974 A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Salzinger, K.
1959 Experimental manipulation of verbal behavior: a review. *Journal of Genetic Psychology*, 61, 65-95.
- Sarason, I.G.
1958 Interrelationships among individual difference variables, behavior in psychotherapy and verbal conditioning. *Journal of Abnormal and Social Psychology*, 56, 339-344.

- Sartre, J.P.
1943 *L'Être et le Néant*. Paris: Gallimard.
- Saslow, G. and Matarazzo, J.D.
1959 A technique for studying changes in interview behavior.
In: Rubenstein, E.A. and Parloff, M.B. (Eds.) *Research in psychotherapy*. Washington: American Psychological Association, 125-159.
- Scaife, M. and Bruner, J.S.
1975 The capacity for joint visual attention in the infant.
Nature, 253, 265-266.
- Schefflen, A.E.
1966 Natural history method in psychotherapy: communicational research. In: Gottschalk, L.A. and Auerbach, A.H. (Eds.) *Methods of research in psychotherapy*. New York: Appleton Century Crofts, 263-289.
- Schefflen, A.E.
1968 Human communication: behavioral program and their integration in interaction. *Behavioral Science*, 13, 44-55.
- Schefflen, A.E.
1973 *Communicational structure: analysis of a psychotherapy transaction*. Bloomington: Indiana University Press.
- Schegloff, E.A.
1968 Sequencing in conversational openings. *American Anthropologist*, 70, 1075-1095.
- Schegloff, E.A. and Sacks, H.
1973 Opening up closings. *Semiotica*, 8, 289-327.
- Schenkein, J.N.
1972 Towards an analysis of natural conversation and the sense of heheh. *Semiotica*, 6, 344-377.
- Schlien, J. (Ed.)
1968 *Research in psychotherapy*, volume 3. Washington: American Psychological Association.
- Schmidt, W. and Hore, T.
1970 Some nonverbal aspects of communication between mother and preschool child. *Child Development*, 41, 888-896.
- Schopen, T. (Ed.)
1977 *Language and their speakers*. Cambridge: Winthrop (in press).
- Sebeok, Th.A. (Ed.)
1968 *Animal communication*. Bloomington: Indiana University Press.

- Segal, M.M. and Webb-Stacy, E.
1975 Rule governed behavior as a psychological process. *American Psychologist*, 30, 541-552.
- Sieglman, A.W. and Feldstein, S. (Eds.)
1977 *Nonverbal behavior and communication*. New York: Hallsted Press.
- Simmel, G.
1921 Sociology of the senses: visual interaction. In: Park, R.E. and Burgess, E.W. (Eds.) *Introduction to the science of sociology*. Chicago: University of Chicago Press, 356-360.
- Skinner, B.F.
1948 *Verbal behavior*. Massachusetts: Harvard University Press.
- Stass, J.W. and Willis, F.N.
1967 Eye-contact, pupil dilation, and personal preference. *Psychonomic Science*, 7, 375-376.
- Stern, D.N.
1974 Mother and infant at play: the dyadic interaction involving facial, vocal and gaze behaviors. In: Lewis, M. and Rosenblum, L.A. *The origins of behavior*, volume 1. New York: Wiley, 187-213.
- Stern, D.N.
1971 A micro-analysis of mother-infant interaction: Behaviour regulating social contact between a mother and her 3½ month-old twins. *Journal of the American Academy of Child Psychiatry*, 10, 501-517.
- Strongman, K. and Champness, B.G.
1968 Dominance hierarchies and conflict in eye-contact. *Acta Psychologica*, 28, 376-386.
- Taylor, H.M.
1975 Americans and Japanese nonverbal communication. In: Neustupný, J.V. (Ed.) *Papers in Japanese Linguistics*, 3, 114-134. *Linguistic Communications* 15, Melbourne, Monash University.
- Thorne, B. and Henley, N. (Eds.)
1975 *Language and sex: difference and domination*. Rowley, Massachusetts: Newbury House.
- Tomkins, S.S.
1963 Shame-humiliation and the taboo on looking. In: Tomkins, S.S. (Ed.) *Affect, imagery, consciousness*, volume 2. New York: Springer Publishing Co., 157-183.
- Tomkins, S.S. and Izard, C.E. (Eds.)
1965 *Affect, cognition and personality*. New York: Springer.

- Trevarthen, C.
1975 Early attempts at speech. In: Lewin, R. *Child alive*. London: Temple Smith, 62-80.
- Triandis, H.C., Vassiliou, V., Vassiliou, G., Tanaka, Y. and Shanmugam, A.V.
1972 *The analysis of subjective culture*. New York: Wiley.
- Tronick, E.D., Als, H. and Brazelton, T.B.
1977 Mutuality in mother-infant interaction. *Journal of Communication*, 27, 2, 74-79.
- Turner, R.
1970 Words, utterances and activities. In: Douglas, J. (Ed.) *Understanding everyday life*. Chicago: Aldine, 176-187.
- Turner, R. (Ed.)
1974 *Ethnomethodology*. Harmondsworth: Penguin.
- Verplanck, W.S.
1955 The control of the content of conversation: reinforcement of statements of opinion. *Journal of Abnormal and Social Psychology*, 51, 668-676.
- Verzeano, M.
1950 Time-patterns of speech in normal subjects. (part 1). *Journal of Speech & Hearing Disorders*, 15, 197-201.
- Verzeano, M.
1951 Time-patterns of speech in normal subjects. (part 2). *Journal of Speech & Hearing Disorders*, 16, 346-350.
- van Vlack, J.
1966 Filming psychotherapy from the viewpoint of a research cinematographer. In: Gottschalk, L.A. and Auerbach, A.H. (Eds.) *Methods of research in psychotherapy*. New York: Appleton Century Crofts, 15-24.
- Vine, I.
1970 Communication by facial-visual signals. In: Crook, J.H. (ed.) *Social behavior in birds and mammals*. London: Academic Press, 279-354.
- Vine, I.
1973 The role of facial-visual signalling in early social development. In: von Cranach, M. and Vine, I. (Eds.) *Social communication and movement*. London: Academic Press, 195-298.
- von Cranach, M. and Ellgring, J.H.
1973 Problems in the recognition of gaze direction. In: von Cranach, M. and Vine, I. (Eds.) *Social communication and movement*. London: Academic Press, 419-443.

von Cranach, M. and Vine, I. (Eds.)

1973 *Social communication and movement*. London: Academic Press.

Watson, O.M.

1970 *Proxemic behavior: a crosscultural study*. The Hague: Mouton.

Werner, H. and Kaplan, B.

1963 *Symbol formation*. New York: Wiley.

Wiemann, J.M. and Knapp, M.L.

1975 Turn-taking in conversations. *Journal of Communication*, 25, 75-92.

Williams, T.R. (Ed.)

1975 *Socialization and communication in primary groups*. The Hague: Mouton.

Yngve, V.H.

1970 On getting a word in edgewise. In: Chicago Linguistic Society, *Papers from the sixth regional meeting*. Chicago, Illinois: Chicago Linguistic Society, 567-578.

Yngve, V.H.

1975 Human Linguistics and Face-to-Face Interaction. In: Williams, T.R. (Ed.) *Socialization and communications in primary groups*. The Hague: Mouton, 313-328.

Zimmerman, D.H. and West, C.

1975 Sex roles, interruptions and silences in conversations. In: Thorne, B. and Henley, N. (Eds.) *Language and sex: difference and domination*. Rowley, Massachusetts: Newbury House.

A P P E N D I C E S

APPENDIX I

SYMBOLS USED IN TRANSCRIPTION

/	Upward intonation
//	Point at which following line interrupts
(0.N)	Pause or gap of n tenths of a second
[	Simultaneous utterances when bridging two lines
]	(At end of simultaneously started or overlapped talk) indicates point at which utterances end <i>vis-à-vis</i> one another
R:	Identification of speaker
(word)	Probably what said, but not clear
(R):	Probably who is speaking, but not clear
()	Something said but not transcribable
<u>but</u>	Accentuation by volume or by intonation
=	Latching, i.e. no gap between 'end' of one and 'start' of another piece of talk (used as transcription facilitator)
--	Untimed pause (generally less than 0.2)
?	Questioning intonation (i.e. not a grammatical mark)
::	Prolongation of syllable preceding. The length of prolongation is given by the relative amount of colons used
-	Cut off prior sound
@	Following words spoken quietly
(())	Contains audio materials other than verbalization, e.g. ((cough))

Appendix II

Listener Responses (pilot-study Australian participants)

For this pilot-study the rules of the L-game were explained by the experimenter to four participants. Duncan's classification for listener responses (see page 65) was used to score the behavior of the listeners.

Type of Listener Responses	Participants (duration of segment)				
	I (1'.42")	II (1'.33")	III (0'.58")	IV (2'.36")	Total (6'.49")
Head Nods (total number)	9	4	5	8	26
multiple N	4	1	-	1	6
double NN	-	1	2	1	4
single N	2	-	2	3	7
1½ N	-	-	-	-	-
½ N ↑	1	2	-	1	4
½ N ↓	1	-	-	1	2
head shake ↔	1	-	1	1	3
Verbal Listener Responses (total number)	12	7	6	14	39
reinforcers { attention	3	4	3	6	} 25
assenting	5	-	1	3	
sentence completion	0	1	-	1	2
brief restatement	2	-	1	1	4
clarifying question	2	2	1	3	8

Appendix III
Frequency of Q-Directed Gaze

I	R	I	R	
	407	2 {	37	51
	355		37	43
	313		36	42
	272		35	41
	215		34	40
	178		33	38
	167		31	35
131	161		29	34
128	159		27	32
127	154		24	31
	152	3 {	24	30
113	142		24	30
112	129	2 {	21	30
111	128		21	26
94	122	2 {	20	24
91	111		20	22
90	102	2 {	19	21
87	94		19	18
	91		16	18
2 {	79	4 {	16	16
	79		16	
	75		16	
	73	2 {	15	15
			15	13
	60		14	11
3 {	58	2 {	13	9
	58		13	7
	58		12	7
	56		12	6
2 {	54	5 {	12	6
	54		12	2
2 {	52		12	2
	52		11	2
	49		10	2
	48		10	1
	44	5 {	10	
	43		10	
	41	2 {	10	
			4	
			4	

Total 2871

f = 68

median = 32

mean = 42.22

s.d. = 34.14

s.m. = 4.14

1.76 seconds

Total 5455

f = 71

median = 53

mean = 76.83

s.d. = 81.93

s.m. = 9.72

3.20 seconds

Appendix IVa
Gaze-Direction During a 5-Minute Period -- Japanese
Per 100 Frames (= 4.17 Seconds)

	I				T			
	Speech		Non-Speech		Speech		Non-Speech	
	Looking	Non-L	Looking	Non-L	Looking	Non-L	Looking	Non-L
1	8	42	-	50	-	-	35	65
2	24	-	5	71	15	55	30	-
3	-	25	5	70	10	-	1	89
4	-	40	8	52	50	-	50	-
5	-	85	-	15	2	13	17	68
6	39	61	-	-	-	-	100	-
7	13	27	60	-	23	17	60	-
8	-	-	57	43	-	80	2	18
9	-	20	58	22	56	4	40	-
10	-	-	92	8	-	100	-	-
11	-	15	55	30	-	45	-	55
12	24	26	25	25	-	50	-	50
13	9	1	-	90	-	-	-	100
14	-	-	-	100	-	-	-	100
15	3	32	49	16	-	50	-	50
16	-	-	-	100	-	100	-	-
17	-	90	-	10	-	-	-	100
18	-	70	-	30	4	26	5	65
19	-	70	27	3	13	7	-	80
20	-	-	16	84	14	86	-	-
21	-	-	42	58	-	90	-	10
22	-	-	98	2	82	18	-	-
23	6	8	84	2	82	-	18	-
24	9	-	81	10	100	-	-	-
25	7	73	12	8	20	-	80	-
26	78	16	4	2	6	-	94	-
27	19	21	52	8	60	-	40	-

	I				T			
	Speech		Non-Speech		Speech		Non-Speech	
	Looking	Non-L	Looking	Non-L	Looking	Non-L	Looking	Non-L
28	16	34	10	40	50	-	50	-
29	19	41	40	-	6	-	94	-
30	8	92	-	-	-	-	100	-
31	98	2	-	-	-	-	100	-
32	65	34	1	-	-	-	100	-
33	44	16	22	18	6	14	70	10
34	23	67	-	10	6	4	90	-
35	36	-	20	44	-	26	31	43
36	-	-	-	100	-	-	-	100
37	55	20	-	25	-	-	72	28
38	60	20	8	12	20	-	74	6
39	-	-	-	100	39	-	12	49
40	-	-	-	100	-	-	-	100
41	-	-	-	100	-	-	-	100
42	45	5	-	50	-	-	28	72
43	20	5	66	9	-	45	11	44
44	19	32	39	10	30	-	70	-
45	-	-	30	70	32	17	24	27
46	-	4	24	72	5	10	61	24
47	11	14	4	71	-	70	-	30
48	20	-	77	3	-	55	-	45
49	10	60	14	16	9	16	-	75
50	-	-	69	31	-	60	-	40
51	-	5	45	50	19	36	10	35
52	-	60	16	24	16	14	15	55
53	-	30	19	51	-	-	72	28
54	-	-	47	53	-	30	-	70
55	-	5	17	78	2	18	10	70
56	-	-	-	100	-	-	-	100
57	50	10	29	11	-	20	-	80

I					T			
Speech		Non-Speech			Speech		Non-Speech	
Looking	Non-L	Looking	Non-L		Looking	Non-L	Looking	Non-L
58	20	-	54	26	-	40	53	7
59	-	10	52	38	26	34	8	32
60	24	21	27	28	-	50	-	50
61	-	11	24	65	8	24	37	31
62	21	79	-	-	-	-	-	100
63	16	54	10	20	-	10	10	80
64	13	22	25	40	38	12	7	43
65	-	-	47	53	-	90	-	10
66	-	-	47	53	43	57	-	-
67	-	8	23	69	13	77	4	6
68	-	5	-	95	-	35	-	65
69	-	-	73	27	35	50	-	15
70	-	10	40	50	70	-	30	-
71	-	15	38	47	39	46	15	-
72	5	-	47	48	35	20	10	35

	2450		4750		2805		4395	
Total Frames	937 (38%)	1513 (62%)	1934 (41%)	2816 (59%)	1084 (39%)	1721 (61%)	1840 (42%)	2555 (58%)
Per 100 Frames	13.01	21.01	26.86	39.11	15.06	23.90	25.55	35.49
In Seconds	0.54	0.88	1.12	1.63	0.63	1.00	1.06	1.48

$$\frac{\text{Gazes List I}}{\text{Total Speech T}} = \frac{1934}{2805} = 0.69$$

$$\frac{\text{Gazes List T}}{\text{Total Speech I}} = \frac{1840}{2450} = 0.75$$

Appendix IVb

Gaze-Direction During a 5-Minute Period — Australian
Per 100 Frames (= 4.17 Seconds)

	R				V			
	Speech		Non-Speech		Speech		Non-Speech	
	Looking	Non-L	Looking	Non-L	Looking	Non-L	Looking	Non-L
1	42	22	14	22	-	-	91	9
2	26	-	59	15	33	-	64	3
3	10	-	65	25	40	-	60	-
4	32	68	-	-	-	-	72	28
5	60	25	15	-	5	-	95	-
6	100	-	-	-	-	-	100	-
7	78	22	-	-	-	-	100	-
8	34	26	28	12	1	9	67	23
9	50	26	-	24	11	4	47	38
10	42	-	56	2	2	48	45	5
11	29	31	40	-	32	-	40	28
12	-	7	7	86	50	-	50	-
13	-	-	100	-	86	14	-	-
14	-	-	64	36	30	70	-	-
15	8	27	9	56	28	2	70	-
16	52	48	-	-	-	-	100	-
17	19	13	32	36	-	60	30	10
18	-	-	66	34	-	100	-	-
19	-	-	100	-	19	81	-	-
20	8	2	85	5	17	50	20	13
21	8	2	75	15	-	75	-	25
22	53	14	8	25	20	-	70	10
23	-	-	86	14	30	65	-	5
24	-	-	90	10	73	27	-	-
25	-	-	72	28	71	29	-	-
26	-	-	55	45	62	38	-	-
27	25	-	70	5	40	20	14	26

	R				V			
	Speech		Non-Speech		Speech		Non-Speech	
	Looking	Non-L	Looking	Non-L	Looking	Non-L	Looking	Non-L
28	-	-	100	-	79	21	-	-
29	20	-	76	4	25	30	24	21
30	25	-	75	-	56	12	32	-
31	-	-	100	-	65	35	-	-
32	10	-	82	8	95	-	5	-
33	95	5	-	-	-	-	100	-
34	84	-	16	-	14	29	32	25
35	-	-	100	-	21	79	-	-
36	-	-	100	-	35	65	-	-
37	-	-	80	20	88	12	-	-
38	-	-	95	5	100	-	-	-
39	-	-	94	6	100	-	-	-
40	9	51	33	7	10	-	90	-
41	100	-	-	-	-	-	100	-
42	15	-	85	-	46	14	32	8
43	-	-	100	-	13	87	-	-
44	-	-	92	8	78	22	-	-
45	-	-	100	-	65	35	-	-
46	-	-	92	8	17	83	-	-
47	-	-	38	62	-	100	-	-
48	-	-	50	50	100	-	-	-
49	-	-	95	5	9	61	-	30
50	30	-	55	15	28	4	45	23
51	-	-	53	47	-	24	-	76
52	-	-	63	37	39	21	-	40
53	2	68	11	19	-	-	100	-
54	61	39	-	-	-	-	100	-
55	100	-	-	-	-	-	100	-
56	83	17	-	-	-	-	100	-
57	81	19	-	-	-	-	100	-

	R				V			
	Speech		Non-Speech		Speech		Non-Speech	
	Looking	Non-L	Looking	Non-L	Looking	Non-L	Looking	Non-L
58	81	-	19	-	54	-	46	-
59	72	14	2	12	-	-	100	-
60	80	20	-	-	-	-	100	-
61	54	46	-	-	-	-	100	-
62	2	34	55	9	1	9	32	58
63	21	49	16	14	24	-	76	-
64	44	56	-	-	-	-	100	-
65	73	27	-	-	-	-	100	-
66	37	10	1	52	-	-	100	-
67	18	8	54	20	27	-	47	26
68	-	-	100	-	54	13	6	27
69	-	-	100	-	64	36	-	-
70	-	-	96	4	64	36	-	-
71	-	-	94	6	4	81	-	15
72	63	7	1	29	-	-	65	35

	2739		4461		3626		3574	
Total	1936	803	3519	942	2025	1601	2967	607
Frames	(71%)	(29%)	(79%)	(21%)	(56%)	(44%)	(83%)	(17%)
Per 100	26.89	11.15	48.88	13.08	28.13	22.24	41.21	8.43
Frames								
In Seconds	1.12	0.46	2.04	0.55	1.17	0.93	1.72	0.35

$$\frac{\text{Gazes List R}}{\text{Total Speech V}} = \frac{3519}{3626} = 0.97$$

$$\frac{\text{Gazes List V}}{\text{Total Speech R}} = \frac{2967}{2739} = 1.08$$

Appendix V
Frequency and Duration of Speaker-Turns

I	T	R	V
			730
349	280	497	530
185	170	317	435
140	150	316	352
126	145	220	277
125	2 { 120	194	225
110	120	175	210
90	115	165	
84	110	147	135
80	105		125
74	100	102	
70	82	76	
70	80	70	
65	70	67	82
5 { 60	4 { 60	64	78
60	60	60	60
60	60	52	54
60	60	30	50
60	2 { 55	2 { 26	40
50	55	26	34
45	2 { 50	25	27
2 { 35	50	3 { 25	26
35	49	25	24
31	45	20	23
2 { 30	2 { 40	2 { 20	2 { 20
30	40	10	20
25	39	2 { 10	18
3 { 25	35		2 { 15
25	32		15
20	30		10
20	30		6
6 { 20	6 { 30		
20	30		
20	30		
3 { 15	26		
15	25		
15	20		
4 { 10	20		
10	20		
10	8 { 20		
10	20		
8	20		
7	20		
2 { 6	20		
6	15		
3 { 5	10		
5	2 { 6		
5	6		
4			
2450	f = 49 2805	f = 48	2739 f = 25 3626 f = 28

Appendix VI
Data for Individual Participants of q-Directed Gaze Only

Japanese			
	q-Gaze Speech	q-Gaze Non-Speech	left
J1	26	10	
J2	-	10	
J3	11	39	
J4	58	-	
J5	1	7	
J6	27	52	
J7	11	22	
			right
J1	-	61	
J2	9	20	
J3	31	39	
J4	18	75	
J5	53	3	
J6	14	35	
J7	14	26	
Australians			
	q-Gaze Speech	q-Gaze Non-Speech	left
A1	82	-	
A2	90	10	
A3	13	52	
A4	11	45	
A5	34	21	
			right
A1	-	100	
A2	10	77	
A3	65	28	
A4	65	28	
A5	28	72	

Left and right refers to the position of the participant on the TV monitor.

Appendix VII

Data for Speech for 7 Japanese Interactions during P_2

	Left				Right			
	Cum.	Freq.	Aver.	Perc.	Cum.	Freq.	Aver.	Perc.
J ₁	75.9	61	1.24	25.30	144.9	67	2.16	48.30
J ₂	101.1	59	1.71	33.70	74.1	62	1.20	24.70
J ₃	168.6	123	1.37	56.20	56.7	70	0.81	18.90
J ₄	106.5	90	1.18	35.50	80.4	72	1.12	26.80
J ₅	56.1	56	1.00	18.70	142.2	85	1.67	47.40
J ₆	85.5	91	0.94	28.50	96.3	82	1.17	32.10
J ₇	91.2	63	1.45	30.40	103.2	75	1.38	34.40

Appendix VIII

Data for Speech for 5 Australian Interactions during P_2

	Left				Right			
	Cum.	Freq.	Aver.	Perc.	Cum.	Freq.	Aver.	Perc.
A ₁	35.1	25	1.40	11.70	147.9	81	1.83	49.30
A ₂	104.1	66	1.58	34.70	90.0	61	1.46	30.00
A ₃	68.4	55	1.24	22.80	86.7	60	1.45	28.90
A ₄	103.2	50	2.06	34.40	106.2	74	1.44	35.40
A ₅	75.6	61	1.24	25.20	103.8	89	1.17	34.60

Appendix IXa
Speaker Turns (Japanese Condition)

H	RW	H	RW
320	615	27	25
	300	25	25
	255	3 { 25	22
165	170	25	20
125	125	2 { 20	20
	105	20	20
103		15	20
100	2 { 95	15	18
	95	5 { 15	15
	90	15	15
	87	15	5 { 15
	85		15
	2 { 80	4 { 10	15
	80	10	10
78		10	3 { 10
70	69	10	10
69			7
2 { 65	56	5	5
65	55	5	5
60	2 { 50	5	5
55	50	5	7 { 5
54	45	5	5
50	3 { 45	5	5
49	45	5	5
45	42	5	
42		4	4
2 { 40		3	3
40	3 { 35	2	
35	35		
35	35		
5 { 35			
35			
35	3 { 30		
30	30		
30	30		
5 { 30			
30			
30			

2226	total number of frames	3158
55	frequency	54
40.47	average number of frames per turn	58.48
6.80	standard error of mean	13.13
1.69	turn duration in seconds	2.44

Appendix IXb
Speaker Turns (Australian Condition)

R	RW	R	RW
600	665	2 { 50	
540	656	2 { 50	
	505	2 { 45	45
280		2 { 45	40
220			40
210		40	40
180	165		39
149	164	35	
120	159	30	30
	120	25	25
	2 { 115		25
	2 { 115		25
94	110		25
	95		25
84	85	2 { 20	20
69		2 { 20	20
60	60		15
		2 { 10	13
		2 { 10	10
		5	5
2991	Total number of frames		3317
25	Frequency		30
119.64	Average number of frames per turn		110.56
30.89	Standard error of mean		32.20
4.99	Turn duration in seconds		4.61

Appendix XI

Gaze-Direction in Japanese Participants in J-J and A-A Interactions (N = 9)

	J-J				J-A			
	Speech		Non-Speech		Speech		Non-Speech	
	g-directed gaze	a-directed gaze	g-directed gaze	a-directed gaze	q-directed gaze	a-directed gaze	q-directed gaze	a-directed gaze
J1	-	65	305	130	13	237	-	250
J2	43	131	99	227	37	333	100	30
J3	57	193	196	54	60	120	105	215
J4	2	68	37	393	378	38	20	64
J5	288	212	-	-	87	313	30	70
J6	263	23	16	198	154	346	-	-
J7	72	193	177	58	57	154	4	285
J8	70	110	132	188	162	188	101	49
J9	57	13	111	319	50	-	365	85

Appendix XII

Gaze-direction in Australian Participants in A-A and A-J Interactions (N = 7)

	A-A				A-J			
	Speech		Non-Speech		Speech		Non-Speech	
	q-directed gaze	a-directed gaze	q-directed gaze	a-directed gaze	q-directed gaze	a-directed gaze	q-directed gaze	a-directed gaze
A1	410	90	-	-	206	23	248	23
A2	-	-	500	-	202	21	238	39
A3	67	23	260	150	72	-	428	-
A4	54	6	226	214	-	-	500	-
A5	327	20	141	12	30	-	470	-
A6	169	161	106	64	10	-	490	-
A7	140	-	360	-	232	138	130	-