THE STRUCTURE OF INFORMAL CONVERSATION: A SEQUENTIAL ANALYSIS

ANDREW PETER THOMAS

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (D.Phil) to the University of York, Department of Psychology.

October 1982

CONTENTS

		Page
Acknowledgements		8
Abstract		9
Introduction		11
Chapter 1	Methodological issues and outline of the research project.	
1.1	Introduction.	29
1.2	Speech segmentation.	29
1.3	Statistical analysis of interpersonal communication.	30
1.3.1	Analysing communication as a process.	31
1.3.2	Preserving the relationship in communication.	32
1.3.3	Analysis of the structure of communication through constraints and rules.	33
Chapter 2	Conversational Exchange Analysis.	
2.1	Introduction.	35
2.2	System requirements.	38
2.3	System framework.	39
2.4	Procedures for dividing conversation into acts.	39
2.4.1	Acts: Transcript division rules.	44
2.5	Act classification.	50
2.5.1	Activity classification.	53
2.5.2	Type classification.	61
2.5.3	Focus classification.	72

		Page
2.6	System reliability.	75
2.7	Example transcript of informal conversation coded using CEA.	76
2.8	Conclusion.	83
Chapter 3	A preliminary analysis of conversation.	
3.1	Introduction.	86
3.2	Method.	89
3.2.1	Subjects.	89
3.2.2	Procedure.	89
3.3	Data analysis and Results.	90
3.3.1	Method of analysis.	90
3.3.2	Test of independence between speech events.	93
3.3.3	The strength of the relationship between speech events.	95
3.4	Discussion.	96
Chapter 4	A sequential analysis of informal conversation using Markov chains.	
4.1	Introduction.	100
4.2	Method.	112
4.2.1	Subjects.	112
4.2.2	Procedure.	112
4.3	Data analysis overview.	113
4.4	A four-code model of conversation.	113
4.4.1	Stationarity.	116
4.4.2	Order of sequential constraint.	121
4.4.3	Homogeneity.	141
4.4.4	A test of adequacy of the four-code model.	145

		Page
4.5	A seven-code model of conversation.	151
4.5.1	Order of sequential constraint.	151
4.5.2	A test of adequacy of the seven-code model.	155
4.6	Replication of the seven-code model results.	156
4.6.1	Order of sequential constraint.	156
4.6.2	A test of adequacy of the replication model.	159
4.6.3	Comparison of dyad groups 1-12 and 13-24.	159
4.6.4	Replication study conclusion.	160
4.7	Discussion.	162
Chapter 5	The characteristics of non-stationary sequences.	
5.1	Introduction.	165
5.2	Method.	166
5.2.1	Sequence generation using Monte-Carlo methods.	167
5.2.2	Calculation of process characteristics.	169
5.2.3	Comparison of observed and generated process characteristics.	170
5.3	Results.	171
5.4	Discussion.	174
Chapter 6	Turn-taking and strategies in informal conversation.	
6.1	Introduction.	177
6.2	Conversational turn-taking.	178

		Page
6.2.1	Method and Results.	184
6.2.2	Discussion.	186
6.3	Conversational strategies.	187
6.4	Conclusion.	194
Chapter 7	Conversational content: A 'Type' analysis.	
7.1	Introduction.	198
7.2	Method and Results.	200
7.3	Discussion.	201
7.4	Conclusion.	217
Chapter 8	A 'chain analysis' study of informal conversation.	
8.1	Introduction.	220
8.2	Method and Results.	223
8.3	Discussion.	233
8.4	Conclusion.	237
Chapter 9	An analysis of conversational structure using Information Theory.	
9.1	Introduction.	240
9.2	Data analysis overview.	244
9.3	Information analyses of conversation coded for speech Activity.	245
9.3.1	Multivariate analysis of conversational structure.	245
9.3.2	Univariate analysis of distributional structure.	249
9.3.3	Univariarate analysis of sequential structure.	251

		Page
9.4	Information analyses of conversation coded for speech Type.	255
9.4.1	Multivariate analysis of conversational structure.	256
9.4.2	Univariate analysis of distributional structure.	258
9.4.3	Univariate analysis of sequential structure.	260
9.5	Discussion.	265
9.6	Conclusion.	270
Chapter 10	Conclusions.	274
Appendices:		
A	Conversational Exchange Analysis inter-observer reliability.	295
B	Controversial topics questionnaire.	301
C	A note on the Anderson-Goodman maximum likelihood ratio statistic.	303
D	Information analyses (Chapter 4).	305
E	Mean inter-event distance process characteristics (Chapter 5 analyses).	317
F	Uncertainty analyses (Chapter 9).	328
G	Computer programs.	337
	1 Sequence stationarity (speaker-specific).	339
	2 Sequence stationarity (transition-specific).	344
	3 First-order analysis.	349
	4 Second-order analysis.	352

		Page
5	Third-order analysis.	356
6	Sample homogeneity.	360
7	N-step contingencies.	372
8	Matrix multiplication.	382
9	Matrix comparator.	385
10	Monte-Carlo generator.	389
11	Inter-event distance.	392
12	Process data comparator.	398
13	'Type' analysis.	402
14	Chain analysis.	406
15	Uncertainty analysis.	409
References		413

ACKNOWLEDGEMENTS

I would like to thank Peter Bull and Derek Roger for their advice and enthusiasm throughout the course of this research project, as well as Olga Sucharyna for her encouragement and invaluable assistance with the reliability studies, John Byrne (University of York, Computer Centre) for his statistical advice, 'Taff' Townsend for his constructive conversation, Marilyn Glicker for typing the manuscripts for the journal articles, Angela Gabb for typing the final draft of the thesis, the Social Science Research Council for financial assistance, and finally, all those who have given up their time to take part in my research.

ABSTRACT

The main aim of this research project was to describe the organisation of informal conversation whilst preserving the process, relational and structural aspects of communication. This required a change in theoretical framework from a traditional, deterministic view of science to a probabilistic, General systems approach, as well as a corresponding change in the methods of analysis.

As 'process' data are interdependent, a method of segmenting speech into analysable units was required. A method was devised (Conversational Exchange Analysis) comprising four sets of rules: one for segmenting conversational speech into units, the remainder for classifying speech according to how information is exchanged in conversation (Activity), together with the content (Type) and referent (Focus) of the speech.

The informal conversations of 24 female dyads were subsequently analysed in order to test the hypothesis that a Markov process was a tenable model for informal conversation. The hypothesis was supported. This, in addition to a subsequently developed first-order model of conversation coded for speech 'Type', was used as a basis for extracting a number of conversational strategies, as well as identifying those speech states instrumental in the turn-taking process. The results are considered to have an important application in social skills training

procedures concerned with the teaching of conversational and general meshing skills.

Finally, through an analysis of conversational constraints, speech events in conversation have been shown to be highly organised, the larger part of the organisation being due to the distribution of events, rather than their sequential arrangement in the speech stream. In addition, it was suggested that the turn-taking function of some speech events was due, in part, to their 'structure-inducing' nature; Offers of information and Replies were found to retain the floor for the speaker whilst Consents, Reactions and Requests relinquished the speaking-turn.

INTRODUCTION

In a recent study by Thomas and Bull (1981), it was suggested that the verbal elements of interaction were an often neglected, though important variable in the study of social interaction. Reviewing the literature concerned with the structure and organisation of conversation, it became apparent that this was indeed the case. For example, in discussing the components of interaction, Trower, Bryant and Argyle (1978) described conversation primarily in terms of either non-verbal behaviours (e.g. head nods, gaze, gesture and body posture) or paralanguage (e.g. changes in pitch and intonation) (pp. 19-22), the verbal elements of conversation being described only in global terms, such as asking and answering questions, greeting and bidding farewell or telling jokes (p. 174), with very little emphasis on their sequential arrangement. Conversational meshing skills were also usually described in the form of non-verbal behaviours, questions and listener responses (Trower et al., 1978, pp. 221-226).

Commenting on the lack of detail in the existing models of interaction, Ellis and Whittington (1981), in their book 'A guide to social skill training', have urged that research is undertaken "which refines and elaborates existing models of social interaction" (p. 196) and have stated that "there is a need for observation and experimental study of the sequencing of such (behavioural) units" (p. 196). Consequently, this research project has been designed to investigate the nature and organisation of the verbal elements of conversation.

The nature of conversation

Conversation, and communication in general, is an act of participation in a relationship (Feldstein & Welkowitz, 1978; Penman, 1980). In its most general sense, a communication process is comprised of a series of behavioural exchanges between people. (This is the working definition of communication that will be used in this chapter, but will be developed in Chapter 3.) The study of the communication process is therefore concerned with the continuing inter-relationship between interacting individuals. As Penman (1980) has noted, "if the prime concern is with the relationship between individuals, then the characteristics of the individuals per se are of minor importance compared with the characteristics of their inter-relationship" (p. 2).

This change in focus, from the study of an individual to a relationship developing in time, requires a concomitant change in approach from, not only the individual to the relational, but also from the static to the dynamic. As relationships are comprised of a continuing process of exchange, an analysis of the inter-relationships between individuals must take into account the time-ordering of events. This of course, has implications for the methods used to collect and analyse the data.

Two approaches to the analysis of inter-personal behaviour

Most interpersonal communication research has been individually oriented (Berger, 1977), focusing on the effects that one person produces or may produce, in another, rather than the reciprocal nature of inter-personal communication. Typically, research of this type has used pairs of interacting subjects, in which the dyad has been the unit of analysis, rather than each of the interacting individuals (Kraemer & Jacklin, 1979). This approach is exemplified by the individually oriented, or 'monadic' studies of small group behaviour in the late 1950's and early 1960's (reviewed in Hare, 1976) such as social facilitation, and the non-verbal communication studies of Argyle and his colleagues in the late 1960's and early 1970's, such as equilibrium theory (e.g. Argyle, 1972). In each of these cases, the analysis of behaviour has been in terms of the consequences of environmental manipulations, the whole process of communication being either ignored, or treated as an intervening variable (Danziger, 1976).

An alternative is the transactional approach, which emphasises the reciprocal nature of the inter-personal communication process, embodying the notion of reciprocity in both theoretical and methodological senses. The transactional approach is usually concerned solely with the analysis of relational communication (Berger, 1977), for which special methods of measuring communication between dyads at the relationship level, rather than the content or referential level, have been devised (e.g. Mark, 1971; Rogers & Farace, 1975). In the analysis of relational communication, the concern is with patterns of control manifested in the relationship, and how these patterns change with time. As Rogers and Farace (1975) point out, such a method of analysis captures the 'process', or time aspect of a progressing communication.

Within the transactional approach a body of research has grown up in which the relationship between people is preserved, but the emphasis has moved away from the analysis of the relationship to the analysis of the communication process at a content level (e.g. Hawes & Foley, 1973; Ellis & Fisher, 1975; Stech, 1975), an approach often referred to as 'process oriented studies' (Berger, 1977).

A change of assumptions

In order to study communication processes, it is not simply a matter of changing the approach from the study of the individual to the study of the relationship, nor a change in emphasis from static to dynamic. As Penman (1980) has noted, such changes, if they are to be effective require "a fundamental epistemological and theoretical reorientation" (p. 4).

In discussing conversation as a communication process, one is essentially referring to an exchange, a reciprocal flow of information that takes place in real time. The two important concepts here are relation and process. In order to preserve these fundamental concepts, inherent in interpersonal communication, a change in both theoretical and methodological assumptions is required.

Taking a conventional empirical approach to the study of interpersonal communication, one has to explain the behaviour of one person in terms of causes and effects, or external stimulation, such as what is being done to that person. But, as has often been noted in the literature (e.g. Ellis & Fisher, 1975; Rogers & Farace, 1975; Penman, 1980), communication is a reciprocal process: as one person is being acted upon, that person is also acting. Where each person in the relationship is simultaneously the stimulus and the response, using mechanistic concepts not only creates conceptual and analytic problems (Penman, 1980), but also problems in trying to attribute causality. As Harré and Secord (1972) point out, according to the conventional conception of cause and effect, the mode of connection is ignored as it is not part of empirical science. But, if the connection between the cause and effect is ignored, what is left of the communication process? (Penman, 1980). Conversely, if one is to look at the connection between the cause and effect, which connection can one look at when all interactants are seen as both cause and effect?

The limitation of conventional empirical methods to account for communication has been recognised for some time. Von Bertalanffy (1962) has expressed explicitly the methodological shortcomings of classical causal analysis when he described conventional empirical methods: "The latter (*the empirical method*) was essentially concerned with two-variable problems, linear causal trains, one cause and one effect, or with few variables at the most. ... One-way causality, the relation between 'cause' and 'effect' or of a pair or a few variables covers a wide field. Nevertheless, many problems, particularly in biology and behavioural sciences, essentially are multivariate problems for which new conceptual tools are needed" (p.2).

Suggesting a change in theoretical framework, Von Bertalanffy (1971) in his book, 'General System Theory', has written, "We may state as characteristic of modern science that this scheme of isolable units acting in one-way causality has proved to be insufficient. Hence the appearance, in all fields of science, of notions like wholeness, holistic, organismic, gestalt, etc., which all signify that, in the last resort, we must think in terms of systems of elements in mutual interaction" (p.44). It would seem appropriate therefore to analyse social situations using a model of communication based on the time ordering of events (Feldstein & Welkowitz, 1978), in which the relationship between elements is construed in terms of constraints, rather than cause and effect (Penman, 1980). Indeed, Hertel (1972) maintains that the major failure of conventional experimental designs "lies in their inability to employ temporal relationships among units" (p.422), and Fisher (1978a) has noted that while the element of time is crucial in communication, it is still the most neglected variable in communication studies.

System theory - An alternative framework

According to Von Bertalanffy (1962), the classical cause and effect approach necessitated the specification of two-variable problems in which a small number of variables, isolated from their environment, were manipulated in order to discover the causal connections between them. He argues that this approach cannot cope with the complexities of living systems, whereas by contrast, system theory addresses itself directly to those complexities, thereby providing a framework for the study of the system itself, rather than the objects of the system.

The concept of a 'system' has been used consistently in the natural sciences to reflect the notion that elements are not isolated, but can be construed as sets of related events (Hall & Fagen, 1975, p.52). However, social scientists view a system to be an interdependent set of elements which is assumed to be more complex than the relations which constitute it; the elements are seen as operating dynamically together according to certain laws or rules in order to produce some overall effect (Allport, 1955). Using this wider definition of a system, system theory has been applied to interpersonal communication, spearheaded by what has come to be known as the 'Palo Alto' group (e.g. Bateson (1967); Jackson (1965) & Watzlawick et al. (1968)), and subsequently used as a guiding theoretical framework by such authors as Ellis and Fisher (1975), Fisher (1970), and more recently by Hawes and Foley (1976) and Penman (1980).

To outline system theory in more detail, Von Bertalanffy (1962) has suggested three basic tenets which are fundamental to the theory, and are considered essential for the study of human systems. Firstly, it is a general science of 'wholeness' (Von Bertalanffy, 1971, p.36), the emphasis being on the study of whole systems and not isolated and independent elements that are in reality related components.

Secondly, in studying communication, by focusing on the system itself and the interrelationship between elements, general systems theory concerns itself with the organisation, or organised complexity of the elements. It is concerned here with the analysis of the complex set of relations between elements in the system, as well as the possible complex hierarchy of relations between systems of different levels of organisation. With the increase in interest in the organisation of systems, a number of methodologies have been developed within system theory, such as the development of cybernetics and information theory, topology and graph theory, and decision and game theory, all of which can be applied to the analysis of the structure of human systems (Rapoport & Horvath, 1968).

Finally, system theory focuses on the dynamic activity of human systems, in which a relatively autonomous self-direction is seen as more important than the behaviourist 'reactive' view of organisms.

System theory and conversation analysis

The analysis of conversation within a system framework has three main focuses: relation, process and structure. In changing the emphasis from an individual to the relationship between the speakers, conversation is seen not simply as a process producing a continuous stream of conversation, but as a process that is generated by, in this case, two people interacting. The analyses in this research preserve the relation between individuals by splitting the speech stream into elements made by the same speaker (within-speaker transitions) and those made by the other speaker (between-speaker transitions). This will be more fully discussed in Chapters 3 and 4.

Process reflects the dynamic nature of interaction, describing the changes which the system undergoes over time, and is one method of describing the organisation of a system (Cushman & Craig, 1976). The time ordering of the elements is therefore essential to the analysis of the sequential organisation of conversation.

A second method for describing the organisation of conversational events is by analysing the structure of the conversational sequences. From the theoretical framework outlined so far, it has been shown that a linear model of cause and effect is inappropriate for the analysis of complexly-organised interrelations. Although cause and effect are redundant, the question of the effects or consequences of elements is nonetheless of interest. A system account considers the range of possible alternatives and then seeks to interpret the sequential arrangement of events on the basis of constraints between elements operating in the system. The structure of conversation will therefore be described in terms of system constraints rather than causes.

As a causal explanation of conversation is inappropriate, so too is any attempt to uncover general laws of behaviour (Cushman & Pearce, 1977; Penman, 1980). Instead, as Berger (1977) has argued, a more satisfactory explanation of social behaviour can be achieved by regarding behaviour as rule-based. As Penman (1980) has indicated, conceptualising the communication process as being based on constraints rather than causes, necessitates a rule-based account of the regularities in communication. Behaviour is therefore seen as rule-following, a person being free to break the rules of social interaction if they wish.

The methodological implications of using a system approach, and techniques for analysing conversation in terms of relation, process and structure are discussed in Chapter 1.

The type of conversation studied

The majority of conversational analyses conducted in a system framework have used 'directed' conversation such as that occurring at interview (e.g. Hawes & Foley, 1973), in committee meetings (e.g. Hawes & Foley, 1976), conflict situations (e.g. Hertel, 1968 - experimentally-induced marital conflict; Gottman et al., 1977b - naturally-occurring marital conflict), group decision-making (e.g. Fisher, 1970; Donohue, Hawes & Mabee, 1981), and psychotherapy sessions (e.g. Frank & Sweetland, 1962; Benjamin, 1979). Very little work has used informal conversation and discussion within the system framework; it is the intention of this research project, therefore, to analyse the conversation of pairs of individuals interacting in conditions as informal as possible, but restricted by the ability to make recordings of suitable quality for subsequent transcription. Rather than using naturally occurring conversation, such as that in snack-bars, or other social gatherings, where extraneous noise may render the recording unintelligible, pairs of individuals were invited to discuss a number of topics in a quiet room equipped with recording apparatus. The discussion topics ranged from religion to feminism and from Student Unions to mercy-killing, the subjects being encouraged to talk about any of the topics that were of interest to both themselves and their partner.

Conclusion and overview of the research project

It has been argued that very little research has been concerned with identifying the nature of verbal elements and their sequential arrangement in informal conversation. This therefore forms the basis of the present research project.

Intrinsic to social interaction are the notions of relation, process and structure, for which conventional cause and effect analytical methods have been shown to be inappropriate. Consequently, a theoretical reorientation is required in which the relationship between individuals and the time ordering of events is acknowledged and has been accomplished by appealing to General systems theory.

By virture of conceptualising conversation as a process, conversation was considered to be composed of a string of concatenous, inter-dependent events. Before analysis of the structure of conversation could proceed, a method of 'segmenting' the speech stream into analysable units was required. Although a number of methods for analysing speech already exist, it was considered that these were inappropriate for the present study by virtue of being either situation-specific, or having fundamental theoretical and methodological flaws. An appropriate method of data collection was therefore devised, together with a method of classifying the speech units according to their conversational content (Chapter 2). Conversational Exchange Analysis (CEA), the method of segmentation and classification developed for this research project, comprises four sets of rules. The first is concerned with segmenting conversation into analysable units, a unit of speech being defined as a 'single thought or idea'. The remaining three sets of rules are concerned with speech classification, coding the speech units along three, conceptually distinct, dimensions. The first, Activity, assesses how speech is made salient in the conversation. For example, is the information asked for or given, agreed with or disputed? The second level of analysis, Type, codes the content of the speech. For example, is the speech unit expressing a belief, telling a story, giving emotional support, or commenting on the progress of the conversation? The final level of analysis, Focus, scores the referent of the speech. For example, is the speaker expressing his own, or someone else's ideas? The Focus level of analysis is not used in this research project, but is included in order to provide a complete system of speech classification for application outside this thesis.

Having demonstrated, from a theoretical viewpoint, the inter-dependent nature of speech events, this assumption has then been tested, and the results form the basis of Chapter 3. Supporting the hypothesis of inter-dependence between speech events, the analyses in Chapter 3 also suggest that there is a small, predictive relationship between speech events.

Typically, analyses of conversation reported in the literature have proceeded from the raw data to conclusions about conversational structure, an underlying model of conversation being assumed, but rarely tested. In cases where tests of the model are carried out, the tests are usually inappropriate. The theme of Chapter 4 was, therefore, concerned with the development of a model of conversation, in which the model assumptions were rigorously tested, using statistical techniques appropriate for a System theory approach.

A Markov process was selected as a model for conversation and tested initially using an eight speech-state process, and finally a 14 speech-state process. The hypothesis that a Markov process was a tenable model for information exchange in informal conversation was supported, and replicated. In addition, the most suitable model was found to be a first-order process, the model being shown to be sufficiently flexible to produce accurate predictions of future events, even when the Markovity assumptions of sample homogeneity and sequence stationarity were violated.

In testing the Markov model, in Chapter 4, some conversations were found to be non-stationary, although this was later found not to affect the predictive ability of the final model of conversation. However, the aim of Chapter 5 was to assess whether the structure of non-stationary conversations differed, when compared to stationary conversations, using three process characteristics, n-step contingencies, and mean and standard deviation inter-event distance.

The technique used in this chapter was to generate a number of stationary conversational sequences, using Monte-Carlo methods, comparing the process characteristics from these, with those derived from the non-stationary sequences. The analyses indicated that the non-stationarity was, in part, due to a change in event-to-event probability of occurrence throughout the non-stationary conversations, as well as a clustering of Requests and Dissents in the first half of the conversations.

The model of conversation derived from the analyses presented in Chapter 4 has then been used, in Chapter 6, to indicate the role that speech content plays in conversational turn-taking. The analyses indicated that Requests and Reactions (listener responses) tended to precede changes in speaker, and Replies, Consents, and again, Reactions tended to follow changes in speaker. In addition, using the theory of directed graphs and the technique of 'condensation' to simplify the conversational system, a number of conversational strategies have been devised.

Up to this point, the analyses have all been carried out for conversations coded for speech Activity. Chapter 7 was, therefore, concerned with the analysis of conversation coded for speech Type. Although no statistical analyses could be performed, due to a paucity of data, an observational approach yielded a number of conversational routines and strategies.

The probabalistic models of conversation and the social strategies presented in Chapters 4, 6 and 7 are considered to have important implications for social skill training procedures based on Argyle and Kendon's (1967) model of social interaction. For example, the cybernetic nature of Argyle and Kendon's (1967) model of social interaction (Trower et al., 1978) makes General System Theory (cybernetics and social skills being special cases of a general system) the "most obvious source of social skill training" (Ellis & Whittington, 1981, p.25). In addition, Ellis and Whittington (1981) have urged that research is undertaken "which refines and elaborates existing models of interaction and which applies such refinements to the procedures of social skill training" (p.196). The conversational strategies based on probabalistic models and outlined in Chapters 6 and 7, would therefore seem ideally suited for application in social skill training procedures concerned with turn-taking and general conversational skills.

One limitation of Markovian methods is their inability to detect speech events that are embedded in larger sequences of conversation. Consequently, chain analysis, a method still within the Markovian approach, has been used in Chapter 8 to analyse the conversations for embedded, and multi-element units. The main conclusion of this analysis is that the conversations appear to be most usefully considered, at both the Activity and Type levels of analysis, in terms of two-event sequences, equivalent to those demonstrated by the Markov analyses presented in Chapters 4, 6 and 7. In addition, the analyses demonstrated a marked lack of embedded speech events, but, by way of extension of the Markov analyses, indicated that a number of two-element chains were cyclic, often forming multi-element chains, six or eight elements in length.

Whereas the preceding analyses have been concerned with the structure of conversation in terms of which speech events are to be found in sequential arrangement, Chapter 9 was concerned with the organisation of conversational events in terms of system constraints. Analysed using Uncertainty statistics, the conversations were found to be highly structured; the ability to predict the identity of the next speech event in the conversation depending (for speech coded for both Activity and Type) to a large extent on the distributional structure of the speech stream - the fact that some speech events occur more often than others, and to a considerably smaller extent on the sequential arrangement of events in the conversations. In addition, it is suggested in the concluding chapter that this technique of analysing system constraints may constitute a useful method of describing the style of a conversation.

Finally, it has been argued in the Conclusion (Chapter 10), that when analysed within a System framework, using appropriate statistical techniques, informal conversation can be seen to be adequately modelled by a first-order Markov process. In addition, informal conversation has been shown to be highly structured. These analyses have enabled two, highly detailed, models of conversation to be produced, one concerned with how information is exchanged, and the other concerned with the type of information exchanged in conversation. The two models have subsequently been used as a basis from which to extract a number of turn-taking and conversational strategies that are considered to have important implications for social skills training procedures concerned with turn-taking, general meshing, and conversational skills.

CHAPTER 1

METHODOLOGICAL ISSUES AND OUTLINE OF THE RESEARCH PROJECT

1.1 Introduction

In the Introduction, a distinction was drawn between 'traditional' and 'system' approaches to the study of interpersonal communication. From a system point of view, interaction and communication were seen, not as being comprised of isolated events, but as a set of related elements that act dynamically together, in order to produce some overall effect. This raises an essential difference between traditional and system approaches. Whereas the traditional view sees events as contiguous, close in proximity but analytically independent, a system view sees events as concatenous, close in proximity, and analytically dependent. As Fisher (1978a) has noted, "If interaction or communication is truly a social system, then the behavioural sequences (which by definition, include more than one person's behaviours) are interdependent and inseparable" (p.213).

1.2. Speech segmentation

As the events in communication are concatenous, by definition there can be no natural units which can be used to describe the communication process. But, to study the process, one has to impose some form of unit onto the conversational sequences.

This process of dividing a concatenous stream of events is known as unitising (Holsti, 1969) or punctuation (Bateson & Jackson, 1964), any form of which must be seen as essentially arbitrary (Penman, 1980). However, the term arbitrary is not meant to imply randomness, but rather to indicate the varying size of speech units produced by different methods of segmentation (Bateson & Jackson, 1964). The consequences of segmenting an essentially continuous stream of behaviour are twofold. Any decision regarding the segmentation is a strategic one (Wilden, 1972), in which the size of the unit is specified in terms of the research aim, and it is essential that the arbitrariness of the segmenting rule is taken into account when interpreting results based on such a decision. Secondly, once determined, the decision must be consistently applied for all streams of behaviour to be analysed.

Before a stream of behaviours can be considered as raw data for analysis, two separate processes are necessary. Firstly, segmentation of the speech stream into discrete units, and secondly, classification of the behaviours in terms of the intended research. The methods by which speech is segmented and subsequently classified are dealt with in depth in Chapter 2.

1.3 The statistical analysis of interpersonal communication

The second methodological issue raised by adopting a system approach concerns the notions that communication takes place in real time and that concatenous events are analytically dependent. As Fisher (1978a) has pointed out, "the nature of the data - ongoing sequences of interactive behaviours - is not easily susceptible to analysis by traditional statistical techniques" (p.213). Although the study of communication does not prohibit the use of conventional inferential statistics, they are frequently inappropriate (Fisher, 1978b, p.97). The techniques appropriate for the analysis of behavioural sequences stem from the statistics used within information theory (Fisher, 1978a, p.213), a development that is closely connected with system theory (Von Bertalanffy, 1971, p.40) and include, as well as information theory (log_2-based) statistics, Markov chains, stochastic probability (Fisher, 1978b,p.97) and graph theoretic methods of analysis (Von Bertalanffy, 1971, p.19).

In discussing the major aspects of a system approach in the Introduction, communication was said to be characterised by process, relation and structure. Process and relation refer here to communication occurring over time, between individuals, and subject to different degrees of variation. Any method of statistical analysis used must, therefore, be capable of preserving the process, relational and structural aspects of communication.

1.3.1 Analysing communication as a process

Unlike a mechanistic view of process, people are not inanimate objects governed fully by natural laws and they do not behave with a one-to-one predictive regularity (Penman, 1980, p.44). While behaviour may be seen as constrained by both external and internal factors, it is not fully determined by these. Thus, the process of communication needs to be seen as stochastic, in the sense that analysis is by assessment of the probabilities of occurrence of a subsequent state from an antecendent state. The important point to note here, is that from a deterministic standpoint, knowledge of the antecedent event allows one to predict the identity of the subsequent state, the if-then statement of conditionality being a linear or quasi-dependent statement. However, a stochastic process implies no such linear relationship between antecedent and subsequent states. A stochastic process simply determines the probability that a subsequent state will follow a specified antecedent state, and states implicitly that, given a reasonable history of past interaction, one can expect one specified event to follow another specified state on a certain percentage of occasions (Fisher, 1978a).

1.3.2 Preserving the relationship in communication

The relationship expressed in communication is preserved in a stochastic analysis by recognising and specifying the individual generating the elements of conversation, rather than conversation being considered as the total output of an interacting group whose members remain unidentified in the analysis. The relationship is preserved in the stochastic analyses by treating the dyad as a pair of interacting individuals, rather than regarding it as a single unit of analysis (see Kraemer & Jacklin, 1979). Process and relation are therefore accommodated in a stochastic process analysis by preserving the relationship between individuals and the time-ordering of events. A full treatment of the application of stochastic process analysis to communication will be given in Chapters 3 and 4.

1.3.3 Analysis of the structure of communication through constraints and rules

The third characteristic of communication, structure, was considered in the Introduction in terms of both constraints and rules. Constraints are said to operate in a system whenever variables occur in a non-random manner, and thus, the presence of any organisation of variables, by definition, indicates the operation of system constraints (Ashby, 1968). The concept of information, in Shannon and Weaver's (1949) terms, is based on the notion of entropy, or disorder. Hence, the theory of information can be used to measure the degree of organisation of events in a system. A full treatment of the application of information theory, to two types of organisation in communication systems, will be given in Chapters 4 and 9.

Finally, organisation of events can also be viewed in terms of rules (Berger, 1977; Penman, 1980). As

Von Bertalanffy (1971) has noted, many system problems concern structural or topological properties of systems. Graph theory, and especially the theory of directed graphs, is useful as it can "elaborate relational structures by representing them in topological space" (p.19), as it is connected mathematically with matrix algebra, and therefore shares a common link with the theory of stochastic probability. A complete treatment of rule-extraction in interpersonal communication using directed graphs will be given in Chapter 6.

CHAPTER 2

CONVERSATIONAL EXCHANGE ANALYSIS†

2.1 Introduction

In Chapter 1, it was argued that events in communication are concatenous, and therefore by definition, there can be no 'natural units' that can be used to describe the communication process. Consequently, in order to investigate interpersonal communication one has to impose some form of unit onto the conversational sequences. It is therefore the aim of this chapter to develop a systematic method of segmenting conversational speech, as well as a content typology by which the units may subsequently be classified.

Before describing the system development, two points need to be made. Firstly, in conceptualising the process of communication in a dyadic system, there are at least three alternative perspectives that can be used; that of each of the participants, and that of an independent observer of the system. Laing and Cooper (1971) have indicated that for the unity of an interacting dyad to be realised, the process must be viewed from the outside, taking an observer's perspective.

---

† The contents of this chapter appear in an abbreviated form in the forthcoming article, 'Conversational Exchange Analysis', by Thomas, A.P., Bull, P.E. and Roger, D., in the Journal of Language and Social Psychology.

Secondly, some systems of observation require the observer to classify behaviour in terms of its impact on the participants (e.g. Bales, 1950), or to assess the intention behind the communication (e.g. Penman, 1980). However, as Dore (1979) has pointed out, an observer can never be certain as to exactly what motivates a person's choice of utterance, or what they intend by use of an utterance. Consequently, Conversational Exchange Analysis (CEA) follows Longabaugh (1963) and Morley and Stephenson (1977) by coding utterances "on the basis of the meaning of the act for the ... relationship as a relationship" (Longabaugh, 1963, p.324), using the utterance form, as well as its role or function at that point in the conversation, to classify the utterance. The segmentation and classification of speech will be more fully covered in section 2.4, 'Procedures for dividing conversation into acts'.

The development of a new classification system was motivated by the shortcomings apparent in existing systems. While there are a number of category systems available for the description of verbal communication, most are inappropriate for informal conversation because these systems have been designed for use only in particular situations (e.g. problem-solving, Bales, 1950; classroom interaction, Pride, 1969; Flanders, 1970; Sinclair & Coulthard, 1975; bargaining and negotiation, Morley and Stephenson, 1977; chronic schizophrenic patients, Longabaugh et al., 1966; marital interaction, Gottman et al, 1977a; Penman, 1980; interpersonal relationships, Borgatta & Crowther, 1965; Mark, 1971; Rogers & Farace, 1975). Of those systems applicable to informal conversation (e.g. Bales, 1950; Crowell & Scheidel, 1961; Scheidel & Crowell, 1964; Danziger & Greenglass, 1970; Wilson, 1974), deficiencies in the system framework or the lack of implementation rules reduces their usefulness. For example, in both Bales' (1950) and Scheidel and Crowell's (1964) systems, different levels of analysis are confounded. Bales' IPA allows one to distinguish between the types of information that can be asked for and given, but does not allow the distinction between the types of information accepted and rejected. Similarly, Scheidel and Crowell distinguish between the type of information given and received, but not the type of information modified.

Although often concerned with the analysis of quite different situations, speech categorisation systems generally fall into one of two classes. The first is concerned with the analysis of task-related conversations, such as how a problem is solved (e.g. Bales, 1950), or the process of bargaining and negotiation (e.g. Morley & Stephenson, 1977); the second is concerned with the relationship between individuals that is expressed in the conversation (e.g. Borgatta & Crowther, 1965; Mark, 1971; Rogers & Farace, 1975; Gottman et al., 1977a; Penman, 1980). The present system, although generalisable to different types of relationships, is of the former type, being concerned with the process of conversation, such as how information is made salient in the conversation and the types of information exchanged (e.g. beliefs, telling stories, giving examples, commenting on the conversation, etc.).

2.2 System requirements

The requirements of CEA, like any method of content analysis, are threefold; the system should be objective, systematic, and have generality (Holsti, 1969). Objectivity requires that each step in the classification process is carried out on the basis of explicit rules. Decisions made are then guided by this set of rules, thus minimising the intrusion of the analyst's subjective predispositions. Similarly, to be systematic, the inclusion or exclusion of categories should be done according to consistently applied rules. Finally, purely descriptive information about communication, unrelated to other attributes is of little value. A classification system requires therefore that the categories have theoretical relevance, or generality. For example, categorising conversational sequences in terms of beliefs, narratives, examples, conclusions, etc., enhances our understanding of conversational structure considerably more than if one simply knows that the interactants spoke about inflation, and last week's football match. Additionally, the CEA system itself has been designed to have generality, being applicable to the analysis of conversation that takes place in a wide variety of situations and reflecting a wide variety of relationships.

2.3 System framework

In developing a set of descriptors by which speech may be subsequently classified, there is always an implicit or explicit decision as to the size of the unit of communication to be categorised (Guetzkow, 1950). Making the decisions explicit through a set of rules ensures objectivity and enhances system reliability. This suggests, as does a study of the literature, that any category system must contain two distinct sets of rules, one defining the unit of communication and the other defining the description to be used (Bales, 1950; Mishler & Waxler, 1968; Wilson, 1974; Morley & Stephenson, 1977). However, as Morley and Stephenson point out, the first set of rules is rarely specified, or available in any detail (e.g. Longabaugh et al., 1966; Wilson, 1974; Gottman et al., 1977a; Penman, 1980), and the second is often only available in summary form, making the published systems very difficult to use. Consequently, the next sections of the chapter are concerned with providing a set of rules for both the division of speech into units, and its subsequent content classification.

2.4 Procedures for dividing conversation into units

There are a number of options available for the division of conversational speech into units. For example, the speech can be defined in terms of an uninterrupted verbal intervention (e.g. Norwine & Murphy, 1938; Jaffe & Feldstein, 1970; Matarrazzo & Wiens, 1972;

Sluzki & Beavin, 1965, reported in Rogers & Farace, 1975), a single classification category (e.g. Langabaugh et al., 1966), or a single thought or idea (e.g. Bales, 1950; Morley & Stephenson, 1977). In the present system, CEA, the unit of analysis selected was a single thought or idea, following Bales (1950) and Morley and Stephenson (1977). This is because a verbal intervention could contain a number of smaller units, each of which could equally well represent a single thought or idea. In such cases, the unit becomes extremely difficult to code as only a single code can be assigned to a number of different thoughts or ideas (Brinich, 1981). A unit based on the categories to be used was rejected simply because it was required that the process of division and classification be kept separate.

The minimal unit of speech conveying a single thought or idea has been variously defined by a number of authors. For example, Wilson (1974) suggests a passage of speech with specific function, Longabaugh et al., (1966) a smallest bit of action, Holsti (1969) a theme, or single assertion about a subject, Penman (1980) a connected flow of behaviour with a single intent of illocutionary form, and Bales (1950) and Fries (1952) a simple sentence. However, in each case, the method of extracting a single theme, or isolating a simple sentence in conversational speech is left unspecified.

Although there have been a number of methods suggested in the literature for defining a sentence in conversational speech, the majority are not entirely satisfactory. For example, the methods suggested by Auld and White (1956) and Penman (1980) define too large a unit. The passage, "He took down a book that he had uh in the bookcase. And uh it was an old book, I remember - a red, red cover" is coded as a single unit by Auld and White (1956, p.277), but would comprise three units in CEA; "He took down a book that he had uh in the bookcase/ And uh it was an old book/ I remember - a red, red cover"/. Similarly, "Perhaps that is what happens, i never though of that", coded as one act by Penman (1980, p.140), would be coded as two acts by CEA; "Perhaps that is what happens/ i never thought of that"/. By segmenting speech more finely, CEA gives a more accurate representation of the individual ideas in conversation.

The methods suggested by Mishler and Waxler (1968) and Danziger and Greenglass (1970) however, define too small a unit. "I think/ (that) they're old enough to have a party" (Mishler & Waxler, 1968), and "she sits/ and talks to him" (Danziger & Greenglass, 1970, p.10) are each two units, but would be coded as single units by CEA. In these examples, the divisions are too fine to extract single thoughts and ideas.

Gottman et al. (1977a) use a purely grammatically-based method in which "a thought unit is usually grammatically separated by commas, ands, buts, and periods" (p.2). However, speech rarely conforms to the simple structure that this would imply. Ordinary conversation is not necessarily grammatically correct

(Brinich, 1981), and in the case of Gottman et al.'s system, CISS, no guidance is given for isolating single thought units.

A psychological unit of speech (after Morley & Stephenson, 1977) conveying a single unit of meaning in the form of a single specific thought or idea is adopted by CEA, the rules for the division of speech into single units being based on the work of Morley and Stephenson (1977), and influenced by Auld and White (1956), Mishler and Waxler (1968) and Danziger and Greenglass (1970). While Morley and Stephenson's rules specify a more useful unit than any other system, a number of their rules, especially those concerning compound sentences have been revised. Compound sentences (e.g. I'll go to the pub with you if you'll buy me a drink) are extremely difficult to divide (Morley & Stephenson, 1977). In the case of conditional clauses, while Morley and Stephenson score them as one act if the conditional clause precedes the main clause and two acts if it follows the main clause, CEA codes all conditional clauses as single units. This is because, for both types of conditional statement, there are two items of information present (going to the pub, and the condition of buying the drink), both of which are necessary components of the same single idea. For example, whereas the following would be scored as two acts by Morley and Stephenson, CEA would score it as a single act; "Tell us next week/, when you've had a get-together"/ (p.198).

The advantage of adopting a psychological unit of speech, where a unit is defined as an idea or thought, is twofold. Firstly, in terms of speech act division, the unit defined is not an arbitrary unit like temporal or whole utterance units, and, having a grammatical basis backed by a set of rules makes the process of act division an easier and more reliable task. For example, Halliday (1970) has noted that connected speech often takes the form of an unbroken series of pitch contours, or tone groups, each tone group representing what the speaker decides to make into a unit of information, and which often corresponds to a clause. However, it is necessary that one makes the distinction between a speech unit that conveys a psychological unit of communication in the form of a single thought and an item of information. For example, I saw John yesterday//, where // represents a tone group, is a single tone group and a single communicative act. However, I saw John// yesterday//, is two tone groups, but only a single communicative act, because the second tone group, 'yesterday' does not make a point in its own right, but only in relation to what has gone before. Similarly, as McNeill (1979) has noted, natural units of speech may correspond to clauses or sentences, although this is not always the case. Such observations nevertheless lend support to a quasi-grammatical basis for the division of conversational speech into units.

Secondly, in terms of speech unit categorisation, as Dore (1979) has indicated, an observer can never be certain as to exactly what motivates a person's choice of utterance, or what they intend by use of an utterance (c.f. Penman, 1980, who uses the notion of intent in his definition of a communicative act). The definition used by CEA allows speech to be categorised as having an illocutionary value based on the utterance form, thereby obviating the need to infer intention. As the structure of language is related to its function (Peck, 1981), a communicative act may therefore be classified with respect to the conversation itself (the utterance form) and the continuing activity within the conversation (its role or function at that point in the conversation).

The complete CEA coding scheme provides rules for identifying independent and dependent clauses, separating linked clauses, and the analysis of incomplete clauses, affirmations, negations and clarifications, nested clauses, relative, conditional and opposition clauses, interruptions, false starts and repetitions.

It is assumed that both tapes and transcripts are used for coding, Waxler and Mishler (1966) and Morley and Stephenson (1977) having shown that this combination is the most reliable.

The following set of rules is used for the division of conversational speech into units representing single thoughts and ideas.

2.4.1 Acts: Transcript division rules

Although an act has been defined as a single thought or idea, it is often helpful to have a system of rules to aid the division of the transcript into single communicative acts. Even though these rules are grammatically based, it is often helpful to think of an independent clause as containing a single thought or idea, and in some cases, an action.

Throughout the system of rules, single thoughts and ideas will be bounded by a back-slash (/).

a) Independent clauses

An independent clause is an utterance containing a subject and a predicate (Quirk & Greenbaum, 1976), although neither have to be stated explicitly, and may include complements and modifiers. For example:

Subject	the GIRL is a clever student/
Predicate	the girl is a CLEVER STUDENT/
Complement	the girl is a clever STUDENT/
Modifier	the girl is a CLEVER student/

An act consists of a single independent clause with or without one or more dependent clauses. Acts may, therefore, take an SVX form where:

S = Subject

V = Verb

X = Object, dependent clause, etc.

b) Linked clauses

Clauses may be linked. An independent clause may be distinguished from a dependent clause in that when two independent clauses are linked, the second may be introduced by a co-ordinating conjunction, of which there are three: and, or, and but (Quirk & Greenbaum, 1976). Dependent clauses however, may be introduced by subordinating conjunctions and pronouns, such as who, what, that, because, so that, if, and when, etc. (Auld & White, 1956). For example:

He came to collect me/ and we went to the disco/
I agree with you/ but I don't think it's true/
You can go if you like/ but I don't want you to/
They arrived yesterday so that they could go to Ripon today /
I'm going to town in an hour if it doesn't rain/
I'll keep on until you agree/
I stayed in so that my sister could go out/

Frequently, dependent clauses may have some part of the SVX formula replaced, for example by the word 'which', as in the following examples:

I'm going to the cinema tonight which has just opened after the flooding/
She was reading the book which was red/

It should be noted that 'but' may sometimes also be used as a subordinating conjunction introducing a dependent clause. For example:

They preach about revolution but don't get very far/
You said you'd see me but you didn't/

The guidelines for separating linked clauses must be used carefully, and it should always be borne in mind that the conversational speech is being segmented into units representing single thoughts. Lists, for example, although separated by 'and' and 'or' are not scored as independent clauses. For example:

> Being on the dole is really unpleasant/ there's the problem of the rent to be paid and food to buy and the bills to be paid as well as the feeling of inferiority you get/

There are occasions when the notion of action is more appropriate than an idea. For example:

> There's a hard core group who really know what they are doing and really getting somewhere/
>
> She sits and talks to him/

Here, in the second example, the thought expressed is concerned with a single action of sitting and talking. However, in the following example, two actions are apparent, going and staying. Hence the utterance is scored as two units of speech;

> Shall we go to the pub/ or shall we stay at home?/

Relative, conditional and opposition clauses are all considered to be dependent clauses by CEA. Each type of clause is shown in the following examples:

> How often do you come across fairy tales which are cruel and frightening?/
>
> You accept cruelty in fairy tales because the way they are written is so other-worldly/
>
> Shall we go to the disco rather than the pub?/

c) Incomplete clauses

In conversational speech, independent clauses are frequently incomplete (i.e. they are not replaced by words such as 'which', etc.), but may still be scored, provided the meaning is clear from the context. There are four main types of incompletion:

i) Ellipsis. (Incomplete subject and incomplete predicate).

Have you got a job?/ Not yet/

ii) Incomplete subject and complete predicate.

What are you doing?/ Singing along/

iii) Complete subject and incomplete predicate.

Do you want to start?/ Shall I?/

iv) Some incomplete clauses are given context by what Mishler and Waxler (1968) call "generally accepted cultural meaning" (p.343).

Good/ Right/ Okay/ Yes/ etc.

d) Affirmation, Negation and Confirmation

i) Affirmations and negations are both counted as single acts if the words following do not amplify or explain the preceding word. For example:

Yes/

No/

That's right/

Yes that's right/

ii) If however, the affirmation or negation stand alone and the words following do not amplify or explain the preceding word, it is counted as a separate act. For example:

I disagree with that/ yes/

Yes/ I agree/

No/ that's the wrong one/

A: You don't like that one?/

B: No/ which one did you choose?/

A: Do you agree?/

B: Yes/ shall we change the topic?/

iii) Tag questions, or confirmations, are not scored as separate acts:

It's your turn <u>isn't it</u>?/

You like chips <u>don't you</u>?/

... it's a question of morals <u>i suppose</u>/

e) Nested clauses

Frequently, one independent clause may be 'nested' or 'embedded' within another independent clause. In this case, each independent clause is scored as a separate act, and the nested clause is marked in the following form; /* text */. For example:

She wasn't put down by her/* what was it? */ ugly sisters and horrible step mother/

f) Interruptions

When a speaker is interrupted, the incomplete point is scored as a single act in the usual manner, but is placed in parenthesis if the speaker has not said enough to make the meaning clear. For example:

A: (I think that ...)

B: but you're wrong/

If the original speaker should continue, only one act is scored for speaker A.

A: I think that ...

B: but you're wrong/

A: ... you know I'm right really/

g) False starts

A false start, or a group of words which do not convey the speaker's meaning is not scored as an act. For example:

I think ... I think I'm right/

I feel ... no you're quite wrong/

Having divided the transcripts into the basic units, each unit is then coded according to its content. The following sections are therefore concerned with the development of the content classification system.

2.5 Act classification

Classifying each act is the process by which the raw data are systematically transformed into units which describe the relevant information content, the rules by which this is accomplished serving as an operational link between the data and any subsequent theory. Coding rules are therefore a central part of any conversational research design; if the classification is ill-founded then any subsequent analysis may be meaningless.

CEA has been designed to meet all the requirements of a content analysis system (Holsti, 1969; Hawes, 1972; Clarke, 1977). It has been designed such that all conceptually different levels of analysis are kept separate (c.f. Longabaugh, 1963; Rogers & Farace, 1975), each level being exhaustive, and each category being mutually exclusive and independent (c.f. Donohue et al., 1981). (Categories are independent in the sense that coding one unit of speech does not predispose the following unit to be coded in any specific way. For example, a Question is not necessarily followed by a Reply.) The CEA system of categories is also intended to be as comprehensive as possible, covering, for example, all the descriptions of language functions (e.g. performatives, regulation, expressing affect, inquiry, metalanguage, et.) identified by Robinson (1972, pp.57-79).

Although a number of systems are presently available for the analysis of informal conversation, these were all found to be inappropriate. Bales' IPA, for example, confounds different levels of analysis by disguising the function of information exchanged with the way it is exchanged. IPA allows the coder to distinguish between the type of information that can be asked for and given, but does not distinguish between the type of information accepted and rejected. A similar objection can be made of the systems described by Danziger and Greenglass (1970) and Rogers and Farace (1975).

Additionally, because IPA is intended to apply generally to the total group interaction process and as each item is categorised in relative isolation, the categories are too general for application to informal conversation. For example, both the introduction of an idea, the synthesis of ideas and drawing a conclusion would all be classified as 'gives orientation', yet the acts serve quite different functions in the conversation.

By contrast, Rogers and Farace (1975), Crowell and Scheidel (1961) and Scheidel and Crowell (1964) define categories that are too small for reliable use. For example, Crowell and Scheidel (1961) suggest categories for modification of an idea that are 'small' and 'large', a distinction that is very difficult and unreliable to make in practice.

The most influential feature of the speech systems reviewed was the use of a number of conceptually different levels of content analysis. This was incorporated into CEA by coding conversational speech on three distinct conceptual levels. These are (1) Activity, which refers to how information is made salient in the interaction, such as, is the information asked for, or given, agreed with or rejected, etc., (2) Type, which refers to the sort of information exchanged, such as beliefs, past experiences, conclusions, comments on the conversation, etc., and (3) Focus, which refers to the referent of the information. For example, one may be referring to one's own opinions, or the opinions of a third party. The construction of these levels of analysis draws extensively from the work of Bales (1950), Crowell and Scheidel (1961), Longabaugh et al. (1966), Danziger and Greenglass (1970), Wilson (1974), Sinclair and Coulthard (1975), and Morley and Stephenson (1977).

2.5.1 Activity classification

There are three major classes of verbal interactive act which make information salient in conversation: the statement, question and mand (Lyons, 1977). Mands refer here to commands, demands, entreaties, etc., and form a sub-set of what might be called directives (Searle, 1976). As the three major classes of act are usually realised by the declarative, interrogative and jussive grammatical forms, respectively (Lyons, 1977), it would seem attractive to classify the salience of speech according to its grammatical form.

However, there is a theoretical problem here (Sinclair & Coulthard, 1975; Myers, 1979). In the literal performance of a mand, a speaker might produce a sentence of the jussive form - shut the door, for example. A problem arises when the illocutionary act is performed non-literally, as in the utterance, Can you shut the door? In this example, the indirect illocutionary act is a mand that has been softened by the interrogative grammatical form, and conveys the illocutionary force of a request. The problem is to decide which illocutionary force is appropriate, if not the one corresponding to the grammatical form.

The lack of fit apparent here between grammar and discourse can be dealt with by what Sinclair and Coulthard call the 'situation', where this refers to the environment, social conventions, shared knowledge of the participants, etc. The three rules specified by Sinclair and Coulthard (1975) may be used to decide when a declarative or interrogative form realise something other than a statement or question. These rules essentially state that if the action referred to in the utterance is physically possible at the time of the utterance, then the utterance is a Directive (Mand). The three rules are as follows:

A) An interrogative form realises a command to do if:

   i) the act contains one of the modal verbs, can, could, will, would, going to, AND

   ii) the subject of the act is also the addressee, AND

   iii) the predicate (verb phrase) describes something which is physically possible at the time of the utterance.

   e.g. A teacher in a classroom may say:

   a) Can you write, Bill/ command

   b) Can Bill write?/ request

   c) Can you swim, Bill?/ request

The first example (a) is a command because it fulfills all the conditions, provided writing materials are available. The second (b) is a request because the subject and addressee are not the same. The third (c) is also a request, assuming that a swimming pool is not available..

B) An interrogative or declarative form realise a command to stop if the act refers to an action or activity which is required at the time.

e.g. A: What are you laughing at?/

B: Nothing/

In this case, the example is a command to stop laughing, if laughing is prohibited at the time, but a request if it is not.

C) An interrogative or declarative form realise a command to do if the act refers to an action which either both, or one party involved, know ought to have been performed and has not been.

e.g. a) Is the book on the table?/ command

b) Is the book on the table?/ request

The first example (a) is a command if all the participants know the answer, a request (example b) if the questioner does not know the answer.

As noted earlier, the Activity of an act is therefore determined not only by its grammatical form, but also the situation, or its role at that point in the conversation.

The Activity dimension of CEA is an extended and much revised form of the speech system presented in Bull and Brown (1977) and used by Thomas and Bull (1981). It is based on Informatives, Elicitations, and Directives, which are the discourse equivalents of Statements, Questions and Mands (Sinclair & Coulthard, 1977). These three categories have subsequently been extended, the work of Bales (1950), Longabaugh et al. (1966), Wilson (1974), and Morley and Stephenson (1977) all being influential. A summary of the Activity dimension may be found in Table 1. (All examples used are taken from genuine conversations.) The Offer, Reply, Consent, Dissent, Modify, Quotes, Insult and Reaction categories are all specialised forms of Informative acts. The Modify category is a specialised form of Consent and Dissent category, when neither are particularly appropriate. For example, 'Fairy tales are cruel but not frightening/', is most appropriately coded as a Modification. The Request category corresponds to Elicitations and the Mand and Threat categories are specialised forms of Directives.

Table 1 Summary of CEA Activity categories

1) Offer refers to speech which initiates conversation by introducing information, opinions, etc. Rhetorical questions and verbal acts made in response to requests which make no attempt to answer the request are also scored as Offers. e.g.

a) I think there's certainly a use for the arts in today's world/

b) I can't stand Maggie and her policies/

2) Reply — refers to speech that is made by the other person in direct response to a Request, where the information asked for is given. If this is not the case, then another category is used.

e.g.

a) A: Oh what do you do?/
   B: I'm an english student/

b) A: Do you fancy a pint at lunch time?/
   B: Mmmm great/

3) Consent — refers to acceptance, agreement, and positive evaluation by which a person specifically endorses what the other person has said.

e.g.

a) Yes, I think you're right/

b) A: That's too difficult/
   B: It is really/

4) Dissent — refers to rejection, disagreement, or negative evaluation by which a person contradicts what the other person has said.

e.g.

a) A: You'll never agree will you?/
   B: Yes I will/

b) A: ... but that implies inferior service/

B: Ah I don't think so/

5) Modify refers to acts that revise a prior idea by affirmation of some essential feature, but suggest a change in some other feature.

e.g.

a) Yes, fairy tales are cruel, but not frightening/

b) A: The student union is always wrong/

B: Well, sometimes/

6) Phatic refers to acts that are similar to Offers in that it refers to speech that initiates conversation, but it does so by introducing information that is conventional and ritualised.

e.g.

a) Hello/

b) How are you?/

7) Quotes refer to acts that are similar to Offers, in that they initiate conversation, but do so by introducing information that is already provided, perhaps by stimulus materials, instructions, etc.

e.g.

a) Fairy tales are cruel and frightening and should be censored/ (Quoting from questionnaire).

b) It says that man can only sin if God exists/

8) Insults refer to acts that are verbally abusive.

e.g.

a) Why don't you push off?/

b) You rat/

9) Reaction refers to vocalisations made by the listener as reactions to what the speaker is saying, and are realised by listener responses (Dittmann & Llewellyn, 1967) and unsuccessful interruptions ("butting-in interruptions", Ferguson, 1977).

e.g.

a) A: ... the government policies and everything/

B: mm/

b) A: ... housing shortage and small council grants/

B: yes/

10) Request refers to speech in which a person is actively asking for information, and is always realised by an interrogative sentence form. To make a request of someone is both to pose the question and in so doing, to give some indication to the addressee that a response is expected by answering the question that is posed. However, one can pose a question which not only does one expect to remain

unanswered, but which one expects to be unanswerable. Such rhetorical questions are like strong statements (Quirk & Greenbaum, 1976), and are therefore usually classified as Offers. In the case of a person replying to a request with an information-seeking request (Arching; Mishler, 1975), the reply would be scored as a Request.

e.g.

A: Are you coming tonight?/

B: Do you think our grants are too small?/

11) Mand refers to speech that has the function of asking for an action to take place and is, therefore, an attempt by the speaker to get the listener to do something (Lyons, 1977, p.130). A Mand expects that some action will take place, and may be realised by the interrogative and imperative grammatical forms.

e.g.

a) I'll expect you tonight/

b) Hold on/

12) Threat is a specialised form of Mand, by which the speaker commits him/herself to some course of action, conditional on the listener doing, or not doing a specified action.

a) If you don't come I'll be really upset/

b) Hold that or I'll spill my tea on you/

2.5.2 Type classification

The Type level of analysis has been designed to extract the sort of information that is exchanged in conversation. The main problem in creating a typology is in maintaining distinctions between categories (Clarke, 1977). Lazarfield (1937) and Guetzkow (1950) both suggest the best method of category design is to begin with large, general categories, which are subsequently re-worked as more reliable and meaningful distinctions are made between smaller, more specific categories. Speech was therefore initially classified into three main types, ideational-informative, social-emotional, and interactive-regulative acts (Myers, 1979). These were subsequently reworked into smaller categories, the Type classifications for CEA being derived from a review of the relevant literature (e.g. Bales, 1950; Crowell & Scheidel, 1961; Mishler & Waxler, 1968; Wilson, 1974; Fishbein & Ajzen, 1975; Penman, 1980), the list of language functions identified by Robinson (1972, pp.57-79), and the analysis of five half-hour informal conversations between pairs of female University students. A summary of the Type categories may be found in Table 2.

Ideational-informative acts are those that exchange ideas and information in the conversation, of which Type categories 1 - 8 broadly represent the information exchanged, and Type categories 9 - 18 represent certain attributes of the conversational process. The second major classification is into social-emotional acts, which may be thought of as regulating interpersonal relationships (categories 19 and 20). The final division into interactive-regulative acts is concerned with the management of the conversation, and is divided into those acts made by the person who has the speaking turn (categories 21 - 24) and those made by the listener (categories 25 - 28).

The typology presented is intended to be as comprehensive as possible without being too complex to use. Inspection of the complete type classification system (Table 2) indeed suggests a high face validity; they seem to represent the processes of which informal conversations are made. This level of analysis has been designed to be highly flexible so that the number of categories may be expanded, or contracted. The system is not therefore research-specific and may be easily applied to the analysis of different situations, relationships, and research aims.

<u>Table 2</u> Summary of CEA Type categories

1) Beliefs/ Information	commit the speaker to 'something being the case', and represents the information or commitment one has to something. A belief may be said

to link an object to some attribute, and although two types of belief can be evidenced - descriptive and inferential (Fishbein & Ajzen, 1975, p.132), they are not distinguished because the distinction is somewhat arbitrary. For example, many attributes of an object that appear to be direct observations cannot be directly perceived. Attributes such as happy, dark, sour, etc., are themselves concepts that have been acquired in the past.

e.g.

a) I believe in fairy tales/

b) I think that life imprisonment sentences are more of a punishment than death/

Note: Two sub-sets of beliefs are distinguished in CEA, those relating to one's beliefs about what the other person has said (Active Recognition) and those concerned with details about oneself, etc. (Personal details).

2) Personal details — are acts that refer to oneself, the other, or real people known to the speakers. Such references may be

to a person's age, name, occupation, etc.

e.g.

a) I'm not politically active at the moment/

b) Yes, I'm a history student too/

3) Intentions

refer to proposals for future action of others and/or self. Intentions are the subjective probability that a person will perform some action, and have the effect of committing the person to some future course of action.

e.g.

a) I'll go and see him tonight/

b) I'll take you to the party tomorrow/

4) Suggestions

refer to proposals for future action, in the same way as Intentions, but come about as a result of the task in hand.

e.g.

a) A: Well i've always fancied the idea of rowing/

B: Why don't you come down with me tomorrow?/

A: Okay, I'll do that/

5) Personal experience	refers to past and present experiences of oneself and one's partner in the conversation. e.g. a) Did you see Waiting for Godot at the Drama Barn last week?/ b) I went to the Royal Ballet at Christmas/
6) Narrative	refers to story-telling and descriptions of past actions and experiences of others. e.g. a) You know Sheila, she became a Christian last year sometime/ b) ... and they blockaded the streets in the Paris commune as well/
7) Subjective feelings	are concerned with descriptions of physical and psychological state of the conversants and people known to the conversants. e.g. a) I feel ill/ b) My girlfriend's not very well at the moment/
8) Attitude	refers to speech which contains an explicit affective, or evaluative, component and represents a person's general feeling of favourableness

or unfavourableness towards some object (Fishbein & Ajzen, 1975, p.216).

e.g.

a) I <u>like</u> fairy tales/

b) I've always <u>loved</u> the opera/

9) Justification/ Explanation — refers to the explicit offering and asking for proof to explain or substantiate a point.

e.g.

a) What does psychosomatic mean?/

b) Well, how would that help you?/

c) A: I wouldn't say that a person who didn't stop another from committing suicide was a murderer/

B: Why not?/

A: <u>because it's not them that killed the person really</u>/

10) Synthesis — refers to acts that organise thoughts and ideas into a coherent whole, and may often be seen as a simplification of something that has gone before.

e.g.

a) What I'm saying is that social problems are really political/

b) So, what are we really saying here?/

11) Summary — refers to acts that restate the major points of the conversation, and may be seen as a summing up of points.

e.g.

a) Okay, we're agreed then that capital punishment is morally wrong/

b) Right, we all agree that abortion is wrong and parents are therefore murderers/

12) Judgement — is used when acts ask for explicit acceptance or rejection of a thought or idea, when a point has been made.

e.g.

a) Do you agree with me?/

b) A: <u>That's true don't you think</u>?/

B: I suppose so/

13) Avoidance — refers to when the speaker reserves a decision or by-passes a question - the speaker avoids commitment or shows hesitancy by being non-committal.

e.g.

a) Maybe/

b) Whatever you say/

c) I suppose so but I don't know/

d) Okay, I agree to disagree/

14) Negation/ Submission	are acts where the speaker defers to the other, and in doing so may negate one's own views. e.g. a) Okay, I give in/ b) Perhaps you're right there/
15) Active Recognition	is used for acts that recognise a point in the discussion, or praise, congratulate, correct, etc. e.g. a) Good point/ b) Right/ c) That's not the point/
16) Salience	is used when an act sums up one's feelings. e.g. a) That brought the point really home to me after seeing him like that/ b) Well, that summed the whole thing up for me/
17) Consequence/ Conclusion	refers to acts that are a result of a previous argument or line of thought and represent the climax or culmination of a series of ideas into a single point. e.g. a) Terrorists are killing innocent people/

they dominate everybody's lives and cause such a senseless loss of life and property/ so surely it's the duty of us to stop them/

b) Well, if you believe that you can't be a Christian/

18) Exemplification refers to acts that give examples.

e.g.

a) Well, look at me, I can't write essays in a week/

b) Look at me, I can't do a mass of effort at the end of a session for an exam/

19) Emotional support refers to acts that request and give emotional support during a conversation.

e.g.

a) Don't take it to heart/

b) Please, will you help me?/

20) Apology refers to the self acceptance of a wrong-doing or a means of expressing guilt.

e.g.

a) I'm sorry/

b) Look, I really didn't mean that/

21) Direction refers to the control aspect of conversation such as comments

concerning instructions and orders, who is to speak next, etc.

e.g.

a) You start/

b) Hang on/

c) Which one interests you?/

22) Meta-statement — is used when acts describe the conversation and its progress and conduct.

e.g.

a) I like this/

b) We have rather similar views on this/

c) How are we doing?/

23) Revision — refers to self-questioning.

e.g.

a) I think that's right/ <u>but let me see, is that right</u>?/

b) Man Ray was a surrealist/ <u>Ooops</u>, <u>no he wasn't</u>/ he was a Dadaist/

24) Restatement — refers to repetitions and requests for clarification and confirmation.

e.g.

a) Pardon?/

b) What?/

c) Sorry/ What was that?/

25) Completion — refers to acts that complete what the other person is saying.

e.g.

a) A: ... and they're suddenly

B: suddenly in a wheel chair/

26) Listener Response — refers to those acts, made by the listener, which operate as signals of the person's continued attention. They usually consist of brief oral sounds (Fries, 1952). Kendon (1967), Dittmann and Llewellyn (1967) and Yngve (1970) have suggested that words such as yeah, um-hmm, I see, good, oh, etc., all operate as listener responses. Such utterances do not constitute a claim for the speaking turn (Duncan, 1972), since they appear to ensure that the speaker who is holding the floor continues to do so (Ferguson, 1977).

e.g.

a) A: Look at all the other people who lead a

A: privileged and sheltered life/ all the

B: mm/

A: aristocracy do/ and I think its not fair

B: yeah/

A: to say students are privileged like that/

27) Unsuccessful interruptions — refers to speech that is made by the LISTENER and correspond to Ferguson's (1977) butting-in interruptions, where no change in speaker occurs.

e.g.

a) A: No I don't agree/ It shouldn't be
B: But i think
A: legalised/

b) A: ... the correspondence between those
A: two isn't as close as you would like
B: well i think
A: to make out/

28) Laughter

2.5.3 Focus classification

A review of the literature and the work of Bugental (1948), Danziger and Greenglass (1970) and Morley and Stephenson (1977) in particular, suggested that a third level of analysis was required, by means of which the referent of the communication could be classified.

In CEA, the Focus dimension is divided into Object and Subject focus, and is used to indicate to whom the act refers, and from whom the description emanates, respectively. Furthermore, a distinction is drawn between personal and non-personal, or institutional referents. For example, 'I think the Gas Board is quite wrong', has a self subject focus (I), and an institution object focus (Gas Board). Focus is only scored when the speech explicitly mentions a referent. A summary of the Focus categories may be found in Table 3.

Although the focus level of analysis is not used in this research for the analysis of conversation, it is included here for the purpose of presenting the CEA system in its entirety.

Table 3 Summary of CEA Subject/Object focus categories.

1) No focus — is used when no explicit focus is mentioned in the speech unit.

e.g.

a) Is it going to be all that inferior?/

b) Yeah/

c) Continuous assessment of work means working at the same pace all the year round/

2) Self — refers to when a person explicitly mentions his/her own beliefs, attitudes, behaviours, etc.

e.g.

Subject: I agree with that/

Object: Are you saying I'm correct?/

3) Partner — is used when the speaker explicitly describes his/her partner's beliefs, past experiences, etc.

e.g.

Subject: You say that I'm wrong/

Object: I think you're quite wrong there/

4) Both — refers to when a person explicitly describes the beliefs, attitudes, etc., of both him/herself and his/her partner.

e.g.

Subject: We don't agree with this then/

Object: Oh, I thought the tea was for us/

5) Other — is used when the beliefs, behaviours, etc., of another person or institution, not involved in the conversation, are explicitly described.

e.g.

Subject: They are quite wrong/
The government is quite wrong/

Object: Well, that belongs to Sheila really/
I'll give it to them when this is over/

6) Hypothetical refers to speech that explicitly describes the beliefs, past experiences, etc., of a hypothetical person, or group, and is typically characterised by the use of the referents one, people, and they.

e.g.

Subject: One can never be free/
They don't know what's good for them/

Object: What would people think?/
They don't know what's good for them/

2.6 System reliability

An important consideration in using any classification system is to demonstrate that the categories can be consistently assigned. CEA contains four sets of rules, one set for the division of speech into acts, and three sets for the assignment of Activity, Type and Focus categories. For each of these sets of rules, a measure of reliability has been calculated, based on the independent coding of five, half-hour tape-recordings and transcripts, totalling 2943 speech units.†

---

† The author would like to thank Olga Sucharyna, Department of Language, University of York, for assisting with the reliability studies.

For the act division rules, a simple percentage error rate was used. This yielded a negligible overall rate of 3.93% inter-observer error. Three hours' training was required to achieve this degree of reliability.

Cohen (1960) has pointed out that the calculation of the proportion of agreements between observers for multi-category, nominal data (e.g. Morley & Stephenson, 1977) is inappropriate because a certain amount is to be expected by chance. Consequently, inter-observer agreement for the remaining three sets of rules for Activity, Type and Focus category assignment have been measured using Cohen's kappa (Cohen, 1960). Kappa is interpretable as the proportion of agreement between observers after the allowance of chance, the upper limit of +1 representing perfect agreement. Inter-observer agreement for the Activity, Type, Subject Focus and Object Focus rules yielded kappas of .969 ($p < .001$), .957 ($p < .001$), .994 ($p < .001$), .981 ($p < .001$), respectively. (The raw analyses may be found in Appendix A.) Four hours' training was required to reach this level of reliability. Such high inter-observer agreement demonstrates that the category system, based on sets of explicit rules, can be applied with a high degree of reliability.

2.7 Example transcript of informal conversation coded using CEA

In order to show a practical application of the speech system, the following example is an extract from a conversation between two teenage girls talking about the National Health Service, which has been transcribed and subsequently coded using CEA. Each unit of speech is marked off in the transcript by a back-slash (/), the coding for the unit being placed, in numerical form, in parenthesis after the back-slash. The coding takes the following form:

(Speaker, Activity code, Type code, Subject focus, Object focus)

the numbers referring to the codings listed in Tables 1, 2 and 3, for Activity, Type and Focus codes, respectively.

A: Pay beds are an essential part of the National Health Service/ (1,7,18,1,1) If people can afford private medical service, they ought to be able to get it/ (1,7,18,1,6)

B: um I agree with that/ (2,1,15,2,1)

A: Oh well I don't/ (1,1,15,2,1) Why?/ (1,10,9,1,1) I mean okay that's not a fair question/ (1,1,22,1,1) but don't you feel that by having a private medical service you're undermining the whole point of the NHS?/ (1,10,1,1,3)

B: I don't think so/ (2,2,1,2,1) I mean the NHS is a bit overcrowded anyway/ (2,1,9,1,5) the waiting list is incredibly long/ (2,1,9,1,1) I think the people who've got money should/ (2,1,1,2,6) I don't know how it actually works/ (2,1,1,2,5) the economics of it/ (2,1,18,1,1) because the way I thought about it was if you've got the money why not spend it to get private treatment from a surgeon or whatever or whoever deals privately rather than crowd out the NHS/ (2,1,9,2,1)

A: But on the other hand there's the surgeon who operates privately in the NHS/ (1,4, 18,1,5)

B: Could he?/ (2,10,1,1,5)

A: Yes he could/ (1,2,1,1,5)

B: Are you sure?/ (2,10,15,1,3)

A: Yes I am sure/ (1,2,1,2,1) don't you believe me?/ (1,10,1,3,2)

B: Oh I'm sorry/ (2,1,20,2,1) yes I do/ (2,2,15,2,1) I just didn't think that could happen/ (2,1,9,2,1)

A: Well take it from first principles/ (1,1,9,1,1)

B: mm/ (2,9,26,1,1)

A: if you had a sort of set up they were aiming for in 1945 or whatever/ (1,1,18,6,6) you know you then get fair treatment for everybody/ (1,1,17,6,6) It assumes basic humanistic principles/ (1,1,1,1,1) that's how I see it anyway/ (1,1,16,1,2) people have a right to live and to decent health/ (1,1,10,1,6)

B: mm/ (2,9,26,1,1)

A: Money is purely incidental to that/ (1,1,1,1,1)

B: Yes but why should my dad say have to wait for an important operation if he can pay and jump the queue?/ (2,10,9,1,5)

A: Pay and jump the queue/ (1,1,25,1,1) that's just the point/ (1,1,15,1,1) what about people who can't pay?/ (1,10,9,1,6)

B: I don't know/ (2,2,13,2,1)

A: If you've got the money to pay then the NHS fails/ (1,1,1,1,5) and lets face it it does fail/ (1,1,1,1,5)

B: It does/ (2,3,15,1,5) yeah/ (2,3,15,1,1) ideally it shouldn't/ (2,1,1,1,5) it does though/ (2,3,15,1,5) I don't see a reason why a person shouldn't use that money though/ (2,10,9,2,6) (sneezes)

A: Bless you/ (1,6,15,1,3)

B: Oh thanks/ (2,6,15,1,1) I've got a bit of a cold I think/ (2,1,7,2,2)

A: Another reason would be that you're taking resources out of the NHS by employing people outside it who could be employed within it/ (1,1,9,1,1)

B: I think most people/ well some people/ (2,1,23,1,1) choose to be employed outside because the pay is better/ (2,1,1,2,6)

A: Yeah/ (1,3,15,1,1) How unfortunate/ (1,1,16,1,1)

B: I don't think those people would move into the NHS if they didn't get enough private patients/ (2,1,1,2,6) they'll always get enough private patients/ (2,1,1,1,6) therefore the question doesn't really come up/ (2,1,17,1,1)

A: Well that's nonsense/ (1,4,15,1,1) if they couldn't operate outside it they would have to operate in the NHS/ (1,1,9,1,6) they would help the overcrowding then/ (1,1,17,1,6)

B: I think they'd prefer to go abroad and get more money/ (2,1,1,2,6)

A: You really think so?/ (1,10,12,1,3)

B: mm/ (2,3,15,1,1) there are an awful lot of surgeons that do go abroad/ (2,1,18,1,5) they earn a lot of money abroad rather than

A: Well how well how/ (1,9,27,1,1)

B: earning a pittance here in the NHS/ (2,1,9,1,5)

A: Well how much does a surgeon get paid by the NHS?/ (1,10,1,1,5)

B: About £20,000/ (2,2,1,1,1)

A: You've got to be joking/ (1,1,16,1,3)

B: You can get a lot more abroad/ (2,1,1,1,6)

A: £20,000 (1,1,16,1,1) my dad works for British Steel and gets about £6,000/ (1,1,2,1,5)

B: Well

he's not skilled I suppose/ (2,1,9,1,5) where do you come from anyway?/ (2,10,2,1,3)

A: Scunthorpe/ (1,2,2,1,1)

B: Oh where's that?/ (2,10,1,1,1) laughs/ (2,9,28,1,1)

A: laughs/ (1,9,28,1,1)

B: I think anyway that because some people are paying for their medical service the waiting list is shorter/ (2,1,1,2,6)

A: Oh rubbish/ (1,4,15,1,1) the waiting list is longer because of private patients/ (1,4,9,1,1)

B: But they're still getting the same sort of service/ (2,1,1,1,6)

A: Well presumably they're not/ (1,4,1,1,6) because they are paying they expect much more in the way of nursing facilities/ (1,1,9,1,6)

B: There's no waiting for major operations/ (2,1,1,1,1)

A: I mean you have to have a ratio of patients to nurses/ (1,1,1,6,1) in any private set-up it's a lot higher than in the NHS/ (1,1,1,1,1)

B: mm/ (2,9,26,1,1)

A: For them it's a luxury to be ill/ (1,1,1,1,6)

B: mm/ (2,4,26,1,1) I still think they're entitled to choose what they do with their money/ (2,1,1,2,6)

A: Yes I agree/ (1,3,15,2,1) I just don't think the system should exist like that/ (1,1,1,2,1) a two-tier system shouldn't exist/ (1,1,18,1,1) have you ever been in hospital?/ (1,10,2,1,3) I have/ (1,1,5,1,2) I had my appendix out when I was at primary school/ (1,1,5, 1,2)

B: oh so did I/ (2,1,5,1,2) anyway I mean if I had lots of money and I needed an urgent operation I would have one/ (2,1,18,2,2) that's if it wasn't a good idea to wait/ (2,1,9,1,1) I wouldn't be affecting waiting lists/ (2,1,1,2,1) I'd use that money/ (2,1,18,2,1) and I don't think I would feel bad about it/ (2,1,1,2,2) that's because it wouldn't have affected anybody else/ (2,1,9,1,1)

A: Well it would have done/ (1,4,15,1,1) as you got a bed then someone else didn't/ (1,1,9,6,6)

B: um but it's a private bed/ (2,5,9,1,1)

A: A bed's still a bed/ (1,4,1,1,1)

B: Well okay/ (2,3,14,1,1) but I still wouldn't feel guilty/ (2,1,1,1,2) that's just the situation everywhere/ (2,1,9,1,1)

A: That's no argument/ (1,4,22,1,1)

B: Well it's true/ (2,4,9,1,1)

A: It's just selfish/ (1,1,8,1,1)

B: Well I don't really know much about it anyway/ (2,1,13,2,1) if I had some figures at my finger tips I could show you that it's right/ (2,1,18,2,2)

A: I don't think so/ (1,4,1,2,1) there have been a lot of cases of this/ (1,1,9,1,1) people jumping the queue by paying for it/ (1,1,18,1,6) the waiting list does effectively get longer/ (1,1,1,1,1) and surgeons are making money out of other people's misery in some cases/ (1,1,1,1,5)

B: Well possibly/ (2,5,14,1,1) I think we agree to disagree/ (2,1,14,2,4)

A: Okay/ (1,3,15,1,1)

B: Well we spent ages on that/ (2,1,22,1,4) it's quite enjoyable isn't it?/ (2,10,22,1,1)

A: Do you think only the rich and idle can afford the luxury of psychological problems?/ (1,10,1,3,1)

B: Hang on/ (2,11,21,1,1) what was that?/ (2,10,24,1,1)

A: I'm not really sure about this/ (1,1,1,2,1) I hate being ill anyway/ (1,1,8,1,2) ...

2.8 Conclusion

Categorising conversational speech into units which describe the information content is the operational link between the data and any subsequent conclusions and theories. It is important therefore, that any system is not only relevant to the research, but is also rigorously derived, conforms to the requirements of an analysis system (e.g. Holsti, 1969) and maintains strict distinctions between conceptually different levels of analysis. In a review of the relevant literature it was apparent that the majority of systems were deficient in at least one of these requirements. CEA has therefore been designed to remedy these shortcomings. Based on a review of the literature and the analysis of a number of tape-recordings, the CEA system comprises four sets of rules. These are used to segment conversation and to classify the resultant units according to their Activity (the way information is exchanged), Type (the information content) and Focus (the referent of the speech), a system which has been shown to be highly reliable.

Originally designed for an analysis of how thoughts and ideas are sequenced in conversation, which is the general theme of the present research, CEA has been designed with sufficient generality for application to conversation arising in a wide variety of situations and relationships. The emphasis on how information is exchanged in conversation and on the types and referent of the information exchanged mean that CEA may be easily applied, for example, to the analysis of interaction cycles (e.g. Flanders, 1970), sequences of social behaviours (e.g. Argyle, 1979; Argyle et al., 1981), conversational gambits (e.g. Keller, 1981), or floor-apportionment (e.g. Thomas & Bull, 1981). Enhancing our understanding of the way conversation is constructed, these types of information also have implications for the teaching of conversational skills. Trower et al.

(1978) for example, have pointed out that it is necessary to know the component behaviours and how they are sequenced, before one can teach conversational skills. CEA allows direct access to these component behaviours and analysed according to the procedures suggested by Argyle (1979), and those detailed in Chapters 1, 3 and 4, a detailed picture of the way conversational behaviours are sequenced could be constructed.

In addition to its use in the analysis of interaction sequences, CEA can also be used to relate the process of social interaction to the outcome of the interaction. For example, given a prisoner's dilemma situation, the outcome of the interaction can be directly related to the processes intrinsic to the interaction (e.g. Jones & Gerard, 1967, pp.562-575). Similarly, given a situation in which negotiation is occurring, one may find that more acceptable outcomes are reached when more personal referents and/or more agreements are present in the interaction.

CEA is therefore a multi-purpose method of conversation analysis, not only being applicable to the analysis of conversation sequencing occurring in a wide variety of situations, and embracing a variety of relationships, but can also be used to either simply describe an interaction, or to relate the process of the interaction to its outcome.

CHAPTER 3

A PRELIMINARY ANALYSIS OF CONVERSATION

3.1 Introduction

Social interaction has, for a number of years, been considered in terms of structured sequences of behaviours. Argyle's (1969, p.168) model of social interaction considers that "each social act is determined by the last act of the other", and Thibaut and Kelly (1959, p.11) have written, "if enough observations were at hand, the elements of a given sequence of acts could be identified on the basis of certain statistical regularities: the elements would be found to occur together repeatedly and to be performed in certain sequential arrangement". Thibaut and Kelly's basic unit of analysis was the "behaviour sequence or set", a unit which "consists of a number of specific motor and verbal acts that exhibit some degree of sequential organisation" (p.11). The important point made by these authors is that in social behaviour, the constituent events or behaviours are not independent of each other, there being a relationship, or degree of organisation between the events.

The sequential organisation of verbal behaviour has also been noted. Argyle (1969, p.115) for example, has suggested that verbal "utterances are generated by other utterances", and Auld and White (1959, p.100) emphasise the "lawfulness and interconnectedness of the events" in verbal communication. In addition, Sacks (1972) has suggested that there is an "overall structural organisation of conversation" which comes into effect at certain points in the conversation. One of these points is getting in and out of the conversation, at which adjacency pairs such as question-answer, request-acceptance or rejection, and compliment-acceptance, etc. are used. Adjacency pairs are said to have two main features: they have the ability to function in a conversation by creating organisation, and secondly, they exhibit relative ordering. For example, a question precedes an answer but not vice-versa, and, given a question, only a part of the range of repertoire responses are admissible. Thus, adjacency pairs contribute to the structure of conversation by organising elements, some of which are constrained by relative ordering. In a similar vein, Argyle (1969, p.115) has also suggested that constraints operate in the conversational system; "different classes of utterance have specific effects on subsequent interaction: questions lead to answers, orders may lead to action, and information may have varied consequences as a result of adding to the data available to others".

On a more empirical level, Scheidel and Crowell (1964) studied problem-solving groups and found consistent patterns of idea development. For example, typical sequences of ideas tended to be, restatement-clarification, substantiation-positive modification or agreement, and clarification-acceptance-further clarification, or extension-restatement-further extension. However, although these sequences were consistently observed, Scheidel and Crowell noted that there was considerable uncertainty in predicting a subsequent event from a preceding event, the discussions being "about 80% as 'unpredictable' as they could be" (p.143). In addition, Gouran and Baird (1972) showed that there were differences in the sequential structures of problem-solving and informal groups. Informal group conversation tended to show a greater degree of organisation than problem-solving discussion (contrary to their expectation). For problem-solving groups, the main organisational feature was for questions to be followed by statements giving information. However, for informal discussion, Gouran and Baird noted more organisational features: agreements were followed more frequently by statements of initiation and more agreement, disagreements were followed more frequently by agreement or disagreement, and questions were followed more frequently by statements giving information.

The general conclusion that may be drawn from this work is that conversational behaviour has an inherent structure, in which units of speech are linked in some ordered fashion. The presence of organisation suggested in the work reviewed implies that the participants in the conversation are behaving non-randomly. The distribution of events in the speech stream is therefore non-random, by which the occurrence of one event is not independent of the occurrence of the previous event or events. Consequently, as there appears to be a dependent relationship between speech events from standpoints which are both theoretical (intrinsic to system theory - Chapter 1) and observational (e.g. the writings of Argyle (1969) and Sacks (1972), etc.), the next step was to test this assumption for informal conversation.

The study was considered to be a pilot investigation of conversational structure. In addition to testing the main assumption of dependence between elements in the speech stream, the study was also used to test the recording system and the method of obtaining data through discussion of controversial topics.

3.2 Method

3.2.1 Subjects

The subjects of the study were 8 female students from York University, aged between 18 and 22 years. Subjects met each other in pairs, no members of each pair being previously acquainted with one another.

3.2.2 Procedure

The study took place in a small, comfortable seminar room, in which each pair of subjects were introduced and allowed to familiarise themselves with the setting, for about five minutes. Subjects were then given a sheet of topical and controversial statements to read through, such as abortion, student's unions, euthanasia, capital punishment, etc., (See Appendix B), and asked to pick out those items that interested them most. They were left alone with each other for about 20 minutes, their conversation being recorded. Beside the subjects' chairs was a coffee table on which were placed directional recording microphones. Each subject was recorded on to a separate track of a 2-track tape-recorder in order to facilitate transcription of the conversation. Subjects were aware that they were being recorded, although, when asked at the end of the conversation, none indicated that this had interfered at all with their conversation. This was probably due to the microphone placing being quite unobtrusive.

Transcripts of each interaction were made and numerically coded using the Activity dimension of Conversational Exchange Analysis (see section 2.5.1). The data were of the event-sequence type (Gottman & Bakeman, 1979), resulting from the encoding of a stream of behaviour as a sequence of events or behaviours usually defined so as to be mutually exclusive and exhaustive.

3.3 Data analysis and results

3.3.1 Method of analysis

The raw data consisted of a sequence of digits, each digit representing the type of utterance made by the individual speakers. One approach to describing the data was simply to conduct a frequency count of the different digits. However, an alternative approach that preserved both the relationship between speakers and the time ordering of the events was to construct a transition table, or matrix. As the data were of the event-sequence type, transitions were considered to be from event to event rather than from time unit to time unit.

In this matrix, each cell corresponds to the frequency with which each of the possible speech acts at event e was followed immediately by each of the speech events at event e+1. If the sequence remained in the same state in successive event units, it is considered to have made a transition from the state to itself. Each digit in the sequence is considered twice, first as the state at even e+1, and then as the state at event e. As the first and last events of the sequence can only participate in one transition each, for k consecutive events, there are k-1, one-step transitions. By convention, the rows of the transition matrix are designated the starting, or antecedent state of each transition (event e), the columns being the outcome, or subsequent state (event e+1). Thus, for a hypothetical two-state system, having states A and B, the transition matrix is shown in Figure 1.

<u>Figure 1</u> Hypothetical two-state frequency transition matrix

		Subsequent state	
		A	B
Antecedent	A	700	300
state	B	120	880

In this example, the A state was observed to follow the A state on 700 occasions, whilst the B state followed the A state only 300 times. Similarly, the B state followed the B state 880 times, whilst the A state only followed the B state on 120 of the occasions.

The time ordering of events has been preserved by using a transition matrix as a data-reduction device, the relationship between speakers being preserved by drawing a distinction between act transitions that occur by the same speaker (within), and those that occur between speakers (between). Cast into transition matrix form, the data obtained from the four interactions recorded for this study are shown in Table 4.

Table 4 Frequency transition matrix and test of independence for four informal conversations combined.

	O^w	O^b	Rp^{wb}	C^w	C^b	D^w	D^b	Rt^w	Rt^b	Rq^w	Rq^b	Total
O	301	98	3	5	55	18	50	-	192	15	15	752
Rp	7	7	5	2	10	0	3	-	9	3	1	47
C	54	21	0	37	4	2	6	-	6	1	2	133
D	62	11	2	9	2	2	2	-	4	7	3	104
Rt	33	155	0	4	0	5	10	28	0	2	7	244
Rq	5	1	31	0	11	2	4	-	3	4	2	63
Total	462	293	41	57	82	29	75	28	214	32	30	1343

Note: (O) Offer, (Rp) Reply, (C) Consent, (D) Dissent, (Rt) Reaction, (Rq) Request. Superscripts w and b refer to within- and between-speaker transitions, respectively.

Chi-square = 1298.28 (45df) $p < .001$

Cramer's V = .438

Assymetric lambda = .1708

3.3.2 Test of independence between speech events

In testing the assumption of independence between elements in the speech stream, one is in effect, testing for independence between rows and columns in a contingency table. The appropriate test, although generally used as a test of association, is the frequency goodness of fit, or chi-square test, validated for use in transition matrix analysis by Bartlett (1951). Expected frequencies have been calculated in the usual manner (row sum x column sum / grand sum) (Fagen & Young, 1978, p.104). A significant chi-square value indicates a dependent relationship between rows and columns. Three points need to be made at this point concerning the sample size, the type of zero cells that occur in the transition matrix, and the problem of zero expected frequencies.

The chi-square technique is quite straightforward as long as the expected values are large. Cochran (1954) and Siegel (1956) have suggested that the minimum requirements for a chi-square analysis are that no cell should contain an expected frequency of less than one and a maximum of 20% of the cells may contain a frequency less than 5. When these conditions are not met, Siegel (1956, p.109) suggests collapsing categories so that small expected frequencies no longer occur. In the present study, very few within-speaker Replies were found to occur, (5 out of 41, all of which were found to be preceded by Replies) which resulted in a number of zero expected frequencies. Consequently, the within-and between-speaker Reply categories have been combined, thereby overcoming the problem.

Secondly, Fagen and Young (1978, p.104) have indicated that the sample size for transition matrix analysis should be between 5 - $10R^2$ events, where R is the number of cells in a rectangular table. Therefore, for 11 possible states, the sample size should be greater than 605 at a minimum, and preferably greater than 1210 events. Summing frequencies for the six speech categories that occurred in the conversations (Offer, Reply, Consent, Dissent, Reaction, and Request) the total number of events produced by the four conversations was 1347, resulting in 1343 transitions. The size of the sample may therefore be considered adequate for the test of independence.

Finally, it is possible to have two types of zero cell in a transition table: sampling zeros and logical zeros. The former are transitions that can occur, but did not do so in the sample selected; the latter are transitions whose occurrence is considered logically impossible (Colgan & Smith, 1978, p.158). In the present analysis, it was considered that transitions to a within-speaker Reaction, other than from itself, would not be logically possible. This is because, by definition, a Reaction is a vocalisation made in response to the other person. Consequently, these transitions were considered to be logically zero and the number of degrees of freedom reduced by one for each zero cell (Kullback, Kupperman, & Ku, 1962).

Applying the chi-square test to the transition matrix shown in Table 4 yielded a chi-square value of 1298.28 with 45 degrees of freedom, a result that is highly significant ($p < .001$), and indicating a dependent relationship between conversational speech acts.

3.3.3 The strength of the relationship between speech events

Blalock (1960) has indicated that if samples are large it is generally fairly easy to establish significance for even a small relationship between variables. For example, given a matrix of sample size N and a significant value of chi-square, if the sample size were to be doubled (2N), keeping the same proportions in each cell, the calculated value of chi-square would be double that of the matrix with sample size N. In effect, when samples are large one is saying very little even though a significant relationship may have been established. Blalock (1960, p.225) has indicated that a more important question concerns the strength of the relationship between two variables.

Two measures, Cramer's V and Goodman and Kruskal's (1954) lambda, have been used to assess the strength of the relationship between speech events. Cramer's V is a measure of association for tables greater than 2 x 2. Ranging between 0 and +1, a large value indicates that a high degree of association exists. The calculated value of Cramer's V for this study's data of $V = .44$ indeed suggests a moderately high degree of association between speech events.

Although both chi-square and Cramer's V indicate a substantial departure from independence between antecedent and subsequent events in the transition matrix, it is possible that there may be little or no predictive association when predicting column categories from row categories. Consequently, asymmetic lambda has been used to measure the percentage improvement in the ability to predict the value of the dependent variable once the value of the independent variable is known. Taking the antecedent events as the independent variable, asymmetric lambda indicates that there is 17.08% improvement in one's ability to predict the type of subsequent act, given that the identity of the antecedent act is known.

3.4 Discussion

The results of the chi-square test indicate a significant departure from independence between rows and columns of the transition matrix, and therefore between antecedent and subsequent events. Gauged using Cramer's V, the relationship between events is shown to be moderately strong, and the use of Goodman and Kruskal's asymmetric lambda suggests a small, predictable relationship between conversational elements in the speech stream. The value of lambda indicates that given a known antecedent event, there is a 17.08% improvement in the ability to predict the identity of the subsequent act, compared to when the identity of the antecedent act is not known.

These results therefore indicate not only a dependent, but also a predictive relationship between speech events, a result that is assumed, but never tested, in the literature. The indication of dependence implies that speakers are behaving non-randomly, and as has already been argued (see section 1.3.3), non-random behaviour indicates the presence of organisation between events. In addition, the indication of a predictive relationship between speech events, which has not been shown before, suggests that it might be possible to construct a predictive model of conversation.

Having indicated the presence of organisation in the speech stream, it may seem a logical step to continue by showing how the elements are organised, with the ultimate aim of deriving sequencing rules. This is the approach that a number of authors such as Hawes and Foley (1973) and Stech (1975) have used. For example, although showing conversation to be very uncertain in terms of predicting a subsequent event from a preceding event (Stech, 1970), Stech (1975), using police-civilian telephone calls as a data base, was able to write out a number of rules for conversational sequence generation. Typically, given that a speech act occurred, he found that the most likely event to occur next was the same event by the same person (Bales, 1950, termed this proactive behaviour). Looking at reactive behaviours, the tendency for one event to be followed by an event by another speaker, Stech found that interaction sequences generally started with Questions or Initiations, which were followed by Answers, Proposals, Information giving, Clarification, Substantiation, and Statement making. Disagreements tended to be followed by more Statement making, and Agreements were followed by either Statement making, or the initiation of a new conversational sequence by Questioning behaviour.

However, in describing the organisation of conversational events based on a simple frequency matrix and test of independence, one is making three basic assumptions. Firstly, it is assumed that the likelihood of a transition from one event to another is the same at any point in the conversation (this is usually referred to as stationarity). Second, the sequences are assumed to be first-order, whereby each event in the speech stream is maximally predicted by a single preceding event. In some cases this assumption may not be justified and more preceding events may be required in order to maximally predict the identity of the subsequent speech state. Finally, the assumption is made that all the conversational sequences are homogeneous; rules derived from the combined sequences are therefore assumed to be the same as those that would be derived from the individual sequences.

Although having demonstrated that a relationship exists between conversational events, the organisation of the speech stream should not be described until the three assumptions of stationarity, order of sequential constraint, and sample homogeneity have been satisfied. Consequently, the next step in assessing the organisation of conversational events is the initial testing of these three assumptions and the subsequent extraction of conversational sequencing rules. This forms the basis of Chapters 4 and 6.

CHAPTER 4

A SEQUENTIAL ANALYSIS OF INFORMAL CONVERSATION USING MARKOV CHAINS†

4.1 Introduction

Having demonstrated a dependent and predictive relationship between speech events in Chapter 3, it was pointed out that a number of authors (e.g. Gouran & Baird, 1972; Hawes & Foley, 1973; Stech, 1975) had inappropriately used such a finding as a basis for describing the organisation of conversation. Before one can begin to describe the structure of the speech stream, the three basic assumptions of stationarity, order of sequential constraint, and sample homogeneity must be satisfied (Hawes & Foley, 1976). As pointed out in Chapter 1, adopting a system theory approach necessitates an alternative methodology to that used by traditional science. Consequently, having shown that a predictive relationship between the speech events exists, the theme of this chapter is the development of a model of conversation that satisfies the assumptions of stationarity, order, and sample homogeneity.

---

† Based on a paper presented to the British Psychological Society, Social Psychology Conference (Oxford, 1981), and an article, 'A sequential analysis of informal conversation using Markov chains', by Thomas, A.P., Roger, D., and Bull, P.E., forthcoming in the British Journal of Social Psychology.

Particular attention is paid to the use of statistical methods compatible with a system framework and appropriate for the analysis of dependent data.

The methods by which the structure and organisation of conversation have been analysed fall into two basic types, representing distinct methodological and theoretical viewpoints (Penman, 1980); deterministic or law based (e.g. Stech, 1975; Nofsinger, 1975, 1976) representing traditional scientific method, and stochastic (e.g. Jaffe & Feldstein, 1970; Hawes & Foley, 1976; Penman, 1980) representing a system framework.

As outlined in the Introduction, traditional methods of analysis used by social scientists are usually law based, individually, or monadically oriented - treating an interacting dyad as if it were a single subject (Kraemer & Jacklin, 1979) - and discontinuous, in that the possibility of variation over time is ignored. By contrast however, a system approach using stochastic methodology views dyadic conversation as a process (actions that are connected over time), comprising relational (between interacting individuals), structural (connected actions that are subject to differing degrees of constraint), and informational (the type of information being communicated) properties.

Whereas mechanistic models of interaction make causal assumptions of the 'if-then' form, stochastic process models view behaviour as being constrained and only partly determined by external or internal factors. The application of a stochastic process assumes that people behave probabilistically rather than in a simple one-to-one, cause and effect fashion (Ashby, 1961), a stochastic process being seen as a process developing in time in a manner controlled by probabilistic laws (Parzen, 1962).

Consequently, the occurrence of any given event in a stochastic process depends, to a variable extent, on those events preceding it. The process yields a distribution of events, each with an associated probability of occurrence and unlike the result of conventional methodologies, it is not possible to predict exactly what behaviour will follow. While subsequent events may therefore be said to be dependent on an antecedent event, this does not imply that events are fully determined, but only that the occurrence of the first event alters the probability of occurrence of the second event.

By way of example, assuming a two state-model of conversation, it is possible to see how the occurrence of an antecedent event alters the probability of occurrence of a subsequent event. Recalling the example of the frequency transition matrix given in section 3.3.1 and shown in Figure 2, this can be construed as a stochastic representation of a hypothetical two-state sequence of behaviours.

Figure 2 Hypothetical two-state frequency transition matrix

		Subsequent state		
		A	B	Total
Antecedent	A	700	300	1000
state	B	120	880	1000
	Total	820	1180	

Normalising the column of row sums by dividing each element by the total number of events gives the resultant vector V. This represents the probability distribution of the two states, where p(i) is the notation for the probability of state i, and i is either A or B in this example. For the hypothetical data:

$$V = \begin{bmatrix} p(A) \\ p(B) \end{bmatrix} = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$$

The elements of V are absolute probabilities and represent the proportion of the total number of speech events spent in each of the speech sates, regardless of the time ordering of the elements.

The matrix itself can also be normalised by dividing each of the row frequencies by the row sum. This yields a matrix P, the elements of which are transitional, or conditional probabilities. The notation here, p_{ij}, represents the probability of being in state j at event e+1, given that the system was in state i at event e.

$$P = \begin{bmatrix} pAA & pAB \\ pBA & pBB \end{bmatrix} = \begin{bmatrix} .70 & .30 \\ .12 & .88 \end{bmatrix}$$

Each row of the matrix is now a probability vector giving the distribution of the possible outcomes of a transition from a specified state i. For example, of all transitions starting in state B, 88% ended in state

B, while 12% ended in state A. The probability matrix is therefore a stochastic matrix, defined as a matrix of non-negative real numbers whose sum of terms in each row is 1 (Jaffe & Feldstein, 1970).

In the above example, given that the first event to occur in a sequence is a B-type event, the probability of a B-type event occurring next is .88. However, if an A-type event were to be the initial event, the probability of a B-type event following would then be .30. Subsequent events are not therefore fully determined, but the occurrence of one event can alter the probability of occurrence of the next event.

The main difference between deterministic and stochastic models lies in the interpretation of a non-occurrence of a probable event. From a law-based position, non-occurrence indicates either observational error or an incorrect statement derived from the law-based principle. Taking a probabilistic point of view, the non-occurrence simply indicates that the system used has a greater degree of complexity and flexibility (Fisher, 1978a). Non-occurrence of probable events undermines the predictive power of deterministic models, whereas stochastic models, based on probabilistic statements, are able to cope with such event-to-event variations. Similarly, stochastic methods do not imply direct linear relationships between events following each other in time. A subsequent event is altered, not determined, by the occurrence of the antecedent event. One event is simply antecedent and the other subsequent; the antecedent event is not a sufficient and/or necessary condition to bring about the subsequent state.

Given that conversation, as defined in the Introduction, is a process that demonstrates relational and structural properties, a model was required that would facilitate a detailed description of the process of conversation. A discrete Markov chain, one of the family of finite stochastic models, was selected as a model for conversation. This is because of its conceptual isomorphism with conversation, a Markov process being concerned with the modelling of variables that are discrete, mutually exclusive and exhaustive, such as speech states, which are interdependent and develop over time in a manner controlled by probabilistic laws.

Of all the stochastic models available for the analysis of conversation (independent, Markov, stationary, Martingale; discussed by Penman, 1980, p.46), the Markov model has been the most widely used. For example, it has been applied to behavioural sequences of primates (Altman, 1965), bird song (Chatfield & Lemon, 1970; Chatfield, 1973), time patterns of dialogue (Jaffe & Feldstein, 1970; Cappella, 1979, 1980; Cappella & Planap, 1981), the process of psychotherapy (Frank & Sweetland, 1962; Benjamin, 1979), experimentally induced marital conflict (Hertel, 1968), group decision making (Fisher, 1970; Donohue, Hawes & Mabee, 1981), mother-infant interaction (Jaffe at al., 1973), interview behaviour (Hawes & Foley, 1973), small group deviance behaviour (Valentine & Fisher, 1974), small group development (Ellis & Fisher, 1975), marital interaction (Parks et al., 1975, 1976), Mishler (1975) has suggested that a finite Markov model may best represent aspects of discourse, such as questioning behaviour, committee interaction (Hawes & Foley, 1976), marital conflict (Gottman, Markman & Notarius, 1977) and communication in self-analytic groups (Krueger, 1979).

The appeal in using a Markov process as a model of conversation lies not only in the assumption that the past has some effect on the present, the effect varying according to the nature of the past event, but also its use both as a descriptive (e.g. Hawes & Foley, 1973) and predictive tool (e.g. Jaffe & Feldstein, 1970). For example, as Raush (1972, p.291) has written, "events influenced by single variables may follow strict deterministic laws. When, however, many variables influence a particular event, outcomes are more likely to be probabilistic. Markov processes seem particularly suited to examining the uncertainties of complex and multiple determination and to providing predictive models for studying the cause of events in our ambiguous, uncertain, natural world".

Strictly, for a process to be Markovian, it must satisfy three basic assumptions: stationarity, order of sequential constraint, and sample homogeneity. Briefly, stationarity requires that the probability of a given event depends only on the time interval and not on the time of occurrence. It therefore assumes that there has been no overall change in the patterns of event-to-event probabilities during the complete time sequence (Cox & Miller, 1965). The assumption of order of sequential constraint refers to the number of events necessary to maximally predict the identity of the next event in the sequence. For example, a first-order sequence requires knowledge of only the immediately preceding event in the sequence to predict maximally the identity of the subsequent event. Finally, homogeneity refers to the assumption that all sequences are similar and that the sample contains no radically different sub-groups.

Reviewing the literature on Markov modelling of conversation, three main types of error emerged. Firstly, conversation was frequently assumed to be Markovian even though some, or all, of the assumptions were not tested. For example, Hertel (1968) and Hawes and Foley (1973) only tested for stationarity, Jaffe and Feldstein (1970) only tested the order of sequential constraint, and Fisher (1970), Valentine and Fisher (1974) and Gottman, Markman and Notarius (1977) all assumed the communication systems under study to be first-order. (The same problem arises in the law-based literature, where a one-to-one predictive correspondence between events is assumed. The work of Nofsinger, 1975, 1976, and Stech, 1970, 1975, is of this type.) Secondly, the statistical methods used to test the Markovity assumptions were often inappropriate, if not invalid. For example, the independence of data items was frequently assumed, but since the relationship between speech events is, by nature, one of dependence (as shown in Chapter 3), this assumption cannot be met. Such statistically inappropriate procedures include using chi-square for assessing stationarity (Hertel, 1968; Lichtenberg & Hummel, 1976), z-scores for assessing goodness of fit (Sackett, 1978; Gottman & Bakeman, 1979), and correlation for assessing sequence stationarity (Hawes & Foley, 1973). Thirdly, statistical requirements were often violated. For example, Donohue, Hawes and Mabee (1981) use an appropriate test (maximum likelihood ratio - Anderson & Goodman, 1957), but use it incorrectly. The test specifies that all transition probabilities must be greater than zero. However, inspection of their composite transition matrix reveals that 23% of cells are zero, thus invalidating the test. Such a violation can bring about a significant change in the results obtained. For example, in testing the assumption of stationarity for the group of subjects as a whole (time-homogeneity), Donohue, Hawes and Mabee (1981) obtained a likelihood ratio (LR) value of 100.49, which was not significant with 90 degress of freedom. Assuming that all individual subject matrices showed identical patterns of zero cells as their composite matrix, and that the cells could all be treated as logically zero (a conservative procedure), the number of degrees of freedom should have been reduced by 23, one degree of freedom per zero cell (a method suggested by Kullback, Kupperman & Ku, 1962). With a revised

67 degrees of freedom, an LR value of 100.49 is significant at the 0.5% level of significance. The group of subjects is not therefore time-homogeneous, contrary to the conclusion drawn by the authors of the article.

Of the few relevant statistically-oriented papers, those of Hawes and Foley (1976), Hewes, Brazil and Evans (1977) and Krueger (1979) are the only ones to have tested the assumptions or order, stationarity and homogeneity, and Jaffe and Feldstein (1970) the assumption of order, in an appropriate manner.

The methods of analysis used in this chapter are intended to rectify these shortcomings. The analytical methods used in this research stem from information theory, the development of which has been closely connected with the development of system theory (Von Bertalanffy, 1971, p.40). The problem of dependence between speech events was overcome by using the Anderson-Goodman maximum likelihood ratio as a test statistic, developed purely for the analysis of sequential processes (Anderson & Goodman, 1957), and Attneave's information statistics, which were devised for the analysis of information transmission (Attneave, 1959). Additional information concerning the maximum likelihood statistic is given in Appendix C.

Preserving the process and structural aspects of conversation by the use of Markov chain and informational methods of analysis, the relationship between individuals has been preserved by the manner in which the data has been organised. In a review of ten studies of the sequential structure of human interaction covering the period 1964-1973, Stech (1975) noted that in only two studies (Bales, 1966, and Stech, 1970) was the relationship between interactants preserved. In both cases, the relationship was expressed by performing separate analyses on transitions between utterances (transitions between one person's statement and the next person's statement), and transitions within utterances (the method by which an individual builds up a string of events into a complete utterance). The same within/between distinction has also been made by Bales et al. (1951) and Lewis (1970a) in his analysis of Bales' (1953) model of group discussion, referred to as Model T5. However, in each case, within and between utterances are treated separately, in which case the relationship is partially destroyed as each matrix only specifies one part of the conversation process. For example, given an antecedent state, it is not possible to state the range of subsequent events that are possible unless one knows in advance whether the next speech act is a within- or between-speaker utterance. Combining the matrices yields a better approximation of the process; given an antecedent event, the complete range of subsequent events, whether they are within- or between-utterances may be specified.

In the present analyses, the relationship between individuals is preserved in two different ways. Either, each interactant is individually specified in the transition matrix, thus:

		Speaker 1		Speaker 2	
		A	B	A	B
Speaker 1	A	.80	.10	.05	.05
	B	.30	.20	.15	.35
Speaker 2	A	.05	.05	.80	.10
	B	.15	.35	.30	.20

or a distinction is drawn between within-speaker and between-speaker transitions, but unlike Bales et al. (1951), Bales (1953), Lewis (1970a), and Stech (1970, 1975), both within- and between-speaker transitions are incorporated into a single matrix, thus:

		Subsequent state			
		Within		Between	
		A	B	A	B
Antecedent state	A	.80	.10	.05	.05
	B	.30	.20	.15	.35

Finally, the aim of this chapter is to assess the patterning of events in the conversational speech stream, whilst preserving the process, relational and structural aspects of conversation. Methods of statistical analysis have been selected that are compatible with a system theory framework and appropriate for the analysis of sequential, dependent data. The study has been designed to answer the following three questions:

1) Can event-to-event transitions in informal conversation be considered to be stationary?
2) Do informal conversations have a common order of sequential constraint, and if so, what is the order?
3) Can a single model describe all the conversational sequences obtained; do all the sequences form a single homogeneous group?

In addition, a number of analyses are also presented that are concerned with the effects of violations of the stationarity and homogeneity assumptions.

4.2 Method

4.2.1 Subjects

The subjects for the study were 48 female students from York University, aged between 18 and 24 years. Subjects met each other in pairs, no members of each pair being previously acquainted with one another.

4.2.2 Procedure

The procedure was identical to that described in Chapter 3 (section 3.2.2).

Transcripts of each of the 24 interactions were made and numerically coded using the Activity dimension of ConversationalExchange Analysis (see section 2.5.1). The 24 conversations produced a total of 14074 speech events (14050 transitions), the data again being of the event-sequence type.

4.3 Data analysis overview

The analysis of the conversations is discussed in three sections. The first section is concerned with the development of a 4-code (8 speaker-state) model of conversation (section 4.4) and is used to test the effects of violation of the assumptions of sequence stationarity and sample homogeneity. The second set of analyses (section 4.5) is concerned with the expansion of the model to include 7 speech codes (14 speaker-states); the third section is a replication study (section 4.6).

4.4 A four-code model of conversation

The likelihood ratio test assumes that all probabilities are greater than zero. In order not to violate this assumption, the initial model was restricted to four speech codes because some dyads did not produce all possible combinations of the Activity codes. The basis of the first set of analyses is therefore a four-code Markov model, representing the Offer/Reply, Consent, Dissent and Request speech codes, for each of the two speakers. The three codes which failed to occur in all combinations were Replies, Modifications and Reactions. Replies and Offers were combined as they both introduce new information, but Modifications and Reactions could not logically be combined with any other speech code and were therefore dropped from the speech stream. This was considered a more acceptable procedure than an arbitrary lumping of categories, especially as Modifications account for only 2.56% of the speech stream, and Reactions, which are listener rather than speaker behaviours, account for only 10.80% of the speech stream. In fact Raush (1972) has warned against arbitrarily lumping categories together as this may confuse, or even mask the structure of the behavioural stream.

Dropping codes from the speech stream results in a decrease in the sample size from 14074 events coded into seven categories to 11919 events in four speech categories. The analyses therefore focus on the four speaker behaviours, Offer/Reply, Consent, Dissent, and Request. By limiting the initial model in this way to 8 speaker-states, the test assumption has not been violated.

The usual procedure in hypothesis testing is to identify the theory under examination with the alternative hypothesis, which is then tested against the null hypothesis. However, in model testing, the model is identified with the null hypothesis, and a violation of the model assumptions tested by seeking rejection of the null hypothesis (Hawes & Foley, 1976). The three assumptions tested were stationarity, order of sequential constraint, and sample homogeneity. They were tested in this order for two reasons. Firstly, testing is usually carried out in this way by convention (e.g. Hawes & Foley, 1976; Anderson & Goodman, 1957 - the authors of the paper deriving the test procedures used in this study). Secondly, the ordering of the tests is the result of a series of logical steps made to isolate potentially different groups of subjects. The testing order may be seen as hierarchical. One might, for example, assess sample homogeneity first, as in conventional hypothesis testing. However, this may give an indication that the sample is homogeneous, but the result may be an artifact. This is because two transition matrices may not be significantly different (i.e. they appear to be generated by the same process), but one may be the result of a stationary process, the other the result of an entirely different non-stationary process. Testing stationarity first allows the possible differentiation of subjects into two groups, stationary and non-stationary. Within these two groups one can then assess the order of sequential constraint (one might conceivably obtain two distinct groups of stationary sequences, one being characterised by a first-order process, the other a second-order process). Finally, within these sub-groups, one can then test for sample homogeneity. The ordering of the test procedures thus represents a logical approach to isolating potentially different groups of conversational sequences.

The test procedures used in this study are all based on the chi-square distribution. A number of the analyses presented have degrees of freedom greater than 100, critical values for which are not generally available. These have been determined using the following formula, suggested by Pearson and Hartley (1970) for determining the critical points of the chi-square distribution when the degrees of freedom are greater than 100.

$$\chi^2 = v\left(1-\frac{2}{9v} + X\sqrt{\frac{2}{9v}}\right)^3$$

where v = number of degrees of freedom

X = + 1.6449 for 5% level of significance

+ 2.3263 for 1% level of significance

+ 3.0902 for .1% level of significance

4.4.1 Stationarity

The assumption of stationarity requires that there is no overall change in transition probabilities across the length of the dyadic sequence. A test of stationarity is considered necessary if analysis is to go beyond just simple transition probability calculations (Penman, 1980). A stationary sequence would be, for example, one in which the probability of occurrence of a Request-Reply transition at the beginning and end of a conversational sequence would not be statistically different. If this was not the case, the transition probabilities would not be stable and the sequence could not be said to be stationary. A stationary Markov chain model states therefore, that the probability of occurrence of any particular event is dependent only on the identity of the past event and not on the point in the sequence at which it occurs.

In a test of the stationarity assumption, using small groups talking for periods of up to two hours, Lewis (1970b) concluded that the average group was characterised by stationary parameters within a session, but the assumption that all groups were so characterised was not accepted. In fact, only 80% of male and 70% of female interactions displayed stationary parameters. Additionally, the assumption of stationarity between sessions could not be accepted.

However, Hawes and Foley (1976) using a sample of three academic committee meetings found that all the meetings had stationary parameters, and in an earlier study of interviewing behaviour found that initial probabilities were very similar to the final probabilities (Hawes & Foley, 1973). In addition, Anderson (1960) found that the rate and density of interviewing communication remained stable over time, as did the proportion of different types of messages (Swinth & Tuggle, 1971).

These results suggest that stationarity may not be a universal finding and should not therefore be assumed, as it may well depend on a number of factors, not least of which are the type of conversation and the amount of time spent talking.

The most direct test of stationarity for each dyad is to consider a sequence of events linked by transition probabilities to be a probability matrix M. This can be divided into smaller matrices M^{e+1}, M^{e+2},.....,M^{e+n}, each of which is considered to be generated by a process not significantly different from the process responsible for generating M. Each individual matrix is then compared to the composite matrix in turn to determine statistical dependence or independence. In each case the null hypothesis states that there will be no significant difference between each individual matrix and the composite matrix M. If the null hypothesis is accepted, then it may be concluded that the process remained stationary across the length of the sequence. A non-significant test statistic therefore supports the hypothesis of sequence stationarity.

In the present study, each dyad was analysed as a separate conversational process using Lewis' (1970b) modification of the Anderson-Goodman (1957) likelihood ratio test. This simply divides the composite sequence into equal blocks: in this case two blocks, representing the first and second halves of each conversational sequence. The method then compares the individual blocks with the composite matrix.

The Anderson-Goodman test states:

$$LR = -2 \sum_{t=1}^{T} \sum_{i,j=1}^{m} n_{ij}^{(t)} (\log_e \hat{p}_{ij} - \log_e \hat{p}_{ij}^{(t)})$$

where LR = Likelihood ratio

$\hat{p}_{ij}$ = estimated probability of an ij transition in the composite matrix

$\hat{p}_{ij}^{(t)}$ = estimated probability of an ij transition in the individual matrix t

$n_{ij}^{(t)}$ = frequency of an ij transition in the individual matrix t

The statistic has a chi-square distribution with degrees of freedom equal to:

$m(m-1)(T-1)$

where m = number of rows/columns in the matrices compared.

T = number of individual matrices compared.

Stationarity was initially assessed with respect to each speaker. For example, speaker A to speaker A, speaker A to speaker B, etc., transitions must all be stable for the overall process to be considered stationary. The results of the analyses, shown in Table 5, indicate that of the 24 dyads, 14 (58.3%) had stationary sequence probabilities and 10 (41.7%) had non-stationary sequence probabilities.

Table 5 Sequence stationarity - speaker specific analysis

Dyad	LR	P
1	74.54	.05
2	44.42	n.s.
3	37.52	n.s.
4	83.26	.05
5	35.27	n.s.
6	41.00	n.s.
7	82.08	.05
8	80.94	.05
9	83.83	.01
10	41.64	n.s.
11	80.25	.05
12	83.36	.05
13	43.89	n.s.
14	42.24	n.s.
15	49.35	n.s.
16	51.31	n.s.
17	28.20	n.s.
18	76.03	.05
19	36.31	n.s.
20	31.09	n.s.
21	36.76	n.s.
22	37.82	n.s.
23	79.27	.05
24	76.01	.05

All statistics have 56 df.
A significant LR value indicates a non-stationary sequence.

As noted earlier, a distinction is drawn between acts that occur by the same speaker in the dyad (within) and those that occur between speakers (between). Re-analysing the sequences for within- and between- speaker transitions, the results of these tests (Table 6) revealed that five dyads (20.8%) had non-stationary within-speaker transitions, and three dyads (12.5%) had non-stationary between-speaker transitions.

Table 6 Sequence stationarity - Within/Between-speaker transitions

	Within-speaker		Between-speaker	
Dyad	LR	P	LR	P
1	22.52	.05	7.36	n.s.
2	15.49	n.s.	6.85	n.s.
3	10.68	n.s.	2.92	n.s.
4	23.39	.025	10.28	n.s.
5	11.16	n.s.	8.57	n.s.
6	7.94	n.s.	7.87	n.s.
7	26.91	.01	6.90	n.s.
8	25.98	.025	17.93	n.s.
9	13.81	n.s.	23.72	.025
10	14.03	n.s.	9.60	n.s.
11	11.31	n.s.	18.43	n.s.
12	9.82	n.s.	31.72	.005
13	10.23	n.s.	24.13	.025
14	10.61	n.s.	11.26	n.s.
15	12.93	n.s.	12.94	n.s.
16	9.50	n.s.	11.11	n.s.
17	5.21	n.s.	4.79	n.s.
18	7.97	n.s.	18.79	n.s.
19	7.51	n.s.	11.59	n.s.
20	6.09	n.s.	7.60	n.s.
21	3.34	n.s.	9.44	n.s.
22	13.36	n.s.	9.79	n.s.
23	10.85	n.s.	14.99	n.s.
24	23.92	.025	12.43	n.s.

All statistics have 12 df.

A significant LR value indicates a non-stationary sequence.

Taking these results of the speaker-specific and within/between analyses together, 13 dyads (54.2%) had stationary sequence probabilities and 11 dyads (45.8%) had non-stationary sequence probabilities. Such results may be seen to confirm the suggestion that the assumption of stationarity may not always be appropriate.

The above analyses were carried out using Programs 1 and 2 (Appendix G) written for the analysis of speaker-specific and within/between speaker stationarity, respectively.

4.4.2 Order of sequential constraint

The order of sequential constraint of a system refers to one aspect of the organisation of the speech events and measures the extent to which constraints operate in the system. In a Markov process, it is assumed that the immediate past has some effect on the present event; the number of preceding events which may be seen as directly affecting the probability of occurrence of the present event defines the order of the sequence.

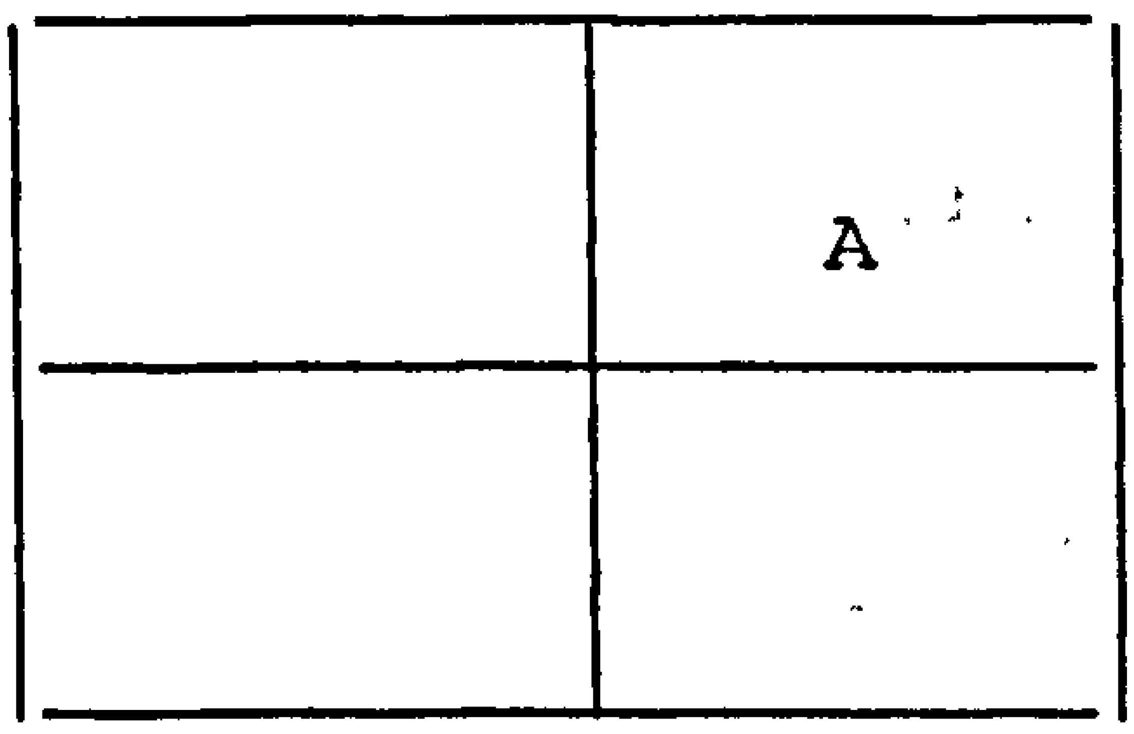
Markov analyses have often been confined to the analysis of first-order sequences, the probability of a transition at event e to e+1. However, Clarke (1977) has pointed out that a more comprehensive model of conversation might be constructed by considering higher order sequence analysis, and while Hawes and Foley's (1976) study of committee interaction suggested a first-order chain was appropriate, Pena's (1975; cited in Penman, 1980) study of classroom behaviour suggested that a second-order chain was a better fit for these data. Parks et al.(1975; in Penman, 1980) and Jaffe and Feldstein (1970) both used first- and second-order chains in their analysis of marital communication and speech bursts, respectively, but noted that whereas a first-order process was generally suitable, some state-to-state transitions were more appropriately treated as second-order. There would appear to be, therefore, no a priori reason to assume a first-order chain for communication (as have some authors, e.g. Hertel, 1968; Hawes & Foley, 1973), since the order of the chain may be dependent on a number of factors, such as the type of communication under study.

As noted in Chapter 1 (section 1.3.3), constraints are said to operate whenever variables (in this case, speech events) occur in a non-random, ordered manner. Consequently, Shannon and Weaver's (1949) concept of information, which is based on the notion of entropy, or disorder, can be used to study the organisation of events in a system. In the present study, an applied form of Information theory (due to Attneave, 1959), has been used to determine the order of sequential constraint by assessing how much new information is contributed by the inclusion of additional past events.

The technical meaning of information used here is not radically different from the everyday meaning, though more precise. Information is something which we gain by reading, or listening, or by directly observing the world about us (Attneave, 1959). A gain in information can only be made about matters in which we are to some extent ignorant, or uncertain. Information may therefore, be defined as that which removes or reduces uncertainty. In 'Through the Looking Glass', Lewis Carroll (1871 - in Gardener, 1970, p.341) gives an excellent example of the uncertainty there is in communication. "It is a very inconvenient habit of kittens (Alice had once made the remark) that, whatever you say to them, they always purr. 'If they would only purr for 'yes', and mew for 'no' or any rule of that sort', she had said, 'so that we could keep up a conversation!' ... On this occasion the kitten only purred, and it was impossible to guess whether it meant 'yes' or 'no'."

The important point here, is that there is a degree of uncertainty in communication. For example, in the case of Alice and the kitten, there is a degree of uncertainty concerning the meaning of the purr, whereas the uncertainty in conversation concerns the identity of the next event in the sequence.

Formalising the concept of information, given the four compartment box in the following example, one may wish to identify which compartment contains the letter 'A'.



The number of (binary - yes/no) questions that one needs to ask is two, the questions being 1) Is it in the top/bottom half of the box? and 2) Is it in the right/left half of the box? The binary digit (bit) being the usual unit of information and uncertainty, one can say that the amount of uncertainty involved in the question, 'Where is the letter A?' amounts to 2 bits. Similarly, 2 bits of information, in the form of two binary questions are required in order to designate a particular cell in the box. This relationship between uncertainty and the number of equally likely alternatives can be expressed thus:

$$H = \log_2 m$$

where H = amount of information or uncertainty, in bits.

$\log_2 m$ = log (to base 2) of the number of equally likely alternatives.

In terms of conversational sequences, the information contained in the speech acts is equal to the log (base 2) of the number of alternative speech events, when all acts are independent and equiprobable. This is known by Attneave (1959) as a zero-order estimate of H. A first order estimate of H assumes sequence events to be independent, but the speech acts to vary in probability of occurrence, and is obtained from:

$$\hat{H}_1 = \sum_{i=1}^{n} \hat{p}_i \log_2 1/\hat{p}_i$$

where $\hat{p}_i$ = the estimated probability of occurrence of act i. Although known by Attneave (1959) and Altman (1965) as a first-order estimate, Jaffe and Feldstein (1970), Chatfield and Lemon (1970), Chatfield (1973) and Hawes and Foley (1976) refer to this as a zero-order estimate. In a similar manner, these same authors refer to Attneave's zero-order estimate as H_{max}. This change in convention, which brings the information theory definition of order in line with the strict definition of a Markov process, has also been used throughout this research. Kemeny and Snell (1960) define a Markov chain as a process which moves through a finite number of states, and for which the probability of entering a certain state depends only on the last state occupied. Consequently, this one-step dependency is referred to as a first-order transition, in common with Jaffe and Feldstein (1970) and others. Similarly, if the probability of entering a certain state depends only on the last two states occupied, this is referred to as second-order. Within this framework the lack of sequential dependency, the usual notion of statistical independence, is zero-order.

A first-order information estimate assesses the amount of uncertainty in predicting a subsequent event on the basis of a single antecedent event and is calculated as follows. Firstly, the average amount of information in a digram (a consecutive pair of acts in the sequence) is calculated, just as if each pair of events were a single symbol (H digram):

$$\hat{H}_{(digram)} = \Sigma \, \hat{p}_{(digram)} \log_2 1/\hat{p}_{(digram)}$$

where the $p_{(digram)}$'s are the proportions of all overlapping pairs of events. A first-order estimate of H is then obtained by subtraction:

$$\hat{H}_1 = \hat{H}_{(digram)} - \hat{H}_0$$

where $\hat{H}_1$ = first order information estimate.

$\hat{H}_0$ = zero-order information estimate.

The rationale for this is as follows. $\hat{H}_0$ is the information content of a symbol considered independently of preceding symbols. The information in a digram must be at least as great as $\hat{H}_0$, if not greater than the amount of information in the first member of the pair of events. $\hat{H}_1$, the amount by which $\hat{H}_{(digram)}$ exceeds $\hat{H}_0$, represents the new information contributed by the second event of the pair. As every event in the sequence (except the first) is the second member of some pair of events, $\hat{H}_1$ may be conceived more generally as the new information added by each symbol in the sequence.

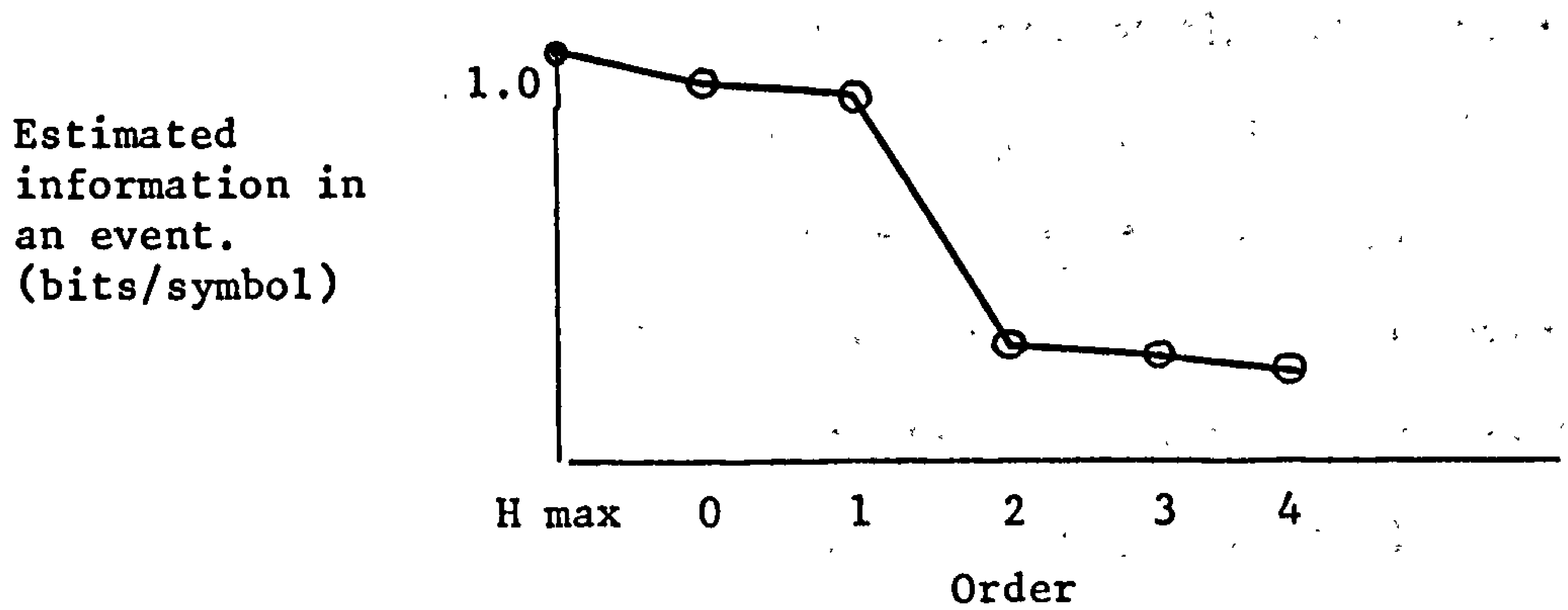
Second and higher-order estimates of H may be made in the same manner:

$$\hat{H}_{(n\text{-gram})} = \hat{p}_{(n\text{-gram})} \log_2 1/\hat{p}_{(n\text{-gram})}$$

$$\hat{H}_n = \hat{H}_{(n\text{-gram})} - \hat{H}_{(n\text{-}1\ \text{gram})}$$

Thus, using this general formula one may obtain estimates of the amount of information contained in increasingly larger numbers of overlapping events in the sequence. Each of the estimates may be plotted on a graph to give a visual indication of the estimated information in an event. The example, shown in Figure 3, is of a hypothetical 4-state behavioural sequence. The Y-axis could equally well be labelled Uncertainty, because when information is maximum, the number of questions that need to be asked to determine the identity of the subsequent event is at a maximum, and uncertainty is therefore at a maximum.

Figure 3 Estimated information in an event for a hypothetical four-state sequence.



From the graph it can be seen that the amount of information for the zero- and first-order estimates is only marginally smaller than $\hat{H}_{max}$. There is very litle deviation from randomness ($\hat{H}_{max}$), since neither symbol nor digram frequencies are disproportionate to any great extent. However, a sharp drop occurs at the second-order point and indicates that a knowledge of the two events which have just occurred considerably reduces the uncertainty in predicting the identity of the next event. Knowing the previous three (order = 3) events reduces uncertainty a little more, and so on. Uncertainty appears to be maximally reduced at the second-order point on the graph. The sequence would therefore be considered to be second-order and refers to the fact that a subsequent act's identity is maximally predicted by knowledge of the two previous events.

So far, the calculated information values refer to the information value of all individual categories, or all paired combinations, or all tripled combinations, etc. The difference between these order estimates gives the amount by which N consecutive events reduces the uncertainty of the event following, or the amount of information which a given event shares with the N events preceding it, and is known as Transmission ($\hat{T}$). So, for example, the difference between $\hat{H}_2$ and $\hat{H}_1$ may be thought of as the amount by which two consecutive events reduces the uncertainty of the one following, compared to when only a single event is used to predict the identity of a consequent event. The differences between each estimate can be tested using a formula suggested by Attneave (1959, p.29) that is based on the chi-square test and exploits the relationship between $\hat{T}$ (the estimate of shared information) and chi-square (the measure of significance of this estimate):

$$\chi^2 = 2(\log_e 2)\ n\hat{T}$$
$$= 1.3863\ n\hat{T}$$

where n = sequence length

$\hat{T}$ = estimate of shared information.

The statistic has (c-1) degrees of freedom for a zero-order analysis, $(c-1)^2$ degrees of freedom for a first-order analysis, and $c^{i-1}(c-1)^2$ degrees of freedom for second and higher-order analyses, where c is the number of categories in the system, and i, in the last formula, is the order of estimation (Chatfield, 1973).

The procedure employed in previous studies has been to obtain the order of the chain for the sequence as a whole (e.g. Hawes & Foley, 1976; Penman, 1980). In this study, however, the order was obtained for within- and between-speaker transitions for each of the 24 dyadic sequences, as well as for the sequence as a whole, ignoring speaker transitions (Overall). Three separate order estimates have therefore been computed: Overall, Within-, and Between-speaker transition estimates. It should be noted that for orders involving a pair or a triple of antecedent acts (second- and third-orders, respectively), the constraint that all antecedent acts should be by the same speaker was imposed. The reason for this is that in the sequence S1 S2 S1 the transition from the pair of antecedent acts S1 S2 could not strictly be considered a within-speaker transition because of the intervening act by the second speaker. Consequently, as the order of sequential constraint increases, the length of the sequence, and therefore the associated N value, becomes smaller. The frequencies for antecedent singles, pairs and triples used in the first-, second-, and third-order approximations were computed from the raw conversational sequences using Programs 3, 4, and 5 (Appendix G), respectively.

The analysis of order, as we have seen, requires three separate computations:

1) The calculation of the amount of information contained in a) a sequence when the events are independent but vary in probability of occurrence ($\hat{H}_0$), b) an antecedent/subsequent unit ($\hat{H}_{digram}$), c) an antecedent pair/subsequent unit ($\hat{H}_{trigram}$), and d) an antecedent triad/subsequent unit ($\hat{H}_{four-gram}$). These n-gram calculations, for each dyadic sequence, are shown in Appendix D.

2) Subtracting the n-gram information measures for each order gives the amount of information in the coded sequences when treated as first-order, second-order and third-order chains. These represent, in uncertainty terms, the amount of uncertainty there is in predicting a subsequent event on the basis of the previous event(s). These calculations for zero-($\hat{H}_0$), first-($\hat{H}_1$), second-($\hat{H}_2$), and third-($\hat{H}_3$) order estimates are given in Appendix D.

3) The difference between successive order estimates gives the Transmission ($\hat{T}$). This represents the amount by which N consecutive events reduces the uncertainty of the event following, or, the amount by which the subsequent event uncertainty is reduced by inclusion of additional past events. Zero-order ($\hat{H}_0$), first-order ($\hat{H}_1 - \hat{H}_0$), second-order ($\hat{H}_2 - \hat{H}_1$), and third-order ($\hat{H}_3 - \hat{H}_2$) estimates, together with their respective χ^2 values, measuring the significance of the difference between orders, are shown in Table 7 (Overall), Table 8 (Within), and Table 9 (Between) for stationary sequences and Table 10 (Overall), Table 11 (Within), and Table 12 (Between) for non-stationary sequences.

Table 7 Order of sequential constraint - stationary sequences

Overall transitions

	zero-order ($\hat{H}_0$) 3df		first-order ($\hat{H}_1-\hat{H}_0$) 9df		second-order ($\hat{H}_2-\hat{H}_1$) 36df		third-order $\hat{H}_3-\hat{H}_2$) 144df	
Dyad	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2
2	.616	300.59[c]	.036	17.57[a]	.045	21.96	.120	58.58
3	.762	351.77[c]	.032	14.77	.034	15.69	.132	60.94
5	.803	647.88[c]	.066	53.25[c]	.030	24.20	.050	40.34
6	.687	572.39[c]	.114	94.98[c]	.028	23.33	.053	44.16
10	.558	331.86[c]	.091	54.12[c]	.066	39.25	.125	74.34
14	.573	512.36[c]	.171	152.90[c]	.025	22.35	.052	46.49
15	.415	319.29[c]	.243	186.96[c]	.078	60.01[b]	.066	50.78
16	.438	350.96[c]	.142	113.78[c]	.053	42.47	.096	76.92
17	.830	629.39[c]	.054	40.95[c]	.015	11.99	.082	62.18
19	.664	557.82[c]	.084	70.57[c]	.046	38.64	.052	43.68
20	.686	277.69[c]	.065	26.31[b]	.060	24.29	.076	30.76
21	.579	394.90[c]	.107	72.98[c]	.038	25.92	.079	53.88
22	.624	363.32[c]	.006	3.49	.116	67.54[b]	.087	50.66

Levels of significance

a $p = .05$

b $p = .01$

c $p = .001$

Table 8 Order of sequential constraint - stationary sequences

Within-speaker transitions

	zero-order ($\hat{H}_0$) 3df		first-order ($\hat{H}_1-\hat{H}_0$) 9df		second-order ($\hat{H}_2-\hat{H}_1$) 36df		third-order ($\hat{H}_3-\hat{H}_2$) 144df	
Dyad	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2
2	.637	144.82[c]	.526	119.59[c]	.522	53.55[a]	.010	0.47
3	.687	136.19[c]	.731	144.91[c]	.127	57.18[a]	.109	17.15
5	.977	392.78[c]	.618	248.45[c]	.047	9.19	.249	25.54
6	.739	274.56[c]	.590	219.20[c]	.265	49.23	.167	16.44
10	.637	197.81[c]	.755	234.40[c]	.012	2.08	.282	27.37
14	.697	218.37[c]	.594	186.10[c]	.280	38.82	.396	24.70
15	.443	133.27[c]	.851	256.00[c]	.177	21.35	0	0
16	.652	232.29[c]	.791	281.82[c]	.417	70.53[c]	.039	3.08
17	1.108	437.77[c]	.305	120.50[c]	.159	37.03	.059	8.34
19	.819	256.59[c]	.717	224.64[c]	.169	22.73	.103	5.71
20	.604	104.67[c]	.749	129.79[c]	.300	19.96	0	0
21	.578	151.44[c]	.607	159.04[c]	.042	4.08	.302	9.21
22	.679	202.38[c]	.563	167.80[c]	.331	55.52[a]	.227	20.77

Levels of significance

a $p = .05$

b $p = .01$

c $p = .001$

Table 9 Order of sequential constraint - stationary sequences

Between-speaker transitions

	zero-order ($\hat{H}_0$) 3df		first-order ($\hat{H}_1$-$\hat{H}_0$) 9df		second-order ($\hat{H}_2$-$\hat{H}_1$) 36df		third-order ($\hat{H}_3$-$\hat{H}_2$) 144df	
Dyad	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2
2	.626	162.28[c]	.132	34.22[c]	.251	31.32	.688	38.15
3	.827	216.68[c]	.169	44.28[c]	.352	41.48	.557	30.11
5	.675	272.30[c]	.040	16.14	.671	138.60[c]	.046	4.27
6	.681	313.43[c]	.166	76.40[c]	.626	115.42[c]	.527	46.11
10	.465	131.50[c]	.030	8.48	1.122	152.43[c]	.158	54.22
14	.541	313.49[c]	.393	227.73[c]	.434	75.81[c]	.227	17.31
15	.446	208.36[c]	.372	173.79[c]	.378	67.59[b]	.455	27.12
16	.319	141.51[c]	.163	72.31[c]	1.072	199.14[c]	.393	35.41
17	.606	219.27[c]	.119	43.06[c]	.947	153.60[c]	.268	24.52
19	.599	314.72[c]	.091	47.81[c]	.966	172.75[c]	.144	11.38
20	.742	170.75[c]	.021	4.83	.354	37.30	.715	24.78
21	.588	246.99[c]	.156	65.53[c]	.490	80.84[c]	.323	21.49
22	.601	169.97[c]	.023	6.50	.733	95.52[c]	.452	34.46

Levels of significance

a $p = .05$

b $p = .01$

c $p = .001$

Table 10 Order of sequential constraint - non stationary sequences

Overall transitions

	zero-order $(\hat{H}_0)$ 3df		first-order $(\hat{H}_1-\hat{H}_0)$ 9df		second-order $(\hat{H}_2-\hat{H}_1)$ 36df		third-order $(\hat{H}_3-\hat{H}_2)$ 144df	
Dyad	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2
1	.636	260.98[c]	.102	41.86[c]	.106	41.88	.144	80.76
4	.622	594.11[c]	.115	109.84[c]	.027	25.79	.040	38.21
7	.699	441.87[c]	.021	13.28	.100	63.22[b]	.108	68.27
8	.590	490.75[c]	.088	73.19[c]	.063	52.40[a]	.090	74.86
9	.407	352.08[c]	.139	120.24[c]	.070	60.55[b]	.103	89.10
11	.583	379.05[c]	.120	78.02[c]	.070	45.51	.102	66.32
12	.413	289.13[c]	.143	100.11[c]	.068	47.60	.119	83.31
13	.508	348.59[c]	.130	89.21[c]	.050	34.31	.086	59.01
18	.589	388.67[c]	.092	60.71[c]	.098	64.67[b]	.123	81.17
23	.591	367.05[c]	.073	45.34[c]	.061	37.88	.155	96.26
24	.412	243.88[c]	.164	97.08[c]	.112	66.29[b]	.107	63.34

Levels of significance

a $p = .05$

b $p = .01$

c $p = .001$

Table 11 Order of sequential constraint - non stationary sequences

Within-speaker transitions

	zero-order $(\hat{H}_0)$ 3df		first-order $(\hat{H}_1-\hat{H}_0)$ 9df		second-order $(\hat{H}_2-\hat{H}_1)$ 36df		third-order $(\hat{H}_3-\hat{H}_2)$ 144df	
Dyad	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2
1	.753	182.68[c]	.513	124.46[c]	.306	45.81	.182	17.66
4	.659	221.99[c]	.479	161.36[c]	.057	7.35	.239	12.26
7	.756	271.44[c]	.567	203.58[c]	.242	54.68[a]	.146	20.85
8	.894	365.61[c]	.619	253.15[c]	.362	86.32[c]	.033	4.85
9	.591	250.71[c]	.659	279.55[c]	.093	21.01	.570	71.12
11	.811	283.32[c]	.472	164.89[c]	.323	64.93[b]	.357	43.55
12	.45	135.99[c]	.797	240.86[c]	.102	13.43	.356	29.12
13	.518	156.55[c]	.751	226.96[c]	.186	28.11	.182	15.14
18	.839	286.12[c]	.718	244.86[c]	.095	19.23	0	0
23	.877	322.18[c]	.251	92.21[c]	.609	149.43[c]	0	0
24	.469	141.09[c]	.667	200.65[c]	.093	16.50	.512	54.65

Levels of significance

a $p = .05$

b $p = .01$

c $p = .001$

Table 12 Order of sequential constraint - non stationary sequences

Between-speaker transitions

	zero-order ($\hat{H}_0$) 3df		first-order ($\hat{H}_1-\hat{H}_0$) 9df		second-order ($\hat{H}_2-\hat{H}_1$) 36df		third-order ($\hat{H}_3-\hat{H}_2$) 144df	
Dyad	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2	$\hat{T}$	χ^2
1	.490	82.19^c	.294	49.32^c	.864	74.26^c	.159	8.82
4	.632	389.88^c	.266	164.10^c	.323	67.17^b	.334	25.93
7	.661	176.60^c	0	0	.850	113.12^c	.324	26.95
8	.424	178.69^c	.153	64.48^c	1.000	170.51^c	.348	31.84
9	.305	134.03^c	.168	73.83^c	1.050	206.69^c	.324	32.34
11	.419	125.47^c	.242	72.46^c	.794	117.77^c	.353	27.89
12	.450	178.42^c	.173	68.59^c	.673	105.43^c	.575	46.79
13	.555	212.35^c	.256	97.95^c	.470	71.02^c	.619	42.91
18	.424	134.60^c	.019	6.03	1.323	183.41^c	.118	9.16
23	.298	75.19^c	.261	65.85^c	.641	77.31^c	.503	34.87
24	.397	115.03^c	.045	13.04	.923	113.88^c	.236	16.69

Levels of significance

a p = .05

b p = .01

c p = .001

Analysed separately for stationary and non-stationary sequences, the results indicated that for stationary sequences, 11/13 (85%) had first-order overall-transitions, 9/13 (69%) had first-order within-speaker transitions, and 10/13 (77%) had second-order between-speaker transitions. For non-stationary dyads, 6/11 (55%) had first-order and 5/11 (45%) had second-order overall-transitions, 7/11 (64%) had first-order and 4/11 (36%) had second-order within-speaker transitions, and 11/11 (100%) had second-order between-speaker transitions.

A general tendency is suggested towards first-order within-speaker and second-order between-speaker transitions, also shown by the mean estimates for the stationary and non-stationary groups (Table 13).

Table 13

Order of sequential constraint

four-code model

	zero-order $(\hat{H}_0)$ 3df			first-order $(\hat{H}_1-\hat{H}_0)$ 9df			second-order $(\hat{H}_2-\hat{H}_1)$ 36df			third-order $(\hat{H}_3-\hat{H}_2)$ 144df		
	$\hat{T}$	χ^2		$\hat{T}$	χ^2		$\hat{T}$	χ^2		$\hat{T}$	χ^2	
Stationary sequences												
Overall	.633	434.38*	(6433)	.093	63.82*	(6433)	.048	32.94	(6433)	.083	56.96	(6433)
Within	.712	215.18*	(2842)	.650	196.44*	(2842)	.215	30.69	(1339)	.139	9.83	(663)
Between	.594	546.56*	(3591)	.144	55.10*	(3591)	.633	100.03*	(1482)	.381	27.47	(676)
Non-stationary sequences												
Overall	.582	400.99*	(5486)	.076	52.36*	(5475)	.072	49.60	(5475)	.111	76.48	(5475)
Within	.697	236.73*	(2751)	.591	200.72*	(2751)	.225	40.72	(1436)	.238	24.75	(825)
Between	.470	160.99*	(2735)	.174	61.03*	(2735)	.807	119.71*	(1177)	.355	27.06	(605)

Chi-square values marked * are significant at $p < .001$.

Values of N are shown in parenthesis.

Chatfield (1973) and Morgan (1976) have noted that the chi-square approximation becomes less tenable as the order increases because the sequences may be of insufficient length. However, treating all the stationary and non-stationary dyads as single sequences (as recommended by Morgan, 1976, p.31) yields identical results, indicating that the estimates of order are indeed acceptable. The maximum reduction in uncertainty regarding a subsequent act's identity, for the four-code model is obtained therefore, from using a single act for within-speaker transitions and a pair of antecedent acts for between-speaker transitions. In general, for between-speaker utterances, the second of the pair of antecedents adds very little new information in terms of conversational sequencing, the pair of events often being identical, and more usually the first event of the pair being an Offer of information.

If a process were completely random, and therefore lacking any order or patterning of events, then each approximation of $\hat{H}$ would equal $\hat{H}_{max}$. Although the value of $\hat{H}$ depends on the number of possible codes, it can be transformed into a number that varies from zero to one by dividing each successive $\hat{H}_n$ by $\hat{H}_{max}$, giving a value for the uncertainty in predicting a consequent event's identity. The complement of uncertainty is stereotypy (Miller & Frick, 1949), and may be calculated by (Gottman & Bakeman, 1979):

$$\hat{S}_n = 1 - \hat{H}_n/\hat{H}_{max}$$

where $\hat{S}_n$ = stereotypy value for the n^{th} order.

$\hat{H}_n$ = information value at order n.

$\hat{H}_{max}$ = maximum amount of information in the system.

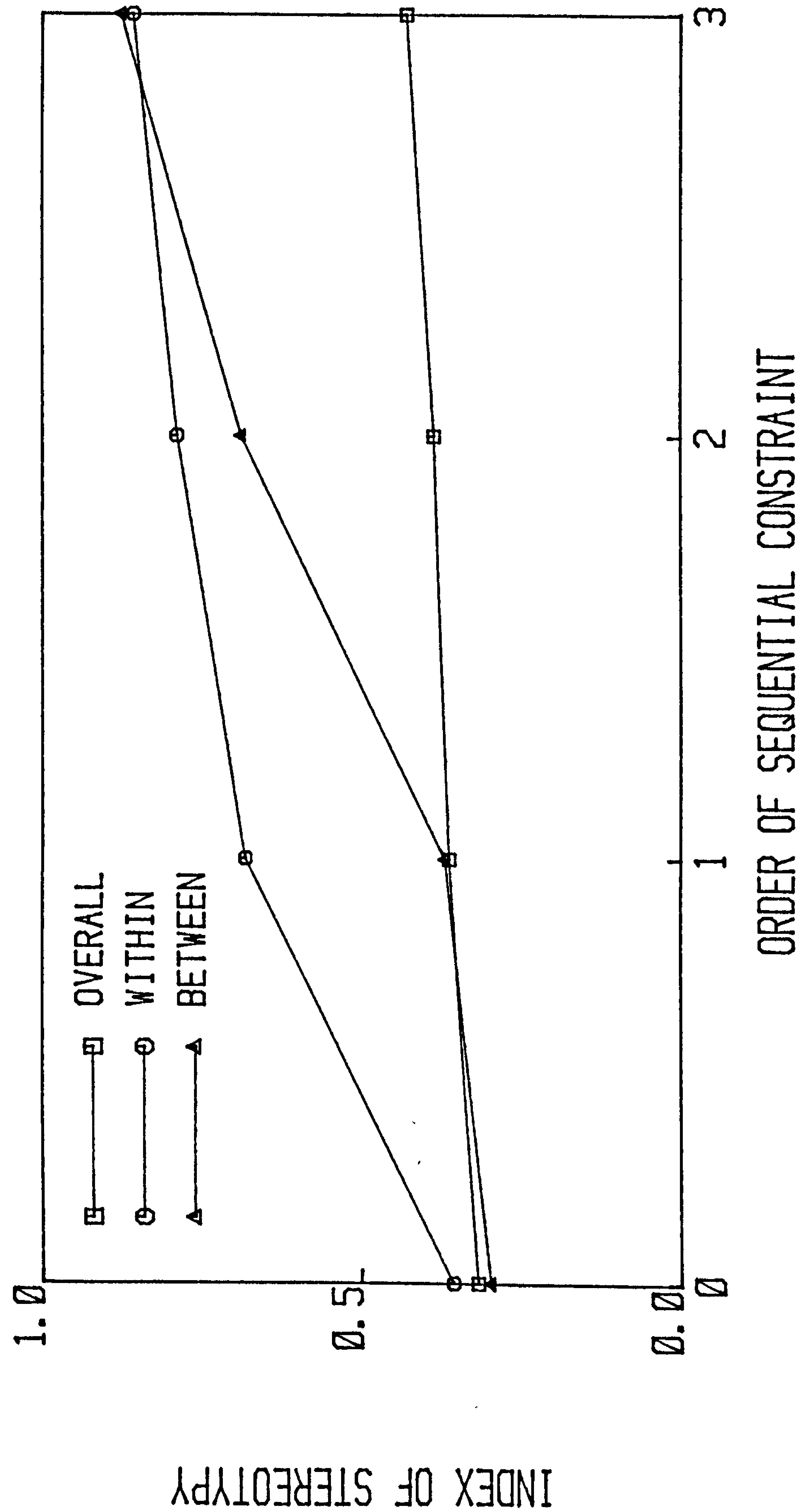
The index varies from zero to one, zero representing a random distribution of speech events, and one representing a completely stereotyped or predictable distribution of events. Stereotypy values have been calculated for the stationary and non-stationary groups (Appendix D) and are plotted in Graphs 1 and 2, respectively. The results suggest a considerable degree of patterning for within- and between-speaker events. The patterning of the speech stream will be more fully discussed in Chapters 6 and 9.

4.4.3 Homogeneity

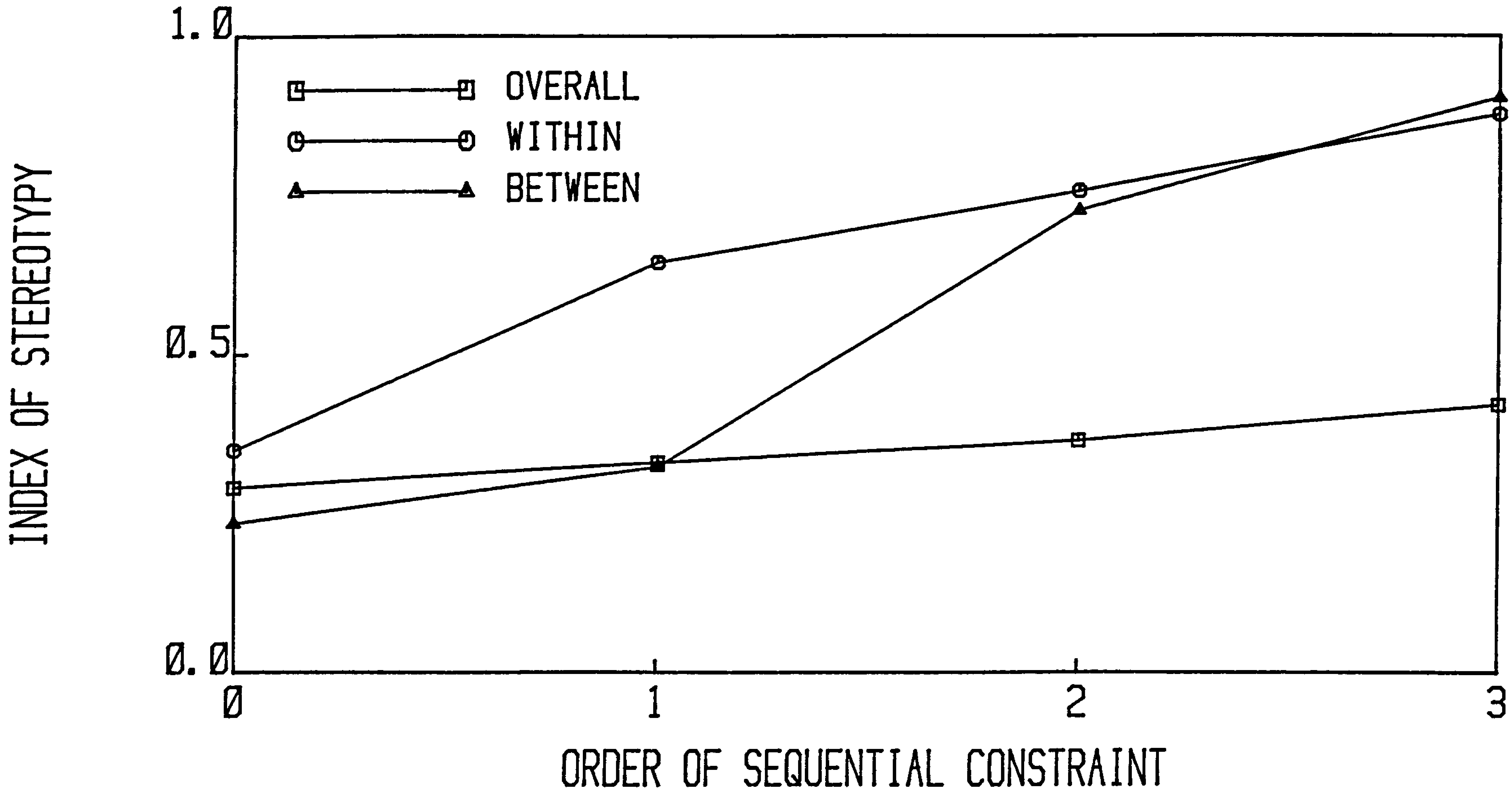
The assumption of homogeneity requires that no dyads or sub-groups of dyads have statistically different transition probabilities. The only known example of a test of homogeneity in the conversation analysis literature is by Hawes and Foley (1976), who found that two meetings of two academic committees were heterogeneous, but three meetings of a single group were homogeneous.

The procedure used to test the assumption of homogeneity in the present study was similar to that used to test the assumption of stationarity, whereby each speaker-specific dyadic matrix was compared in turn to a composite matrix derived from the summation of all the

GRAPH 1 INDEX OF STEREOTYPY
STATIONARY SEQUENCES



GRAPH 2 INDEX OF STEREOTYPY
NON-STATIONARY SEQUENCES



the dyadic matrices. (Program 6, Appendix G, was written specifically for the testing of sample homogeneity). The homogeneity assumption states that there are no significantly different transition matrices between the dyads, the null hypothesis being tested using the Anderson-Goodman maximum likelihood statistic (Anderson & Goodman, 1957). The degrees of freedom were calculated using:

$$m(m-1)(T-1)$$

where m = number of rows/columns in the transition matrices.

T = number of dyadic matrices compared.

The results (Table 14) indicated that the 24 dyads were not homogeneous. Extracting the stationary dyads, the homogeneity assumption was again tested, and showed that the 13 stationary dyads were also not homogeneous (Table 14).

TABLE 14

Homogeneity

Stationary and non-stationary sequences

Transition	LR	df	p
Overall	981.12	276	<.001
Within-speaker	429.31	276	<.001
Between-speaker	855.95	276	<.001

Stationary sequences

Transition	LR	df	p
Overall	512.84	144	<.001
Within-speaker	209.14	144	<.001
Between-speaker	459.37	144	<.001

4.4.4 A test of adequacy of the four-code model

The previous analyses suggest that the model-to-data fit is better for some dyadic sequences than for others. However, an important test of the adequacy of a stochastic model is how well is predicts future outcomes (Jaffe & Feldstein, 1970; Cappella, 1976); for predictive purposes the Markov chain has been shown to be fairly robust to violation of assumptions (Bartholomew, 1973, pp. 55-94; Hewes & Evans-Hewes, 1974). Such a test of adequacy has been used in this study. One procedure for testing the predictive power of a Markov model is to compare observed and predicted probability matrices. It can be shown for example, that the first-order Markov chain prediction for a matrix of n-step transition probabilities is equal to the observed one-step matrix raised to the n-th power (Finkbeiner, 1966; Jaffe & Feldstein, 1970):

$$M^{(n)} = (M^{(1)})^n$$

where n = number of transition steps.

$M^{(n)}$ = predicted matrix at the n^{th} power.

$M^{(1)}$ = observed one-step matrix.

The n-step, or lag transition matrices, representing the distribution of probabilities n events along the sequence may be computed from the sequences of raw data and compared to the corresponding powers of the one-step matrix. If the model predictions fit the data well, the elements of the observed matrices should not differ significantly from the values predicted by matrix multiplication, for each of the powers.

The terms 'n-step' and 'lag' have been used here. In the example in Figure 4, the first sequence shows transitions that are between adjacent events; they are known as one-step or lag 0 transitions. In the second example, the transition is between two events separated by an intermediate event. In this case, they are referred to as a two-step, or lag 1 transition.

<u>Figure 4</u> Illustrative example of sequential step (lag).

One-step transition (lag 0)

In each transition state i and state j occur in adjacent event units.

State sequence	=	A	B	A	A	B	A	B
1st transition (A to B)		i →	j					
2nd transition (B to A)			i →	j				
3rd transition (A to A)				i →	j			
4th transition (A to B)					i →	j		
etc.								

Two-step transition (lag 1)

In each transition state i is separated from state j by two event units.

State sequence	=	A	B	A	A	B	A	B
1st transition (A to A)		i ⟶		j				
2nd transition (B to A)			i ⟶		j			
3rd transition (A to B)				i ⟶		j		
4th transition (A to A)					i ⟶		j	
etc.								

Three programs were written for the test of adequacy, and may be found in Appendix G. The first program (Program 7) computes the observed n-step transition matrices and the second (Program 8) raises the observed one-step matrix to the n^{th} power. The third program (Program 9) is a comparator program that tests the difference between observed and predicted matrices at each of the n-steps, using the Anderson-Goodman likelihood ratio as a test statistic.

Using the likelihood ratio test to test for differences between observed speaker-specific and predicted speaker-specific matrices upto the tenth power (an arbitrary cut-off point, but considered sufficient by Cappella, 1975), no differences were found for any of the dyads, both stationary and non-stationary. The results of these ten tests for each of the 24 dyads are presented in Table 15.

Table 15   Comparison of observed and predicted transition matrices.

Lags 1 to 5.

	L.R. values				
Dyad	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5
1	29.57	34.37	29.94	41.98	35.55
2	29.42	25.72	30.95	28.11	31.57
3	21.08	20.93	29.68	16.68	21.07
4	27.91	21.62	27.37	28.48	24.86
5	21.56	14.80	19.76	29.05	22.15
6	35.08	29.16	27.19	36.97	37.74
7	28.44	28.69	27.97	32.38	32.76
8	29.28	52.05	38.80	32.42	40.50
9	35.01	40.20	28.74	33.48	23.85
10	38.42	27.21	26.44	21.95	28.74
11	26.89	35.64	41.51	36.75	31.01
12	23.76	35.40	39.78	37.68	34.36
13	24.14	30.99	38.19	30.54	26.81
14	20.76	20.45	30.55	19.30	21.87
15	30.63	33.62	42.94	45.20	37.80
16	41.28	46.16	29.47	30.05	41.40
17	15.85	34.40	33.42	24.79	27.16
18	28.21	39.39	36.26	26.40	29.64
19	44.27	36.27	32.71	33.68	32.86
20	24.48	26.93	26.55	27.31	23.13
21	19.61	21.07	21.46	19.25	24.86
22	22.23	21.94	18.96	29.31	26.74
23	25.91	41.59	37.03	44.72	35.57
24	31.10	38.32	36.06	53.82	35.12

All statistics with 56 df.
All comparisons between observed and predicted matrices are not significant.

Table 15 continued. Comparison of observed and predicted transition matrices.

Lags 6 to 10.

Dyad	L.R. values				
	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10
1	27.70	38.64	30.98	21.06	25.41
2	27.46	25.42	34.06	25.76	19.34
3	24.76	22.14	25.00	22.63	23.51
4	22.23	27.42	23.72	32.67	30.70
5	17.71	13.66	28.51	33.90	43.53
6	20.53	22.53	34.46	21.17	29.91
7	30.05	18.82	31.14	29.87	30.15
8	32.98	31.66	29.18	34.37	47.93
9	30.89	33.78	37.87	26.63	25.00
10	24.58	29.89	29.44	37.28	32.91
11	26.24	33.91	30.79	27.83	31.91
12	39.28	36.21	32.73	26.59	35.47
13	31.16	33.59	30.79	27.64	30.93
14	21.23	23.04	20.71	27.09	17.94
15	36.08	39.08	26.62	28.39	32.66
16	38.74	19.35	25.38	35.54	27.86
17	19.35	22.12	28.87	20.81	26.05
18	34.72	41.62	34.81	37.11	25.36
19	32.50	29.84	27.07	28.49	25.77
20	24.94	32.61	19.44	23.53	28.39
21	23.94	13.52	21.06	34.79	19.76
22	31.81	28.88	21.84	16.69	28.23
23	36.95	46.07	39.73	24.78	25.96
24	47.67	38.71	47.91	31.33	43.32

All statistics with 56 df.
All comparisons between observed and predicted matrices are not significant.

Such a result supports the appropriateness of a first-order process for the dyadic sequences. Since future outcomes were accurately predicted for the non-stationary sequences, the implication is that the four-code model is sufficiently flexible to minimise the effects of non-stationarity. Stationarity is not therefore considered to be an important factor when predicting future outcomes.

Earlier it was shown that the conversational sequences were hetergeneous. However, comparison of observed and predicted transition matrices for all 24 dyads treated as a single group (this is the same test that is recommended and used by Jaffe, Feldstein & Cassotta, 1967) indicated no significant differences for any of the powers (Table 16). Such a result indicates that although the sequences were initially shown to be heterogeneous, accurate predictions of future events may nonetheless be made.

<u>Table 16</u> Comparison of observed and predicted transition matrices for 24 dyads combined. Lags 1 to 10.

Lag	LR	df	p
1	6.58	56	n.s.
2	7.14	56	n.s.
3	5.78	56	n.s.
4	4.71	56	n.s.
5	4.30	56	n.s.
6	4.51	56	n.s.
7	3.27	56	n.s.
8	2.58	56	n.s.
9	3.08	56	n.s.
10	2.38	56	n.s.

The Anderson-Goodman test requires that all probabilities are greater than zero. In order that this assumption was not violated the initial model of conversation was restricted to four speech codes (eight speaker-states). However, as the Markov model is known to be fairly robust to violation of assumptions, and as the assumptions of sequence stationarity and sample homogeneity have been shown to be of little importance in predicting future outcomes, these tests would not, therefore, have to be performed for a larger model. Consequently, the next section is concerned with the development and subsequent replication of a seven-code (14 speaker-states) model of informal conversation.

4.5 A seven-code model of conversation

The expanded model used seven speech codes (Offer, Reply, Consent, Dissent, Modify, Reaction, Request) and in order to provide a replication of the findings, the analyses were initially confined to half the sample. These twelve dyads (1-12) were randomly transcribed and coded and so represent a random selection of dyads from the complete sample. The analyses of order and tests of adequacy were initially performed on dyads 1-12 as a group, and then for the remaining group of dyads (13-24).

4.5.1 Order of sequential constraint

Order of sequential constraint was assessed in the same manner as for the four-code model, the results (Table 17 - the information calculations are shown in

Appendix D) being similar, but not identical to those of the four-code model. In addition to the overall- and within-speaker transitions, between-speaker transitions were also found to be first-order.

In the same manner as the four-code analyses of order, an index of stereotypy has been calculated (Appendix D) and plotted graphically (Graph 3), and again indicates a large degree of patterning of events in the speech stream. The patterning of the speech stream will be more fully discussed in Chapters 6 and 9.

TABLE 17

Order of sequential constraint

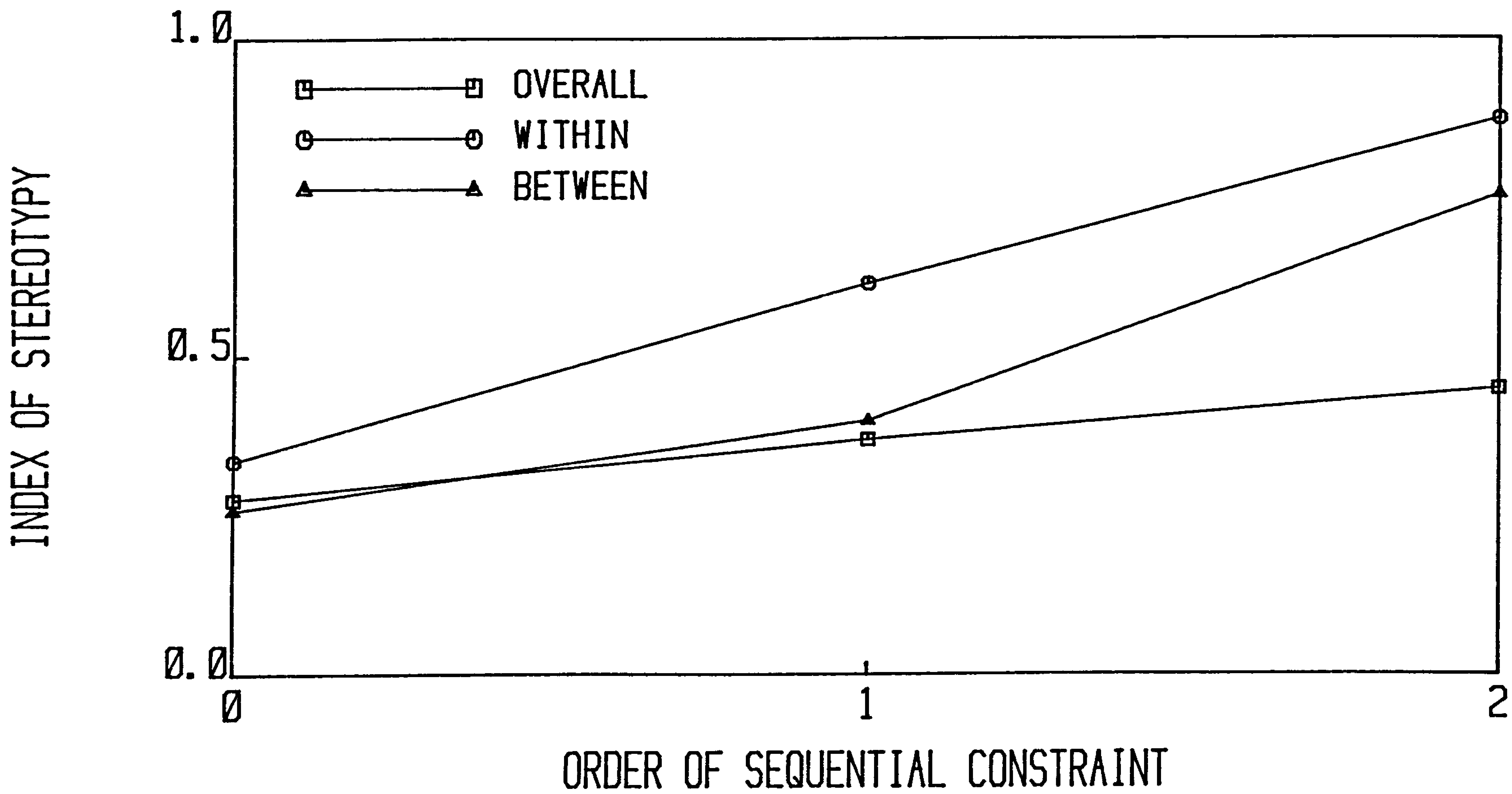
Seven-code model (dyads 1-12)

	zero-order (H_o) 6df			first-order (H_1-H_o) 36df			second-order (H_2-H_1) 252df		
	T	χ^2		T	χ^2		T	χ^2	
Overall	.77	591.41*	(6609)	.27	208.01*	(6609)	.22	168.43	(6609)
Within	.94	324.73*	(2977)	.79	272.20*	(2977)	.72	118.70	(1427)
Between	.72	302.51*	(3632)	.40	171.22*	(3632)	1.00	180.20	(1560)

Chi-square values marked * are significant at $p < .001$.

Values of N are shown in parenthesis.

GRAPH 3 INDEX OF STEREOTYPY
DYADS 1-12



4.5.2 A test of adequacy of the seven-code model

A test of adequacy of the model was carried out in the same way as for the four-code model, by comparison of observed speaker-specific and predicted speaker-specific matrices, for the first ten powers. The Anderson-Goodman test assumes that all probabilities are greater than zero. However, for the seven code analyses it was considered that transitions to a within-speaker Reaction, other than from itself, would not be logically possible. This is because, by definition, a Reaction is a vocalisation in response to the other person. Consequently, these transitions were considered to be logically zero (Colgan & Smith, 1978) and the number of degrees of freedom were reduced by one for each zero cell (Kullback, Kupperman & Ku, 1962).

The results (Table 18) showed no significant differences between observed and predicted outcomes for any of the powers, suggesting that a first-order model is tenable for the seven-code data.

Table 18 Comparison of observed and predicted transition matrices for dyads 1-12 combined. Lags 1 to 10

Lag	LR	df	p
1	15.47	170	n.s.
2	18.91	170	n.s.
3	19.10	170	n.s.
4	15.45	170	n.s.
5	13.95	170	n.s.
6	12.03	170	n.s.
7	12.35	170	n.s.
8	11.22	170	n.s.
9	9.75	170	n.s.
10	8.78	170	n.s.

4.6 Replication of the seven-code model results

Using the second half of the sample (dyads 13-24) as the replication data base, three types of analysis were performed: 1) order of sequential constraint, 2) a test of adequacy of the model, by comparison of observed and predicted transition matrices, and 3) comparison of observed transition matrices (lags 1-10) between dyads 1-12 and 13-24, to find if the two groups represented the same population.

4.6.1 Order of sequential constraint

Assessing the order of sequential constraint yielded results identical to those of the first group (dyads 1-12), and may be found in Table 19. Again an index of stereotypy has been calculated (Appendix D) and plotted graphically (Graph 4), and indicates, again, a large degree of patterning of events in the speech stream.

TABLE 19

Order of sequential constraint

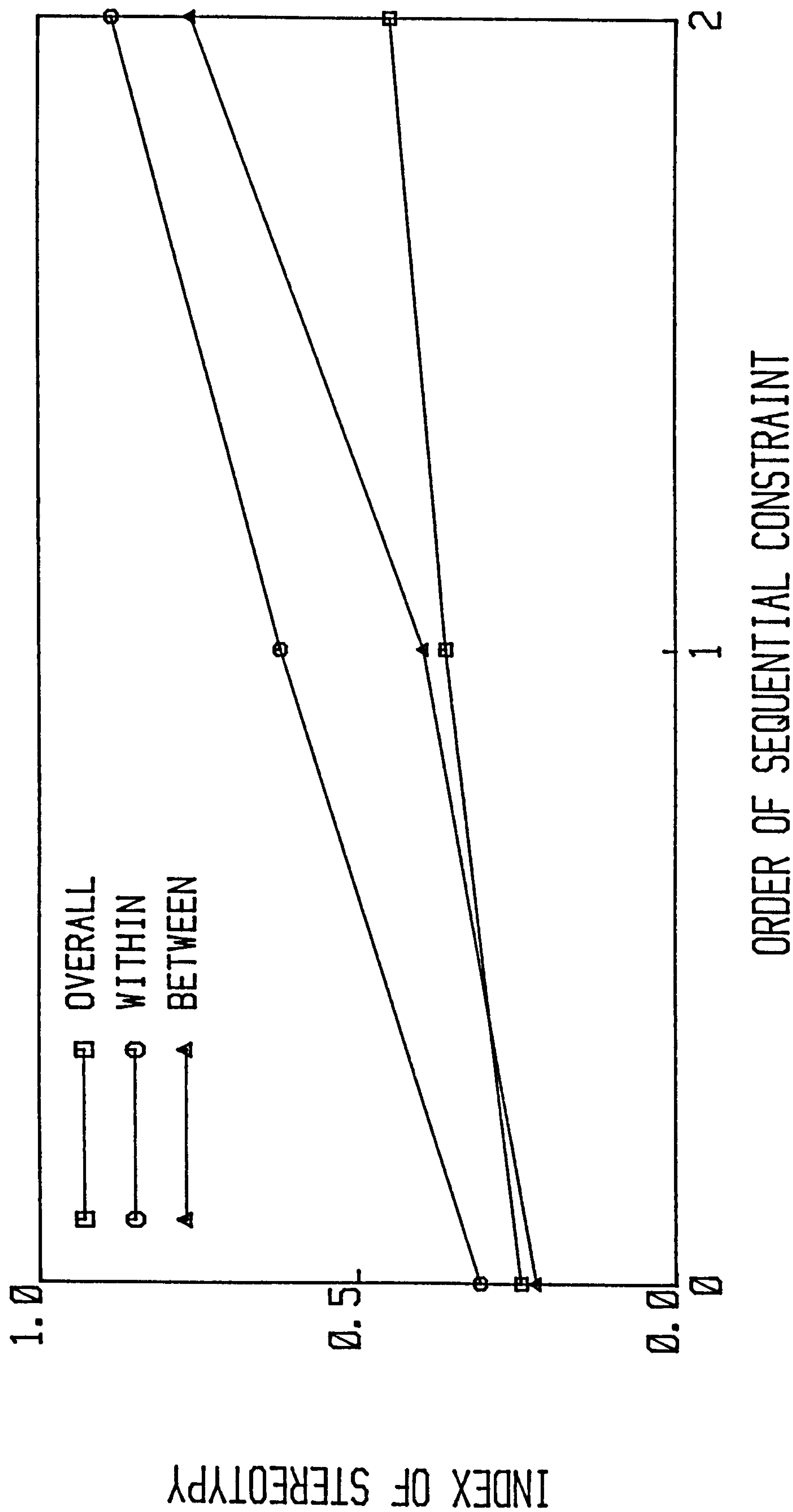
Seven-code model (dyads 13-24)

	zero-order (H_o) 6df			first-order (H_1-H_o) 36df			second-order (H_2-H_1) 252df		
	T	χ^2		T	χ^2		T	χ^2	
Overall	.684	589.88*	(7465)	.333	287.18*	(7465)	.247	213.01	(7465)
Within	.864	313.41*	(3140)	.886	321.40*	(3140)	.737	115.32	(1354)
Between	.617	308.28*	(4325)	.494	246.82*	(4325)	1.030	172.98	(1454)

Chi-square values marked * are significant at $p < .001$.

Values of N are shown in parenthesis.

GRAPH 4 INDEX OF STEREOTYPY
DYADS 13-24



4.6.2 A test of adequacy of the replication model

The second test, a test of adequacy of the model by comparison of observed speaker-specific and predicted speaker-specific transition matrices indicated no significant differences (Table 20), again suggesting that a first-order model is tenable for the seven-code data.

Table 20 Comparison of observed and predicted transition matrices for dyads 13-24 combined. Lags 1 to 10.

Lag	LR	df	p
1	21.06	170	n.s.
2	26.13	170	n.s.
3	26.57	170	n.s.
4	21.13	170	n.s.
5	19.38	170	n.s.
6	18.02	170	n.s.
7	16.50	170	n.s.
8	15.65	170	n.s.
9	14.98	170	n.s.
10	13.61	170	n.s.

4.6.3 Comparison of dyad groups 1-12 and 13-24

Finally, comparison of the observed (lag 1-10) matrices for the two dyad groups (1-12 and 13-24) showed no significant differences (Table 21) and indicates that the two groups of dyads represented the same population.

Table 21 Comparison of observed transition matrices for dyad groups 1-12 and 13-24. Lags 1 to 10.

Lag	LR	df	p
1	29.04	170	n.s.
2	26.46	170	n.s.
3	28.27	170	n.s.
4	23.95	170	n.s.
5	25.03	170	n.s.
6	23.80	170	n.s.
7	24.61	170	n.s.
8	21.76	170	n.s.
9	23.01	170	n.s.
10	22.24	170	n.s.

4.6.4 Replication study conclusion

The overall conclusion drawn from these analyses was that the two groups of dyads had the same order of sequential constraint and represented the same population, the results of the analysis of dyads 1-12 being replicable. Consequently, the two groups of dyads were combined and a probability and frequency distribution for the seven code model constructed. This is shown in transition matrix form in Table 22 and constitutes a stochastic representation of the way information is exchanged in informal dyadic conversation.

TABLE 22

A seven-state model of informal dyadic conversation

Within-speaker transitions

	O	Rp	C	D	M	Rt	Rq
O	.43 (2947)	.003 (24)	.01 (30)	.004 (30)	.004 (25)	0 (0)	.03 (207)
Rp	.15 (116)	.29 (240)	.03 (25)	.03 (24)	.01 (8)	0 (0)	.04 (27)
C	.24 (558)	.02 (36)	.15 (388)	.02 (31)	.04 (112)	0 (0)	.04 (91)
D	.22 (146)	.02 (10)	.01 (4)	.25 (167)	.02 (14)	0 (0)	.02 (16)
M	.27 (96)	.008 (3)	.04 (11)	.02 (4)	.16 (56)	0 (0)	.06 (23)
Rt	.13 (199)	.003 (7)	.04 (57)	.03 (47)	.01 (14)	.02 (32)	.03 (43)
Rq	.07 (74)	.004 (4)	.001 (1)	.001 (1)	.001 (1)	0 (0)	.11 (106)

Between-speaker transitions

	O	Rp	C	D	M	Rt	Rq
O	.12 (783)	.007 (53)	.16 (1317)	.03 (187)	.01 (61)	.16 (1237)	.03 (242)
Rp	.17 (137)	.005 (4)	.12 (93)	.02 (17)	.01 (10)	.10 (81)	.06 (49)
C	.37 (927)	.009 (23)	.04 (113)	.005 (13)	.01 (15)	.02 (55)	.03 (97)
D	.13 (81)	.04 (24)	.09 (63)	.08 (53)	.01 (9)	.08 (38)	.05 (32)
M	.13 (45)	.01 (3)	.15 (50)	.04 (16)	.02 (7)	.09 (39)	.02 (7)
Rt	.65 (1001)	.01 (16)	.01 (14)	.02 (23)	.01 (22)	.01 (18)	.03 (40)
Rq	.07 (71)	.37 (384)	.23 (255)	.05 (44)	.01 (6)	.04 (33)	.05 (49)

The transition matrices show the probability of occurrence for all possible speech-code transitions. Due to rounding they do not sum to 1. N values are shown in parenthesis.

Note: Offer (O), Reply (Rp), Consent (C), Dissent (D), Modify (M), Reaction (Rt), Request (Rq).

4.7 Discussion

In order to assess the appropriateness of a Markov process as a model of conversation, three questions were posed at the close of the Introduction to this chapter concerning the assumptions of sequence stationarity, order of sequential constraint, and sample homogeneity. In general, the sequences tended towards being first order, stationarity was not a universal finding (confirming Lewis, 1970b), and neither the stationary nor non-stationary groups could be considered to be homogeneous.

Markovity has frequently been assumed in the literature, even though assumptions have remained untested (e.g. Hawes & Foley, 1973), or have been assessed using inappropriate methods (e.g. Lichtenberg & Hummel, 1976; Sackett, 1978) or test assumptions have been violated (e.g. Donohue, Hawes & Mabee, 1981). Using appropriate statistics, the present study, in similarity to Hawes and Foley (1976) and Lewis (1970b), has demonstrated that these assumptions are not always satisfied. However, the results of the model adequacy tests suggest that a Markov process is flexible enough to absorb at least some violation of the stationarity and homogeneity assumptions. Consequently, having analysed the conversational data by methods that preserve the process, relational and structural aspects of conversation, the seven-state and replication results indicate that a first order Markov process is a tenable model for information transmission in informal conversation, when coded using CEA. The transition matrices shown in Table 22 may, therefore, be considered to be an adequate stochastic representation of how information is exchanged in informal conversation.

This is an important result as it confirms the nature of the conversational model implicit, yet usually untested or tested inappropriately, in most of the stochastic analyses of conversation reviewed. Such a model is also assumed in the deterministic literature, usually in the form of a Markov-like process of a single event determining the occurrence of the next event (e.g. Auld & White, 1956; Scheidel & Crowell, 1964; Gouran & Baird, 1972; Hawes & Foley, 1973; Stech, 1975).

It is important to note at this point that the order of sequential constraint of a conversational sequence is very much dependent on the type of conversation analysed and the methods used to segment and classify the speech stream. For example, the present analysis, and much of the work reviewed, takes a 'thought' or an 'idea' as the unit of analysis, the majority of analyses indicating the order of sequential constraint to be first order. However, segmenting the speech stream into smaller units, such as individual words, may well yield higher orders of sequential constraint (Attneave, 1959). Similarly, using methods of analysis whereby subjects are asked to 'reconstitute' speech (e.g. Clarke, 1975) may yield higher orders of sequential constraint.

The overall conclusion therefore, is that a first-order Markov process is an appropriate model for informal conversation, and, as indicated in the Introduction to this chapter (Section 4.1), such a process can be used as a tool for describing the organisation of conversational events. Indeed, the stereotypy indices shown in Graphs 3 and 4 indicate a moderate degree of patterning of events in the speech stream. The next step therefore, is to identify the speech events which repeatedly occur together, and those events important in the conversational turn-taking mechanism, as this type of information can be used to construct conversational strategies that may enhance our present knowledge of conversational interaction; this forms the basis of Chapter 6.

CHAPTER 5

THE CHARACTERISTICS OF NON-STATIONARY SEQUENCES

5.1 Introduction

In Chapter 4, it was demonstrated that dyads often produce conversational sequences with non-stationary transition probabilities. In the sample used, 11/24 of the four-state sequences were found to be non-stationary, although the non-stationarity was later found to be an unimportant factor when predicting future outcomes. The theme of this chapter is concerned with assessing whether those sequences found to be non-stationary in Chapter 4, differ in any other respect from stationary sequences. Any differences may give an indication as to how non-stationarity affects the general character of the conversation sequence.

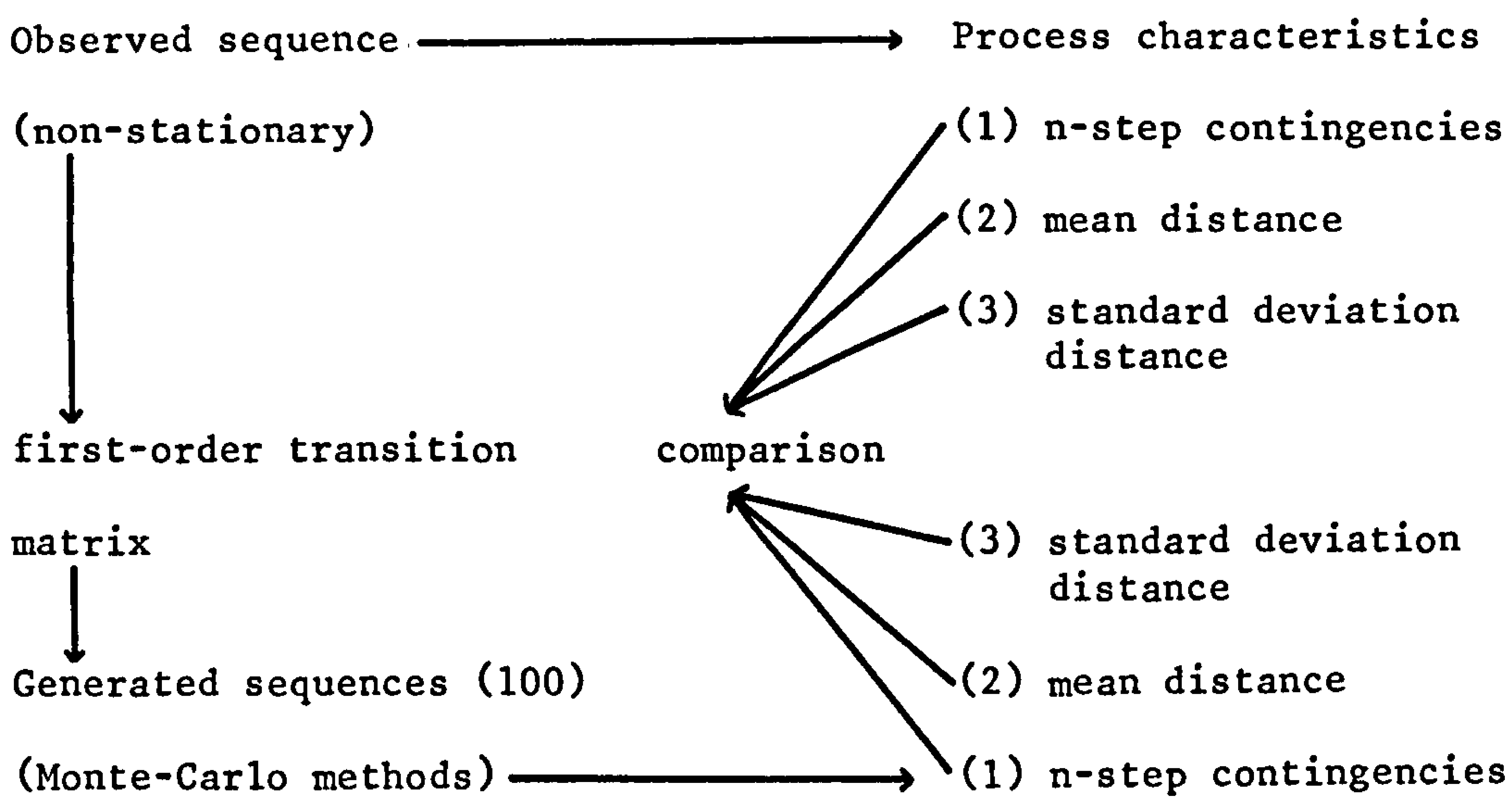
Kemeny and Snell (1960) have indicated that, in addition to a transition matrix, sequences may be described by a number of process characteristics, which may be used to outline specific interactions in terms of the sequential relationships between events (Hertel, 1968). Kemeny and Snell (1960) specify a number of process characteristics that can be calculated from a sequence of events, the most meaningful, in psychological terms (Hawes & Foley, 1976), being: (1) mean distance - the mean of the number of transition steps necessary to move from one state to any other, (2) standard deviation distance - the standard deviation of the number of transition steps necessary to move from one state to any other, and (3) n-step contingencies - the probability of moving from one state to another in a specified number of steps. (Hawes & Foley (1976) have suggested that the mean (1) and standard deviation (2) distance measure reflect the length of conversational phases, and the n-step contingencies (3) reflect the intensity of the phases.)

The aim of this chapter is two fold: 1) the assessment of the conversational sequences found to be non-stationary in Chapter 4 in terms of the three process characteristics specified, and 2) specification of the effects of non-stationarity on the character of the conversations.

5.2 Method

The method by which comparison of process characteristics derived from stationary and non-stationary sequences was accomplished, is as follows. Initially, for each non-stationary dyadic sequence, three process characteristics were calculated: n-step contingencies, and mean and standard deviation distance between events. Using the first-order transition matrix of the observed non-stationary sequences, one hundred stationary sequences of the same length were then generated by Monte-Carlo methods, and again the three process characteristics were calculated from the sequences. The observed and mean generated process characteristics were then compared using the Anderson-Goodman likelihood ratio test as a test of difference (Anderson & Goodman, 1957). A flow diagram of the method used is shown in Figure 5.

Figure 5 Method used for comparison of observed and generated process characteristics.



5.2.1 Sequence generation using Monte-Carlo methods

Sequence generation started initially from the first-order speaker-specific transition matrix of the non-stationary sequence.

In this matrix the rows indicate the probability of occurrence of those events which were observed to follow a given antecedent event. Each sequence was generated using the information provided by each row of the matrix, determination of each event in the generated sequence being a two-stage process. Firstly, a random number was generated, and secondly, this was used to select one of the events in a given row of the transition matrix.

Random number generation was accomplished using a computer sub-routine, producing numbers in the range 0 - 999, each number being used to select a particular speech code in the following manner. The probabilities of each subsequent act given in the transition matrix were transformed into non-overlapping ranges of numbers in the range 0 - 999, the size of the range being determined by the size of the probability (adjustments were made to the size of the number range, if the row probabilities did not sum to exactly 1.0). An example of the range transformation is shown in Figure 6.

<u>Figure 6</u> Range transformation and event selection.

Matrix probabilities

		Speaker 1		Speaker 2	
		A	B	A	B
Speaker 1	A	.2	.3	.3	.2
Ranges		0-199	200-499	500-799	800-999

Each category of speech was selected when a random number fell in range. For example, assuming a conversation started with speaker 1 uttering act type A, if the random number generated was 936, the subsequent event selected would be a B-type act from speaker 2. This subsequent act would then become the antecedent act for the next subsequent act selection. Since each act was determined only by the row probabilities of its antecedent and the row probabilities were not changed during the sequence generation, the sequence generated was purely Markovian and stationary.

Three constraints were placed on the generated sequences and were included in order to facilitate a more precise model of the observed sequence. The three constraints were:

(1) Each of the 100 sequences generated were the same length as the observed sequence.

(2) The start state of the generated sequence was the same as that of the observed sequence.

(3) The end state of the generated sequence was the same as that of the observed sequence.

These constraints ensured that the generated sequence bore some relation to the observed sequence being modelled. For example, if the generated sequence was started in any speech state and was of a different length, the simulation would have nothing in common with the sequence being modelled.

The Monte-Carlo generator program listing is given in Appendix G (Program 10).

5.2.2 Calculation of process characteristics

The following three process characteristics were calculated for both the observed and generated sequences.

(1) Mean distance between all possible two-event combinations of states. Program 11 (Appendix G) was written to calculate mean distance measures, and does so by calculating the number of intervening events between each two-act combination of events throughout the sequence.

(2) Standard deviation distance between all possible two-act combinations of states. (Program 11, Appendix G).

(3) N-step contingencies for all two-act combinations of states. Program 7 (Appendix G) was used to calculate the probability of moving from one state to another in a specified number of steps (n-step contingency), for step sizes 1 to 10.

For the generated sequences, the mean of the 100 sequence process characteristics was calculated and used in the subsequent observed/generated sequence comparison.

5.2.3 Comparison of observed and generated process characteristics

For each non-stationary dyadic sequence, comparisons between observed and generated process characteristics were made using the Anderson-Goodman likelihood ratio test as a test statistic. The test was used in the same manner as for the tests of homogeneity in Chapter 4 (Section 4.4.3). A listing of the comparator program may be found in Appendix G (Program 12). For each test, the null hypothesis stated that there would be no statistically significant difference between the observed and generated sequence process characteristics. Similar to the tests of homogeneity, a significant LR value indicates that observed and generated process characteristics were statistically different. A non-significant LR value indicates that even when the sequence has been shown to be non-stationary, the process characteristics do not differ significantly from those one would expect to be generated by a stationary Markov chain.

5.3 Results

The results of the observed/generated process characteristic comparisons are shown in Tables 23 (n-step contingencies) and 24 (mean and standard deviation distance measures), and indicate that for all of the non-stationary dyads, there appears to be no significant difference between n-step contingencies (lags 1 to 10) derived from observed non-stationary and generated stationary sequences. The results of the distance measure analyses indicate that 6/11 non-stationary sequences have mean inter-event distances and 4/11 non-stationary sequences have standard deviation inter-event distances that differ significantly from those derived from generated stationary sequences.

Table 23 Comparison of n-step contingencies derived from observed and generated conversational sequences.

	L.R. values				
Dyad	step 1	step 2	step 3	step 4	step 5
1	27.72	32.87	32.04	43.74	37.38
4	31.03	22.97	31.39	28.61	24.98
7	30.82	30.75	28.08	31.74	33.57
8	31.69	53.82	42.76	35.97	39.61
9	35.11	40.79	35.41	34.50	22.55
11	34.86	40.24	34.09	31.99	23.32
12	25.95	39.76	43.75	41.21	39.12
13	26.11	33.17	40.96	33.09	26.54
18	28.22	42.33	38.35	31.21	29.47
23	26.29	44.72	35.97	43.99	33.89
24	32.34	34.94	32.82	52.75	34.15

	step 6	step 7	step 8	step 9	step 10
1	29.42	36.48	28.19	21.59	25.89
4	24.17	27.32	26.00	31.29	31.36
7	29.47	18.62	32.23	28.03	28.89
8	34.13	32.43	31.85	34.23	51.37
9	32.45	32.99	37.34	26.83	26.30
11	34.60	35.37	36.25	26.08	26.76
12	46.11	39.72	36.66	28.49	36.54
13	33.03	34.82	33.91	30.11	33.40
18	33.34	43.31	34.22	36.12	26.81
23	33.12	43.23	38.65	24.90	27.47
24	48.35	40.36	50.62	32.77	44.87

All comparisons with 56 df.
All comparisons not significant.

Table 24 Comparison of mean and standard deviation distance measures derived from observed and generated conversational sequences.

	L.R. values	
Dyad	Mean distance	S.D. distance
1	84.16**	63.07
4	193.34***	185.39***
7	41.35	62.66
8	85.77**	83.98**
9	49.19	56.28
11	99.96***	76.98*
12	31.09	32.64
13	414.29***	362.97***
18	23.33	28.56
23	81.03*	60.29
24	30.19	20.32

All comparisons with 56 df.

L.R. values marked * are significant at $p < .05$, ** $p < .01$, *** $p < .001$.

5.4 Discussion

In Chapter 4 (Section 4.5.2), it was demonstrated that a transition matrix calculated by matrix multiplication did not differ significantly from the observed transition matrix derived from the conversational sequence, even when the sequence of events had been shown to be non-stationary. From this, it was concluded that a Markov model was sufficiently robust to withstand at least some violation of the stationarity assumption. The present n-step contingency analyses indicate that transition matrices, at various event-step sizes, derived from a non-stationary sequence and from simulated stationary sequences based on the same transition matrix also show no significant differences. The conclusion one may draw from such a result is that although a conversational sequence may not be stationary, overall, event-to-event transition probabilities are statistically no different from those found in a stationary sequence.

However, in 6/11 cases significant differences in sequence structure are apparent. The difference between stationary and non-stationary sequences appears to lie not simply in changes in the probability of occurrence of a specified antecedent/subsequent transition over time, but in the manner in which the transitions are arranged in the sequence. For example, inspection of the mean inter-event distances for each of the six sequences which exhibited significantly different sequence characteristics (Appendix E), it would appear that for the Dissent and Request states, inter-event distances with all other states are almost invariably greater for the generated sequences. This suggests that, whereas in the generated sequences the Dissent and Request states are evenly distributed in the sequence, this is not the case for the observed sequence. In the observed conversational sequence, these states are more likely to be found in clusters (which is indicated by the considerably smaller observed inter-event distances - Appendix E), and perhaps towards one end of the conversation (which is indicated by the sequence being non-stationary).

The results of the process characteristic analyses therefore suggest that, whereas the probability of antecedent/subsequent event transitions in a non-stationary sequence do not differ from those found in a stationary sequence (n-step contingency analyses), the non-stationarity is due in a large proportion of cases (55%) to the manner in which events are distributed in the conversation. Stationary sequences are therefore characterised by event-to-event transitions being approximately equally distributed across the length of the conversation, whereas non-stationary sequences are more likely to be characterised by clusters of events, typically Dissents and Requests that occur towards one end of the conversation. Generally, Requests and Dissents tended to be concentrated in the first half of the conversation. Of the 196 Requests in the six sequences, 77.04% occurred in the first half of the conversation. Similarly, of the 177 Dissents in the six sequences, 67.23% occurred in the first half of the conversation. The same would also appear to be true for those non-stationary sequences whose distance measures did not differ significantly from those derived from the simulated sequences as inspection of the process characteristic matrices (Appendix E) indicates a similar pattern of deviations, although considerably less marked, and less consistent.

The overall conclusion that can be drawn from this study of non-stationary conversational sequences using Monte-Carlo simulation techniques and involving comparison of three process characteristics, is that the character of non-stationary sequences tends to be different from that of stationary sequences. Although non-stationary and stationary sequences differ, it tends to be not simply due to the antecedent/subsequent transition probabilities changing over time, but in addition, the manner in which events are distributed in the conversational sequence. Whilst stationary sequences have a uniform distribution of events, non-stationary sequences are characterised by phases, or a clustering of events; typically, Dissents and Requests tend to be concentrated in the first part of the conversation.

CHAPTER 6

TURN-TAKING AND STRATEGIES IN INFORMAL CONVERSATION

6.1 Introduction

The analyses presented in Chapter 4 have indicated that the events in informal conversation have a moderate degree of organisation (assessed by the information analyses and stereotypy indices), the stochastic analyses indicating that the 24 conversations combined can be conceived of as a first-order Markov process. These analyses however, yield only very general information about the organisation of conversation. Consequently, the theme of this chapter is to use the stochastic model presented in Chapter 4, as a basis for describing the organisation of conversational events in more detail. Specifically, the transition matrices presented in Table 22 (Chapter 4), have been used to identify the speech acts which repeatedly occur together and those acts important in the turn-taking process. This information has then been used to suggest a number of conversational strategies.

Two sets of analyses are presented in this chapter. The first is concerned with the role of speech content in conversational turn-taking, and examines the relationship by comparing observed with expected frequencies for speech events occurring prior to a change of speaker. The second set of analyses use graphical methods, interpreting the model of conversation derived in Chapter 4 (Table 22) in terms of conversational strategies.

6.2 Conversational turn-taking

There is a considerable literature concerned with the role of a variety of behaviours considered to be 'signals', 'cues', and 'predictors' of a change in speaker. For example, Kendon (1967) has suggested that conversational interchanges between a speaker and a listener are regulated, in part, by gaze, although this hypothesis has attracted considerable criticism in recent years (e.g. Rutter et al., 1978; Beattie, 1978; Lazzerini et al., 1978).

The possible role of posture in turn-taking has also been considered. For example, Scheflen (1964) has been interested in the way non-verbal markers structure conversation, and has suggested that body motion and speech are integrated at three different levels of organisation: these he calls 'point', 'position', and 'presentation'. A 'point' corresponds to making a point in discussion and tends to be indicated by a change in head posture. If the point is in question form, the accompanying change in posture is a raising of the head (Scheflen, 1964; Thomas & Bull, 1981). Similarly, in a study of pre-school children, DeLong (1974) noted that two kinesic signals (a leftward movement of the head and a dropping of the head and/or arms and hands), consistently indicated the termination of a child's utterance.

Using a multi-channel approach to the analysis of turn-taking, Duncan (1972) was able to formulate a system of turn-taking signals and rules, the signals taking a variety of forms, such as syntax, intonation pattern, paralanguage and body motion. He found that the display of a 'turn-yielding' signal, consisting of the termination of any hand gesticulation and/or a drop in paralinguistic pitch usually relinquished the speaking turn and the display of a 'speaker-state' signal, usually consisting of a shift away in head direction accepted and secured the speaking turn (Duncan, 1974; Duncan & Niederehe, 1974; Duncan & Fiske, 1977; Duncan, Brunner & Fiske, 1979). By way of extension to Duncan's system, Wiemann and Knapp (1975) have demonstrated that a number of cues are available as 'turn-requesting' signals, gaze and head nods being the most important.

In a review of the turn-taking literature, it became apparent that turn-taking cues were generally considered to be non-verbal (e.g. gaze, posture, intonation, or pitch) with very little work concerned with the role of speech content in the turn-taking process. In fact all the research reviewed simply divided the speech into 'utterances', representing a person's complete speaking turn, and ignored the content of the utterances. However, in discussing the use of 'adjacency pairs' as a method of 'getting in and out of' a conversation, Sacks (1972) acknowledges that the content of speech does have a role to play in floor apportionment. The most obvious example would be a question, which functions by offering the speaking-turn to another person. Similarly, in a study of the role of posture change in interaction, Thomas and Bull (1981) demonstrated not only that there was a strong relationship between the type of speech uttered and specific changes in head posture, but that these were also related to changes of speaker. They noted that "it is therefore unlikely that turn-taking signals (in any modality) would have been successfully demonstrated in past research, because, by segmenting speech only into 'long utterances', rather than specific types of speech, the finer relationships were lost"; they argued that speech content, although often neglected, may have an important role to play in the process of floor-apportionment.

Stech (1975) was concerned with constructing a grammar for conversation, based on speech content; he tested a number of hypotheses, derived from a set of propositions, concerning the location of four types of speech in conversational sequences. The four types of speech were statements, questions, agreements and disagreements. The general conclusions drawn from the anlyses were that (1) closed questions were followed by agreements and disagreements, (2) agreements and disagreements tended to follow closed questions and statements, (3) questions and statements represented initiator acts in sequences, and (4) statements and agreements represented closing acts in sequences.

Although Stech's results provide some useful information about conversational sequencing, there are a number of methodological problems with his work. For example, Stech used segments of conversation taken from the Watergate transcripts (two over the telephone and one face to face), three marital decision-making transcripts from three different couples, and an unspecified number of role-played police-student telephone calls. Three points concerning the nature of the data base are important here. Firstly, none of the samples are particularly representative of naturally occurring conversation: the police-student calls and the Watergate interactions are by nature highly structured, resulting from the interrogative nature of the situations. Secondly, each sample, consisting of a number of separate interactions, was assumed to be homogeneous, although no tests of sample homogeneity were reported, and thirdly, it may not be appropriate to assume that telephone and role-play interactions are the same as naturally occurring face-to-face interactions. For example, a number of authors have indicated that telephone and face-to-face conversations differ in several respects; Rutter and Stephenson (1979) have suggested that they differ in content, style and outcome and an experiment by Stephenson, Ayling and Rutter (1976) comparing the two communication media found that telephone conversations tended to be more task oriented and more depersonalised than face-to-face conversations. In an earlier study, Rutter and Stephenson (1977) found telephone conversations to be less spontaneous, and to contain fewer interruptions (and therefore longer utterances) than face-to-face conversations. Similarly, Wilson (1974) found that there were differences in the content of the conversations, telephone conversations being characterised by more agreements and disagreements of opinion and more phatic utterances whilst Stephenson, Ayling and Rutter (1976) and Morley and Stephenson (1977) found that there was more offering of information and less praise in telephone calls. It would seem inappropriate therefore, to assume that telephone conversations are representative of naturally occurring informal conversations.

Role-played conversations pose similar problems, particularly with regard to their artificiality. Mehrabian (1968), for example, has questioned whether we are at liberty to assume that role-play corresponds to what people do in a naturally occurring situation, and from a review of a number of studies in which role-play and deception were compared, Miller (1972) concluded that the main objection to role-play was that subjects often did not know what behaviours to display for a given situation. Similarly, the use of role-played conversation is open to the same criticism.

In Stech's study, speech was coded into four categories, statements, questions, agreements and disagreements, for which inter-rater reliabilities of between 86% and 96% agreement are given. In all cases, however, coding was based only on transcripts which were not marked with intonation patterns. Both Waxler and Mishler (1966) and Morley and Stephenson (1977) have indicated that category assignment is considerably enhanced by the use of audio tapes, because there are many instances where intonation and pitch change are important for the classification of speech. Given the following example:

A: Jane's going to Bradford tomorrow/
B: Is she/

B's response would be difficult to classify using just a transcript as it could be classified as either a Request or a Reaction. However, given a transcript plus an audio tape, classification is made much simpler, a low-falling pitch corresponding to a Reaction, whereby no explicit appeal is made to the listener (Gimson, 1962), and a high-rising pitch corresponding to a Request (Gimson, 1962). Similarly, agreement and disagreement are not always easily discriminable, Stech giving the example of 'unh - huh' for both an agreement and a disagreement. In the following example, using a transcript alone, classification of B's response would not be possible:

A: Well, teachers do get paid too much/ and they ...
B: unh - huh/

However, knowing the pitch change associated with B's response considerably simplifies the classification of the speech, as agreements and disagreements tend to be associated with different changes in pitch. Agreement and satisfaction tend to be indicated by a large step-up in pitch, disagreement and dismay with a large step-down in pitch (Crystal, 1969). Consequently, although Stech reports a high degree of consistency in category assignment, without the use of an audio tape in addition to the transcript, it is possible that although category assignment was consistent, it was consistently incorrect.

The analyses to be presented in this section are aimed at examining the relationship between speech content and observed changes in speaker and are designed to overcome the shortcomings apparent in Stech's work.

6.2.1 Method and Results

Two analyses have been performed on the Chapter 4 sample of 24 dyadic conversations: the first is concerned with identifying those speech categories which consistently precede speaker changes, the second with identifying those speech codes which consistently follow a change in speaker. Statistical assessment was by means of the chi-square goodness of fit test (with Yates' correction for tables with 1 degree of freedom).

One possible application of the goodness of fit test would be to determine whether the location of speech categories prior to and following a change of speaker departed from equiprobable occurrence. However, because there were more between-speaker acts, comparison of observed values with an equiprobable distribution would be incorrect. Consequently, as the within- and between-speaker acts were observed to occur in the ratio of 1 : 1.306 for the present data, expected frequencies have been calculated using the same ratio.

The results of the analyses performed are summarised in Tables 25 and 26 for speech codes which precede and speech codes which follow changes in speaker, respectively.

Table 25 Analysis of speech codes which precede a change in speaker.

Speech code	Observed frequencies		Expected frequencies		χ^2
	No change	Speaker change	No change	Speaker change	
Offer	3301	3880	3114.05	4066.95	19.71*
Reply	440	391	360.36	470.64	30.69*
Consent	1216	1243	1066.35	1392.65	36.84*
Dissent	357	300	284.91	372.09	31.76*
Modify	193	167	156.11	203.89	14.98*
Reaction	399	1134	664.79	868.21	186.93*
Request	187	842	446.23	582.77	264.88*

Chi-square values marked * are significant at $p < .001$.

Table 26 Analysis of speech codes which follow a change in speaker.

Speech code	Observed frequencies		Expected frequencies		χ^2
	No change	Speaker change	No change	Speaker change	
Offer	4136	3045	3114.05	4066.95	591.60**
Reply	324	507	360.36	470.64	6.30*
Consent	554	1905	1066.35	1392.65	433.81**
Dissent	304	353	284.91	372.09	2.14
Modify	230	130	156.11	203.89	60.92**
Reaction	32	1501	664.79	868.21	1061.86**
Request	513	516	446.23	582.77	17.38**

Chi-square values marked * are significant at $p < .025$ and ** are significant at $p < .001$.

6.2.2 Discussion

Stech's (1979) finding that statements (which approximate to Offers in CEA) usually preceded and followed a change in speaker was not replicated in the present study, Offers tending to precede and follow within-speaker transitions. It is suggested that this discrepancy is due to a fault in the coding scheme adopted by Stech in which unsolicited (Offers) and solicited (Replies) information were not distinguished. Distinguishing between these two methods of information transmission, it is apparent that it is only solicited information in the form of Replies which tend to follow a change in speaker.

Taken together, the two analyses indicate that Requests and Reactions tend to precede changes in speaker and Replies, Consents and Reactions tend to follow changes in speaker. Dissents, although usually preceding within-speaker transitions, are equally likely to follow within- as well as between- speaker transitions. The overall conclusion that may be drawn from the analyses is that although speech content is an active variable in floor-apportionment, its role is limited,Requests and Reactions tending to precede changes in speaker, and Replies, Consents and Reactions tending to follow changes in speaker.

It is interesting to note that although speech content has only a limited role in turn-taking, this does not reduce its importance. For example, knowing that a Request has just occurred may be a better predictor of a change in speaker than signals in other modalities, Requests being associated with a change in speaker on 81.8% of occasions.

6.3 Conversational strategies

Von Bertalanffy (1971) has noted that many system problems concern both structural and topological properties of systems. For this reason, graphs, and in particular directed graphs, provide a useful tool to "elaborate relational structures by representing them in topological space" (p.19). Similarly, Kaufmann (1972) has suggested that the analysis of Markov chains may be simplified by using graphs, a point that has been demonstrated by ethologists (e.g. Altman, 1965; Van Hoof, 1973; Sustare, 1978) and psychologists alike (e.g. Hawes & Foley, 1973; Rogers & Farace, 1975; Gottman et al., 1977b).

In general, a graph is defined as a figure consisting of points, called vertices, and line segments connecting vertices, called the edges of a graph (Ore, 1963). If the direction of movement between vertices is specified, the graph becomes more specialised, and is known as a 'directed graph', or 'digraph'. When applied to the analysis of Markov chains, digraphs provide a basis for identifying patterns of relationships between the system states. One particular type of graph, known as a state diagram, is particularly suited to the analysis of the sequential ordering of states in a system (Sustare, 1978; Van Hooff, 1979), and may be used as a basis from which to infer probabalistic rules and strategies (Van Hooff, 1979; Penman, 1980).

Construction of a state diagram proceeds from a transition matrix, and, like Markov analyses, requires that all category states are discrete and mutually exclusive. In a state diagram, a box or circle is drawn for each state, and an arrow is drawn connecting two boxes when a transition has been found to occur between those two states. The direction of the arrow indicates the sequential ordering of the states, always pointing towards the state following the transition. Recurrent arrows (from one state back to itself) are drawn when series of events are recognisable as separate occurrences of the same state in succession.

Systems with even quite a small number of possible states can become diagrammatically very complex, making interpretation very difficult. In the present analysis a technique known as 'condensation' (Harary et al., 1965) has been used to reduce the complexity of the state diagram, condensation being accomplished by excluding those transitions whose probability of occurrence is less than chance. The condensed state diagram, based on the model of conversation presented in Table 22 (Chapter 4) is shown in Figure 7.

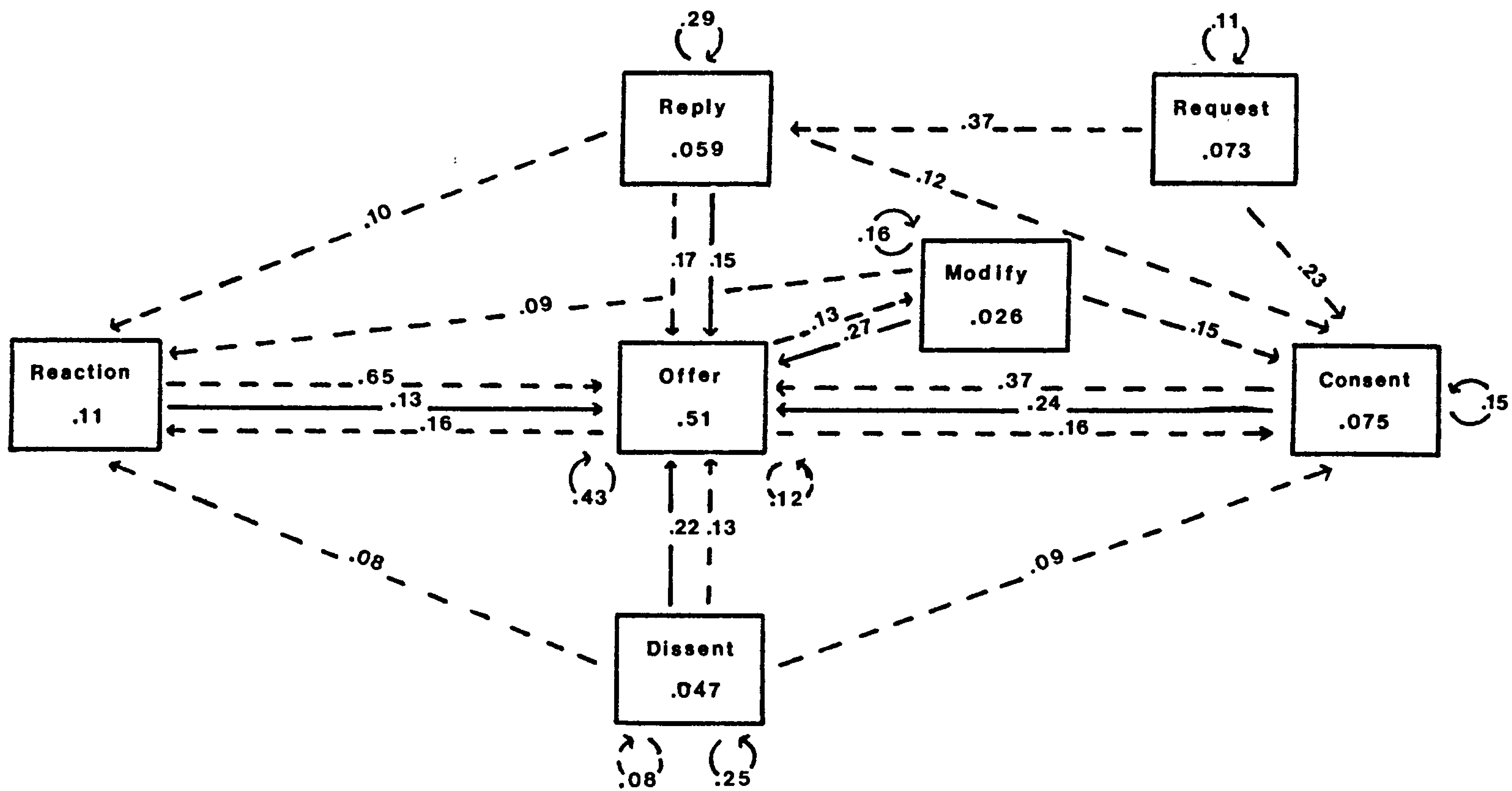
In the following discussion of the state diagram, in which a number of conversational strategies are outlined, several terms are used to describe the states represented by the 14-state model of conversation. The list below presents these terms and their definitions.

(1) Probable transitions have a probability of occurrence greater than would be expected if all possible transitions were equally probable. In the 14-state model, transition probabilities greater than .077 are considered probable (one state out of a possible thirteen) for all states, except Reactions where probable transitions are greater than .072 (one state out of fourteen).

(2) Major states are acts having more than one probable transition to them and less probable transitions from them.

(3) Minor states are acts having more probable transitions from them than to them.

Figure 7 Condensed state diagram of the seven-state model of informal conversation.



Note: Within-speaker transitions are marked ———, and between-speaker transitions are marked - - - -.

Probabilities inside state boxes refer to the state usage.

(4) Initiator states have their most probable transitions to a major or a minor state, but the probability of returning to the initial state (other than from itself) is less than .077 (.072 for Reactions).

Distinguishing between major and minor states, the Offer, Consent and Reaction categories operate as major states, and the Reply, Dissent, Modify and Request categories as minor states. Accounting for 80% of all speech acts in the sample, the major states form the basic core of conversation, in which information is Offered and either agreed with (Consent) or acknowledged (Reaction), and adequately demonstrates that the basic process of conversation and discussion is simply concerned with the imparting of information and ideas. This basic system of information exchange is diversified by the minor states, all of which serve to return the conversation quickly to the basic Offer speech state.

The seven speech categories, as well as reflecting how information is made salient in conversation, also have a variety of other functions. For example, the Offer, Reply, and Dissent categories tend to perpetuate themselves by retaining the floor for the speaker (pro-active tendency - Bales, 1955). In contrast, the Consent, Reaction and Request categories are reactive (Bales, 1955) as they relinquish the speaker-turn. It is therefore possible to suggest two conversational strategies, one for retaining the listener state, the other for gaining the speaker-turn. To retain the listener state, the listener should use a Reaction when the speaker stops talking. This will result in a switch to the original speaker, Reactions functioning in the same manner as listener responses (Dittman & Llewellyn, 1967).

A claim for the floor may be made by an Offer or Consent, the latter being the more probable. Whereas an Offer of information secures the turn for the new speaker, unless immediately followed by a within-speaker Offer, a Consent will relinquish the turn and return the floor to the original speaker. The most effective strategy therefore, for securing the speaking-turn is to Offer information. Should the new speaker begin the turn with a Consent, the floor is thrown open for either party to claim, finally being secured by an Offer of information.

The analyses also suggest that there are three initiator states, Request, Dissent and Modify, each of which function in different ways. As expected, a Request functions by offering the floor, the outcome of which tends to be either a Reply or a Consent (corresponding to the outcomes of open and closed questions, respectively). Reply tends to follow Reply, but when this ceases the claim for the turn is thrown open for negotiation, as the next most probable events following a Reply are within- (.15) and between-speaker Offers (.17) and between-speaker Consents (.12). If the listener takes the turn with an Offer (the marginally more probable strategy), then the exchange of the floor is effected and secured. If, however, the turn is taken with a Consent, the speaker turn, as described above, tends to be relinquished. Similarly, if a Consent follows a Request, the speaking turn is relinquished and restored to the original speaker. Depending on the outcome of a Request, two different patterns of floor-exchange emerge; Replies tend to have a long term effect, by retaining the floor for the new speaker, whereas Consents have a short-term effect by quickly returning the floor to the original speaker. In both cases, where the conditions of the Request are fulfilled, the most appropriate strategy for retaining the floor is to Offer new information.

The second initiator state is the Dissent category and represents a method of breaking into the interaction. For example, a Dissent is immediately followed by a speaker-switch, the turn being retained either by continuing to Dissent or by Offering information. Should the turn be lost, four outcomes are probable; two tend to retain the floor for the new speaker (Offer, Dissent), the remaining two return the floor to the original speaker (Consent, Reaction). Dissenting is, therefore, an effective way of gaining the floor. In cases where the floor is lost, the most effective strategy in regaining it is to Offer or Dissent, as Consenting immediately relinquishes the turn.

The final initiator state, Modifications, are similar to Dissents in that they tend to retain the floor, but in contrast, are only more likely to do so if the system quickly moves to Offering information. This is due to the cyclic within-speaker Modify state and the between-speaker Modify-Offer transition having similar probabilities of occurrence. Consequently, in order to retain the speaking turn after Modifying information, it is important to Offer new information quickly, as Modifications tend to throw the floor open for negotiation.

6.4 Conclusion

Using the detailed stochastic model of informal conversation presented in Chapter 4 (Table 22) as a basis, the present chapter has been concerned with identifying the role of speech content in the process of floor-apportionment and the extraction of a number of conversational strategies. The results of the turn-taking (Section 6.2) and strategy analyses (Section 6.3) have three main implications. Firstly, in addition to the analyses presented in Chapter 4, the turn-taking and strategy analyses provide additional evidence for the suggestion that the turn-taking mechanism is a first-order process (Sacks, Schegloff & Jefferson, 1974), by demonstrating the operation of single antecedent Requests, Reactions (Section 6.2), Consents and Dissents (Section 6.3) in the floor-apportionment process. Secondly, the results re-affirm the suggestion of Thomas and Bull (1981) that speech content does have an important, though often neglected, role to play in conversational floor-apportionment.

Thirdly, the model of conversation presented in Chapter 4 and the strategies presented in this chapter for gaining, retaining and relinquishing the speaker-turn have important implications for social skills training; they begin to provide information about the organisation of social behaviours that Ellis and Whittington (1981) have indicated is so necessary before successful social skills training can occur.

Trower et al. (1978) have noted that in order to retrain those deficient in some aspect of social skill, "we need to know how elements are combined in normal social interaction" (p.15). The present analyses not only add to our understanding of the sequential arrangements in conversation, but also demonstrate that strategies, based on speech content, have important implications for turn-taking and general meshing skills. For example, in social skills training, meshing skills have usually been in the form of non-verbal behaviours, questions and listener responses (e.g. Trower, Bryant & Argyle, 1978), but, as the present analyses suggest, speech content is also an important variable in conversational meshing. Since both verbal and non-verbal behaviours may be active, social skills training procedures would be enhanced by taking account of speech content. In addition, the following strategies might be included in a social skills training programme concerned with the organisation of conversation.

(1) Given a Request by one speaker, it is generally expected that a change of speaker will occur,

the new speaker tending to either Reply (45.6% of occasions) or Consent (30.3% of occasions) to the Request.

(2) Reactions, on 74% of occasions, result in a switch to the original speaker, and represent a method of ensuring that the other person continues to speak. It is not, therefore, a particularly effective method of obtaining the floor whilst another person is speaking.

(3) The most effective method for a listener to obtain the speaker-turn is to Dissent. Representing a way of breaking into the conversation, Dissenting gives the listener a better than one in two chance of gaining the floor (the floor is lost on 45% of occasions after a Dissent). Moving to an Offer state after Dissenting ensures the turn is retained by the new speaker.

(4) Having obtained the floor, the best way of retaining it is to continue to Offer information, Offers retaining the floor on 46% of occasions. Although this may seem rather low, of the 54% of occasions where a between-speaker transition occurs, 66% are accounted for by Consents and Reactions, which function by returning the floor to the original speaker.

(5) If, for any reason, the floor is thrown open for negotiation, the speaker-turn may be secured by an Offer of information and relinquished by either a Request for information or a Consent.

(6) The least stable speech state is that of Modification, securing the speaking-turn on 56% of

occasions. Its instability arises from the cyclic within-speaker Modify transition and the between-speaker Modify-Offer transition having similar probabilities of occurrence. To ensure that the turn is retained, it is important that the speaker moves to the highly stable Offer speech state, as Modifications tend to throw the floor open for negotiation.

In summary, the analyses which have been presented in both Chapter 4 and the present chapter indicate that speech content is an active variable in the turn-taking process. In addition, by using a graphical method, a number of conversational strategies have been extracted from the model of informal conversation presented in Chapter 4 (Table 22). These strategies provide information about the way conversation and speaker-turns are structured and are considered to have important implications for turn-taking and general meshing skills.

CHAPTER 7

CONVERSATIONAL CONTENT: A 'TYPE' ANALYSIS

7.1 Introduction

The analyses of conversation presented so far have all been concerned with the Activity dimension of CEA and have looked at the way in which information is exchanged in conversation. For example, the analyses presented in Chapters 4 and 6 have suggested that the basic building blocks of conversation are the Offer, Consent and Reaction speech states. Accounting for over 80% of all speech events, conversation may be seen as a process of information exchange (Offer), which is either actively agreed with (Consent) or acknowledged (Reaction). The analyses to be presented in this chapter are concerned with examining the types of information that are exchanged in conversation, using the Type dimension of CEA (see Section 2.5.2) to classify conversational speech.

Typically, conversational research at the content level of analysis has been concerned with the identification of specific phases that the discursive or conversational interaction passes through. In small-group research, the work of Bales (1950), Bennis and Shepard (1956), Tuckman (1965) and Fisher (1970) all support the notion that group development occurs by passing through a number of different phases. For example, Fisher (1970), in studying group decision-making, noted four phases: Orientation (characterised by agreement), Conflict (dispute and disagreement), Emergence (characterised by ambiguity and some dispute), and Reinforcement (where argument is of secondary importance, the phase consisting primarily of interpretation and agreement). Similarly, Ellis and Fisher (1975), in a study of classroom interaction, noted three separate phases: (1) interpersonal conflict over individual differences, (2) confrontation, in which disfavouring comments were the norm, and (3) substantive conflict, in which interpretation, support and dispute all occurred. In a review of some fifty models, Tuckman (1965) considered that decision-making and discussion were generally characterised by forming, storming, norming and performing. Although the number and labels for the phases differ from investigator to investigator, the phases themselves are similar in nature. Generally, a group orients itself, generates ideas, evaluates these ideas, engages in conflict, chooses, and finally implements its best idea.

Phase studies, by definition, identify relatively coarse constellations of behaviours. They are important to the study of interaction as they describe the basic processes that discussions pass through, but there is the persistent problem of determining when the communication system leaves one phase and enters the next (Hawes & Foley, 1973). In order to enhance our knowledge of the way thoughts and ideas are organised in conversation, the analysis of conversational content needs to be conducted at a finer level, not simply identifying conversational phases, but the individual behaviours which make up conversation. This forms the basis of the present chapter.

7.2 Method and Results

The analyses presented are based on the combined conversations of the 24 dyads used in the seven-state Activity analyses in Chapter 4, coded using the Type dimension of CEA (see Section 2.5.2).

No formal analysis of the data has been carried out for the following reason. As noted in Chapter 4 (Sections 4.1 and 4.4), the Anderson-Goodman likelihood ratio test requires that all transition probabilities are greater than zero. As the Type analysis matrices contain 58.7% zero cells, any form of statistical testing would be invalid. In such cases, two possible courses of action have been used in the past to reduce the number of zero cells; the specification of a priori zeros (Colgan & Smith, 1978) and the collapsing of categories (Siegel, 1956). However, such procedures are considered inappropriate for the present analysis for two reasons: (1) the specification of logical zeros requires that certain transitions are logically impossible (Colgan & Smith, 1978), which is not the case for the Type data, and (2) cells should only be combined if the new combination is meaningful, which again, is not the case with these data. In addition, as Raush (1972) has pointed out, any arbitrary lumping of categories may confuse, or even mask the underlying structure of the behavioural stream. Consequently, neither of these two procedures for reducing the number of zero cells has been used.

Although no formal analysis has been carried out, a first-order Markov process has been assumed in the following analyses. This is considered an acceptable assumption as a Markov model has been shown to be robust to violations of the underlying assumptions (Bartholomew, 1973; Hewes & Hewes-Evans, 1974). The Type data from the 24 conversations have therefore, simply been cast into first-order transition matrices. These may be found in frequency and probability forms in Tables 27 and 28 (within-speaker transitions) and Tables 29 and 30 (between-speaker transitions), respectively. (Program 13, Appendix G, was written specifically for the Type analyses).

Although statistical analysis of the data is preferred, it has been demonstrated that this is clearly not possible. Any comments made about the sequencing of information in conversation are therefore made on the basis of an explorative, observational analysis.

7.3 Discussion

In the following discussion of the transition matrices, several terms are used to describe the states represented by the 52-state model. The list below presents these terms and their definitions.

(1) Probable transitions are those whose probability of occurrence is greater than would be expected if all possible transitions were

Table 27 Frequency table for within-speaker Type transitions

	B	P	P.exp	N	S.F.	A	J	Syn	Sum	Jdg	Av	Neg	A.R.	Sal	Con	Ex	E.S.	Ap	D	M	Rev	Res	Com	L.R.	U.I.	L
B	1566	15	35	24	0	10	161	1	11	26	12	3	44	3	96	288	0	0	24	32	52	11	1	0	0	0
P	12	16	3	0	0	2	0	0	0	0	1	0	0	0	3	1	0	0	1	1	1	1	0	0	0	0
P.exp	24	3	113	20	1	2	2	0	1	0	1	0	2	0	3	11	0	1	2	1	5	0	0	0	0	0
N	30	3	6	162	0	2	0	0	0	3	1	0	1	1	0	3	0	0	0	0	2	0	0	0	0	0
S.F.	3	0	0	0	4	1	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
A	12	0	3	2	0	8	1	0	0	0	1	0	2	0	0	4	0	0	0	0	0	0	0	0	0	0
J	87	1	3	1	0	0	160	0	1	1	1	1	3	0	15	47	0	0	0	3	8	0	0	0	0	0
Syn	2	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Sum	3	0	0	0	0	0	1	0	11	1	0	0	2	0	0	0	0	0	2	0	1	1	0	0	0	0
Jdg	7	0	0	0	0	1	1	0	0	6	0	0	0	0	1	4	0	0	0	0	1	0	0	0	0	0
Av	36	3	1	0	0	3	0	0	0	2	4	0	2	0	1	8	0	1	3	2	0	0	0	0	0	0
Neg	13	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A.R.	466	14	25	11	2	3	49	3	6	20	9	6	267	2	21	137	1	4	31	32	10	3	4	12	2	1
Sal	6	1	1	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	1	1	0	0	0	0	0
Con	45	2	3	1	0	0	8	0	1	2	1	1	2	0	34	23	0	0	0	3	4	0	0	0	0	0
Ex	279	1	11	15	0	4	43	0	1	10	3	1	12	2	45	325	0	0	7	9	14	3	0	0	0	0
E.S.	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Ap	2	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0
D	20	2	1	0	0	0	0	0	0	0	1	0	2	0	1	23	0	1	13	9	2	0	0	0	0	0
M	33	1	1	1	0	0	5	0	1	2	1	0	4	1	1	20	0	2	6	18	0	0	0	0	0	0
Rev	35	0	2	4	0	0	5	0	0	2	2	0	0	0	6	20	0	0	2	0	11	1	0	0	0	0
Res	12	1	0	0	0	0	5	0	1	1	0	0	4	0	0	5	0	0	2	0	0	4	0	0	0	1
Com	5	0	0	0	0	0	2	0	0	0	0	0	8	0	0	0	0	0	0	0	0	1	0	0	0	0
L.R.	152	1	10	2	1	2	16	0	2	5	5	0	38	0	6	38	0	0	6	0	3	0	3	18	0	0
U.I.	3	0	0	0	0	0	0	0	0	0	0	0	3	0	0	2	0	0	0	0	0	0	0	0	0	0
L	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Note: (B) Belief, (P) Personal detail, (P.exp) Personal experience, (N) Narrative, (S.F.) Subjective Feelings, (A) Attitude, (J) Justify/Explain, (Syn) Synthesis, (Sum) Summary, (Jdg) Judgement, (Av) Avoidance, (Neg) Negation/Submission, (A.R.) Active Recognition, (Sal) Salience, (Con) Conclusion, (Ex) Example, (E.S.) Emotional Support, (Ap) Apology, (D) Direction, (M) Metastatement, (Rev) Revision, (Res) Restatement, (Com) Completion, (L.R.) Listener response, (U.I.) Unsuccessful Interruptions, (L) Laughter.

Table 28 Probability table for within-speaker Type transitions

	B	P	P.exp	N	s.f.	A	J	Syn	Sum	Jdg	Av	Neg	A.R	Sal	Con	Ex	E.S.	Ap	D	M	Rev	Res	Com	L.R.	U.I.	L
B	.29	003	006	004	0	002	03	0002	002	005	002	0005	008	0005	018	05	0	0	004	006	01	002	0002	0	0	0
P	09	12	023	0	0	015	0	0	0	0	008	0	0	0	023	008	0	0	008	008	008	008	0	0	0	0
P.exp	07	009	34	06	003	006	006	0	003	0	003	0	006	0	009	033	0	003	006	003	015	0	0	0	0	0
N	086	009	018	47	0	006	0	0	0	009	003	0	003	003	0	009	0	0	0	0	006	0	0	0	0	0
s.f.	18	0	0	0	24	06	06	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
A	16	0	04	03	0	11	01	0	0	0	01	0	03	0	0	05	0	0	0	0	0	0	0	0	0	0
J	10	001	004	001	0	0	19	0	001	001	001	001	004	0	018	056	0	0	0	004	009	0	0	0	0	0
Syn	20	0	0	0	0	0	10	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Sum	05	0	0	0	0	0	017	0	18	017	0	0	03	0	0	0	0	0	03	0	017	017	0	0	0	0
Jdg	037	0	0	0	0	005	005	0	0	03	0	0	0	0	005	02	0	0	0	0	005	0	0	0	0	0
Av	38	03	01	0	0	03	0	0	0	02	04	0	02	0	01	08	0	01	03	02	0	0	0	0	0	0
Neg	45	0	0	0	0	0	07	0	0	0	0	07	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A.R.	20	006	011	005	0009	001	022	001	003	009	004	003	12	0009	009	06	0004	002	014	014	004	001	002	005	0009	0004
Sal	21	034	034	034	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	034	034	0	0	0	0	0
Con	11	005	007	002	0	0	02	0	002	005	002	002	005	0	08	06	0	0	0	007	009	0	0	0	0	0
Ex	18	0006	007	009	0	003	028	0	0006	006	002	0006	008	001	029	21	0	0	005	006	009	002	0	0	0	0
E.S.	0	0	0	0	2	0	2	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Ap	15	0	0	0	0	0	15	0	0	0	0	0	08	0	0	0	08	0	08	0	08	0	0	0	0	0
D	10	01	005	0	0	0	0	0	0	0	005	0	01	0	005	12	0	005	066	045	01	0	0	0	0	0
M	13	004	004	004	0	0	02	0	004	008	004	0	015	004	004	08	0	008	023	069	0	0	0	0	0	0
Rev	25	0	014	028	0	0	035	0	0	014	014	0	0	0	04	14	0	0	014	0	077	007	0	0	0	0
Res	15	013	0	0	0	0	064	0	013	013	0	0	05	0	0	064	0	0	03	0	0	05	0	0	0	013
Com	09	0	0	0	0	0	034	0	0	0	0	0	04	0	0	0	0	0	0	0	0	017	0	0	0	0
L.R.	11	0007	007	001	0007	001	011	0	001	004	004	0	027	0	004	027	0	0	004	006	002	0	002	013	0	0
U.I.	11	0	0	0	0	0	0	0	0	0	0	0	11	0	0	07	0	0	0	0	0	0	0	0	0	0
L	08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	08	0	0	0	0	0	0	0	0	0	0

Note: To improve clarity the decimal point has been omitted from the above table.

(B) Belief, (P) Personal detail, (P.exp) Personal experience, (N) Narrative, (s.f.) Subjective feelings, (A) Attitude, (J) Justify/Explain, (Syn) Synthesis, (Sum) Summary, (Jdg) Judgement, (Av) Avoidance, (Neg) Negation/submission, (A.R.) Active recognition, (Sal) Salience, (Con) Conclusion, (Ex) Example, (E.S.) Emotional support, (Ap) Apology, (D) Direction, (M) Metastatement, (Rev) Revision, (Res) Restatement, (Com) Completion, (L.R.) Listener response, (U.I) Unsuccessful interruptions, (L) Laughter.

Table 29 Frequency table for between-speaker Type transitions

	B	P	P.exp	N	S.F.	A	J	Syn	Sum	Jdg	Av	Neg	A.R.	Sal	Com	Ex	E.S.	Ap	D	M	Rev	Res	Com	L.R.	U.I.	L
B	1150	7	19	9	0	7	64	1	4	53	15	7	180	3	14	98	0	1	17	33	8	19	28	591	19	3
P	9	32	2	1	0	0	1	0	0	1	2	0	10	0	2	2	0	0	1	2	0	4	0	17	1	1
P.exp	17	2	13	1	0	2	4	0	0	1	0	0	32	0	0	4	0	0	0	1	0	2	1	63	0	0
N	17	1	1	1	0	3	0	0	0	1	0	0	32	0	2	1	0	0	0	0	0	2	0	71	1	1
S.F.	1	0	1	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
A	8	1	0	2	0	5	1	0	1	2	0	0	7	1	0	2	0	0	0	0	0	0	0	11	0	0
J	83	0	1	1	0	0	40	0	0	10	2	1	175	2	3	14	1	2	1	5	0	2	3	161	1	0
Syn	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	3	0	0
Sum	5	0	0	0	0	0	1	1	5	0	0	0	18	1	1	0	0	0	2	1	0	0	0	3	0	0
Jdg	30	1	0	1	0	0	12	0	0	2	5	0	99	0	1	5	0	0	0	2	1	0	0	6	0	1
Av	10	1	0	1	0	0	1	0	0	0	1	0	2	0	0	3	0	0	1	1	0	0	0	8	0	0
Neg	0	0	0	0	0	0	0	0	0	0	0	3	2	0	0	1	0	0	0	2	0	0	0	4	0	0
A.R.	529	8	21	22	1	11	97	3	10	16	8	0	86	3	63	174	0	0	20	29	5	1	2	26	0	0
Sal	2	0	0	1	0	0	1	0	0	1	0	0	3	1	0	0	0	1	2	1	0	0	0	8	1	0
Con	44	2	2	1	0	0	9	0	0	1	2	0	101	0	4	6	1	0	3	3	1	2	0	94	2	0
Ex	122	2	7	1	0	1	14	0	1	4	0	0	267	1	3	41	0	0	3	5	0	8	15	256	1	3
E.S.	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Ap	2	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0
D	26	1	2	0	0	0	1	0	0	0	2	0	38	0	0	8	0	0	23	6	0	8	0	8	0	0
M	23	0	2	0	0	0	4	0	1	3	1	1	57	0	2	5	0	0	8	35	0	1	0	21	0	1
Rev	12	0	0	0	0	0	3	0	0	1	0	0	15	0	0	5	1	0	0	0	3	0	0	12	0	0
Res	10	1	1	0	0	0	2	0	0	1	0	0	15	0	0	3	0	0	2	1	1	1	0	14	0	0
Com	9	0	0	0	0	0	1	0	0	0	0	0	22	0	1	4	0	0	1	0	0	2	0	2	0	0
L.R.	494	5	44	62	2	6	120	1	2	8	8	2	18	6	79	198	0	0	6	11	7	0	1	9	0	0
U.I.	13	1	0	0	0	0	0	0	0	0	1	0	1	0	1	3	0	0	0	0	0	0	0	0	0	0
L	3	1	0	0	0	1	0	0	0	1	0	0	0	0	0	3	0	0	1	0	0	0	0	0	0	0

Note: (B) Belief, (P) Personal detail, (P.exp) Personal experience, (N) Narrative, (S.F.) Subjective feelings, (A) Attitude, (J) Justify/Explain, (Syn) Synthesis, (Sum) Summary, (Jdg) Judgement, (Av) Avoidance, (Neg) Negation/Submission, (A.R.) Active Recognition, (Sal) Salience, (Con) Conclusion, (Ex) Example, (E.S.) Emotional Support, (Ap) Apology, (D) Direction, (M) Metastatement, (Rev) Revision, (Res) Restatement, (Com) Completion, (L.R.) Listener Response, (U.I.) Unsuccessful interruptions, (L) Laughter.

Table 30 Probability table for between-speaker Type transitions

	B	P	P.exp	N	S.F.	A	J	Syn	Sum	Jdg	Av	Neg	A.R	Sal	Con	Ex	E.S	Ap	D	M	Rev	Res	Com	L.R	U.I	L
B	21	001	003	002	0	0013	012	0002	0007	01	003	0012	16	0005	003	02	0	0002	003	006	0015	003	005	11	003	0005
P	07	24	015	008	0	0	008	0	0	008	015	0	076	0	015	015	0	0	008	015	0	031	0	13	008	008
P.exp	051	006	039	003	0	006	012	0	0	003	0	0	096	0	0	012	0	0	0	003	0	006	003	19	0	0
N	05	003	003	003	0	009	0	0	0	003	0	0	092	0	006	003	0	0	0	0	0	006	0	204	003	003
S.F.	059	0	059	0	18	0	0	0	0	0	0	0	0	0	0	059	0	0	0	0	0	0	0	0	0	0
A	11	014	001	027	0	068	014	0	014	027	0	0	095	014	0	027	0	0	0	0	0	0	0	15	0	0
J	099	0	0	001	0	0	048	0	0	012	002	001	21	002	003	017	001	002	001	006	0	002	003	19	001	0
Syn	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	30	0	0
Sym	083	0	0	0	0	0	017	017	084	0	0	0	30	017	017	0	0	0	033	017	0	0	0	05	0	0
Jdg	16	005	0	005	0	0	064	0	0	011	027	0	34	0	005	027	0	0	0	011	005	0	0	032	0	005
Av	11	011	0	011	0	0	011	0	0	0	011	0	021	0	0	032	0	0	011	011	0	0	0	084	0	0
Neg	0	0	0	0	0	0	0	0	0	0	0	103	069	0	0	034	0	0	0	069	0	0	0	14	0	0
A.R.	23	004	009	009	0004	005	043	001	004	007	004	004	038	001	028	076	0	0	009	013	002	0004	0008	011	0	0
Sal	069	0	0	034	0	0	034	0	0	034	0	0	103	034	0	0	0	034	069	034	0	0	0	0	034	0
Con	108	005	005	002	0	0	022	0	0	002	005	0	25	0	009	015	002	0	007	007	002	005	0	23	005	0
Ex	079	001	005	0006	0	0006	009	0	0006	003	003	0	173	0006	002	027	0	0	002	003	0	005	01	17	0006	002
E.S.	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0
Ap	15	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	077	0	0	0	0	0	0
D	13	005	01	0	0	0	005	0	0	0	01	0	192	0	0	04	0	0	12	034	0	04	0	04	0	0
M	088	0	008	0	0	0	015	0	004	001	004	004	22	0	008	019	0	0	031	134	0	004	0	08	0	004
Rev	085	0	0	0	0	0	02	0	0	007	0	0	11	0	0	035	007	0	0	0	02	0	0	085	0	0
Res	13	013	013	0	0	0	026	0	0	013	0	0	192	0	0	038	0	0	026	013	013	013	0	05	0	0
Com	16	0	0	0	0	0	017	0	0	0	0	0	38	0	017	069	0	0	017	0	0	034	0	034	0	0
L.R.	35	004	03	004	001	004	086	0007	001	006	006	001	013	004	057	142	0	0	004	008	005	0	001	006	0	0
U.I.	46	036	0	0	0	0	0	0	0	0	036	0	036	0	036	11	0	0	0	0	0	0	0	0	0	0
L	25	085	0	0	0	083	0	0	0	083	0	0	0	0	0	25	0	0	083	0	0	0	0	0	0	0

Note: To improve clarity the decimal point has been omitted from the above table.

(B) Belief, (P) Personal detail, (P.exp) Personal experience, (N) Narrative, (S.F.) Subjective feelings, (A) Attitude, (J) Justify/Explain, (Syn) Synthesis, (Sum) Summary, (Jdg) Judgement, (Av) Avoidance, (Neg) Negation/Submission, (A.R.) Active recognition, (Sal) Salience, (Con) Conclusion, (Ex) Example, (E.S.) Emotional support, (Ap) Apology, (D) Direction, (M) Metastatement, (Rev) Revision, (Res) Restatement, (Com) Completion, (L.R.) Listener response, (U.I.) Unsuccessful interruption, (L) Laughter.

equally probable; transition probabilities greater than .019 (one out of 52 possible states) are considered probable.

(2) Major states are those acts having more than one probable transition to them and less probable transitions from them.

(3) Minor states are those acts having more probable transitions from them than to them.

(4) Initiator states are those acts that have their most probable transitions to a major state or a minor state, but the probability of returning to the initial state (other than from itself) is less than .019 (one state out of a possible 52 states).

(5) A cyclic state is one whose most probable transition is to itself.

Before continuing the discussion, an important point to note is that eight speech-states occurred less than 52 times in the total sample of 24 conversations; this has the effect that even a single transition from them becomes 'probable'. Because this is a basic sampling problem of codes occurring fairly infrequently, they are excluded from the remainder of the discussion. The speech states concerned are the Subjective feelings, Synthesis, Negation/Submission, Salience, Emotional support, Apology, Unsuccessful interruptions, and Laughter categories.

Distinguishing between major and minor states, the Belief (35.9%), Narrative (2.6%), Explanation (6.3%), Active recognition (17.0%), Example (11.56%), and Listener response (10.45%) categories operate as major states, and the remaining categories as minor states (figures in parenthesis refer to the percentage of the total sample of speech units accounted for by each code). Accounting for approximately 84% of all the speech units, the major states form the basic core of information exchanged in conversation. This takes the form of beliefs being given (Belief), points explained (Explanation), examples given (Example), and stories told (Narrative), as well as points in the conversation being recognised for their pertinence (Active recognition), or simply being acknowledged (Listener response). The distinction noted in the Activity analysis (Section 4.7) of Offer, Consent and Reaction states forming the 'core' of conversation is again broadly suggested here, in which Beliefs, Explanations, Examples and Narratives are generally associated with Offers, Active recognitions with Consents, and Listener responses with Reactions. The process of conversation and discussion may, therefore, again be seen to be concerned with the imparting of information. This usually takes the form of one's beliefs, which are occasionally embellished by Explanation, Examples and story telling, to which the listener adds recognition of the points made. The basic system of information exchange is diversified by the minor states.

The basic core of conversation is composed of the major states which tend to be highly interconnected.

(Figure 8). Taking the example of within-speaker transitions, they tend to consist of inter-communicating within-speaker Beliefs, Examples and Explanations, Narratives generally preceding Beliefs rather than Examples or Explanations. For example:

A: I'm a christian but I don't believe in God/ (Belief) erm I adopt a christian way of life but don't believe you see/ (Explanation) I go and help out at a local hospital/ (Example) I try to help my friends as well/ (Example) I like to think that I do help someone/ (Belief)

A: I had a friend who worked in a home for the mentally handicapped/ (Example) She used to be very patient with them/ (Narrative) and used to take them for long walks and play with them in the park/ (Narrative) She was absolutely marvellous with them/ (Belief)

Turning to between-speaker transitions, Listener responses tend to be followed by Beliefs and Examples, and Active recognitions by Beliefs, Examples and Explanations. For example:

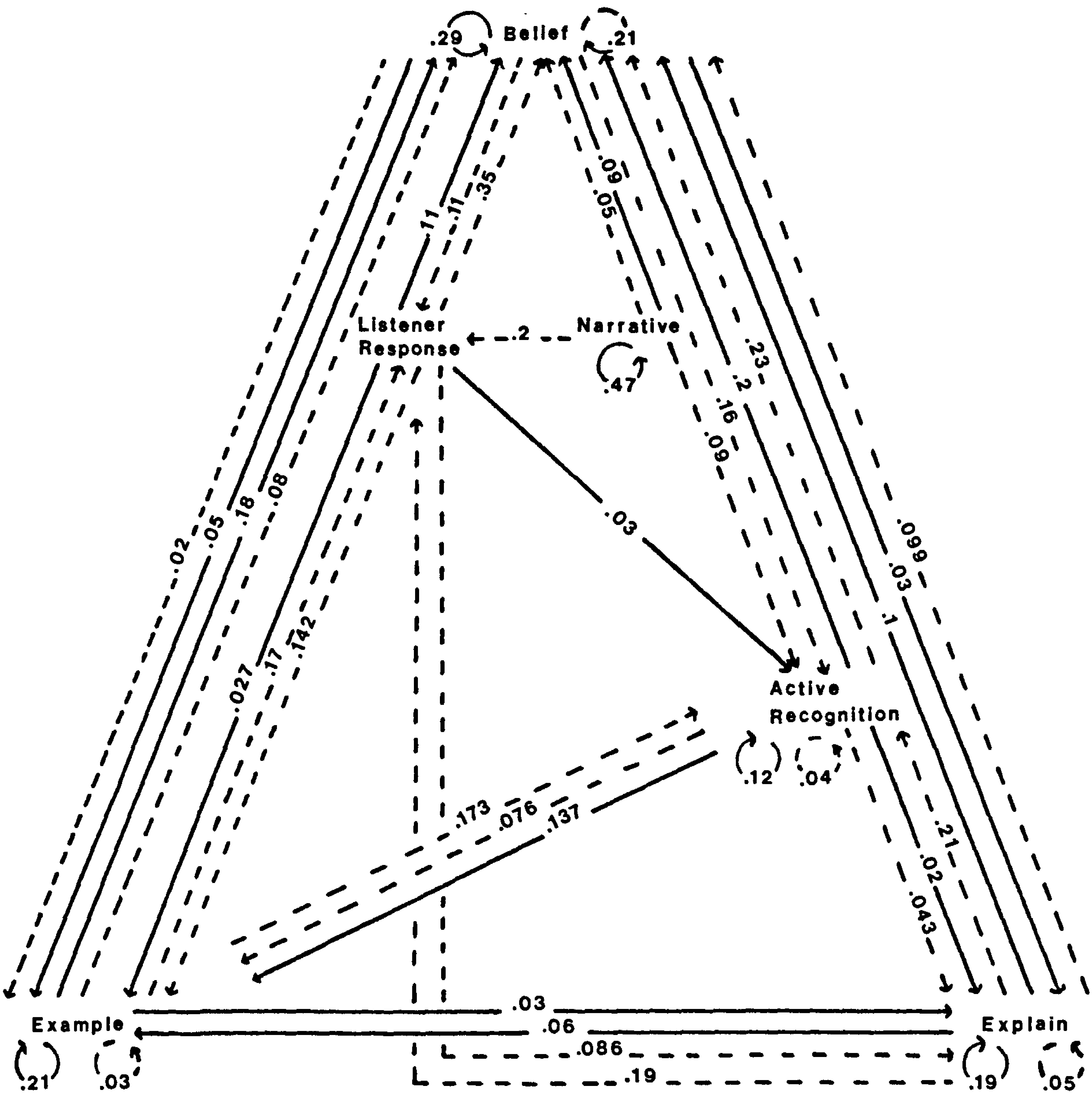
A: uh-huh/ (Listener response)

B: I think they should have the vote/ (Belief)

A: mm/ (Listener response)

B: like here in Britain/ (Example)

Figure 8 Diagram showing the inter-connections between the major Type states.



Key

Within-speaker transitions ——————→

Between-speaker transitions - - - - -→

A: That's right/ (Active recognition)

B: If they didn't it would be undemocratic/ (Explanation)

These patterns suggest that conversation is composed of phases comprising within-speaker Beliefs, Examples and Explanations, which are embellished by within-speaker Narratives, and punctuated by between-speaker Listener responses and Active recognitions. When a speaker-switch occurs, between-speaker transitions generally follow the pattern of Beliefs followed by either Beliefs or Examples, again punctuated by Listener responses and Active recognitions, and between-speaker Explanations generally being followed by Active recognitions or Beliefs. Therefore, given that a speaker-switch occurs, the loss of the turn tends to be transient if associated with an Active recognition or Listener response, but more permanent if associated with any other speech code. For example, a short-term speaker switch:

A: He really wanted to die/ (Belief) He had to go through a whole legal system from a hospital bed/ (Narrative)

B: uh-huh/ (Listener response)

A: that's before they'd allow him to die and turn off the life support machine/ (Belief)

and a long-term speaker switch:

A: there's some people who are so severely deformed/ (Belief) you know I suppose it costs so much to keep them alive anyway/ (Belief)

B: but I mean just because it costs something you can't just kill someone off/ (Belief) I don't think that's right at all/ (Belief) they have a life to lead/ (Belief) they can do a lot too if they are helped/ (Belief) but that costs money which they don't have/ (Belief)

Of all the speech states, only four are truly cyclic, having themselves as the most probable subsequent state. Three of these, Beliefs, Narratives and Examples are major states, again demonstrating the stable nature that these states provide for conversation. The fourth cyclic state is that of Personal experience, a result that is not too surprising as personal experiences may take a number of speech units to describe adequately. For example:

A: I went to that party last night/ (Personal experience) I went with Sarah and Caroline/ (Personal experience) It was really good/ (Personal experience) Got really drunk/ (Personal experience)

The analysis further suggests that there are three initiator states: Summary, Revision and Completion, It is clear why this is so for Revisions and Completions, as these are by nature very transient states, enabling one to revise an idea:

A: i think that's right/ <u>but would that be right</u>?/

and complete an idea:

A: ... and they're suddenly/

B: <u>suddenly in a wheelchair/</u>

Such states may occur anywhere in the conversation, but inevitably lead back to the main theme of the conversation.

Similarly, although the usage of the Summary state is very low (.45% of the total sample of speech events) and does not tend to be very predictable, it is generally preceded by between-speaker Beliefs and followed by a between-speaker Active recognition. The occurrence of the Summary state in conversation may be a useful method of distinguishing between the end of one conversational phase and the beginning of the next. For example:

A: well i think capital punishment is right/ (Belief)

B: Well we're agreed then that capital punishment is okay for murder but not for things like rape, theft and burglary/ (Summary)

A: yes that's right/ (Active recognition)

B: How do you feel about women's rights ...

In addition, there are a number of other regularities that can be extracted from the Type matrices. For example, a Personal detail by one speaker tends to precede a Personal detail by the other speaker ($p = .24$). For example:

A: I'm reading history/ (Personal detail)

B: oh I'm a biologist/ (Personal detail)

A: I'm boyfriendless at the moment/ (Personal detail)

B: Me too/ (Personal detail) He gave me up about a month ago/ (Explanation) It really depresses me sometimes/ (Subjective feeling)

Such a result is, perhaps, to be expected from the self-disclosure literature. For example, Argyle (1972, p.118) has suggested that intimate disclosure is increased when two people are isolated, and that if A discloses to B, B will disclose to A, Naegele (1958) indicating that when people come to 'trust' each other more, a reciprocity of self-disclosure occurs. Additionally, Jourard (1971) has indicated that the more intimate information a person discloses about him/herself, the greater the likelihood that the other will reciprocate. Consequently, talking about information personal to oneself tends to elicit information of a personal nature from the other person. Such a sequencing of events may constitute a useful social strategy. For example, in getting to know another person, it is important that personal information is given, in order to be reciprocated. By initially demonstrating a willingness to give personal information about oneself, the other person reciprocates with similar information. Such a strategy may be useful in a social situation in which it is required to 'bring a person out' in order to discuss personal matters with them.

Metastatements appear to operate in a similar manner to Personal details. For example, the most probable outcomes of a Metastatement are a between-speaker Active recognition (.22), a between-speaker Metastatement (.134) and a within-speaker Belief (.13). In making a comment about the conversation and its progress, the speaker is directly responded to, either in the form of an Active recognition:

A: I'm enjoying this/ (Metastatement)
B: Right/ (Active recognition)
Let's press on/ (Direction)

or a further Metastatement:

A: Hey this is good fun/ (Metastatement)
B: Well i find it really boring/
(Metastatement)

Such a regularity could form the basis of a social strategy. Rather than directly asking a person whether he/she is enjoying him/herself, a strategy that may lead to a socially acceptable but nevertheless untrue reply, a response may be obtained simply by commenting on the situation oneself. This is illustrated in the following, hypothetical example:

A: Do you like doing this?/ (Request, Metastatement)
B: erm yes fine/ (Consent,

(Active recognition)

A: This is okay you know/ (Offer, Metastatement)

B: Well i'm not so sure/ (Dissent, Active recognition)

Directions refer to the control aspects of conversation and account for approximately 1.4% of all the speech events; they operate as a kind of crossroads in the conversation, where a number of courses of action are available to the speakers. The most probable subsequent events after a Direction are between-speaker Active recognitions (.19), Beliefs (.13) and Directions (.12), and within-speaker Examples (.12) and Beliefs (.10). The usual course of action is for the speaker who uses a Direction to retain long-term control over the floor. For example:

A: Why don't we continue with abortion/ (Direction)

B: Okay/ (Active recognition)

A: I don't like the idea at all/ (Belief) It makes me feel very strange/ (Belief)

However, Directions are also a good point at which long-term speaker switches can occur, as the probability of a within-speaker Belief (.20) after an Active recognition is only slightly smaller than a between-speaker Belief (.23). For example:

A: We could do the housing one/ (Direction)

B: Right/ (Active recognition) I'm all in favour of landlords renting to anyone/ (Belief) I think it's wrong they won't rent to someone who's not the same colour, or a different religion/ (Belief)

Alternatively, there may be a dispute about who has the control of the conversation. For example:

A: Let's go on with the religious one/ (Direction)
B: We could do the housing one/ (Direction)

The dispute may be settled in either of the two ways illustrated above.

Finally, Completions, which are acts whereby the listener completes what the other person is saying, have a tendency to act as a subtle method of taking the floor from the other person. For example:

A: If you don't give them treatment they're suddenly/ (Belief)
B: suddenly in a wheelchair/ (Completion)
A: Right/ (Active recognition)
B: and I really feel that's wrong/ (Belief) especially when only a little money is needed/ (Belief)

Completions correspond to overlap interruptions, identified by Ferguson (1977). An important aspect of Completions in conversation would therefore lie in their ability to bring about a long-term change in speaker.

7.4 Conclusion

The results of the Type analyses presented in Tables 27 - 30 are the product of conceptualising conversation as a stochastic process, and have enabled a detailed picture of information exchange to be obtained. The significance of these results is twofold. Firstly, the analyses add to our theoretical understanding of conversation by demonstrating that information exchange is a highly structured process. Secondly, it is considered that the analyses may have useful practical application in social skills training. For example, conversation has been shown to comprise a small number of major states concerned with the giving of one's beliefs, and explaining why one holds such beliefs, embellished by the giving of examples and story-telling. This major conversational routine may be diversified by a number of other minor states, all of which return the conversation back to the stable major conversational routine. In addition to describing the basic flow of conversation, the analyses also suggest the following strategies that might be included in a social skills programme concerned with the organisation of conversational content.

(1) By giving away personal information about oneself, there is a greater than chance probability (.24) of receiving personal information from the other person. In giving such information a person obliges the other to reciprocate, thereby eliciting information of a personal nature and

giving the other person a chance to take part in the conversation.

(2) Metastatements operate in a similar manner, by bringing the other person into the conversation. They also have an additional function of eliciting information about a person's feelings towards the conversation, without asking directly for the information.

(3) Directions function as branching points in the conversation where negotiation for the conversational floor can be conducted.

(4) A subtle method of obtaining the speaking turn is to complete (Completion) what the other person is saying. In a large proportion of cases (38%), this leads to a recognition of the point made by the original speaker, followed by a speaker-switch. The new speaker then proceeds by giving his/her own views about the topic under discussion.

(5) Given that a speaker-switch occurs, the loss of the turn tends to be transient if associated with an Active recognition or Listener response, but more permanent if associated with any other speech state. Consequently, in order to retain the speaking-turn, the most effective strategy is to give one's own beliefs about the topic, as an Active recognition or Listener response will result in the loss of the floor to the original speaker.

Finally, the analyses presented in this chapter have not been subject to any statistical testing, and should therefore be treated as purely explorative. However, their importance lies in the information that they provide about the manner in which information is organised in conversation, and the way speaker-turns are structured. In addition, from the social strategies outlined in this discussion, the analyses can be seen to have useful practical application in social skills training procedures concerned with conversational exchange.

CHAPTER 8

A 'CHAIN ANALYSIS' STUDY OF INFORMAL CONVERSATION

8.1 Introduction

The analyses of conversation presented so far have all made use of Markovian methods, which lead most directly to the study of two-event transitions. However, although a first-order process has been shown to be an adequate representation of the conversations studied (Chapter 4), it is possible that within the first-order speech stream, some speech events may occur in sequences which are substantially longer than two events. Although it is possible to construct longer sequences by adding series of two-event units together, this procedure may lead to an over-simplification of the underlying structure of conversation. There is evidence, for example, of embedding in long sequences, from both linguistic (e.g. Chomsky, 1957) and social psychological work (e.g. Birdwhistell, 1970; Goffman, 1972; Argyle et al., 1981) which may be masked if such a procedure were used. For example:

A^1 "What'll ya have?"

B^2 "Ya got those almond things?"

A^2 "Not today, honey"

B^1 "Black coffee and a toasted muffin"

Goffman (1972) p.180

would be represented in a first-order Markov process as three separate transitions:

(1) Request $\xrightarrow[\text{speaker}]{\text{between}}$ Request

(2) Request $\xrightarrow[\text{speaker}]{\text{between}}$ Reply

(3) Reply $\xrightarrow[\text{speaker}]{\text{between}}$ Reply

A more appropriate unit of analysis, in which the embedding is preserved, would be a four-event sequence taking the following form:

- Request A^1
  - Request B^2
  - Reply A^2
- Reply B^1

Markov processes cannot be successfully used to analyse such data, but there are ways of dealing with sequences longer than two events within the Markovian approach. One method, for example, is lag sequential analysis (Sackett, 1978). Using Gottman and Bakeman's (1979) example, imagine that an experimenter observed a young child and recorded the following four behaviours: take, hit, cry and give. Each behaviour is, in turn, designated the criterion behaviour, and the probabilities with which the behaviours follow it at different lags are computed. If a behaviour is sequentially independent of the criterion, then its conditional probabilities, at various lags, should be about the same as its simple unconditional probability, the extent of any deviation being gauged with a z-score. Such a procedure allows the extraction of sequences of behaviours of the form, take/hit/cry, and take/any behaviour/cry. However, the problem with this method is in measuring the deviations between observed and expected probabilities. Gottman and Bakeman (1979), who themselves advocate this method of analysis note that "because dyadic states in successive time intervals (or simply successive dyadic states in the case of event sequence data) are likely not independent in the purest sense, it seems most conservative to treat the resulting z simply as an index or score and not to assign p-values to it" (p.190). The appeal of this method is thus limited by the need either to violate the assumption of independence implicit in the z-score, or to use some arbitrary deviation criterion.

One alternative method is 'chain analysis' (Dawkins, 1976), a method initially applied to the analysis of grooming in flies (Dawkins & Dawkins, 1976), and more recently by Graham et al. (cited in Argyle et al., 1981) to the analysis of the situation of paying a visit to the doctor.

Chain analysis proceeds with the raw data, consisting of digits representing conversational events, in the order in which they occurred. A chain analysis program scans the data, counting frequencies of doublets, or pairs of events. When it has found the commonest pair of events (e.g. a transition from event type 1 to event type 5), the frequency is recorded and the data scanned again, this time replacing all occurrences of that doublet by a single symbol which stands for that pair of events. The pair of events represented by 1, 5 is now regarded as a higher-order unit, and the iterative process repeated, this time including the event type 1 to event type 5 transition as a single element in the sequence. If this is repeated many times, larger units, and more elaborate hierarchical structures, may be revealed.

The aim of this chapter is to re-analyse the sample of conversational sequences, using the method of chain analysis suggested by Dawkins (1976), in order to see if the first-order sequences contain chains of conversational behaviours comprising more than two events.

8.2 Method and Results

The analyses to be presented are based on the sample of 24 dyads used in the Activity analyses in Chapter 4, coded using both the Activity and Type dimensions of CEA (Sections 2.5.1 and 2.5.2, respectively).

The conversational sequence, for each dyad, was formed into an N x N transition matrix, where N = 7 for the Activity chain analysis and N = 26 for the Type chain analysis. The transition matrix was then analysed using the chain analysis method suggested by Dawkins (1976) to find chains longer than two events in the conversational sequence. (The chain analysis program, written for this chapter, may be found in Appendix G, Program 14.) Chain analysis works by finding and recording the highest cell frequency in the matrix with which one event is followed by another, and then re-numbering this pair of events as a new two-item unit. The data are then re-scanned and every occurrence of the pair of events making up the new unit is deleted and replaced by the new identification number. The matrix is then extended to incorporate the new unit by deleting the old frequency for one event following another, and recording the frequencies with which the new two-event unit precedes, and is followed by, other events in the sequence. The process continues, treating larger units (three-, four-event units, etc.) in the same way. For example, if the next highest frequency in the matrix is recorded for the new two-event unit following a single event, then a new three-event unit would be created. This would be re-numbered in the original data, the matrix extended and the process repeated. An example of the chaining process is given in Figure 9.

As the chaining process is iterative, and the sequences to be scanned become progressively shorter as events and chains are deleted and replaced by single symbols, it is possible for the sequence to be shortened to a single symbol, representing the entire conversational sequence. Before this occurs, some criterion is required at which the chaining process stops.

Figure 9 An example of the chaining process.

Original data 0 1 2 1 0 1 0 1 1 0 1

Original matrix

	0	1	2
0	1	4	0
1	2	2	1
2	0	1	0

The highest cell frequency of 4 in the above matrix corresponds to a transition from event type 0 to an event type 1. Transition 0 → 1 is therefore re-numbered as '3', and the matrix extended, thus,

Modified data 3 2 1 3 3 1 1 3

Modified matrix

	0	1	2	3
0	0	0	0	0
1	1	0	0	2
2	0	1	0	0
3	0	1	1	1

The highest cell frequency of 2 in the above matrix corresponds to a transition from event type 1 to an event type 3, or 1 → (0 → 1). Transition 1 → 3 is therefore re-numbered as '4', and the matrix extended, thus,

Modified data 3 2 4 3 1 4

and the process of chaining is repeated.

Up to this point, two chains have been evidenced in the hypothetical example,

0 → 1 with a frequency of occurrence of 4

1 → (0 → 1) with a frequency of occurrence of 2

There appears to be no objective method of establishing a criterion point in the literature. Dawkins and Dawkins (1976) use an index of mutual replaceability, based on the Spearman Rank Correlation Coefficient, "which expresses the similarity between two acts with respect to the distribution of acts which precede and follow them" (p.749), as a stopping criterion. However, as the authors point out, chaining is stopped at an arbitrary value of the index, there being nothing objective about the criterion selected. As an alternative, Dawkins (1976) simply stopped the chaining process when no doublets which occurred more than twice could be found.

Chaining was stopped in the present analyses when the number of times a given chain occurred in a single dyadic sequence became less than five. The choice of five as a criterion was arbitrary in the sense that it is not based on any objective measure, but it did appear to give the best compromise between excluding all units longer than two events from the analysis, and retaining a large number of idiosyncratic multi-event units. It is recognised that this is not an entirely adequate solution to the problem of specifying a stopping criterion, but choosing a cut-off frequency of 5 per dyadic sequence does give, for these data, the best compromise between giving a general picture of how speech events are chained together, and a highly complex picture consisting of a large number of idiosyncratic chains of speech events.

Two additional decisions were made regarding the chaining process. Firstly, as both the Activity (Chapters 4 and 6) and Type (Chapter 7) analyses had demonstrated that most speech states were cyclic, all immediate repetitions of the same Activity speech state by the same speaker were deleted from the conversational sequence before the chaining process started. In the case of the Type analyses, both the Activity and Type codes had to be the same as those in the previous event before deletion was allowed. The chain analyses were therefore concerned with extracting chains of speech-state phases, rather than individual speech states.

Secondly, the distinction between 'within-' and 'between-speaker' transitions was maintained. This constraint was imposed in order that the speech chains were constructed with respect to the dyad, rather than an 'overall' dyad, thereby retaining the relationship between the speakers.

The results of the chain analyses for each of the 24 dyadic sequences were combined to form composite results, which may be found in Tables 31 and 32, for the Activity and Type analyses, respectively.

Table 31 Activity chain analysis results.

All chains are arranged in frequency order. All transitions between speech events are 'between-speaker', unless otherwise stated.

<table>
<tr><th>Chain</th><th>Total frequency of occurrence</th></tr>
<tr><td>(Offer→Consent)</td><td>1107</td></tr>
<tr><td>(Offer→Reaction)</td><td>1035</td></tr>
<tr><td>|→((Offer→Reaction))→|</td><td>333</td></tr>
<tr><td>(Offer→Offer)</td><td>252</td></tr>
<tr><td>(Request→Reply)</td><td>245</td></tr>
<tr><td>(Request→Consent)</td><td>205</td></tr>
<tr><td>|→((Offer→Consent))→|</td><td>174</td></tr>
<tr><td>(Consent + $\xrightarrow[\text{speaker}]{\text{within}}$ Offer)</td><td>120</td></tr>
<tr><td>Offer→(Offer→Consent)</td><td>99</td></tr>
<tr><td>(Reaction→Offer)</td><td>80</td></tr>
<tr><td>(Offer→Request)</td><td>60</td></tr>
<tr><td>(Consent→Offer)</td><td>54</td></tr>
<tr><td>(Offer→Dissent)</td><td>53</td></tr>
<tr><td>(Offer→Consent)→Offer</td><td>44</td></tr>
<tr><td>(Offer→Reaction)→Offer</td><td>44</td></tr>
<tr><td>(Reply→Reaction)</td><td>38</td></tr>
<tr><td>Offer→(Offer→Reaction)</td><td>34</td></tr>
<tr><td>(Offer→Reaction)→(Offer→Consent)</td><td>32</td></tr>
<tr><td>(Dissent→Dissent)</td><td>18</td></tr>
<tr><td>(Reply→Consent)</td><td>18</td></tr>
<tr><td>|→((Reaction→Offer))→|</td><td>17</td></tr>
</table>

Table 31 continued.

Chain	Total frequency of occurrence
(Offer→Consent)→(Offer→Offer)	15
|→(((Offer→Reaction)))→|	14
(Offer→Request)→Consent	13
(Dissent→Offer)	11
(Request→Dissent)	10
Offer $\xrightarrow[\text{speaker}]{\text{within}}$ (Request $\xrightarrow[\text{speaker}]{\text{between}}$ Reply)	10
(Offer $\xrightarrow[\text{speaker}]{\text{within}}$ Request)	8
|→((Offer→Offer))→|	7
|→((Offer→Reaction))→|→Offer	7
(Offer→Consent)→Reaction	7
|→((((Offer→Reaction))))→|	7
|→((Reply→Reaction))→|	7
(Offer $\xrightarrow[\text{speaker}]{\text{within}}$ Consent)	6
(Offer→Consent)→(Offer→Reaction)	6
|→((Request→Reply))→|	6
(Offer→Reply)	5
|→(((Offer→Consent)))→|	5
Offer→(Request→Consent)	5
(Offer→Offer)→(Consent→Offer)	5
(Offer→Consent)→|→((Offer→Reaction))→|	5
(Consent→Offer)→(Reaction→Offer)	5
(Modify→Consent)	5
Request→(Consent→Offer)	5

Table 32 Type chain analysis results.

All chains are arranged in frequency order. All transitions between speech events are 'between-speaker', unless otherwise stated.

Chain	Total frequency of occurrence
(Belief→Active recognition)	787
(Belief→Listener response)	453
(Belief→Belief)	358
(Example→Active recognition)	155
(Example→Listener response)	134
(Active recognition→Belief)	111
(Listener response→Belief)	108
(Explain→Listener response)	107
(Explain→Active recognition)	101
(Belief→Example)	94
(Active recognition $\xrightarrow[\text{speaker}]{\text{within}}$ Belief)	90
(Belief $\xrightarrow[\text{speaker}]{\text{within}}$ Example	81
(Example→Belief)	66
(Belief $\xrightarrow[\text{speaker}]{\text{within}}$ Explain)	55
(Belief→Active recognition)→Belief	54
(Judgement→Active recognition)	43
(Listener response→Example)	35
(Belief→Listener response)→Belief	35
|→((Belief→Listener response))→|	31

Table 32 continued.

Chain	Total frequency of occurrence
(Conclusion→Active recognition)	28
(Belief $\xrightarrow[\text{speaker}]{\text{within}}$ Belief)	25
(Belief→Explain)	25
Belief→(Belief→Active recognition)	24
(Conclusion→Listener response)	23
(Narrative→Listener response)	23
(Belief $\xrightarrow[\text{speaker}]{\text{within}}$ Conclusion)	19
(Personal experience→Listener response)	17
|→((Belief→Active recognition))→|	16
Belief→(Belief→Listener response)	15
(Metastatement→Active recognition)	11
Belief→(Belief→Belief)	10
|→((Example→Listener response))→|	9
Example→(Listener response→Example)	8
(Narrative→Listener response)	8
(Explain $\xrightarrow[\text{speaker}]{\text{within}}$ Belief)	7
(Belief→Listener response)→(Explain→Listener response)	6
(Active recognition→Example)	6
(Listener response→Conclusion)	6
Explain→(Listener response→Belief)	6
(Example $\xrightarrow[\text{speaker}]{\text{within}}$ Belief) $\xrightarrow[\text{speaker}]{\text{between}}$ Active recog.	6

Table 32 continued.

Chain	Total frequency of occurrence
(Completion→Active recognition)	5
(Narrative→Active recognition)	5
(Listener response→Narrative)	5
(Belief→Unsuccessful interruption)	5
|→((Explain→Listener response))→|	5
(Belief→Judgement)	5
(Explain→Belief)	5
(Avoidance→Belief)	5
Belief —within speaker→ (Explain —between speaker→ Active rec.)	5
(Belief —between speaker→ Active rec.) —within speaker→ (Belief —between speaker→ Active recognition)	5

8.3 Discussion

Before discussing the results of the chain analyses, it should be noted that the analyses, unlike those of Chapters 4, 6 and 7, are not concerned simply with individual speech states, but in many cases, with repetitions of the same speech event by the same speaker.

Taking the Activity chain analysis first, inspection of Table 31 indicates that the sample of 24 conversations comprise a large number of two-element chains, and a relatively small number of multi-element chains. The majority of the chains consisted primarily of Offers and either Consents or Reactions, more usually as two-element chains:

(Offer —— between ——→ Consent)	N = 1107
(Offer —— between ——→ Reaction)	N = 1035
(Offer —— between ——→ Offer)	N = 252
(Consent —— within ——→ Offer)	N = 120
(Reaction —— between ——→ Offer)	N = 80
(Consent —— between ——→ Offer)	N = 54

and less frequently in multi-element chains:

|→(Offer-between→Reaction)→|	N = 333
|→(Offer-between→Consent)→|	N = 174
Offer-between→(Offer-between→Consent)	N = 99
(Offer-between→Consent)-between→Offer	N = 44
(Offer-between→Reaction)-between→Offer	N = 44
(Offer-between→Reaction)-between→ →(Offer-between→Consent)	N = 32

(Offer-between→Consent)-between→

→(Offer-between→Offer) N = 15

|→|→(Offer-between Reaction)→|→| N = 14

Similar to the previous Activity analyses (Chapters 4 and 6), these results again suggest that the basic core of conversation, made by the dyads in the sample, is concerned with the imparting of information and ideas, and the subsequent agreement or acknowledgement by the other member of the dyad. In most cases, both the two-element and the multi-element chains demonstrate the interconnectedness of the Offer, Consent and Reaction speech phases. For example, after a phase of Offering, the most frequent occurrence is a between-speaker Consent phase, or a between-speaker Reaction, a between-speaker Offer phase being the least frequent of the three alternatives. Following both Consent and Reaction phases, the most frequent occurrence is a between-speaker Offer phase. Consequently, taking the Activity analyses of Chapters 4 and 6 and the present analysis together, the results indicate that not only do Offers and Consents tend towards cyclic within-speaker transitions, but when treated as phases, Offering phases are most usually followed by between-speaker Consent and Reaction phases, which are in turn most usually followed by phases of between-speaker Offering.

In addition, two other main two-element chains have been shown to occur:

(Request-between→Reply) N = 245

(Request-between→Consent) N = 205

and confirm the findings of the Markov analyses in Chapters 4 and 6.

Turning now to the Type chain analysis, the results shown in Table 32 are essentially similar to those obtained in Chapter 7. Six speech states, Beliefs, Examples, Explanations, Narratives, Active recognitions, and Listener responses predominate the speech stream, again indicating that the conversations sampled were concerned mainly with the imparting of one's beliefs and their subsequent acknowledgement in the form of either Active recognitions or Listener responses:

(Belief-between→acknowledgement)	N = 1240
(Example-between→acknowledgement)	N = 289
(Explain-between→acknowledgement)	N = 208
(Narrative-between→acknowledgement)	N = 28

The flow of the conversation tends to be embellished with Examples, Explanations and Narratives, either within- or between-speaker:

(Belief-between→Example)	N = 94
(Belief-within→Example)	N = 81
(Belief-between→Explain)	N = 25
(Belief-within→Explain)	N = 55
(Belief-within→Conclusion)	N = 19

In most cases, acknowledgement in either Active recognition or Listener response form, returns the conversation to a between-speaker Belief phase:

(acknowledgement-between→Belief) N = 219

(acknowledgement-between→Example) N = 41

(acknowledgement-between→Conclusion) N = 6

(acknowledgement-between→Narrative) N = 5

and less frequently to a within-speaker Belief phase:

(acknowledgement-within→Belief) N = 90

Multi-element chains, of which there were very few, confirm the interconnectedness of the major Type states. The most common chains comprised Beliefs and either Active recognitions or Listener responses:

(Belief-between→Active recognition)-between→Belief	N = 54
Belief-between→(Listener response-between→Belief)	N = 35
|→(Belief-between→Listener response)→|	N = 31
(Belief-between Active→recognition)	N = 16

etc.

In common with the Activity chain analysis, the Type chain analysis demonstrates a pattern of information exchange essentially similar to that given by the Markov analyses presented in Chapter 7. At the Activity level of analysis, the conversations sampled are essentially concerned with the Offering of information, and subsequent agreement or acknowledgement, the speech stream being embellished by a comparatively small number of Request/Reply and Request/Consent phases. At the Type level of analysis, a similar pattern of information exchange is seen, in which the information exchanged is generally in the form of Beliefs, but in a number of instances in the form of Examples, Explanations and Narratives.

8.4 Conclusion

There are two main conclusions that can be drawn from these chain analyses of conversation. Firstly, for the type of conversation studied and the method of speech classification used, the conversations would appear to be most usefully considered, at both the Activity and Type levels of analysis, in terms of two-event sequences, equivalent to those demonstrated by the Markov analyses presented in Chapters 4, 6 and 7. The analyses demonstrate quite clearly that a first-order Markov process is quite adequate to describe the conversations analysed. In addition, the analyses also indicate a marked lack of embedded speech events, suggesting that the conversations proceeded by each speaker giving their views about a topic, and then passing the speaking-turn to the other member of the dyad in a strictly organised fashion, rather than in an embedded question and answer form, as outlined in the Introduction (Section 8.1) to this chapter.

Secondly, the analyses confirm the Markov analyses (Chapters 6 and 7) in the suggestion that, at the Activity level of analysis, the conversations studied consist essentially of Offers of information and acknowledgement in the form of Consents and Reactions, and at the Type level of analysis the conversations consist essentially of Beliefs, Examples, Explanations, Narratives, and acknowledgements in the form of Active recognitions and Listener responses.

Although a number of multi-element chains were found, when repeated speech events made by the same speaker had been accounted for, these chains provided essentially similar information about the structure of the conversations studied as the Markov analyses presented in Chapters 4, 6 and 7. However, the multi-element chains did indicate that for a number of speech states, two-element phases were often cyclic, at both the Activity and Type levels of analysis. For example:

→(Offer-between→Reaction)→	N = 333
→(Offer-between→Consent)→	N = 174
→|→(Offer-between→Reaction)→|→	N = 14
→|→(Offer-between→Consent)→|→	N = 5
→(Belief-between→Listener response)→	N = 31
→(Belief-between→Active recognition)→	N = 16

The chain analyses presented in this chapter have demonstrated that, for the type of conversations studied, the dyadic sequences would appear to be most usefully considered in terms of two-event chains, equivalent to those demonstrated by the previous Markov analyses. However, when repeated speech events by the same speaker are deleted, and the sequences are construed in terms of phases rather than single speech events, the chain analyses have shown that there are occasions when, in addition to single states, two-phase chains may also be cyclic. Such results lend support to the notion that the conversations studied consist essentially of Offer, Consent and Reaction phases at the Activity level of analysis, and phases of Belief-, Example-, Explanation-, and Narrative-giving, followed by acknowledgement in the form of Active recognitions and Listener responses, at the Type level of analysis.

CHAPTER 9

AN ANALYSIS OF CONVERSATIONAL STRUCTURE USING INFORMATION THEORY

9.1 Introduction

In both the Introduction and Chapter 1, communication was discussed in terms of a number of characteristics. One of these was 'structure', and refers to the organisation of events in the conversational system. For communication to occur in conversation, the participants must behave in a non-random fashion or there would be no exchange of information. If the participants behave non-randomly, by definition, their behaviour is structured in some manner. Whenever variables, and in this case the interactants, behave non-randomly, constraints are said to be operating as the presence of organisation (i.e. non-randomness) indicates the existence of constraints (Ashby, 1968). This implies that as constraints are operating, the flow of speech events is always predictable, to some extent at least.

Whereas the analyses of Chapters 4 and 6 have been concerned with establishing the types of constaint that operate in informal conversation (i.e. Replies tend to follow Requests, etc.), the present chapter is concerned with estimating the degree to which these constraints operate, as well as the source of the constraint. For example, constraints may operate simply because some speech events occur more frequently than others (distributional structure), or alternatively, the source of the constraint may be due to the tendency of some events to occur in predictable strings (sequential structure).

The analysis of conversational structure through system constraints may be carried out using Shannon's (1948) Theory of Communication as a theoretical framework and forms the subject of this chapter.

Shannon's Theory rests on two main assumptions: (1) information is associated with a selection process, and (2) this process is statistical in nature in the sense that it involves probabilities of events (Frick, 1968). From these two assumptions, it is possible to see how the stochastic nature of the conversational process is subsumed within the Theory of Communication.

Information encompasses the concepts of choice and uncertainty, information being concerned with a speaker's choice, and designed to reduce the listener's uncertainty. In its general sense therefore, information and uncertainty are quantitatively equivalent: the less uncertainty there is about the occurrence of an event, the less information there is in the event, information only existing if there is some a priori uncertainty about the occurrence of an event (Garner, 1962). Consequently, if uncertainty can be measured, the information present may be taken to be simply the decrease in uncertainty (Garner, 1962). Typically, maximum information (maximum uncertainty) exists when every event in a finite set of events has equal probability of occurrence, minimum information (minimum uncertainty) only occurring when the subsequent event is perfectly predictable (i.e. when the subsequent event is always the same).

In mathematical terms, the information in a set of events is measured as a $\log_2$ function of the number of possible events and expressed as bits of information. Therefore, the greater the number of possible events, the greater the amount of information present and the greater the uncertainty of the subsequent event. The information in a set of events may be calculated by:

$$\hat{H} = - \sum_{k=1}^{r} \hat{p}(k) \log_2 \hat{p}(k)$$

where r is the number of states

$\hat{p}(k)$ is the probability of the k^{th} event

Shannon (1948)

which is equivalent to the measure of uncertainty U, used by Garner and McGill (1956) and Garner (1962);

$$\hat{U} = - \sum_{k=1}^{n} \hat{p}(k) \log_2 \hat{p}(k)$$

where n is the number of states

$\hat{p}(k)$ is the probability of the k^{th} event

Garner (1962)

This statistic can be used to measure the constraints operating in a conversational system directly, where 'constraint' has been defined by Garner (1962) as "the amount of inter-relatedness, or structure, of a system of variables as measured in informational terms" (p.145). This type of analysis has been used a number of times in the conversational literature, for example by Stech

(1970), Gouran and Baird (1972) and Hawes and Foley (1973), but the analyses, being univariate, tend to provide only a general picture of the overall structure of conversation. Whereas univariate analyses can indicate the amount of uncertainty associated with a specific speech state, they are not however, applicable to the analysis of the degree to which constraints are operating in the speech stream, or to identifying the source of the constraint. However, multivariate information analyses, which have been used by both Stech (1975) and Penman (1980), do not have such limitations.

The main advantage of multivariate information analyses is that the statistic can be partitioned in a manner analogous to that used in ANOVA (Garner & McGill, 1956) to provide estimates of distributional, sequential and error uncertainty, which can in turn be used to estimate the source and degree of the constraints operating in the conversational system. For example, in a review of the literature using multivariate information analyses to assess the structure of conversation, Stech (1975) noted that the predictability of conversational acts due to the actual distribution of acts in the conversations ranged from 3.4% to 41.9%, and predictability due to the sequential structure of the speech acts ranged from 3.8% to 29.8%, the sample of conversations being drawn from discussion groups, classroom interaction and police-civilian telephone calls.

In this chapter, the informal conversations are examined in terms of speech Activity and speech Type, using both multivariate and univariate information analyses. The former are used to describe the degree and source of the constraints between speech events, and the latter are used to describe the degree of uncertainty in predicting the distribution of subsequent events, given a specified antecedent event.

9.2 Data analysis overview

The same 24 dyadic conversations analysed in Chapters 4, 6, 7 and 8 have been used in the following analyses. The data have been analysed using the methodology specified by Hake and Garner (1951), in which information statistic values are calculated separately for each dyadic sequence and then averaged over all the dyads in the sample.

Three different analyses have been performed on the conversations coded for both the Activity and Type of speech uttered: (1) a multivariate analysis of conversational structure, by which the degree and source of the constraint between speech events has been specified, (2) a univariate analysis of the distribution of speech events in the sample of conversations, and (3) a univariate analysis of the sequential structure of the conversations.

Program 15 (Appendix G) was written to perform the multivariate uncertainty analyses.

9.3 Information analyses of conversation coded for speech activity

9.3.1 Multivariate analysis of conversational structure

The rationale and derivation of the following Uncertainty statistics may be found in Garner and McGill (1956) and Garner (1962). An illustrative example is given for the analysis of a first-order process in which the following notation is used:

y = the criterion variable - subsequent event.

x = the predictor variable - antecedent event.

In a first-order process, there are two variables under study, a subsequent or criterion variable (y), and an antecedent or predictor variable (x). Multivariate information analysis partitions the uncertainty of the criterion variable into a number of components:

$$\hat{U}_{(y)} = \hat{U}_{(y:x)} + \hat{U}_{x(y)}$$

where $\hat{U}_{(y)}$ is the total amount of uncertainty in the subsequent event. The smaller this uncertainty, the more predictable the subsequent event. The upper limit of this uncertainty, often referred to as the 'nominal' uncertainty, is given by:

$$\hat{U}_{(y)\ max} = \log_2 k$$

where k is the number of possible speech states.

$\hat{U}_{x(y)}$ is the error uncertainty (Garner & McGill, 1956) or 'noise' (Donohue, Hawes & Mabee, 1981); in a similar manner to ANOVA it represents the amount of uncertainty remaining in the criterion event (y) after the predictable uncertainty has been accounted for and indicates the extent to which the subsequent event is not constrained by the antecedent event. $\hat{U}_{(y:x)}$ is the predictable uncertainty, usually referred to as the multiple contingent uncertainty (Garner & McGill, 1956), and measures the amount of information in the subsequent event which is contingent on the antecedent event. The smaller the multiple contingent uncertainty, the smaller the sequential dependencies in the speech stream.

Turning now to the analysis of conversation; a matrix of events can provide two estimates of the predictability, or structure of the speech stream, the simpler being the <u>distributional constraint</u>. This is mathematically equivalent to the complexity of the system (Garner, 1962) and measures the extent to which uncertainty is decreased due to the actual probabilities of event occurrence not being equally distributed (Garner, 1962), or, the degree to which the sequence is structured by virtue of some events occurring more frequently than others. Therefore, the more equiprobable the events are, the more complex the system.

The second estimate of structure is <u>sequential constraint</u>. This represents the degree to which a set of events tend to occur in predictable strings, and refers to the degree of organisation of the speech events and the extent to which constraints are operating sequentially in the speech stream.

The results of the basic uncertainty analyses, shown in Table 33, indicate that the mean amount of uncertainty in the conversational system is 73.96% for overall-, 56.36% for within-, and 78.66% for between-speaker transitions. Of this, the mean amount of uncertainty accounted for by constraints operating between the antecedent and subsequent events is 14.55% for overall-, 33.00% for within-, and 23.60% for between-speaker transitions, expressed as a percentage of the total uncertainty in the conversational system. The results for the individual dyads are shown in Appendix F.

Table 33 Mean system Uncertainty - speech Activity.

Transition	Uncertainty source			
	Nominal	Total	Error	Multiple contingent
	%	% Nominal	% Total	% Total
Overall	100	73.96	85.45	14.55
Within	100	56.36	67.00	33.00
Between	100	78.66	76.40	23.60

At this point, it is interesting to note the extent to which the total predictability of the system is attributable to each type of constraint operating in the conversational system. Distributional constraint, defined as the difference between the nominal and actual uncertainty (Garner, 1962), has been calculated by expressing the difference between the total system uncertainty shown in Table 33 and the nominal uncertainty, as a percentage of the nominal uncertainty in a seven-state system. Sequential constraint, the difference between the total system uncertainty and the error uncertainty (Garner, 1962), is equivalent to the multiple contingent uncertainty given in Table 33. Expressed as a percentage of the nominal uncertainty, the multiple contingent uncertainty represents the degree of sequential constraint in the conversational system. Summing the distributional and sequential constraints yields the total amount of constraint operating between events in the conversational system. The results of this analysis are shown in Table 34.

Table 34 Mean distributional and sequential structure - speech Activity.

Transition	Structure		
	Distributional (Complexity)	Sequential (Organisation)	Total
Overall	26.04%	10.76%	36.80%
Within	43.64%	18.60%	62.24%
Between	21.34%	18.56%	39.90%

Overall, the results suggest that the speech sequences are structured to a moderate degree, the extent of the predictability accounted for by the distribution of events in the speech stream being approximately double that of the sequential arrangement of events for within-speaker transitions and approximately equal for between-speaker transitions. The greater predictability of within-speaker transitions is accounted for mainly by an increase in distributional structure, and indicates that within-speaker speech events are more predictable, and less complex, as there is a smaller range of probable subsequent within-speaker speech events. At the Activity level of analysis, the speech sequences may therefore be seen as being moderately complex (a moderate amount of uncertainty being accounted for by the distribution of events) and fairly low in organisation (a fairly small amount of uncertainty being accounted for by the sequential arrangement of events).

9.3.2 Univariate analysis of distributional structure

In the previous section, it was demonstrated that although the sequential constraint operating between antecedent and subsequent events was similar for both within- and between-speaker transitions, the predictability attributable to the distribution of the speech events was approximately double for within-speaker transitions. This section is designed to assess which, if any, of the antecedent acts contribute to the high distributional structure.

The distributional structure of the conversations, coded for speech Activity, has been analysed by simply calculating the frequency with which each speech category occurs as a within- or between-speaker event for the 24 dyads combined, and expressing this as a percentage of the total number of within- or between-speaker events, respectively. The results of this analysis are shown in Table 35.

Table 35 Distributional structure of conversation coded for speech Activity.

Speech category	Within-speaker N	Within-speaker %	Between-speaker N	Between-speaker %
Offer	4136	67.88	3045	38.27
Reply	324	5.32	507	6.37
Consent	554	9.09	1905	23.94
Dissent	304	4.99	353	4.44
Modify	230	3.77	130	1.64
Reaction	32	.53	1501	18.86
Request	513	8.42	516	6.48

The results indicate that the high degree of distributional structure for within-speaker transitions is accounted for mainly by Offers, which form 67.88% of the within-speaker transitions. Between-speaker transitions are distributed more widely across the speech categories, Offers (38.27%), Consents (23.94%) and Reactions (18.86%) together forming 81.07% of the between-speaker transitions. The high within-speaker distributional structure shown in Table 34, may therefore be seen to be due to the Offer speech category, between-speaker transitions tending to be more widely distributed over the Offer, Consent and Reaction categories.

9.3.3 Univariate analysis of sequential structure

This section is intended to assess which antecedent events contribute highly to the sequential structure of the speech stream, giving a possible indication of those acts that may be seen as 'structure - inducing'.

The sequential structure of conversation has been assessed by intially calculating the amount of information in a transition matrix row, using the univariate information statistic, H:

$$\hat{H} = - \sum_{k=1}^{r} \hat{p}_{(k)} \log_2 \hat{p}_{(k)}$$

This statistic yields the amount of information in each of the seven categories of speech. The sequential structure of events corresponds to the relative uncertainty (R.U.) in predicting a subsequent event and is determined by:

$$R.U. = \frac{\text{Actual Uncertainty in an antecedent event}}{\text{Nominal (Maximal) Uncertainty}}$$

For example, if the R.U. is large for a given antecedent event, the range of subsequent events is large, and the antecedent event cannot therefore be said to be structure-inducing. However, if the R.U. is small for a given antecedent event, the range of predictable subsequent events is small, and the antecedent event may, therefore, be seen as structure-inducing.

Four separate analyses have been performed for the 24 dyads combined: (1) overall transitions, in which the type of transition (within- or between-speaker) is ignored, (2) within-speaker transitions, in which the R.U. has been calculated with the knowledge that a within-speaker transition follows, (3) between-speaker transitions, in which the R.U. has been calculated with the knowledge that a between-speaker transition follows, and (4) a within- plus between-speaker transition analysis, in which the R.U. has been calculated with the knowledge that either a within- or between-speaker transition may follow. The results of these analyses are shown in Table 36. The information statistic calculations may be found in Appendix F.

<u>Table 36</u> Relative Uncertainty of a subsequent event following a specified antecedent event.

Speech Activity	Transition type			
	Overall R.U.%	Within R.U. %	Between R.U.%	Within+Between R.U. %
Offer	67.5	24.7	76.1	64.3
Reply	82.9	50.3	79.2	73.4
Consent	64.5	71.3	49.9	70.9
Dissent	78.3	59.3	84.7	78.4
Modify	84.8	65.7	81.8	80.8
Reaction	43.8	68.2	32.5	50.6
Request	72.3	44.0	72.7	67.9

In summary, the results suggest that for within-speaker transitions, the Offer, Reply and Request categories are the most structure-inducing, and the

Consent and Reaction categories are the most structure-inducing of the between-speaker transitions, as they have the smallest subsequent event relative uncertainties. This, of course, assumes that one knows the transition type (within- or between-speaker) for the next speech event. If this information is not available, then predicting the identity of the next event regardless of speaker (Overall transitions) becomes relatively uncertain, Reactions having the lowest associated subsequent event uncertainty. Similarly, attempting to predict the identity of the next speech event when the transition type can be either within- or between-speaker is a very uncertain task, Reactions again being associated with the lowest amount of consequent event uncertainty.

These results indicate that there is a considerable degree of uncertainty in predicting the identity of a consequent event, given a specified antecedent event, and reflect the low sequential constraints operating between speech events that were shown in Table 34. However, the uncertainty can be reduced if the transition type is known (either within- or between-speaker) for the Offer, Reply, Consent, Reaction and Request categories of speech.

Whereas the last set of analyses was concerned with the regularity of a subsequent event following a specified antecedent event, the analyses to be presented next are concerned with the regularity of an antecedent event preceding a subsequent event (i.e. the consistency with which a specified subsequent event is preceded by the same antecedent event).

Using the same measure of Uncertainty as before, three analyses are presented for (1) overall transitions, in which the type of transition is not taken into account, (2) within-speaker transitions, in which the R.U. has been calculated with the knowledge that the preceding antecedent was a within-speaker event, and (3) between-speaker transitions, in which the R.U. has been calculated with the knowledge that the preceding antecedent was a between-speaker event. The results are shown in Table 37. The information statistic calculations may be found in Appendix F.

<u>Table 37</u> Relative Uncertainty of an antecedent event preceding a specified subsequent event.

Speech Activity	Transition type		
	Overall R.U.%	Within R.U.%	Between R.U.%
Offer	67.3	50.4	74.3
Reply	70.1	44.2	43.0
Consent	70.0	50.2	59.3
Dissent	78.9	55.8	77.9
Modify	81.9	73.6	80.0
Reaction	35.6	0	40.8
Request	81.3	80.9	79.6

The results again reflect the low sequential structure of the conversations and indicate considerable uncertainty in speech event predictability. However, it is apparent from Table 37 that within-speaker subsequent events are generally preceded by a smaller range of antecedent events than between-speaker events, as the relative uncertainties are generally smaller for within-speaker events. In particular, Offers, Replies, Consents and Dissents would appear to be associated with less uncertainty concerning the preceding within-speaker antecedent event, as are between-speaker Replies and Reactions.

It should be noted that even though within-speaker Reactions have zero uncertainty concerning the identification of the preceding antecedent event, they are not considered to be structure-inducing. This is due to the manner in which the Reaction category is scored. Transitions to a within-speaker Reaction, other than from itself, are considered logically impossible (Section 4.5.2 as by definition, a Reaction is a vocalisation made in response to the other person. Consequently, the only within-speaker transition to a Reaction is from a Reaction, and therefore, the relative uncertainty in predicting the identity of the previous antecedent event is zero.

9.4 Information analyses of conversation coded for speech Type.

9.4.1 Multivariate analysis of conversational structure

Using the same information measures as the Activity analyses, a multivariate analysis of conversational structure for conversation coded for speech Type has been performed, by which the degree and source of the constraint between speech events may be specified.

The results of the basic uncertainty analysis for the 24 dyad group are shown in Table 38, and indicate that the amount of uncertainty in the conversational system is 60.98% for overall-, 55.68% for within-, and 58.21% for between-speaker transitions. Of this, the mean amount of uncertainty accounted for by constraints operating between the antecedent and subsequent events is 16.47% for overall-, 28.09% for within-, and 22.55% for between-speaker transitions, expressed as a percentage of the total uncertainty in the conversational system. The results for the individual dyads are shown in Appendix F.

Table 38 Mean system Uncertainty - speech Type

Transition	Uncertainty source			
	Nominal %	Total % Nominal	Error % Total	Multiple contingent % Total
Overall	100	60.98	83.53	16.47
Within	100	55.68	71.91	28.09
Between	100	58.21	77.45	22.55

Distinguishing between the two types of constraint that operate in the conversational system (Table 39), the amount of uncertainty accounted for by the distribution of events is 39.02% (Overall transitions), 44.32% (within-speaker), and 41.79% (between-speaker), and the amount of uncertainty accounted for by the sequential arrangement of speech events is 10.04% (overall transitions), 15.64% (within-speaker), and 13.13% (between-speaker).

<u>Table 39</u> Mean distributional and sequential structure - speech Type.

Transition	Structure		
	Distributional (Complexity)	Sequential (Organisation)	Total
Overall	39.02%	10.04%	49.06%
Within	44.32%	15.64%	59.96%
Between	41.79%	13.13%	54.92%

These results are essentially similar to those of the Activity analyses, and indicate that, at the Type level of analysis, the speech sequences may be seen as being moderately complex (a moderate amount of uncertainty being accounted for by the distribution of events) and fairly low in organisation, as only a fairly small amount of uncertainty is accounted for by the sequential organisation of events.

9.4.2 Univariate analysis of distributional structure

In the same manner as the Activity analysis, this section has been designed to assess which, if any, of the antecedent events contribute highly to the distributional structure of the speech stream. The distributional structure has been analysed by calculating the frequency with which each category of speech occurs as a within- or between-speaker event, for the 24 dyads combined, and expressing this as a percentage of the total number of within- and between-speaker events, respectively. The results of this analysis are shown in Table 40.

The results indicate that there are five speech categories that are the main contributors to the distributional structure: Beliefs (38.95%), Explanations (5.99%), Active Recognitions (16.26%), Examples (11.00%), and Listener Responses (9.96%); they closely correspond to the major Type states evidenced in Chapter 7. However, by way of extension of the Chapter 7 results, the present analysis also indicates that the Belief, Personal Experience, Narrative, Explanation, Revision and Example categories tend to be within-speaker events, and the Active Recognition, Listener Response and Completion categories tend to be between-speaker events.

Table 40 Distributional structure of conversation coded for speech Type.

Speech category	Distributional structure					
	Overall		Within		Between	
	N	%	N	%	N	%
Belief	5473	38.95	2854	46.84	2619	32.91
Personal Detail	130	.93	64	1.05	66	.83
Personal Experience	334	2.38	218	3.58	116	1.46
Narrative	348	2.48	244	4.00	104	1.31
Subjective feelings	15	.11	9	.15	6	.08
Attitude	74	.53	38	.62	36	.45
Explain	842	5.99	466	7.65	376	4.73
Synthesis	10	.07	4	.07	6	.08
Summary	60	.43	36	.59	24	.30
Judgement	187	1.33	81	1.33	106	1.33
Avoidance	94	.67	43	.71	51	.64
Negation	29	.21	14	.23	15	.19
Active Recognition	2284	16.26	398	6.53	1886	23.70
Salience	27	.19	9	.15	18	.23
Conclusion	410	2.92	234	3.84	176	2.21
Example	1546	11.00	965	15.84	581	7.30
Emotional Support	5	.04	2	.03	3	.04
Apology	13	.09	9	.15	4	.05
Direction	193	1.37	102	1.67	91	1.14
Metastatement	159	1.84	120	1.97	139	1.75
Revision	142	1.01	116	1.90	26	.33
Restatement	77	.55	25	.41	52	.65
Completion	58	.41	8	.13	50	.63
Listener Response	1400	9.96	30	.49	1370	17.22
Unsuccessful Interruptions	28	.19	2	.03	26	.33
Laughter	12	.09	2	.03	10	.13

9.4.3 Univariate analysis of sequential structure

This section examines which antecedent events contribute highly to the sequential structure of the speech stream, using the same univariate information measures as the previous Activity analyses.

Three separate analyses have been performed for the 24 dyads combined: (1) within-speaker transitions, in which the R.U. has been calculated with the knowledge that a within-speaker transition follows, (2) between-speaker transitions, in which the R.U. has been calculated with the knowledge that a between-speaker event follows, and (3) a within-plus between-speaker transition analysis, in which the R.U. has been calculated with the knowledge that either a within- or a between-speaker transition may follow. The results of these analyses are given in Table 41. The information statistic calculations may be found in Appendix F.

Table 41 Relative Uncertainty of a subsequent event following a specified antecedent Type event.

Speech category	Transition type		
	Within	Between	Within + Between
	R.U. %	R.U. %	R.U. %
Belief	46.66	54.34	59.35
Personal Details	54.68	62.65	65.47
Personal Experience	46.76	51.92	57.57
Narrative	28.49	43.17	45.03
Subjective Feelings	45.06	38.12	51.56
Attitude	53.31	62.44	65.52
Explain	44.23	51.27	56.98
Synthesis	40.89	20.66	42.91
Summary	49.61	52.85	59.22
Judgement	49.72	42.93	44.93
Avoidance	51.76	56.42	59.42
Negation	21.74	46.59	43.56
Active Recognition	59.39	58.68	66.22
Salience	49.76	67.69	65.69
Conclusion	55.85	50.25	58.75
Example	49.25	50.38	58.61
Emotional Support	33.72	33.72	33.72
Apology	53.19	32.38	54.12
Direction	55.04	58.25	63.82
Metastatement	59.91	58.63	65.42
Revision	56.06	53.46	62.07
Restatement	60.19	59.72	66.89
Completion	35.08	44.49	49.45
Listener Response	62.93	55.59	59.45
Unsuccessful Interruptions	33.21	35.72	43.99
Laughter	21.27	50.44	48.98

In summary, the results suggest that there is a fairly high degree of uncertainty in predicting a subsequent event from all of the antecedent categories, However, the following speech states may be seen as more structure-inducing than any others, as they have the lowest subsequent event relative uncertainties: Narratives, Negations, Emotional Support, Unsuccessful Interruptions, Laughter and Completions for within-speaker transitions, and Subjective Feelings, Synthesis, Emotional Support, Apology, and Unsuccessful Interruptions for between-speaker transitions. With the exception of Narratives, all of these transitions occurred very infrequently (less than 26 times in the complete sample - see Chapter 7 (Section 7.2) for a discussion of this sampling problem), and therefore the low relative uncertainties are more likely due to a basic sampling problem than their being of a structure-inducing nature. However, the within-speaker Narrative category is the exception, and is highly structure-inducing, the most probable subsequent event being a within-speaker Narrative, occurring on 47% of occasions (Chapter 7).

The previous set of analyses was concerned with the consistency with which a specified antecedent event was followed by the same subsequent event. The series which follows is concerned with the regularity with which an antecedent event precedes a specified subsequent event.

Using the same measure of Uncertainty as before, two analyses are presented for (1) within-speaker transitions in which the R.U. has been calculated with the knowledge that the preceding event was a within-speaker event, and (2) between-speaker transitions in which the R.U. has been calculated with the knowledge that the preceding event was a between-speaker event. The results of these analyses are given in Table 42. The information statistic calculations may be found in Appendix F.

The results are essentially the same as the previous set of analyses, subsequent events tending to be associated with a large range of antecedent events. Again, the outstanding exceptions are within-speaker Narratives, which are preceded by themsleves on 66% of occasions. In addition, Active Recognitions tend to have fairly predictable antecedent events, being preceded by themselves on 67.1% of occasions, and by Beliefs and Listener Responses on 11.1% and 9.5% of occasions, respectively.

Table 42 Relative Uncertainty of an antecedent event preceding a specified subsequent Type event.

Speech category	Transition type	
	Within	Between
	R.U. %	R.U. %
Belief	52.65	57.16
Personal Details	63.10	57.27
Personal Experience	50.68	56.23
Narrative	38.44	40.72
Subjective Feelings	43.80	31.04
Attitude	65.33	55.87
Explain	52.04	56.76
Synthesis	17.25	38.12
Summary	54.65	48.93
Judgement	60.89	54.51
Avoidance	66.99	64.69
Negation	47.17	45.66
Active Recognition	39.61	59.70
Salience	46.74	56.78
Conclusion	54.38	43.70
Example	56.97	54.74
Emotional Support	21.27	33.72
Apology	43.80	31.91
Direction	63.29	64.80
Metastatement	59.69	63.55
Revision	59.08	50.89
Restatement	51.51	61.25
Completion	29.91	34.99
Listener Response	54.95	57.34
Unsuccessful Interruptions	0	32.32
Laughter	21.27	50.44

9.5 Discussion

The analyses, summarised in Tables 33 and 34, indicate that the conversations, at the Activity level of analysis, are relatively highly structured. The amount of Uncertainty accounted for by within- and between-speaker transitions was 62.24% and 39.90%, respectively. Within-speaker transitions are therefore more predictable, and more highly structured than between-speaker transitions.

Distinguishing between distributional and sequential structure, the amount of Uncertainty in the conversational system accounted for by the sequential arrangement of events is approximately equal for both types of speaker transitions (Within-speaker = 18.60%, Between-speaker = 18.56%). However, looking at the distribution of events, within-speaker transitions account for approximately double the amount of Uncertainty in the speech system (Within-speaker = 43.64%, Between-speaker = 21.34%). The structure of within-speaker transitions is, therefore, fairly simple, as indicated by the relatively large amount of uncertainty accounted for by the distribution of events, a single type of event, Offers, tending to dominate the within-speaker speech stream (Table 35). In addition, the small amount of uncertainty accounted for by the sequential arrangement of events indicates that there is little organisation of within-speaker events in the speech stream. In contrast, between-speaker transitions are more complex (as indicated by the relatively small amount of uncertainty accounted for by the distribution of events), Offers, Consents, and Reactions being the main contributors to the distributional structure (Table 35). Similar to within- speaker transitions, between-speaker transitions are low in organisation (Table 34).

The analyses presented in Table 36, concerned with the sequential structure of conversation, indicate that those events whose subsequent events are the most easily predicted are within-speaker Offers, Replies, and Requests, and between-speaker Consents and Reactions. These categories may be seen as 'structure inducing', as they tend to be sequentially arranged in the speech stream.

In Chapter 6, it was indicated that the Offer and Reply categories tended to be proactive, functioning by retaining the speaker turn, and the Consent, Reaction and Request categories tended to be reactive as they function by relinquishing the speaker turn. It is suggested here, that these categories gain their proactive/ reactive function because they are structure-inducing, the specific function being determined by whether the event is within- or between-speaker structure inducing. However, there is one exception. Whereas the Offer and Reply categories are proactive and within-structuring and the Consent and Reaction categories are reactive and between-structuring, the Request category is exceptional by being both reactive and within-structuring. The way in which Requests function may be elucidated by inspection of Table 22 (Chapter 4). Requests function by offering the speaker-turn, and in this sense are reactive, the most likely subsequent events being Replies and Consents. In comparison with within-speaker transitions from a Request, the outcome of a between-speaker transition is more variable, or uncertain, as there are two probable outcomes, Reply or Consent. There is less uncertainty associated with a within-speaker transition as there is only a single probable transition to the Request state. Consequently, Requests are reactive, as they function by offering the floor, and are also within-structuring, for if the turn, when offered, is not immediately taken, the most probable subsequent event is another Request by the same speaker.

The final set of Activity analyses was concerned with the consistency with which one event is preceded by another. The results (Table 37) again reflect the low sequential structure of the conversations, within-speaker events usually being preceded by a smaller range of probable events than between-speaker events. The within-speaker speech events with the least antecedent uncertainty (Offer, Reply, Consent, and Dissent) all have the same state as the most probable antecedent (Table 22), thus indicating the stable nature of these speech states. Once entered into, these states tend to be self-perpetuating. In the case of between-speaker events, the Reply and Reaction categories are associated with the least antecedent uncertainty. Inspection of Table 22 indicates that in the case of Replies, the antecedent is usually a Request, and an Offer in the case of Reactions.

Taken together, these analyses allow one to differentiate between states which tend to have the same state as both antecedent and subsequent event, and states whose subsequent events are very variable. Within-speaker Offers and Replies tend to be of the former type, within-speaker Consents and Dissents, the latter type. In conversational terms, this indicates that on entering the Offer and Reply states, it is most probable that the preceding and next events are the same. However, for Consents and Dissents, although the preceding event was most probably the same, the next event is liable to be uncertain. The implication of these results is that the information-giving categories, Offer and Reply, are more stable and reflect 'long term' phases in conversation, whereas the Consent and Dissent states represent 'short term' phases, in which an agreement or disagreement is made prior to the giving of more information (Table 22).

To summarise the Activity uncertainty analyses, the conversations have been shown to be quite highly structured, within-speaker transitions accounting for considerably more uncertainty than between-speaker transitions, the majority of this being accounted for by the distribution of events in the speech stream. A univariate analysis has indicated that the within-speaker Offer, Reply, and Request categories, and the between-speaker Consent and Reaction categories are 'structure inducing', and it is suggested that the proactive function of Offers and Replies, and the reactive function of Consents and Reactions are gained from this ability to structure the speech stream. Additionally, the univariate uncertainty analyses indicate that within-speaker Offers and Replies represent 'long term' phases in conversation (both antecedent and subsequent phases tend to be the same) and Consents and Dissents represent 'short term' phases in conversation (the antecedent event tends to be the same, but the subsequent event is more variable).

Turning now to the analysis of conversation coded for speech Type, Table 39 indicates quite a high degree of structure in the speech stream, within- and between-speaker transitions accounting for approximately equal amounts of uncertainty, the distribution of events in the speech stream again accounting for the larger part of the conversational structure.

Distinguishing between distributional and sequential structure, the amount of uncertainty accounted for in the conversational system by the sequential arrangement of events is 15.64% and 13.13%, and by the distribution of events is 44.32% and 41.79% for within- and between-speaker events, respectively. The structure of the speech stream, as indicated by the amount of uncertainty accounted for by the distribution of events is therefore, fairly simple, although more complex than the Activity dimension of speech. In addition, there is little organisation of events in the speech stream, indicated by the fairly low amount of uncertainty accounted for by the sequential arrangement of events.

To summarise the Type uncertainty analyses, the conversations have been shown to be quite highly structured, the majority of the structure for both within- and between-speaker transitions being accounted for by the distribution of events in the speech stream. Confirming the results of Chapter 7, Beliefs, Explanations, Active Recognitions, Examples, and Listener Responses are again shown to be the major conversational states, by accounting for the larger part of the distributional structure of conversation (Table 40). In addition, the results presented in Tables 41 and 42 indicate that Narratives are the main 'structure-inducing' state, tending to be preceded and followed by the same Narrative state, and implies that Narratives form 'long term' phases of conversation, in which story-telling is the prime intention.

9.6 Conclusion

As noted in the Introduction to this chapter, for communication to occur in conversation, the participants must behave in a non-random fashion, or there could be no exchange of information. Non-random behaviour implies some form of structure, and, as Ashby (1968) has pointed out, the presence of structure indicates the existence of constraints in the conversational system. The analyses presented in this chapter indicate a large degree of structuring of the speech stream; in so far as people communicate with each other, what person A says to person B may therefore be considered to constrain person B's response.

The degree to which a system is seen to be structured is dependent not only on the type of conversation analysed, but also the method of speech analysis used. For example, Gouran and Baird (1972) showed that there were differences in structure between informal and problem-solving groups, and Stech (1975), in a review of the literature covering the period 1964 - 1972, showed that total structural estimates varied between 7.2% and 51.4% for a sample of conversations comprising discussion groups, classroom interaction, and police-civilian telephone calls. In addition, Stech (1968 - reported in Stech, 1975) noted that the structural estimate varied by up to 30%, according to the method used for classifying and coding the conversational speech. However, these sources of variation aside, the estimates of conversational structure presented in this chapter have a number of important implications for the models of conversation presented in Chapters 4 and 7 and their application in the construction of social strategies.

Firstly, the uncertainty analyses presented in this chapter indicate that there is a considerable degree of constraint operating in the speech system, and that although subsequent event predictability is high, subsequent events are nevertheless quite variable. Although the existence of constraints indicates the reciprocal nature of conversational interaction, reciprocation is a probabilistic tendency and not a foregone conclusion - the data presented suggesting that there are always possibilities for alternative directions, for both within- and between-speaker transitions. This implies that a conversational rule never specifies an absolute one-to-one relationship between speech events, but rather a 'range of probable outcomes'. In using the models of informal conversation presented in Chapters 4 and 7 in a practical application, such as social skills training procedures, it is important therefore, to bear in mind the probabilistic relationship between conversational events. Although the models of conversation developed in Chapters 4 and 7 have been used to extract a number of rules and strategies (Chapters 6 & 7), it should be noted that these rules are simply 'conventional practices' (Argyle, 1979, p.24) having a specified probability of occurrence, and as such, may be broken.

The second point to note is that the analyses presented in this chapter indicate that the ability to predict subsequent speech events depends, for speech coded for both Activity and Type, to a large extent on the distributional structure of the speech stream - the fact that some speech events occur more often than others. Consequently, when constructing rules to express the sequential ordering of conversational events, some rules will primarily be the result of the distribution of events in the speech stream, and others, the result of the sequential constraints operating in the conversational system. In the present analysis, an Offer followed by a within-speaker Offer would be a rule of the former type, and a Request followed by a Consent, the latter.

On a more speculative note, it is suggested that the methods of analysis employed in this chapter may be used to assess the style of conversation. For example, as discussed earlier, it is known that the estimates of conversational structure are dependent on the type of conversation studied and the method of speech classification used (Stech, 1975). It is therefore suggested that by using the same method of speech classification throughout, the measures of system constraint employed in this chapter may be used to estimate the 'relative-structuring' of conversations taking place in a variety of situations and expressing different relationships. Such an approach could potentially yield information about the style of an interaction, enabling for example, the formality of a conversation to be empirically determined and quantified.

Finally, the analyses or distributional structure emphasise the findings of Chapters 6 and 7, demonstrating that the Offer, Consent, and Reaction categories at the Activity level of analysis, and the Belief, Explanation, Active Recognition, Example and Listener Response categories at the Type level of analysis, are the main conversational states. In addition, the present analyses have indicated that a number of speech states are structure-inducing. For example, a number of two-event sequences have been demonstrated in Chapters 4, 6, 7 and 8, such as Request-Reply, Request-Consent, Offer-Reaction, Consent-Offer, Reaction-Offer, and Narrative-Narrative transitions which have been shown to reflect the sequential arrangements that exist in the conversations.

CHAPTER 10

CONCLUSIONS

In this final chapter, the main findings of the thesis will be reviewed, and their implications discussed.

Reviewing the literature concerned with the structure and internal organisation of conversation, and in particular, the model of social interaction proposed by Argyle and Kendon (1967), it became increasingly apparent that little was known about both the role and nature of the verbal elements in conversation, and their sequential arrangement. Consequently, an analysis of conversational structure, in terms of speech content, became the main theme of this research project.

Many of the early studies of conversation took a 'monadic' approach to the analysis of conversation and were concerned only with the effect that one person produced in another rather than with the reciprocal nature inherent in interpersonal communication. By contrast, it was argued at the outset of this project that the study of a communication process, and in this case informal conversation, should be concerned with the analysis of the inter-relationship between individuals, taking into account the time-ordering of events.

In studying communication processes it is not simply a matter of changing the approach from the study of the individual to the study of the relationship, nor a change in emphasis from the static to the dynamic, but rather, a change in theoretical orientation, as well as a corresponding change in methodology that is required. This reorientation could be achieved by using General systems theory as a theoretical framework, which had the added advantage of subsuming within it Argyle and Kendon's (1967) model of social interaction.

By virtue of conceptualising conversation as a process, conversation was considered to be composed of a string of concatenous, interdependent events. Consequently, before analysis of the structure of conversation could proceed, a method of 'segmenting' the speech stream into analysable units was required.

Although a number of methods for analysing speech already existed, these were thought to be inappropriate for the present study: either they were too situation-specific, or they possessed fundamental theoretical and methodological flaws. Consequently, an appropriate alternative method of data collection was devised, as well as a system for classifying the speech units according to their conversational content (Chapter 2). In the design of the classification scheme, particular emphasis has been placed on the theoretical foundation of speech segmentation and classification. Like any method of content analysis, the requirements of the new system were threefold: the system should be (1) objective (each step in the classification process being carried out on the basis of explicit rules, thereby minimising the analyst's subjective predispositions), (2) systematic (the inclusion or exclusion of categories should be carried out according to consistently applied rules), and (3) have generality (the categories should be theoretically relevant as well as being applicable to the anlysis of different types of relationships and situations).

Conversational Exchange Analysis (CEA), the method of segmentation and classification developed for this research project, comprises four sets of rules. The first is concerned with segmenting conversation into analysable units, the unit of speech used being defined as a 'single thought or idea'. The remaining three sets of rules are concerned with speech classification, coding the speech units along three conceptually distinct dimensions. The first, Activity, assesses how speech is made salient in the conversation. For example, is the information asked for or given, agreed with or disputed? The second, Type, codes the content of the speech. For example, the speech may be expressing a belief, telling a story, giving emotional support, or commenting on the conversation. The final level of analysis, Focus, scores the referent of the speech. For example, is the speaker expressing his/her own ideas, or the ideas of someone else? This last level of analysis has not been used in this research project, but was included in the CEA system in order to give a generality of application outside the bounds of this thesis.

In the Introduction, the interdependent nature of speech events in conversation was demonstrated from a theoretical standpoint. This assumption was then tested in Chapter 3 using four dyadic conversations, coded using the Activity dimension of CEA. The hypothesis of interdependence was supported, the analyses additionally indicating that there was a small predictive relationship between speech events in informal conversation.

A review of the literature had indicated that a number of authors had used such a finding of speech event dependence as a basis for describing the organisation of events in conversation. However, it was pointed out in Chapter 4 that this not only assumed a particular model of conversation that remained untested, but also ignored the possibility that a sample of conversations may not be homogeneous and may also have structures that change over time. The theme of Chapter 4 was, therefore, concerned with the development and testing of a stochastic model of informal conversation.

By virtue of its conceptual links with General systemstheory, and encompassing the notions of process, relation and structure, a Markov process was selected as a potential model of conversation for subsequent development and testing. Reviewing the literature on Markov modelling, three main types of error emerged.

Firstly, conversation was frequently assumed to be Markovian, even though some, or all, of the Markovity assumptions were not tested. Secondly, the methods used to test the assumptions were often inappropriate, if not invalid; finally, statistical requirements were often invalidated. To rectify these shortcomings, an eight-state Markov process was tested for suitability as a model of informal conversation, with particular emphasis being placed on the correct use of statistics appropriate to a General systems approach. The analyses demonstrated that such a model was indeed acceptable; in addition, they indicated that a first-order Markov process was sufficiently flexible to produce accurate predictions of future events, even when the Markovity assumptions of sample homogeneity and sequence stationarity were violated. A 14-state Markov model was subsequently developed and replicated.

During the testing of the four-state Markov model of conversation in Chapter 4, some conversations were found to be non-stationary, although this was later found not to affect the predictive ability of the final model of conversation. However, the aim of Chapter 5 was to assess whether the structure of non-stationary conversations differed from the stationary conversations on three process characteristics considered by Hawes and Foley (1976) to have psychological meaning: n-step contingencies, and mean and standard deviation inter-event distances.

The technique used in Chapter 5 was to generate by means of Monte-Carlo methods, a number of conversations with stationary parameters from the non-stationary first-order matrices, comparing the process characteristics derived from the stationary sequences with those from the non-stationary sequences. The results of these analyses suggested that in a large proportion of cases there were significant differences in the distribution of speech categories in the conversations. Typically, whereas speech events in stationary conversations were approximately equally distributed across the length of the conversation, non-stationary sequences were more likely to be characterised by clusters of Dissents and Requests occurring in the first half of the conversation. A similar trend was indicated for those non-stationary conversations whose distance measures showed no significant differences, although the clusters were considerably less marked, and less consistent.

A serendipitous finding, supported and refined by these results is the work of Calabrese (1975); he has suggested that questioning behaviour tends to occur mainly at the beginning of the interaction, decreasing as the interaction continues. The present analysis indicates that although this may often be the case, it is not a universal finding, tending to be more marked in some conversations than others. By way of extension, the analyses have also shown that disagreement tends to be unevenly distributed through the conversations.

On a more speculative note, it is suggested that the methods of analysis used here may have an important practical application in the analysis of a different aspect of conversational interaction, notably interaction style. Both Calabrese (1975) and the present analyses have suggested that questioning behaviour tends to be concentrated in one part of the conversation, and the present analyses have indicated that this also tends to be true for Dissents as well. Giles and St. Clair (1979, p.140) have noted that increasing the concentration of questions in a conversation changes the character of the interaction from one of informality to formality. Consequently, if the concentration of specific speech events gives rise to a conversation of a specific character, the method of comparing distances between speech events, as in the present study, is one that has a potential use in the analysis of conversational style.

Returning to the discussion of the model of conversation presented in Chapter 4, the aims of Chapter 6 were twofold: (1) identification of the role of speech content in conversational turn-taking, and (2) the extraction of conversation strategies. Firstly, the turn-taking analyses indicated that Requests and Reactions tended to precede changes in speaker, and Replies, Consents, and again, Reactions tended to follow changes in speaker. These data indicated that speech content had a specific, though limited, role in the floor-apportionment mechanism. Secondly, using the theory of directed graphs and a simplification technique known as 'condensation', a number of conversational strategies were devised. These are summarised below.

1) Given a Request by one speaker, it is generally expected that a change of speaker will occur, the new speaker tending to either Reply or Consent to the Request.
2) The consequence of Replying is to secure the speaking-turn, Replies tending to have a long-term effect, functioning by retaining the floor for the new speaker.
3) By contrast, Consents by the new speaker have a short-term effect on the floor-apportionment mechanism, generally returning the floor to the original speaker.
4) Reactions result in a switch to the original speaker in a large proportion of cases (74%), and therefore represent a method of ensuring that the other person continues to speak.
5) Dissents are the most effective method of obtaining the speaking-turn, securing the floor on 55% of occasions and represent a way of breaking into the conversation.
6) After Dissenting, an Offer of information by the same speaker retains the speaking-turn for the new speaker.

7) Having obtained the speaking turn, the most effective method of retaining the floor is to continue to Offer information.

8) If, for any reason, the floor is thrown open for negotiation, the speaker-turn may be secured by an Offer of information and relinquished by either a Request for information or a Consent.

9) Frequently, one may agree only in part with a point made in conversation. These are classified as Modifications and are the least stable of all the speech states. The consequence of a Modification is to return the floor to the original speaker, although the turn may be retained by the new speaker by quickly Offering new information.

Whereas Chapters 4 and 6 were concerned with the analysis of conversation coded for speech Activity, Chapter 7 was concerned with the construction of a model of conversation for speech coded according to the 'Type' of information exchanged. Owing to the paucity of data no statistical analyses could be performed, but an observational approach yielded a number of conversational routines and strategies. These are summarised below.

1) The main conversational routines, for speech coded according to the Type of information exchanged, comprise within-speaker Beliefs, Examples and Explanations, which are embellished

by within-speaker Narratives, and punctuated by between-speaker Listener Responses and Active Recognitions.

2) By giving away personal information about oneself (Personal Details), one predisposes the other speaker to reciprocate with information of a similar nature. This may be seen as confirmation of the suggestion by both Jourard (1971) and Argyle (1972) that if one person discloses personal information to another, the disclosure is reciprocated.

3) Comments about the conversation in the form of Metastatements operate in a similar manner to Personal Details; they pre-dispose the other person to reciprocate information and could form the basis of a social strategy. For example, rather than directly asking a person whether he/she is enjoying him/herself, a strategy that may lead to a socially acceptable but nevertheless untrue reply, a response may be obtained simply by commenting on the situation oneself.

   e.g. A: Hey, I'm enjoying this/ (Metastatement)
   B: Well, I find it rather tedious/ (Metastatement)

4) Directions function as a sort of crossroads in the conversation, where a number of options are open to the speakers. The usual course of events is for the speaker who gives the Direction to retain long-term control over the conversation.

e.g. A: Why don't we continue with this/
(Direction)
B: Fine/
(Active Recognition)
A: Well, it seems to me .../ (Belief)

However, Directions are also a good point at which long-term speaker-switches can occur.

e.g. A: We could do the abortion one/ (Direction)
B: Great/
(Active recognition)
I'm totally against this because .../
(Belief)

Sometimes, negotiation occurs for the floor:

A: Shall we go on with the mercy-killing?/
(Direction)
B: I'd prefer moving on to the Student's
Union one/ (Direction)
A: Okay/ (Active recognition)

and may be resolved in either of the two ways illustrated above.

5) Completions function as a subtle method of obtaining the speaker-turn; in 18% of cases they result in an immediate speaker-switch, and in 38% of cases they lead to an Active recognition by the original speaker, followed by a phase of Offering of information by the new speaker.

6) Given that a speaker-switch occurs, the loss of the floor tends to be short-term if associated with an Active recognition or a Listener response,

but more permanent if associated with any other speech state. Consequently, in order to retain the floor, the most effective strategy is to give one's Beliefs about a topic as soon as possible, an Active recognition or Listener response by itself resulting in the loss of the floor to the original speaker.

The analyses presented in Chapters 4, 6 and 7 have three main implications. First, they confirm the suggestion of Sacks, Schegloff and Jefferson (1974) that conversational floor-apportionment is a first-order process, whereby the identity of the next event in the conversation may be maximally predicted on the basis of just the previous event. Second, they re-affirm the suggestion of Thomas and Bull (1981) that speech content does have an important role to play in the management of conversation, and third, in terms of speech content, the social strategies extracted from the models of conversation enhance our knowledge of the process of interaction.

On a more speculative note, the strategies would appear to have a practical application in a social skills programme concerned with the teaching of conversational and general meshing skills. For example, from a theoretical standpoint, Argyle and Kendon's (1967) model of social interaction is essentially an application of General systems theory to social interaction, and constitutes a 'cybernetic paradigm for skill acquisition' (Ellis & Whittington, 1981, p.25). As this is the case, then the social strategies that have been extracted within a system framework are ideally suited for application in social skill training procedures.

From a practical point of view, whilst the analyses provide detailed information concerning the most frequently used speech states, as well as their more usual sequential ordering, the social strategies provide precise and detailed methods of prolonging the interaction, retaining the listening state, offering the speaking turn, and breaking into and retaining the speaking state. This type of information is essential if the socially unskilled are to be taught with any degree of success. For example, to be socially skilled implies the use of social rules, even though these rules may be used 'automatically', and without the knowledge of the user (Argyle, 1979). A common characteristic of socially unskilled conversation is that social rules appear to be continually broken or go unheeded (Trower et al., 1978). As a consequence, the communication becomes unpredictable and is viewed unfavourably by others (Kiesler, Kiesler & Pallack, 1967), which only serves to impair the communication still further. As Trower et al. (1978) have written, "Training ... is concerned with effective plans and strategies to achieve desired goals, and involves the reorganisation of behavioural events into a different sequence. In order to do this kind of analysis and retraining, we need to know how elements are combined in normal social interaction" (p.15). The analyses and social strategies presented in this research project would, therefore, appear to be well suited for application in a social skills training programme concerned with meshing and general conversation skills, from both theoretical and practical points of view.

A limitation of Markovian methods is their inability to detect units that are 'embedded' within larger units. A method still within the Markovian approach, chain analysis, was thus used to analyse the conversations for embedded, and multi-element speech units (Chapter 8). The main conclusion of this analysis indicated that the conversations appeared to be most usefully considered, at both the Activity and Type levels of analysis, in terms of two-event sequences, equivalent to those demonstrated by the Markov analyses presented in Chapters 4 and 7. In addition, the analyses demonstrated a marked lack of embedding, but, by way of extension of the Markov analyses, indicated that a numberoof two-element chains were cyclic, often forming multi-element chains up to six or eight speech events in length. Such results lend additional support to the notion that the conversations studied consist essentially of phases of Offering information interspersed by phases of Consents and Reactions, at the Activity level of analysis, and phases of giving Beliefs, Examples, Explanations and Narratives, punctuated by Active Recognitions and Listener Responses, at the Type level of analysis.

Whereas the preceding analyses were concerned with the structure of conversation in terms of which speech events were to be found in sequential arrangement, Chapter 9 was concerned with the organisation of conversational events in terms of system constraints. Analysing the organisation of conversation using Uncertainty statistics, the general finding was that there was a large degree of structuring in the conversations. To the extent that the presence of structure indicates the existence of constraints in the conversational system, it has been demonstrated that what person A says to B may be considered, at least to some degree, to constrain person B's response. In particular, at the Activity level of analysis, within-speaker Offers, Replies and Requests, and between-speaker Consents and Reactions have been shown to be 'structure-inducing', having the effect of constraining the type of speech to follow (for both within- and between-speaker transitions). At the Type level of analysis, Narratives were found to be the only structure-inducing speech-state.

It is important to note that the degree to which a system is constrained depends on the type of conversation analysed (e.g. Gouran & Baird, 1972) and the method of speech analysis used (e.g. Stech, 1975). Consequently, the estimates of conversational structure presented are strictly limited to the present research project; nevertheless, these results are considered to have two important implications. Firstly, the uncertainty analyses indicate that although there is a considerable degree of constraint operating in the conversations, subsequent events are nevertheless quite variable. Consequently, although the existence of constraints indicates the reciprocal nature of the conversational interactions, reciprocation must be considered as a probabilistic tendency and not as a foregone conclusion; there are always possibilities for alternative directions to be taken, for both within- and between-speaker transitions. A conversational rule can never therefore specify an absolute one-to-one relationship between speech events, but more appropriately should specify a range of 'probable' outcomes. Consequently, the rules and strategies that have been specified on the basis of the Markov analyses are rules in the sense that they may be construed as 'conventional practices' (Argyle, 1979, p.24) and by virtue of their probabalistic nature may be broken.

Secondly, on a speculative note, it is suggested that the technique of measuring the structure of conversation through system constraints may have a useful application in determining the character or style of conversation. For example, as discussed in Chapter 9 (Section 9.6), Gouran and Baird (1972) and Stech (1975) have noted considerable variation in the estimates of structure for a sample of conversations ranging from informal discussion to formal interviews. Measures of both distributional and sequential constraint may, therefore, prove to be useful in empirically determining the character or style of a conversation.

Having discussed the implications of the results of this work for future research, the methods of analysis used throughout this research project are also considered to have wider application. Although primarily intended for use in this research project, it is maintained that CEA (Chapter 2) represents a useful tool for the analysis of a wide range of conversations, including informal and formal discussion, general conversation one might hear at work, in a cafe, canteen, around the dinner table, etc. CEA may also be applied to speech occurring in more 'context-bound' situations, although it should be recognised that context-specific systems (e.g. negotiation - Morley & Stephenson, 1977, etc.) may be more appropriate for the analysis of speech occurring in such situations.

In addition to its use in the analysis of interaction sequences, CEA may also be used to relate the process of an interaction to its outcome. For example, given a situation in which different social strategies are used to maximise a person's gain (e.g. prisoner's dilemma, Jones & Gerard, 1967, pp.562-575; bargaining and negotiation, Morley & Stephenson, 1977), the process of the interaction may be described in terms of speech content by CEA and related to its outcome.

In the literature there is an indication that the content and style of a conversation may be related. For example, as discussed earlier, Giles and St. Clair

(1979) have shown that there is a clear relationship between the amount of questioning that occurs early in a conversation and the perceived formality of the interaction. Similarly, Raush et al. (1970) have observed that two groups of newly-wed couples could be distinguised on the basis of their differential use of referents: where pregnancy occurred early in marriage, the couples tended to use the referent 'I' in conversation, whereas matched couples, for whom early pregnancy did not occur, tended to use more 'we' referents. Raush et al. speculated that where pregnancy occurred early in marriage the couples were concerned more with establishing marital and familial roles rather than developing the interpersonal aspects of the relationship. Although their conclusion is speculative, Raush et al. have demonstrated that the style of an interaction may nevertheless be manifested in the content of conversation. By virture of categorising speech at a number of distinct conceptual levels and comprising expansive sets of speech descriptors, CEA would therefore appear to be well suited to the analysis of conversational style through speech content.

Whilst considering future research, two additional projects, although somewhat tangential to the main theme of the present research, are suggested. The first concerns sex differences in social interaction. There is a considerable literature indicating that the communicative behaviour of males and females differs in many respects. For example, women have been found to engage in more person-oriented gaze than men (Argyle, 1972), Look at the other more while speaking (Exline & Winters, 1965), use more intensifiers (e.g. very, too, etc.), 'empty adjectives' (e.g. divine, charming, etc.), polite forms of speech, gestures and hedges (Lakoff, 1973), and tend to favour more the use of standard speech (Labov, 1966). As the present research was based entirely on conversations between females, the analysis of conversations between males and between males and females therefore seem to be a potentially illuminating line of research that would enhance and refine the analyses presented in this thesis.

The second suggested project concerns using a multi-modal approach to the analysis of interaction. There are clearly many modes of communication besides that of speech content, such as gaze direction, posture and paralanguage, that need to be taken into account when building a model of social interaction. Although there has been some work in this area (e.g. Kendon, 1967, 1972, 1978; DeLong, 1974; Duncan, 1972, 1974; Duncan & Niederehe, 1974; Duncan & Fiske, 1977; Duncan, Brunner & Fiske, 1979; Wiemann & Knapp, 1975; Thomas & Bull, 1981), the methods of analysis have often been of a global nature and the role of speech content in interaction has frequently been ignored (cf. Thomas & Bull, 1981). Combining recording schemes for other communication modalities with the extensive system of speech categories presented in CEA would certainly yield a more comprehensive, multi-channel model of social interaction.

Finally, although the statistical methods of analysis used throughout this research project (e.g. Anderson-Goodman maximum likelihood statistic, $\log_2$ based information statistics and Uncertainty statistics) are not original, an additional contribution of this research project lies in the explanation and demonstration of the appropriate usage of these statistics, as well as the provision of a package of easily-used computer programs designed to carry out such analyses (Appendix G).

The results of this research project would thus appear to have considerable implications for future research, not only with respect to the application of the social strategies to social skills training, but also to the wider study of verbal interaction using the methods of conversational and statistical analyses specified. From the results of this project, it has been argued that when analysed within a General systems framework using appropriate statistical techniques, informal conversation can be seen to be adequately modelled by a first-order Markov process. The models of conversation derived from the sequential analyses, contribute to our understanding of the interaction process and it is now possible to provide a number of specific conversational strategies with potential application to a social skills training programme concerned with conversational and general meshing skills. Moreover, the specification of an appropriate theoretical framework within which to study communication, and the provision of a method of conversational analysis now make it possible to extend conversational research to the investigation of a wide range of conversations and situations, as well as relating the process of the interaction to its outcome. Exploring the implications of these results, it has been possible to proceed beyond the original aim of this research project, by indicating a number of proposals for future research. However, the main aims of this project were threefold: (1) to provide detailed information about the elements of informal conversation and their sequential arrangement, (2) the specification of a theoretical framework and (3) development of a methodology appropriate for conversational research. To the extent that this has been successful, and a substantial body of data has been provided specifying the verbal elements of informal conversation and their sequential arrangement, the essential purpose of this research has been fulfilled.

APPENDIX A

CONVERSATIONAL EXCHANGE ANALYSIS

INTER-OBSERVER RELIABILITY

The calculation of reliability figures using Cohen's (1960) kappa requires that two judges independently categorise a sample of the data, from which the degree and significance of agreement can then be determined. Kappa has three assumptions: (1) the coding of one unit is independent of the coding of any subsequent items, (2) the categories are nominal, independent, mutually exclusive, and exhaustive, and (3) the judges operate independently.

Kappa is directly interpretable as the proportion of joint judgements on which there is agreement, after chance agreement has been excluded, its upper limit of +1 indicating perfect inter-judge agreement. The coefficient is simply the proportion of chance-expected disagreements which do not occur, or alternatively, it is the proportion of agreement between observers after chance agreement has been removed from consideration.

Having formed the judgements into an agreement matrix form, where the rows represent the experimenter's judgements and the columns the judgements of an independent observer, the coefficient may be calculated by:

$$\kappa = \frac{p_o - p_c}{1 - p_c} \quad \text{or} \quad \kappa = \frac{f_o - f_c}{N - f_c}$$

where p_o/f_o = the proportion/frequency of units on which the judges agreed.

p_c/f_c = the proportion/frequency of units for which agreement is expected by chance.

N = the total number of judgements made.

The significance of the agreement is determined by evaluating the normal curve deviate (Cohen, 1960, p.44):

$$z = \frac{\kappa 1 - \kappa 2}{\sqrt{\sigma\kappa 1^2 + \sigma\kappa 2^2}}$$

using

$$z = \frac{\kappa}{\sigma\kappa o}$$

where κ = the calculated value of kappa

$$\kappa_o = \sqrt{\frac{p_c}{N(1-p_c)}} \quad \text{or} \quad \sqrt{\frac{f_c}{N(N-f_c)}}$$

where p_c/f_c = the proportion/frequency of units on which the judges agreed.

N = the total number of judgements made.

For each of the four sets of rules - Activity, Type, Subject focus, and Object focus, a measure of reliability has been calculated based on a sample of 2943 units, representing the independent coding of five, half-hour tape-recordings and transcripts.

Activity reliability

		Independent observers' judgements						
		O	Rp	C	D	M	Rt	Rq
	O	1528 (807.4)	5	7	2	4		
	Rp	3	145 (81)	5				
	C	2	438 (70)			2	8	
Experimenters' judgements	D	1			153 (8)	3		
	M	1		1	2	69 (2)		
	Rt			11	1		399 (57)	
	Rq	2					151 (8)	

Note: (O) Offer, (Rp) Reply, (C) Consent, (D) Dissent, (M) Modify, (Rt) Reaction, (Rq) Request.

Values in parenthesis are frequencies of agreement expected by chance.

$f_o = 2883$ $f_c = 1033.4$

$\kappa = .969$

$z = 78.15$ $p < .001$

Type reliability

Independent observers' judgements

Experimenters' judgements	B	P	P.exp	N	S.F	A	J	Syn	Sum	Jdg	Av	Neg	A.R	Sal	Con	Ex	E.S	Ap	D	M	Rev	Res	Com	L.R	U.I	L
B	909 (297.9)					10							1		2	9										
P		24 (.2)																								
P.exp	1		42 (.6)													1										
N				77 (2.1)											1											
S.F.					7 (.02)										1	4										
A	1					14 (.076)																				
J	16						291 (32.4)																			
Syn								0 (0)																		
Sum									3 (.003)																	
Jdg										22 (.16)																
Av						1					13 (.06)															
Neg												4 (.005)														
A.R.	6												434 (70.19)											16		
Sal													1	3 (.004)												
Con	2														70 (1.8)											
Ex	5			2		4										375 (51.15)										
E.S.																	5 (.008)									
Ap																		8 (.02)								
D																			46 (.72)							
M	1												1							61 (1.3)						
Rev	1																				26 (.24)					
Res																						14 (.07)				
Com																							2 (.001)			
L.R.													16											386 (54.9)		
U.I.																									2 (.001)	
L																										1 (.0003)

Note: (B) Beliefs, (P) Personal details, (P.exp) Personal experience, (N) Narrative, (S.F) Subjective feelings, (A) Attitude, (J) Justification, (Syn) Synthesis, (Sum) Summary, (Jdg) Judgement, (Av) Avoidance, (Neg) Negation/Submission, (A.R) Active recognition, (Sal) Salience, (Con) Consequence, (Ex) Example, (E.S) Emotional support), (Ap) Apology, (D) Direction, (M) Metastatement, (Rev) Revision, (Res) Restatement, (Com) Completion, (L.R) Listener Response,(U.I) Unsuccessful interruptions, (L) Laughter.
Values in parenthesis are frequencies of agreement expected by chance.
$f_o = 2839$ $f_c = 513.93$ $\kappa = .957$ $z = 232.1$ $p < .001$

Subject focus reliability

Experimenters' judgements	Independent observers' judgements:	N	S	P	B	O	H
	N	1854 (1174)		3		5	4
	S		596 (121)				
	P			28 (.3)			
	B				53 (.95)		
	O					92 (2.9)	
	H	6					302 (49.1)

Note: (N) No focus, (S) Self, (P) Partner, (B) Both, (O) Other, (H) Hypothetical

Values in parenthesis are frequencies of agreement expected by chance.

fo = 2933 fc = 1331.15

$\kappa = .994$

$z = 59.16$ $p < .001$

Object focus reliability

		Independent observers' judgements					
		N	S	P	B	O	H
	N	2234 (1705)					1
	S		118 (4.7)				
Experimenters' judgements	P	4		59 (1.2)			
	B				15 (.08)		
	O					124 (5.7)	7
	H	7				3	371 (49.1)

Note: (N) No focus, (S) Self, (P) Partner, (B) Both, (O) Other, (H) Hypothetical.

Values in parenthesis are frequencies of agreement expected by chance.

$fo = 2921$ $fc = 1768.78$

$\kappa = .981$

$z = 43.41$ $p < .001$

APPENDIX B

CONTROVERSIAL TOPICS QUESTIONNAIRE

1. Mercy-killing should be legalised.
2. Examinations stimulate students to greater effort by providing the opportunity to compete for rewards, and should be maintained as an integral part of the education system.
3. In view of the current housing shortage, a landlord should rent to the first eligible applicant, regardless of race, creed or colour.
4. Capital punishment is the only deterrent to acts of violence.
5. Most illnesses are psychosomatic; people exaggerate minor problems in order to gain attention or pity.
6. In realistic terms, the campaign by women's organisations for the right to work has resulted in the exchange of a woman's potentially free time for a man's world of industrial and business drudgery.
7. A person who witnesses, but makes no attempt to prevent the suicide of another, is effectively a murderer.
8. Scenes of sex and violence depicted in films and on television have a corrupting influence on society, and should be censored.
9. Students lead a privileged and sheltered existence, and their political views are consequently naive and irrelevant.

10. Babies born with deformities are likely to become a burden to society and they should not therefore be allowed to survive.

11. Man can only sin if God exists.

12. Pay beds are an essential part of a National Health Service. If people can afford private medical service, they ought to be able to get it.

13. The University Student Union should provide an entertainment service only, and should not involve itself in political affairs.

14. Terrorist acts cause a senseless loss of life, challenge law and order and waste public resources. There can be no justification for terrorism.

15. Only the rich and idle can afford the luxury of psychological problems.

16. Racial minorities are losing their separate identities. What is needed is not integration but a sense of pride in belonging to a minority.

17. Examinations enable the teacher or examiner to assess the student's capacity and progress, and to discard those unlikely to benefit from the education they receive.

18. Fairy tales and comics are frequently cruel and frightening, and children should not be exposed to them.

19. There is no use for the arts in today's world, and education should be concentrated on science instead.

20. Most experiments on animals are carried out not from a concern for the relief of human suffering but for the sake of commercial profit and the advancement of career.

APPENDIX C

A NOTE ON THE ANDERSON-GOODMAN MAXIMUM LIKELIHOOD RATIO STATISTIC

A Markov model states that the probability of state j following state i, is c, and this is constant for every paired instance of states i and j. The model however, does not say what value c will take. c is known as a parameter of the model, the problem of finding this value usually being known as 'estimation'. Restle (1971) has indicated that when estimating parameters, it is preferable to use the data one has in the fullest possible way, as methods of estimation that use all the data are known to improve the chances of getting good agreement between the model and the data. At this point, one might assume that this reduces the possibility of rejecting the model. However, if insufficient methods are used to estimate parameters and a bad fit is obtained, then it is not possible to tell whether the fault lies in the model or in the estimated parameter values. Conversely, it is possible for ineffficient methods to give parameter values that appear to fit the model better than if all the data were taken into account.

A class of statistics called 'sufficient statistics' exists which use every aspect of the data that is relevant for estimating the value of a parameter, and it is a theorem of statistics that maximum likelihood estimates are functions of sufficient statistics (Restle, 1971). The Anderson-Goodman test is such a statistic, and by definition, is therefore a sufficient statistic. The benefit, therefore, of using the Anderson-Goodman maximum likelihood ratio as a method of analysis is that the estimates to which it leads are based on sufficient statistics, while any other method may, or may not, use all the data in an optimum way. As the parameter estimates are optimally derived, then a good or bad fit of the model to the data may be seen as a comment on the model, and not due to the methods used to estimate the model parameters. Consequently, the statistical methods used in this research (the Anderson-Goodman test) have the benefit of being sufficient statistics, the model of conversation developed having optimally derived parameter estimates.

APPENDIX D

INFORMATION ANALYSES

N-gram calculations (Four-state data)

Overall transitions

Dyad	H max	Amount of information in an n-gram $\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram	$\hat{H}$ four-gram
1	2.0	1.364	2.626	3.782	4.794
2	2.0	1.384	2.732	4.035	5.218
3	2.0	1.238	2.444	3.616	4.656
4	2.0	1.378	2.641	3.877	5.073
5	2.0	1.197	2.328	3.429	4.480
6	2.0	1.313	2.512	3.683	4.801
7	2.0	1.301	2.581	3.761	4.833
8	2.0	1.410	2.722	3.981	5.150
9	2.0	1.593	3.047	4.431	5.712
10	2.0	1.442	2.793	4.078	5.238
11	2.0	1.417	2.714	3.941	5.066
12	2.0	1.587	3.031	4.407	5.664
13	2.0	1.492	2.854	4.166	5.392
14	2.0	1.427	2.683	3.914	5.093
15	2.0	1.585	2.927	4.191	5.389
16	2.0	1.562	2.982	4.349	5.620
17	2.0	1.170	2.286	3.387	4.406
18	2.0	1.411	2.730	3.951	5.049
19	2.0	1.336	2.588	3.794	4.948
20	2.0	1.314	2.563	3.752	4.865
21	2.0	1.421	2.735	4.011	5.208
22	2.0	1.376	2.746	4.000	5.167
23	2.0	1.409	2.745	4.020	5.140
24	2.0	1.588	3.012	4.324	5.529

N-gram calculations (Four-state data)

Within-speaker transitions

Dyad	H max	Amount of information in an n-gram $\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram	$\hat{H}$ four-gram
1	2.0	1.247	1.981	2.409	2.655
2	2.0	1.363	2.200	2.515	2.820
3	2.0	1.313	4.715	5.170	5.516
4	2.0	1.341	2.203	3.008	3.574
5	2.0	1.023	1.429	1.787	1.896
6	2.0	1.261	1.932	2.338	2.577
7	2.0	1.244	1.921	2.356	2.645
8	2.0	1.106	1.593	1.718	1.810
9	2.0	1.409	2.159	2.816	2.903
10	2.0	1.363	1.971	2.567	2.881
11	2.0	1.189	1.906	2.300	2.337
12	2.0	1.550	2.303	2.954	3.249
13	2.0	1.482	2.213	2.758	3.121
14	2.0	1.303	2.012	2.441	2.474
15	2.0	1.557	2.263	2.792	3.321
16	2.0	1.348	1.905	2.045	2.146
17	2.0	0.892	1.479	1.907	2.276
18	2.0	1.161	1.604	1.952	2.300
19	2.0	1.181	1.645	1.940	2.132
20	2.0	1.396	2.043	2.390	2.737
21	2.0	1.422	2.237	3.010	3.481
22	2.0	1.321	2.079	2.506	2.706
23	2.0	1.123	1.995	2.258	2.521
24	2.0	1.531	2.395	3.166	3.425

N-gram calculations (Four-state data)

Between-speaker transitions

Dyad		Amount of information in an n-gram			
	H max	$\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram	$\hat{H}$ four-gram
1	2.0	1.510	2.726	3.078	3.271
2	2.0	1.374	2.616	3.607	3.910
3	2.0	1.173	2.177	2.829	2.924
4	2.0	1.368	2.470	3.249	3.694
5	2.0	1.325	2.610	3.224	3.792
6	2.0	1.319	2.472	2.999	2.999
7	2.0	1.339	2.678	3.167	3.332
8	2.0	1.576	2.999	3.422	3.497
9	2.0	1.695	3.222	3.699	3.852
10	2.0	1.535	3.040	3.423	3.648
11	2.0	1.581	2.920	3.465	3.657
12	2.0	1.550	2.927	3.631	3.760
13	2.0	1.445	2.634	3.353	3.453
14	2.0	1.459	2.525	3.157	3.562
15	2.0	1.554	2.736	3.540	3.889
16	2.0	1.681	3.199	3.645	3.698
17	2.0	1.394	2.669	2.997	3.057
18	2.0	1.576	3.133	3.367	3.483
19	2.0	1.401	2.711	3.055	3.255
20	2.0	1.258	2.495	3.378	3.546
21	2.0	1.412	2.668	3.434	3.877
22	2.0	1.399	2.775	3.418	3.609
23	2.0	1.702	3.143	3.943	4.240
24	2.0	1.603	3.161	3.796	4.195

Order estimate calculations (Four-state data)

Overall transitions

Dyad	Amount of information in an N^{th} order estimate				
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$	$\hat{H}_3$
1	2.0	1.364	1.262	1.156	1.012
2	2.0	1.384	1.348	1.303	1.183
3	2.0	1.238	1.206	1.172	1.040
4	2.0	1.378	1.263	1.236	1.196
5	2.0	1.197	1.131	1.101	1.051
6	2.0	1.313	1.199	1.171	1.118
7	2.0	1.301	1.280	1.180	1.072
8	2.0	1.410	1.322	1.259	1.169
9	2.0	1.593	1.454	1.384	1.281
10	2.0	1.442	1.351	1.285	1.160
11	2.0	1.417	1.297	1.227	1.125
12	2.0	1.587	1.444	1.376	1.257
13	2.0	1.492	1.362	1.312	1.226
14	2.0	1.427	1.256	1.231	1.179
15	2.0	1.585	1.342	1.264	1.198
16	2.0	1.562	1.420	1.367	1.271
17	2.0	1.170	1.116	1.101	1.019
18	2.0	1.411	1.319	1.221	1.098
19	2.0	1.336	1.252	1.206	1.154
20	2.0	1.314	1.249	1.189	1.113
21	2.0	1.421	1.314	1.276	1.197
22	2.0	1.376	1.370	1.254	1.167
23	2.0	1.409	1.336	1.275	1.120
24	2.0	1.588	1.424	1.312	1.205

Order estimate calculations (Four-state data)

Within-speaker transitions

Dyad	Amount of information in an N^{th} order estimate				
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$	$\hat{H}_3$
1	2.0	1.247	0.734	0.428	0.246
2	2.0	1.363	0.837	0.315	0.305
3	2.0	1.313	0.582	0.455	0.346
4	2.0	1.341	0.862	0.805	0.566
5	2.0	1.023	0.406	0.358	0.109
6	2.0	1.261	0.671	0.406	0.239
7	2.0	1.244	0.677	0.435	0.289
8	2.0	1.106	0.487	0.125	0.092
9	2.0	1.409	0.750	0.657	0.087
10	2.0	1.363	0.608	0.596	0.314
11	2.0	1.189	0.717	0.394	0.037
12	2.0	1.550	0.753	0.651	0.295
13	2.0	1.482	0.731	0.545	0.363
14	2.0	1.303	0.709	0.429	0.033
15	2.0	1.557	0.706	0.529	0.529
16	2.0	1.348	0.557	0.140	0.101
17	2.0	0.892	0.587	0.428	0.369
18	2.0	1.161	0.443	0.348	0.348
19	2.0	1.181	0.464	0.295	0.192
20	2.0	1.396	0.647	0.347	0.347
21	2.0	1.422	0.815	0.773	0.471
22	2.0	1.321	0.758	0.427	0.200
23	2.0	1.123	0.872	0.263	0.263
24	2.0	1.531	0.864	0.771	0.259

Order estimate calculations (Four-state data)

Between-speaker transitions

Dyad	Amount of information in an N^{th} order estimate				
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$	$\hat{H}_3$
1	2.0	1.510	1.216	0.352	0.193
2	2.0	1.374	1.242	0.991	0.303
3	2.0	1.173	1.004	0.652	0.095
4	2.0	1.368	1.102	0.779	0.445
5	2.0	1.325	1.285	0.614	0.568
6	2.0	1.319	1.153	0.527	0.000
7	2.0	1.339	1.339	0.489	0.165
8	2.0	1.576	1.423	0.423	0.075
9	2.0	1.695	1.527	0.477	0.153
10	2.0	1.535	1.505	0.383	0.225
11	2.0	1.581	1.339	0.545	0.192
12	2.0	1.550	1.377	0.704	0.129
13	2.0	1.445	1.189	0.719	0.100
14	2.0	1.459	1.066	0.632	0.405
15	2.0	1.554	1.182	0.804	0.349
16	2.0	1.681	1.518	0.446	0.053
17	2.0	1.394	1.275	0.328	0.060
18	2.0	1.576	1.557	0.234	0.116
19	2.0	1.401	1.310	0.344	0.200
20	2.0	1.258	1.237	0.883	0.168
21	2.0	1.412	1.256	0.766	0.443
22	2.0	1.399	1.376	0.643	0.191
23	2.0	1.702	1.441	0.800	0.297
24	2.0	1.603	1.558	0.635	0.399

N-gram calculations (Four-state data)

Stationary and Non-stationary groups

Stationary sequences

Transition	Amount of information in an n-gram				
	H max	$\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram	$\hat{H}$ four-gram
Overall	2.0	1.367	2.641	3.867	5.010
Within	2.0	1.288	1.926	2.349	2.633
Between	2.0	1.406	2.668	3.297	3.545

Non-stationary sequences

Transition	Amount of information in an n-gram				
	H max	$\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram	$\hat{H}$ four-gram
Overall	2.0	1.418	2.760	4.030	5.189
Within	2.0	1.303	2.015	2.502	2.751
Between	2.0	1.530	2.886	3.435	3.629

Order estimate calculations (Four-state data)

Stationary and Non-stationary groups

Stationary sequences

Transition	Amount of information in an N^{th} order estimate				
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$	$\hat{H}_3$
Overall	2.0	1.367	1.274	1.226	1.143
Within	2.0	1.288	0.638	0.423	0.284
Between	2.0	1.406	1.262	0.629	0.248

Non-stationary sequences

Transition	Amount of information in an N^{th} order estimate				
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$	$\hat{H}_3$
Overall	2.0	1.418	1.342	1.270	1.159
Within	2.0	1.303	0.712	0.487	0.249
Between	2.0	1.530	1.356	0.549	0.194

Stereotypy calculations (Four-state data)

Stationary and Non-stationary groups

Stationary sequences

Transition	Order of sequential constraint 0	1	2	3
Overall	0.317	0.363	0.387	0.429
Within	0.356	0.681	0.789	0.858
Between	0.297	0.369	0.689	0.876

Non-stationary sequences

Transition	Order of sequential constraint 0	1	2	3
Overall	0.291	0.329	0.365	0.421
Within	0.349	0.644	0.757	0.876
Between	0.235	0.322	0.726	0.903

N-gram calculations (Seven-state data)

Dyad groups 1 - 12 and 13 - 24

Dyad group 1 - 12

Transition	Amount of information in an n-gram			
	H max	$\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram
Overall	2.807	2.037	3.804	5.351
Within	2.807	1.867	2.944	3.301
Between	2.807	2.087	3.774	4.461

Dyad group 13 - 24

Transition	Amount of information in an n-gram			
	H max	$\hat{H}$ probability	$\hat{H}$ digram	$\hat{H}$ trigram
Overall	2.807	2.123	3.913	5.456
Within	2.807	1.943	3.000	3.320
Between	2.807	2.190	3.886	4.552

Order estimate calculations (Seven-state data)

Dyad groups 1 - 12 and 13 - 24

Dyad group 1 - 12

Transition	Amount of information in an N^{th} order estimate			
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$
Overall	2.807	2.037	1.767	1.547
Within	2.807	1.867	1.077	0.357
Between	2.807	2.087	1.687	0.687

Dyad group 13 - 24

Transition	Amount of information in an N^{th} order estimate			
	H max	$\hat{H}_o$	$\hat{H}_1$	$\hat{H}_2$
Overall	2.807	2.123	1.790	1.543
Within	2.807	1.943	1.057	0.320
Between	2.807	2.190	1.696	0.666

Stereotypy calculations (Seven-state data)

Dyad groups 1 - 12 and 13 - 24

Dyad group 1 - 12

Transition	Order of sequential constraint 0	1	2
Overall	0.274	0.369	0.449
Within	0.335	0.616	0.873
Between	0.257	0.399	0.755

Dyad group 13 - 24

Transition	Order of sequential constraint 0	1	2
Overall	0.244	0.362	0.450
Within	0.308	0.623	0.886
Between	0.220	0.396	0.762

APPENDIX E

MEAN INTER-EVENT DISTANCE

PROCESS CHARACTERISTICS

Dyad 1

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	2.1	15.7	26.6	29.9	6.4	13.5	28.5	43.0
		2.2	13.2	36.8	33.7	6.8	8.1	25.3	40.6
	C	3.1	9.5	20.4	29.2	9.5	19.0	27.7	45.2
		3.3	9.8	30.9	30.8	6.2	9.0	26.0	47.4
	D	4.0	20.6	17.6	43.5	9.4	18.5	12.3	49.4
		3.3	12.5	31.3	33.2	6.8	8.6	18.9	47.1
	Rq	5.0	15.3	22.2	22.6	2.2	8.7	28.1	24.0
		4.0	11.7	36.9	28.9	3.8	11.8	27.1	46.3
Speaker 2	O/Rp	3.8	7.0	27.8	29.8	4.4	17.2	36.6	39.6
		4.0	7.6	33.1	33.3	4.7	10.4	27.3	41.1
	C	2.6	12.3	21.5	31.0	3.5	10.0	21.1	31.8
		2.6	12.1	35.0	34.5	6.3	10.0	22.7	40.9
	D	3.8	35.5	9.6	21.5	9.6	24.7	11.9	36.5
		3.7	11.2	33.0	31.3	6.1	10.9	21.5	42.7
	Rq	1.3	12.3	23.4	21.8	5.5	10.8	44.2	48.6
		2.2	12.7	35.8	32.5	5.5	7.9	26.0	44.6

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 84.16 $p < .01$

Dyad 4

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	4.9	6.0	38.5	11.5	3.6	12.4	40.1	36.2
		5.1	6.4	138.1	11.7	3.3	10.8	142.8	51.2
	C	4.9	5.1	34.3	12.0	2.5	14.8	34.4	30.6
		6.2	5.1	134.3	10.6	2.3	14.6	141.2	54.3
	D	1.6	4.6	79.0	14.6	4.0	17.0	56.5	30.6
		3.6	7.4	124.2	11.8	3.2	11.7	120.0	45.6
	Rq	6.8	5.6	41.4	9.5	2.0	14.4	49.5	31.1
		6.7	5.4	133.4	10.2	1.5	15.1	144.5	54.7
Speaker 2	O/Rp	6.4	4.2	35.7	8.4	2.5	14.9	37.3	32.9
		6.8	4.5	133.8	8.9	2.5	15.4	142.2	54.9
	C	3.9	5.7	28.5	11.4	3.4	13.1	42.8	31.6
		4.3	6.3	139.8	11.5	3.2	13.1	139.9	154.8
	D	8.0	2.5	34.5	24.0	2.0	10.5	32.3	32.5
		8.1	3.6	85.3	10.9	3.1	14.2	48.9	43.2
	Rq	2.7	7.6	54.3	7.8	3.9	19.5	61.0	45.5
		2.8	7.2	129.5	12.5	3.9	11.7	146.4	67.6

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 193.34 $p < .001$

Dyad 7

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	3.0	16.8	24.2	18.4	4.5	10.5	46.4	25.4
		3.0	13.4	32.8	24.8	4.5	10.9	56.3	21.1
	C	3.2	11.4	23.8	17.6	3.5	12.7	57.6	24.8
		3.8	11.4	30.6	24.8	3.9	12.2	56.9	23.4
	D	4.1	17.1	28.6	20.4	3.8	13.2	49.9	22.3
		3.4	13.7	30.2	23.5	4.1	10.6	59.2	21.4
	Rq	4.9	13.8	19.6	22.1	2.4	14.2	45.5	24.6
		5.2	12.2	29.7	26.1	2.4	12.4	56.7	24.9
Speaker 2	O/Rp	5.6	15.6	24.0	16.4	2.7	15.1	51.4	30.7
		4.8	11.0	31.4	22.4	2.7	12.5	56.7	25.1
	C	3.5	15.4	28.6	16.9	3.4	11.5	51.6	32.7
		4.1	12.4	31.2	24.2	3.6	11.4	61.5	24.1
	D	7.2	9.1	14.6	10.6	2.7	12.8	71.5	51.8
		5.9	11.6	33.2	22.2	2.4	11.0	69.8	26.1
	Rq	3.8	12.2	18.2	16.7	4.0	16.3	53.4	19.7
		3.9	13.2	33.3	22.0	4.1	12.1	56.9	23.9

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 41.35 n.s.

Dyad 8

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	5.3	16.0	50.0	14.5	5.0	11.8	46.6	24.9
		5.6	14.2	34.1	12.2	4.1	16.9	30.5	20.2
	C	8.6	13.1	70.0	13.5	2.2	19.1	68.6	23.9
		9.2	13.1	32.9	10.9	2.2	21.4	32.2	25.7
	D	7.2	15.2	31.8	25.4	5.1	11.2	18.8	20.5
		7.8	13.6	32.9	11.9	4.1	20.1	22.9	24.3
	Rq	8.9	11.6	45.3	10.4	1.5	16.8	48.5	23.3
		10.2	13.2	32.3	10.6	1.5	22.9	32.3	28.6
Speaker 2	O/Rp	9.5	11.9	60.6	11.1	2.0	20.6	55.2	25.3
		9.8	12.6	32.3	8.9	2.0	22.4	32.9	28.3
	C	7.7	12.9	42.9	13.9	2.8	16.2	52.4	24.4
		8.9	13.5	37.2	9.9	2.7	17.9	34.0	28.9
	D	13.6	11.1	29.1	15.9	2.3	15.3	28.3	21.1
		10.8	10.8	35.2	10.8	2.4	24.1	27.0	28.6
	Rq	4.6	21.4	52.4	15.0	5.3	21.5	23.8	22.8
		5.1	14.1	30.6	11.7	4.3	19.1	30.7	28.4

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 85.77 $p < .01$

Dyad 9

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	2.9	20.4	26.2	31.5	8.3	15.6	29.4	6.8
		2.9	19.9	19.9	51.6	7.5	12.9	22.7	7.9
	C	7.2	15.2	22.1	33.5	4.3	11.2	24.5	14.3
		4.7	15.8	17.3	46.8	5.4	12.4	21.1	10.2
	D	4.1	28.5	12.3	46.4	7.7	16.8	31.7	5.7
		4.8	19.1	13.2	54.3	6.4	13.5	22.9	8.8
	Rq	7.8	28.5	15.2	43.6	2.1	15.0	28.1	14.1
		5.3	15.6	19.1	46.1	2.9	14.8	24.1	10.8
Speaker 2	O/Rp	7.2	19.2	23.1	36.9	3.8	15.0	29.4	13.5
		5.9	14.9	16.1	45.3	3.9	14.4	23.2	11.3
	C	4.7	23.9	26.6	28.2	6.9	12.7	27.6	9.8
		3.6	18.2	18.9	52.6	6.3	13.7	22.7	9.8
	D	6.2	32.2	9.2	46.2	10.1	25.6	15.3	9.2
		6.5	18.3	14.4	48.3	5.4	15.3	14.8	12.3
	Rq	2.1	18.9	23.4	33.1	8.4	18.7	28.5	10.1
		2.4	19.8	19.4	53.5	7.3	14.2	23.1	9.9

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 49.19 n.s.

Dyad 11

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	2.4	11.9	60.1	38.9	9.7	23.9	49.9	7.1
		2.4	12.7	37.5	32.1	7.4	23.3	51.9	8.1
	C	3.0	9.7	70.9	36.3	6.5	16.7	67.2	7.2
		4.0	9.5	37.6	33.1	5.1	19.7	52.1	10.8
	D	3.5	18.5	16.5	56.9	20.3	26.0	23.1	7.9
		2.7	12.4	31.5	31.7	7.4	24.5	50.8	6.6
	Rq	5.1	4.9	107.8	32.0	3.0	20.8	103.7	9.5
		4.7	10.3	138.6	30.4	2.9	21.3	58.5	12.1
Speaker 2	O/Rp	4.8	8.4	73.7	34.9	4.2	21.2	69.9	9.1
		5.2	7.6	39.5	27.1	4.0	21.5	57.5	11.7
	C	3.2	12.4	59.5	38.4	6.3	21.4	60.2	10.6
		3.7	10.7	36.9	31.4	5.0	21.6	58.3	10.7
	D	6.2	5.5	16.5	40.0	16.1	26.9	23.6	6.4
		5.6	9.7	27.9	26.9	5.1	22.2	44.4	11.1
	Rq	2.2	12.5	58.2	37.3	9.2	22.8	52.4	9.5
		2.3	12.7	37.9	31.7	6.9	23.4	51.3	9.9

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 99.96 $p < .001$

Dyad 12

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	2.8	15.1	52.1	13.5	7.2	13.7	55.1	9.9
		2.8	15.1	53.2	14.1	7.4	10.1	58.9	8.1
	C	4.4	11.8	53.6	13.7	7.2	10.6	61.6	10.3
		3.6	11.8	55.3	13.1	6.0	12.4	63.1	9.4
	D	7.6	16.9	26.1	11.5	3.8	11.3	58.9	6.9
		4.9	16.5	31.1	13.5	6.9	10.8	59.1	8.8
	Rq	6.1	9.1	38.0	11.4	3.9	16.8	34.6	10.7
		4.5	12.9	54.3	11.8	4.0	13.3	56.9	9.8
Speaker 2	O/Rp	6.6	10.2	42.0	8.0	4.6	14.6	37.8	13.6
		4.9	11.7	54.4	7.9	4.5	13.7	58.5	10.8
	C	2.8	11.9	56.9	15.2	8.9	9.3	64.3	10.6
		3.1	15.1	52.7	14.3	7.1	9.6	64.6	9.5
	D	10.5	10.5	19.7	5.4	3.8	18.6	40.6	10.7
		5.6	9.9	51.4	8.7	4.5	14.9	44.3	14.3
	Rq	2.8	14.2	60.4	12.4	6.8	15.6	55.1	9.3
		2.5	15.2	50.5	13.9	7.2	11.4	60.9	10.4

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 31.09 n.s.

Dyad 13

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	4.1	8.0	64.9	15.4	6.3	10.7	12.3	16.4
		4.1	8.8	79.1	13.8	6.2	10.7	75.0	15.0
	C	5.6	6.1	57.7	13.8	3.8	13.0	12.0	33.4
		6.9	6.3	73.7	11.3	3.6	13.2	77.4	18.2
	D	1.2	13.4	64.5	12.0	9.0	9.4	10.0	17.8
		1.1	9.6	73.8	14.8	6.8	11.4	69.6	61.8
	Rq	7.7	8.6	50.3	10.7	2.4	14.3	21.0	18.8
		7.6	6.7	73.0	10.6	2.5	14.7	73.6	36.2
Speaker 2	O/Rp	8.3	5.9	48.0	9.0	3.1	16.0	10.3	19.5
		8.1	5.4	73.5	8.2	3.1	15.5	74.7	36.5
	C	8.0	9.2	47.8	14.0	5.3	9.4	20.0	18.7
		6.2	7.9	79.0	12.4	4.9	9.8	79.9	26.5
	D	13.0	3.0	31.0	6.0	5.0	10.0	1.0	18.0
		6.6	2.1	32.4	8.8	4.1	9.9	7.6	13.2
	Rq	3.5	6.4	72.7	13.8	5.5	10.2	10.0	17.6
		4.0	8.1	82.0	14.5	6.1	11.1	78.2	17.3

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 414.29 $p < .001$

Dyad 18

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	2.9	10.2	57.1	17.5	5.9	15.7	32.5	19.6
		2.9	12.3	59.9	18.0	5.8	16.1	48.2	17.9
	C	3.7	8.1	53.6	15.1	4.9	12.3	41.0	14.2
		4.7	8.2	57.6	17.2	4.5	18.6	52.4	18.1
	D	4.0	11.9	50.3	13.8	4.9	3.8	54.8	17.2
		5.3	13.3	52.5	19.1	6.3	5.6	66.5	21.9
	Rq	6.0	8.9	45.5	16.9	4.7	16.2	31.7	16.1
		6.2	10.5	58.9	16.7	2.6	19.7	44.9	20.9
Speaker 2	O/Rp	5.8	7.0	56.3	13.9	3.2	19.2	38.1	16.3
		5.8	8.9	56.7	15.3	3.2	19.5	51.4	20.2
	C	5.5	9.6	59.9	15.7	3.2	13.4	56.7	17.0
		4.8	11.6	67.0	16.8	5.5	14.0	51.5	20.1
	D	4.8	5.8	40.3	11.3	8.7	21.1	17.7	20.1
		7.1	8.3	62.1	10.0	4.1	21.6	38.3	20.8
	Rq	2.5	8.2	59.9	14.9	4.9	16.1	52.3	19.8
		2.8	11.1	73.0	18.9	6.0	15.3	52.3	20.1

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 23.33 n.s.

Dyad 23

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	2.1	21.9	27.5	45.9	12.7	9.1	26.4	34.4
		2.1	12.0	30.7	57.7	9.8	10.3	25.2	26.6
	C	4.3	9.3	33.2	43.3	5.0	12.9	33.3	28.7
		3.6	9.5	25.1	58.2	6.3	11.1	24.1	29.0
	D	5.4	17.0	15.2	37.5	5.4	10.3	32.9	26.5
		3.9	11.9	20.5	62.1	6.6	11.5	21.9	42.6
	Rq	5.9	8.1	24.3	49.3	3.6	16.0	17.8	30.3
		4.1	7.6	24.0	59.3	2.9	10.5	29.9	28.6
Speaker 2	O/Rp	4.6	9.3	34.0	42.7	5.5	13.1	34.8	39.1
		4.1	9.5	25.5	52.7	5.9	11.1	21.4	28.7
	C	3.8	15.8	36.8	42.5	6.2	8.8	35.1	32.9
		3.5	10.2	24.8	59.3	7.9	9.2	23.8	26.6
	D	3.5	21.6	13.1	33.7	12.6	12.2	20.8	27.8
		3.6	10.1	21.3	66.9	9.1	9.8	21.8	66.3
	Rq	2.6	15.3	25.3	55.2	7.9	8.6	25.7	27.8
		2.1	12.8	25.7	62.5	10.1	10.7	35.6	28.9

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 81.03 $p < .05$

Dyad 24

Observed/Generated mean distance

		Speaker 1				Speaker 2			
		O/Rp	C	D	Rq	O/Rp	C	D	Rq
Speaker 1	O/Rp	3.7	7.8	21.9	48.5	5.0	12.8	28.3	26.4
		3.7	10.6	15.2	45.3	5.3	14.2	25.0	20.1
	C	4.2	8.4	17.2	56.6	4.2	13.5	34.1	26.3
		5.1	8.4	14.3	47.4	3.7	16.4	24.8	22.8
	D	3.8	15.4	9.0	46.7	6.8	13.2	14.4	26.9
		4.9	11.3	10.2	50.9	5.8	16.6	20.9	22.1
	Rq	8.4	6.3	15.4	46.7	1.4	17.8	26.2	26.0
		6.3	7.6	18.6	52.6	2.1	17.6	40.4	25.9
Speaker 2	O/Rp	7.2	8.8	15.5	43.9	3.1	15.8	31.1	32.7
		5.8	7.5	13.7	42.1	3.3	17.1	25.1	23.7
	C	5.8	9.3	17.8	47.0	4.7	12.9	23.8	22.5
		4.5	10.3	14.9	42.5	4.6	14.3	28.0	32.6
	D	7.9	18.4	5.6	45.3	5.4	16.5	11.3	22.2
		6.3	10.4	11.8	47.4	5.9	16.2	14.4	24.3
	Rq	2.4	7.4	15.2	43.2	4.9	9.1	26.7	20.2
		2.8	11.4	33.8	54.2	4.8	16.5	39.6	21.2

Note: (O/Rp) Offer/Reply, (C) Consent, (D) Dissent, (Rq) Request

L.R. = 30.19 n.s.

APPENDIX F

UNCERTAINTY ANALYSES

Multivariate uncertainty analyses for conversation coded for speech Activity.

Overall transitions.

Dyad	Uncertainty						
	Nominal	Total		Error		Contingent	
	bits	bits	% Nominal	bits	% Total	bits	% Total
1	2.807	1.955	69.64	1.695	86.71	.259	13.29
2	2.807	2.147	76.49	1.823	84.89	.324	15.11
3	2.807	1.859	66.22	1.651	88.79	.208	11.21
4	2.807	1.921	68.43	1.658	86.31	.263	13.69
5	2.807	1.877	66.84	1.597	85.12	.279	14.88
6	2.807	1.901	67.69	1.606	84.49	.295	15.51
7	2.807	2.204	78.52	1.966	89.19	.238	10.81
8	2.807	2.024	72.11	1.706	84.26	.319	15.74
9	2.807	2.196	78.24	1.873	85.27	.323	14.73
10	2.807	2.026	72.15	1.798	88.77	.227	11.23
11	2.807	2.196	78.23	1.917	87.29	.279	12.71
12	2.807	2.065	73.57	1.801	87.21	.264	12.79
13	2.807	2.250	80.15	1.870	83.11	.380	16.89
14	2.807	1.972	70.26	1.630	82.65	.342	17.35
15	2.807	2.186	77.85	1.706	78.03	.480	21.97
16	2.807	2.151	76.62	1.824	84.79	.327	15.21
17	2.807	1.688	60.14	1.445	85.56	.244	14.44
18	2.807	2.087	74.33	1.802	86.35	.285	13.65
19	2.807	1.964	69.97	1.644	83.71	.319	16.29
20	2.807	2.017	71.83	1.746	86.59	.270	13.41
21	2.807	2.129	75.85	1.881	88.35	.248	11.65
22	2.807	2.347	83.61	1.990	84.80	.357	15.20
23	2.807	2.335	83.18	1.993	85.37	.342	14.63
24	2.807	2.332	83.09	1.950	83.60	.383	16.40
Mean	2.807	2.076	73.96	1.774	85.45	.302	14.55

Note: Rows do not sum due to rounding.

Within-speaker transitions

Dyad	Uncertainty						
	Nominal	Total		Error		Contingent	
	bits	bits	% Nominal	bits	% Total	bits	% Total
1	2.807	1.392	49.59	.877	63.01	.515	36.99
2	2.807	1.623	57.79	1.286	79.25	.337	20.75
3	2.807	1.430	50.95	1.119	78.29	.311	21.71
4	2.807	1.526	54.36	1.106	72.49	.419	27.51
5	2.807	1.198	42.69	.796	66.43	.402	33.57
6	2.807	1.260	44.88	.969	76.93	.291	23.07
7	2.807	1.668	59.42	1.259	75.45	.409	24.55
8	2.807	1.414	50.36	.913	64.59	.501	35.41
9	2.807	1.764	62.84	1.185	67.19	.579	32.81
10	2.807	1.548	55.13	1.116	72.09	.432	27.91
11	2.807	1.499	53.41	1.020	68.03	.479	31.97
12	2.807	1.728	61.55	1.116	64.59	.612	35.41
13	2.807	2.082	74.16	1.289	61.94	.792	38.06
14	2.807	1.561	55.59	1.069	68.48	.492	31.52
15	2.807	1.725	61.44	.947	54.91	.778	45.09
16	2.807	1.706	60.75	.957	56.13	.748	43.87
17	2.807	.929	33.10	.639	68.79	.290	31.21
18	2.807	1.578	56.19	.816	51.75	.761	48.25
19	2.807	1.179	42.00	.940	79.76	.239	20.24
20	2.807	1.540	54.86	1.071	69.53	.469	30.47
21	2.807	1.838	65.49	1.272	69.19	.566	30.81
22	2.807	1.994	71.04	1.208	60.59	.786	39.41
23	2.807	1.841	65.57	1.299	70.61	.541	29.39
24	2.807	1.945	69.28	1.175	60.39	.770	39.61
Mean	2.807	1.582	56.36	1.060	67.00	.522	33.00

Note: Rows do not sum due to rounding.

Between-speaker transitions

Dyad	Uncertainty						
	Nominal	Total		Error		Contingent	
	bits	bits	% Nominal	bits	% Total	bits	% Total
1	2.807	2.062	73.46	1.546	74.94	.517	25.06
2	2.807	2.349	83.67	1.743	74.19	.606	25.81
3	2.807	1.955	69.45	1.569	80.29	.385	19.71
4	2.807	1.973	70.29	1.478	74.91	.495	25.09
5	2.807	2.165	77.12	1.707	78.87	.458	21.13
6	2.807	2.084	74.22	1.491	71.55	.593	28.45
7	2.807	2.452	87.35	2.027	82.65	.425	17.35
8	2.807	2.243	79.89	1.629	72.62	.614	27.38
9	2.807	2.335	83.17	1.808	77.42	.527	22.58
10	2.807	2.216	78.94	1.788	80.68	.428	19.32
11	2.807	2.440	86.92	1.806	74.01	.634	25.99
12	2.807	2.077	73.98	1.546	74.44	.531	25.56
13	2.807	2.149	76.56	1.512	70.35	.637	29.65
14	2.807	1.959	69.77	1.333	68.03	.626	31.97
15	2.807	2.189	77.96	1.439	65.76	.749	34.24
16	2.807	2.246	80.00	1.797	80.01	.449	19.99
17	2.807	2.044	72.82	1.594	77.97	.450	22.03
18	2.807	2.279	82.19	1.920	84.26	.359	15.74
19	2.807	2.169	77.27	1.569	72.37	.599	27.63
20	2.807	2.128	75.79	1.607	75.52	.521	24.48
21	2.807	2.086	74.31	1.626	77.94	.460	22.06
22	2.807	2.378	84.71	1.962	82.51	.416	17.49
23	2.807	2.607	92.86	2.031	77.92	.576	22.08
24	2.807	2.397	85.39	1.960	81.78	.437	18.22
Mean	2.807	2.208	78.66	1.687	76.40	.521	23.60

Note: Rows do not sum due to rounding.

Univariate uncertainty analyses for conversation coded for speech Activity.

Relative Uncertainty of a subsequent event following a specified antecedent event.

Speech category	Transition type							
	Overall		Within		Between		Within + Between	
	bits	R.U.%	bits	R.U.%	bits	R.U.%	bits	R.U.%
Offer	1.895	67.5	.693	24.7	2.136	76.1	2.448	64.3
Reply	2.327	82.9	1.412	50.3	2.223	79.2	2.794	73.4
Consent	1.810	64.5	2.001	71.3	1.401	49.9	2.699	70.9
Dissent	2.198	78.3	1.665	59.3	2.378	84.7	2.985	78.4
Modify	2.380	84.8	1.844	65.7	2.296	81.8	3.076	80.8
Reaction	1.229	43.8	1.914	68.2	.912	32.5	1.926	50.6
Request	2.029	72.3	1.235	44.0	2.041	72.7	2.585	67.9

Relative Uncertainty of an antecedent event preceding a specified subsequent event

Speech Category	Transition type					
	Overall		Within		Between	
	bits	R.U.%	bits	R.U.%	bits	R.U.%
Offer	1.889	67.3	1.415	50.4	2.086	74.3
Reply	1.968	70.1	1.241	44.2	1.207	43.0
Consent	1.965	70.0	1.409	50.2	1.665	59.3
Dissent	2.215	78.9	1.568	55.86	2.187	77.9
Modify	2.299	81.9	2.066	73.6	2.246	80.0
Reaction	.999	35.6	0	0	1.145	40.8
Request	2.282	81.3	2.271	80.9	2.234	79.6

Note: Nominal uncertainty for Overall, Within-, and Between-transitions is 2.807 bits (7 speech categories) and 3.807 bits (14 speech categories) for Within + Between transitions.

Multivariate uncertainty analyses for conversation coded for speech Type.

Overall transitions.

Dyad	Uncertainty						
	Nominal	Total		Error		Contingent	
	bits	bits	% Nominal	bits	% Total	bits	% Total
1	4.700	2.418	50.86	1.991	82.34	.427	17.66
2	4.700	2.663	56.00	2.361	88.66	.302	11.34
3	4.700	2.591	54.49	2.172	83.84	.419	16.16
4	4.700	2.823	59.38	2.382	84.36	.442	15.64
5	4.700	2.632	55.36	2.401	91.19	.232	8.81
6	4.700	3.062	64.40	2.522	82.35	.541	17.65
7	4.700	3.068	64.52	2.462	80.26	.605	19.74
8	4.700	2.989	62.86	2.353	78.73	.636	21.27
9	4.700	2.948	61.99	2.466	83.66	.482	16.34
10	4.700	2.523	53.07	2.109	83.57	.415	16.43
11	4.700	2.858	60.11	2.305	80.65	.553	19.35
12	4.700	2.901	62.42	2.499	86.14	.402	13.86
13	4.700	2.956	62.17	2.556	86.46	.400	13.54
14	4.700	2.953	62.11	2.449	82.91	.504	17.09
15	4.700	2.791	58.71	2.383	85.38	.408	14.62
16	4.700	2.957	62.18	2.506	84.74	.451	15.26
17	4.700	3.227	67.86	2.512	77.86	.714	22.14
18	4.700	3.117	65.56	2.422	77.70	.695	22.30
19	4.700	3.031	63.74	2.592	85.51	.439	14.49
20	4.700	2.879	60.54	2.333	81.06	.545	18.94
21	4.700	3.008	63.26	2.485	82.61	.523	17.39
22	4.700	2.711	57.00	2.338	86.27	.372	13.73
23	4.700	2.752	57.87	2.383	86.60	.368	13.40
24	4.700	2.919	61.38	2.469	84.62	.449	15.38
Mean	4.700	2.866	60.98	2.394	83.53	.472	16.47

Note: Rows do not sum due to rounding.

Within-speaker transitions

Dyad	Uncertainty						
	Nominal	Total		Error		Contingent	
	bits	bits	% Nominal	bits	% Total	bits	% Total
1	4.700	1.820	37.87	1.169	64.27	.650	35.73
2	4.700	2.429	50.52	2.045	84.20	.384	15.80
3	4.700	2.204	45.85	1.504	68.21	.701	31.79
4	4.700	2.545	52.94	1.915	75.23	.631	24.77
5	4.700	2.235	46.49	1.713	76.63	.522	23.37
6	4.700	2.899	60.32	2.179	75.17	.720	24.83
7	4.700	2.917	60.68	1.978	67.80	.939	32.20
8	4.700	2.759	57.41	1.786	64.70	.974	35.30
9	4.700	2.638	54.86	1.883	71.41	.754	28.59
10	4.700	2.177	45.29	1.593	73.16	.584	26.84
11	4.700	2.617	54.44	1.814	69.31	.803	30.69
12	4.700	2.738	56.95	1.934	70.63	.804	29.37
13	4.700	2.815	58.56	2.143	76.14	.672	23.86
14	4.700	2.917	60.68	2.130	73.04	.787	26.96
15	4.700	2.460	51.18	1.849	75.14	.612	24.86
16	4.700	2.762	57.46	1.965	71.14	.797	28.86
17	4.700	3.032	63.07	1.851	61.04	1.181	38.96
18	4.700	2.914	60.61	1.847	63.39	1.067	36.61
19	4.700	2.819	58.66	2.159	76.59	.659	23.41
20	4.700	2.600	54.09	1.810	69.62	.789	30.38
21	4.700	3.035	63.13	2.309	76.09	.725	23.91
22	4.700	2.419	50.32	1.719	71.09	.699	28.91
23	4.700	2.569	53.45	2.048	79.69	.522	20.31
24	4.700	2.485	51.70	1.830	73.65	.655	26.35
Mean	4.700	2.617	55.68	1.882	71.91	.735	28.09

Note: Rows do not sum due to rounding.

Between-speaker transitions

Dyad	Uncertainty						
	Nominal	Total		Error		Contingent	
	bits	bits	% Nominal	bits	% Total	bits	% Total
1	4.700	2.501	52.02	1.854	74.12	.647	25.88
2	4.700	2.553	53.09	2.017	79.03	.535	20.97
3	4.700	2.539	52.83	2.033	80.04	.507	19.96
4	4.700	2.768	57.59	2.079	75.09	.689	24.91
5	4.700	2.622	54.53	2.271	86.64	.350	13.36
6	4.700	2.878	59.86	2.103	73.07	.775	26.93
7	4.700	2.870	59.70	2.281	79.46	.589	20.54
8	4.700	2.789	58.01	1.966	70.48	.823	29.52
9	4.700	2.889	60.09	2.259	78.22	.629	21.78
10	4.700	2.531	52.65	2.039	80.58	.492	19.42
11	4.700	2.589	53.86	1.998	77.15	.592	22.85
12	4.700	2.735	56.89	2.138	78.17	.597	21.83
13	4.700	2.789	58.02	2.115	75.83	.674	24.17
14	4.700	2.733	56.86	2.031	74.29	.703	25.71
15	4.700	2.646	55.04	1.992	75.29	.654	24.71
16	4.700	2.796	58.16	2.254	80.63	.542	19.37
17	4.700	2.983	62.06	2.336	78.29	.647	21.71
18	4.700	2.888	60.08	2.173	75.24	.715	24.76
19	4.700	2.949	61.33	2.255	76.48	.694	23.52
20	4.700	2.747	57.13	2.017	73.43	.729	26.57
21	4.700	2.674	55.62	1.895	70.88	.779	29.12
22	4.700	2.634	54.79	2.272	86.24	.362	13.76
23	4.700	2.634	54.79	2.202	83.58	.432	16.42
24	4.700	2.920	60.74	2.297	78.67	.623	21.33
Mean	4.700	2.736	58.21	2.119	77.45	.617	22.55

Note: Rows do not sum due to rounding.

Univariate uncertainty analyses for conversation coded for speech Type.

Relative Uncertainty of a subsequent event following a specified antecedent event.

Speech category	Transition type: Within		Between		Within + Between	
	bits	R.U.%	bits	R.U.%	bits	R.U.%
Belief	2.193	46.66	2.554	54.34	3.383	59.35
Personal Details	2.570	54.68	2.945	62.65	3.732	65.47
Personal Experience	2.198	46.76	2.431	51.92	3.282	57.57
Narrative	1.339	28.49	2.029	43.17	2.567	45.03
Subjective feelings	2.118	45.06	1.792	38.12	2.939	51.56
Attitude	2.506	53.31	2.935	62.44	3.735	65.52
Justify/Explain	2.079	44.23	2.410	51.27	3.248	56.98
Synthesis	1.922	40.89	0.971	20.66	2.446	42.91
Summary	2.332	49.61	2.484	52.85	3.376	59.22
Judgement	2.337	49.72	2.018	42.93	2.561	44.93
Avoidance	2.433	51.76	2.652	56.42	3.387	59.42
Negation	1.022	21.74	2.189	46.59	2.483	43.56
Active Recognition	2.792	59.39	2.758	58.68	3.775	66.22
Salience	2.339	49.76	3.182	67.69	3.745	65.69
Conclusion	2.625	55.85	2.362	50.25	3.349	58.75
Example	2.315	49.25	2.368	50.38	3.341	58.61
Emotional Support	1.585	33.72	1.585	33.72	1.922	33.72
Apology	2.500	53.19	1.522	32.38	3.085	54.12
Direction	2.587	55.04	2.738	58.25	3.638	63.82
Metastatement	2.816	59.91	2.756	58.63	3.729	65.42
Revision	2.635	56.06	2.513	53.46	3.538	62.07
Restatement	2.829	60.19	2.807	59.72	3.813	66.89
Completion	1.649	35.08	2.091	44.49	2.819	49.45
Listener Response	2.958	62.93	2.613	55.59	3.389	59.45
U. Interruptions	1.561	33.21	1.679	35.72	2.508	43.99
Laughter	1.000	21.27	2.371	50.44	2.792	48.98

Note: Nominal uncertainty for Within- and Between-transitions is 4.700 bits (26 speech categories) and 5.700 bits (52 speech categories) for Within + Between-transitions.

Relative Uncertainty of an antecedent event preceding a specified subsequent event.

Speech category	Transition type: Within		Transition type: Between	
	bits	R.U.%	bits	R.U.%
Belief	2.475	52.65	2.687	57.16
Personal Details	2.966	63.10	2.692	57.27
Personal Experience	2.382	50.68	2.643	56.23
Narrative	1.807	38.44	1.914	40.72
Subjective feelings	2.059	43.80	1.459	31.04
Attitude	3.071	65.33	2.626	55.87
Justify/Explain	2.446	52.04	2.668	56.76
Synthesis	.811	17.25	1.792	38.12
Summary	2.569	54.65	2.300	48.93
Judgement	2.862	60.89	2.562	54.51
Avoidance	3.149	66.99	3.041	64.69
Negation	2.217	47.17	2.146	45.66
Active Recognition	1.862	39.61	2.806	59.70
Salience	2.197	46.74	2.669	56.78
Conclusion	2.556	54.38	2.054	43.70
Example	2.678	56.97	2.573	54.74
Emotional Support	1.000	21.27	1.585	33.72
Apology	2.059	43.80	1.500	31.91
Direction	2.975	63.29	3.046	64.80
Metastatement	2.806	59.69	2.987	63.55
Revision	2.777	59.08	2.392	50.89
Restatement	2.421	51.51	2.879	61.25
Completion	1.406	29.91	1.645	34.99
Listener Response	.871	54.95	2.695	57.34
U. Interruptions	0	0	1.519	32.32
Laughter	1.000	21.27	2.371	50.44

Note: Nominal uncertainty for Within- and Between-transitions is 4.700 bits (26 speech categories).
Listener Responses and Unsuccessful Interruptions may only be associated with the Reaction Activity category. There, fore, similar to within-speaker Reactions, the range of within-speaker categories that may precede a Listener Response and an Unsuccessful Interruption is limited, in this case, to Active Recognitions, Listener Responses and Unsuccessful Interruptions. The nominal uncertainty is, therefore, 1.585 bits (3 speech categories).

APPENDIX G

COMPUTER PROGRAMS

All the following programs have been written in the BASIC programming language for execution on a DEC System 10 computer.

The programs expect data to be input from disk files in a 'one datum per line' format. There are two types of program: (1) those that can be used for 'Activity' analyses only, and (2) those that can be used for both 'Activity' and 'Type' analyses. The data format for each program is specified below.

Activity programs	1	2	3	4	5	6	7	8	9	10	11	12
data format	a	a	a	a	a	a	d	a	b	c	d	e

Activity/Type programs	13	14	15
data format	f	g	h

Data formats

a) A three digit number of the form;

Code	Speaker	Activity	Type
Permissible range	1 or 2	0-9	0

b) i) Frequency data from Program 7

ii) Probability data from Program 8

c) Probability data from Program 1

d) The same format as (a). In addition, it is assumed that for data input from Program 10, there are 100 sequences, N items in length.

e) Frequency data from Program 7 (N-step contingency comparisons) and Program 11 (Inter-event distance comparisons).

f) A four digit number of the form;

Code	Speaker	Activity	Type
Permissible range	1 or 2	0-6	01-28

g) A four digit number of the form;

Code	Speaker	Activity	Type
Permissible range	1 or 2	0-9	0-29

for both Activity and Activity + Type analyses.

h) i) Activity analysis.

A three digit number of the form;

Code	Speaker	Activity	Type
Permissible range	1 or 2	0-9	0

ii) Type analysis.

A four digit number of the form;

Code	Speaker	Activity	Type
Permissible range	1 or 2	0-9	0-28

```
00010 REM PROGRAM 1     SEQUENCE STATIONARITY
00020 REM
00030 REM SPEAKER-SPECIFIC TRANSITIONS
00040 REM A/K/K1= RAW DATA/NO. OF CODES/MATRIX SIZE
00050 REM N1/C= SEQUENCE LENGTH/COUNTER
00060 REM P/P1/P2= OVERALL/FIRST-/SECOND-HALF MATRIX
00070 REM H/H1/H2= P/P1/P2 ROW SUMS
00080 REM R/R1/R2= TOTAL/FIRST-/SECOND-HALF L.R.
00090 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00100 REM W/W5/X= FREQ./PROB./ROW TOTAL PRINT MATRIX
00110 REM Z= MATRIX PRINT SELECTOR
00120 REM B$/R$= MATRIX/RUN TITLE
00130 REM U$/V$= INPUT/OUTPUT FILE NAME
00140 DIM H(20),H1(20),H2(20),P(20,20),P1(20,20)
00150 DIM P2(20,20),W(20,20),W5(20,20),X(20)
00160 REM SET STORES= 0
00170 FOR I=1 TO 20
00180 FOR J=1 TO 20
00190 LET P(I,J)=P1(I,J)=P2(I,J)=W(I,J)=W5(I,J)=0
00200 NEXT J
00210 NEXT I
00220 FOR J=1 TO 20
00230 LET H(J)=H1(J)=H2(J)=X(J)=0
00240 NEXT J
00250 LET C=N1=R=R1=R2=0
00260 REM OPEN FILES
00270 PRINT "INPUT FILE NAME" ;
00280 INPUT U$
00290 PRINT
00300 PRINT "OUTPUT FILE NAME"  ;
00310 INPUT V$
00320 REM OPEN PARAMETER FILE
00330 LET V1$="PARA.DAT"
00340 FILE #1,U$;#2,V$;#3,V1$
00350 MARGIN #2,132
00360 SCRATCH #2,3
00370 REM MATRIX SIZE
00380 PRINT
00390 PRINT "NUMBER OF ACTIVITY CODES USED"
00400 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00410 PRINT
00420 PRINT "NUMBER OF CODES USED"  ;
00430 INPUT K
00440 LET K1=K+K
00450 PRINT
00460 PRINT "SEQUENCE LENGTH"  ;
00470 INPUT N1
00480 PRINT
00490 PRINT "TITLE"  ;
00500 INPUT R$
00510 REM INPUT DATA
00520 LET L=1
00530 INPUT #1,A
00540 LET C=1
00550 LET M=1
00560 GOSUB 2600
00570 IF END #1 THEN 830
```

```
00580 LET C=C+1
00590 INPUT #1,A
00600 LET M=L+1
00610 GOSUB 2600
00620 REM ADD TO MATRICES
00630 REM OVERALL MATRIX
00640 LET X=N(1)+1
00650 LET Y=N(2)+1
00660 IF S(1)=1 THEN 680
00670 LET X=X+K
00680 IF S(2)=1 THEN 700
00690 LET Y=Y+K
00700 LET P(X,Y)=P(X,Y)+1
00710 REM FIRST HALF MATRIX
00720 IF C> INT(N1/2) THEN 750
00730 LET P1(X,Y)=P1(X,Y)+1
00740 GOTO 790
00750 IF C= INT((N1/2)+1) THEN 790
00760 REM SECOND HALF MATRIX
00770 LET P2(X,Y)=P2(X,Y)+1
00780 REM SHIFT INPUT STACK
00790 LET Q=1
00800 LET S(Q)=S(Q+1)
00810 LET N(Q)=N(Q+1)
00820 GOTO 570
00830 REM ROW SUMS
00840 FOR I=1 TO K1
00850 LET H(I)=H1(I)=H2(I)=0
00860 NEXT I
00870 FOR I=1 TO K1
00880 FOR J=1 TO K1
00890 LET H(I)=H(I)+P(I,J)
00900 LET H1(I)=H1(I)+P1(I,J)
00910 LET H2(I)=H2(I)+P2(I,J)
00920 NEXT J
00930 NEXT I
00940 REM LIKELIHOOD RATIO
00950 LET R=R1=R2=0
00960 FOR I=1 TO K1
00970 FOR J=1 TO K1
00980 IF P(I,J)=0 THEN 1010
00990 LET E=LOGE(P(I,J)/H(I))
01000 GOTO 1020
01010 LET E=0
01020 IF P1(I,J)=0 THEN 1050
01030 LET F=LOGE(P1(I,J)/H1(I))
01040 GOTO 1060
01050 LET F=0
01060 IF P2(I,J)=0 THEN 1090
01070 LET G=LOGE(P2(I,J)/H2(I))
01080 GOTO 1100
01090 LET G=0
01100 LET R1=R1+(P1(I,J)*(E-F))
01110 LET R2=R2+(P2(I,J)*(E-G))
01120 NEXT J
01130 NEXT I
01140 LET R=-2*(R1+R2)
```

```
01150 REM PRINT OUTPUT
01160 PRINT #2,DATE$
01170 PRINT #2
01180 PRINT #2,"SOURCE FILE= ",U$
01190 PRINT #2
01200 PRINT #2,R$
01210 PRINT #2
01220 PRINT #2,"SEQUENCE LENGTH=";N1
01230 PRINT #2
01240 LET Z=0
01250 LET Z=Z+1
01260 ON Z GOTO 1270 ,1780 ,1440 ,1800 ,1610 ,1820 ,2380
01270 FOR I=1 TO K1
01280 FOR J=1 TO K1
01290 LET W5(I,J)=0
01300 NEXT J
01310 NEXT I
01320 FOR I=1 TO K1
01330 FOR J=1 TO K1
01340 LET W(I,J)=P(I,J)
01350 IF H(I)=0 THEN 1370
01360 LET W5(I,J)=P(I,J)/H(I)
01370 NEXT J
01380 NEXT I
01390 FOR J=1 TO K1
01400 LET X(J)=H(J)
01410 NEXT J
01420 LET B$="TOTAL SEQUENCE, SPEAKER 1"
01430 GOTO 1840
01440 FOR I=1 TO K1
01450 FOR J=1 TO K1
01460 LET W5(I,J)=0
01470 NEXT J
01480 NEXT I
01490 FOR I=1 TO K1
01500 FOR J=1 TO K1
01510 LET W(I,J)=P1(I,J)
01520 IF H1(I)=0 THEN 1540
01530 LET W5(I,J)=P1(I,J)/H1(I)
01540 NEXT J
01550 NEXT I
01560 FOR J=1 TO K1
01570 LET X(J)=H1(J)
01580 NEXT J
01590 LET B$="FIRST HALF-SEQUENCE, SPEAKER 1"
01600 GOTO 1840
01610 FOR I=1 TO K1
01620 FOR J=1 TO K1
01630 LET W5(I,J)=0
01640 NEXT J
01650 NEXT I
01660 FOR I=1 TO K1
01670 FOR J=1 TO K1
01680 LET W(I,J)=P2(I,J)
01690 IF H2(I)=0 THEN 1710
01700 LET W5(I,J)=P2(I,J)/H2(I)
01710 NEXT J
```

```
01720 NEXT I
01730 FOR J=1 TO K1
01740 LET X(J)=H2(J)
01750 NEXT J
01760 LET B$="SECOND HALF-SEQUENCE, SPEAKER 1"
01770 GOTO 1840
01780 LET B$="TOTAL SEQUENCE, SPEAKER 2"
01790 GOTO 2110
01800 LET B$="FIRST HALF-SEQUENCE, SPEAKER 2"
01810 GOTO 2110
01820 LET B$="SECOND HALF-SEQUENCE, SPEAKER 2"
01830 GOTO 2110
01840 PRINT #2,B$
01850 PRINT #2
01860 REM COLUMN INDICATOR
01870 FOR J=1 TO K
01880 PRINT #2,TAB(10*J);J-1;
01890 NEXT J
01900 PRINT #2
01910 PRINT #2
01920 REM PRINT MATRIX
01930 FOR I=1 TO K1
01940 PRINT #2
01950 IF I>K THEN 1980
01960 PRINT #2,I-1;
01970 GOTO 1990
01980 PRINT #2,I-(K+1);
01990 FOR J=1 TO K
02000 PRINT #2,TAB(10*J);W(I,J);
02010 NEXT J
02020 PRINT #2
02030 FOR J=1 TO K
02040 PRINT #2,TAB(10*J);W5(I,J);
02050 NEXT J
02060 PRINT #2
02070 NEXT I
02080 PRINT #2
02090 PRINT #2
02100 GOTO 1250
02110 PRINT #2,B$
02120 PRINT #2
02130 REM COLUMN INDICATOR
02140 FOR J=K+1 TO K1
02150 PRINT #2,TAB(10*(J-K));J-(K+1);
02160 NEXT J
02170 PRINT #2
02180 PRINT #2
02190 REM PRINT MATRIX
02200 FOR I=1 TO K1
02210 PRINT #2
02220 IF I>K THEN 2250
02230 PRINT #2,I-1;
02240 GOTO 2260
02250 PRINT #2,I-(K+1);
02260 FOR J=K+1 TO K1
02270 PRINT #2,TAB(10*(J-K));W(I,J);
02280 NEXT J
```

```
02290 PRINT #2,TAB(10+(10*(J-K)));X(I)
02300 FOR J=K+1 TO K1
02310 PRINT #2,TAB(10*(J-K));W5(I,J);
02320 NEXT J
02330 PRINT #2
02340 NEXT I
02350 PRINT #2
02360 PRINT #2
02370 GOTO 1250
02380 PRINT #2
02390 PRINT #2
02400 PRINT #2,"SEQUENCE STATIONARITY, USING THE"
02410 PRINT #2,"ANDERSON-GOODMAN L.R. TEST"
02420 PRINT #2
02430 PRINT #2,"L.R. HAS CHI-SQUARE DISTRIBUTION"
02440 PRINT #2
02450 LET K8=(K1*(K1-1))*(2-1)
02460 PRINT #2,"LIKELIHOOD RATIO =";R;"WITH";K8;"DF"
02470 REM PRINT PARAMETER FILE
02480 FOR I=1 TO K1
02490 FOR J=1 TO K1
02500 IF H(I)=0 THEN 2520
02510 LET W5(I,J)=P(I,J)/H(I)
02520 NEXT J
02530 NEXT I
02540 FOR I=1 TO K1
02550 FOR J=1 TO K1
02560 PRINT #3,W5(I,J)
02570 NEXT J
02580 NEXT I
02590 STOP
02600 REM RAW DATA CONVERSION ROUTINE
02610 LET S(M)=N(M)=T(M)=0
02620 IF A<100 THEN 2660
02630 LET S(M)=S(M)+1
02640 LET A=A-100
02650 GOTO 2620
02660 IF A<10 THEN 2700
02670 LET N(M)=N(M)+1
02680 LET A=A-10
02690 GOTO 2660
02700 LET T(M)=A
02710 RETURN
02720 END
```

```
00010 REM PROGRAM 2     SEQUENCE STATIONARITY
00020 REM
00030 REM OVERALL/WITHIN-/BETWEEN-SPEAKER TRANSITIONS
00040 REM A/K/K1= RAW DATA/NO. OF CODES/MATRIX SIZE
00050 REM N1/C= SEQUENCE LENGTH/COUNTER
00060 REM P/P1/P2= OVERALL/FIRST-/SECOND-HALF MATRIX
00070 REM H/H1/H2= P/P1/P2 ROW SUMS
00080 REM R/R1/R2= TOTAL/FIRST-/SECOND-HALF L.R.
00090 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00100 REM W/W5/X= FREQ./PROB./ROW TOTAL PRINT MATRIX
00110 REM Z= MATRIX PRINT SELECTOR
00120 REM B$/R$= MATRIX/RUN TITLE
00130 REM U$/V$= INPUT/OUTPUT FILE NAME
00140 DIM H(10),H1(10),H2(10),P(10,10),P1(10,10)
00150 DIM P2(10,10),W(10,10),W5(10,10),X(10)
00160 REM SET STORES= 0
00170 FOR I=1 TO 10
00180 FOR J=1 TO 10
00190 LET P(I,J)=P1(I,J)=P2(I,J)=W(I,J)=W5(I,J)=0
00200 NEXT J
00210 NEXT I
00220 FOR J=1 TO 10
00230 LET H(J)=H1(J)=H2(J)=X(J)=0
00240 NEXT J
00250 LET C=N1=R=R1=R2=0
00260 REM OPEN FILES
00270 PRINT "INPUT FILE NAME" ;
00280 INPUT U$
00290 PRINT
00300 PRINT "OUTPUT FILE NAME"  ;
00310 INPUT V$
00320 FILE #1,U$;#2,V$
00330 MARGIN #2,132
00340 SCRATCH #2
00350 REM MATRIX SIZE
00360 PRINT
00370 PRINT "NUMBER OF ACTIVITY CODES USED"
00380 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00390 PRINT
00400 PRINT "NUMBER OF CODES USED"  ;
00410 INPUT K1
00420 PRINT
00430 PRINT "SEQUENCE LENGTH"  ;
00440 INPUT N1
00450 PRINT
00460 PRINT "TITLE"  ;
00470 INPUT R$
00480 LET K=0
00490 LET K=K+1
00500 IF K=4 THEN 2270
00510 FOR I=1 TO 10
00520 FOR J=1 TO 10
00530 LET P(I,J)=P1(I,J)=P2(I,J)=W(I,J)=W5(I,J)=0
00540 NEXT J
00550 NEXT I
00560 FOR I=1 TO 10
00570 LET H(I)=H1(I)=H2(I)=X(I)=0
```

```
00580 NEXT I
00590 LET L=1
00600 INPUT #1,A
00610 LET C=1
00620 LET M=1
00630 GOSUB 2280
00640 IF END #1 THEN 910
00650 LET C=C+1
00660 INPUT #1,A
00670 LET M=L+1
00680 GOSUB 2280
00690 REM ADD TO MATRICES
00700 REM OVERALL MATRIX
00710 LET X=N(1)+1
00720 LET Y=N(2)+1
00730 ON K GOTO 780 ,740 ,760
00740 IF S(1)=S(2) THEN 780
00750 GOTO 870
00760 IF S(1)<>S(2) THEN 780
00770 GOTO 870
00780 LET P(X,Y)=P(X,Y)+1
00790 REM FIRST HALF MATRIX
00800 IF C> INT(N1/2) THEN 830
00810 LET P1(X,Y)=P1(X,Y)+1
00820 GOTO 870
00830 IF C= INT((N1/2)+1) THEN 870
00840 REM SECOND HALF MATRIX
00850 LET P2(X,Y)=P2(X,Y)+1
00860 REM SHIFT INPUT STACK
00870 LET Q=1
00880 LET S(Q)=S(Q+1)
00890 LET N(Q)=N(Q+1)
00900 GOTO 640
00910 REM ROW SUMS
00920 FOR I=1 TO K1
00930 LET H(I)=H1(I)=H2(I)=0
00940 NEXT I
00950 FOR I=1 TO K1
00960 FOR J=1 TO K1
00970 LET H(I)=H(I)+P(I,J)
00980 LET H1(I)=H1(I)+P1(I,J)
00990 LET H2(I)=H2(I)+P2(I,J)
01000 NEXT J
01010 NEXT I
01020 REM LIKELIHOOD RATIO
01030 LET R=R1=R2=0
01040 FOR I=1 TO K1
01050 FOR J=1 TO K1
01060 IF P(I,J)=0 THEN 1090
01070 LET E=LOGE(P(I,J)/H(I))
01080 GOTO 1100
01090 LET E=0
01100 IF P1(I,J)=0 THEN 1130
01110 LET F=LOGE(P1(I,J)/H1(I))
01120 GOTO 1140
01130 LET F=0
01140 IF P2(I,J)=0 THEN 1170
```

```
01150 LET G=LOGE(P2(I,J)/H2(I))
01160 GOTO 1180
01170 LET G=0
01180 LET R1=R1+(P1(I,J)*(E-F))
01190 LET R2=R2+(P2(I,J)*(E-G))
01200 NEXT J
01210 NEXT I
01220 LET R=-2*(R1+R2)
01230 REM PRINT OUTPUT
01240 PRINT #2
01250 PRINT #2
01260 PRINT #2
01270 PRINT #2,DATE$
01280 PRINT #2
01290 PRINT #2,"SOURCE FILE= ",U$
01300 PRINT #2
01310 PRINT #2,R$
01320 PRINT #2
01330 LET Z=0
01340 LET Z=Z+1
01350 ON Z GOTO 1360 ,1550 ,1740 ,2160
01360 FOR I=1 TO K1
01370 FOR J=1 TO K1
01380 LET W5(I,J)=0
01390 NEXT J
01400 NEXT I
01410 FOR I=1 TO K1
01420 FOR J=1 TO K1
01430 LET W(I,J)=P(I,J)
01440 IF H(I)=0 THEN 1460
01450 LET W5(I,J)=P(I,J)/H(I)
01460 NEXT J
01470 NEXT I
01480 FOR J=1 TO K1
01490 LET X(J)=H(J)
01500 NEXT J
01510 IF K=1 THEN B$="TOTAL SEQUENCE (OVERALL)"
01520 IF K=2 THEN B$="TOTAL SEQUENCE (WITHIN)"
01530 IF K=3 THEN B$="TOTAL SEQUENCE (BETWEEN)"
01540 GOTO 1920
01550 FOR I=1 TO K1
01560 FOR J=1 TO K1
01570 LET W5(I,J)=0
01580 NEXT J
01590 NEXT I
01600 FOR I=1 TO K1
01610 FOR J=1 TO K1
01620 LET W(I,J)=P1(I,J)
01630 IF H1(I)=0 THEN 1650
01640 LET W5(I,J)=P1(I,J)/H1(I)
01650 NEXT J
01660 NEXT I
01670 FOR J=1 TO K1
01680 LET X(J)=H1(J)
01690 NEXT J
01700 IF K=1 THEN B$="FIRST HALF-SEQUENCE (OVERALL)"
01710 IF K=2 THEN B$="FIRST HALF-SEQUENCE (WITHIN)"
```

```
01720 IF K=3 THEN B$="FIRST HALF-SEQUENCE (BETWEEN)"
01730 GOTO 1920
01740 FOR I=1 TO K1
01750 FOR J=1 TO K1
01760 LET W5(I,J)=0
01770 NEXT J
01780 NEXT I
01790 FOR I=1 TO K1
01800 FOR J=1 TO K1
01810 LET W(I,J)=P2(I,J)
01820 IF H2(I)=0 THEN 1840
01830 LET W5(I,J)=P2(I,J)/H2(I)
01840 NEXT J
01850 NEXT I
01860 FOR J=1 TO K1
01870 LET X(J)=H2(J)
01880 NEXT J
01890 IF K=1 THEN B$="SECOND HALF-SEQUENCE (OVERALL)"
01900 IF K=2 THEN B$="SECOND HALF-SEQUENCE (WITHIN)"
01910 IF K=3 THEN B$="SECOND HALF-SEQUENCE (BETWEEN)"
01920 PRINT #2,B$
01930 PRINT #2
01940 REM COLUMN INDICATOR
01950 FOR J=1 TO K1
01960 PRINT #2,TAB(10*J);J-1;
01970 NEXT J
01980 PRINT #2
01990 PRINT #2
02000 REM PRINT MATRIX
02010 FOR I=1 TO K1
02020 PRINT #2
02030 PRINT #2,I-1;
02040 FOR J=1 TO K1
02050 PRINT #2,TAB(10*J);W(I,J);
02060 NEXT J
02070 PRINT #2,TAB(10+(10*J));X(I)
02080 FOR J=1 TO K1
02090 PRINT #2,TAB(10*J);W5(I,J);
02100 NEXT J
02110 PRINT #2
02120 NEXT I
02130 PRINT #2
02140 PRINT #2
02150 GOTO 1340
02160 PRINT #2
02170 PRINT #2
02180 PRINT #2
02190 PRINT #2,"SEQUENCE STATIONARITY, USING THE"
02200 PRINT #2,"ANDERSON-GOODMAN L.R. TEST"
02210 PRINT #2
02220 PRINT #2,"L.R. HAS CHI-SQUARE DISTRIBUTION"
02230 PRINT #2
02240 LET K8= (K1*(K1-1))*(2-1)
02250 PRINT #2,"LIKELIHOOD RATIO =";R;"WITH";K8;"DF"
02260 GOTO 490
02270 STOP
02280 REM RAW DATA CONVERSION ROUTINE
```

```
02290 LET S(M)=N(M)=T(M)=0
02300 IF A<100 THEN 2340
02310 LET S(M)=S(M)+1
02320 LET A=A-100
02330 GOTO 2300
02340 IF A<10 THEN 2380
02350 LET N(M)=N(M)+1
02360 LET A=A-10
02370 GOTO 2340
02380 LET T(M)=A
02390 RETURN
02400 END
```

```
00010 REM PROGRAM 3     FIRST-ORDER ANALYSIS
00020 REM
00030 REM EXTRACTS ANTECEDENT SINGLE/SUBSEQUENT UNITS
00040 REM A1/K= RAW DATA/NO. OF CODES USED
00050 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00060 REM U$/W$= INPUT/OUTPUT FILE NAME
00070 REM V$/V1$/V2$= RAW DATA OUTPUT FILES
00080 REM P= OVERALL MATRIX
00090 REM P1= SAME SPKR. ANT/SAME SPKR. SUB MATRIX
00100 REM P3= DIFFERENT SPKR. ANT/SUB MATRIX
00110 DIM P(10,10),P1(10,10),P3(10,10)
00120 LET V$="SO.DAT"
00130 LET V1$="SW.DAT"
00140 LET V2$="SB.DAT"
00150 PRINT "INPUT FILE NAME"  ;
00160 INPUT U$
00170 PRINT
00180 PRINT "OUTPUT FILE NAME";
00190 INPUT W$
00200 FILE #1,U$;#2,W$;#3,V$;#4,V1$;#5,V2$
00210 MARGIN #2,132
00220 SCRATCH #2,3,4,5
00230 PRINT
00240 PRINT "NUMBER OF CODES USED (MAX 10)"
00250 PRINT "NOTE: IF 5 CODES USED, CODES 0-4 ASSUMED"
00260 PRINT
00270 PRINT "NUMBER OF ACTIVITY CODES USED"  ;
00280 INPUT K
00290 LET L=1
00300 INPUT #1,A1
00310 LET M=1
00320 GOSUB 1380
00330 IF END #1 THEN 550
00340 INPUT #1,A1
00350 LET M=L+1
00360 GOSUB 1380
00370 REM ADD TO MATRICES
00380 LET X=N(1)+1
00390 LET Y=N(2)+1
00400 REM OVERALL MATRIX
00410 LET P(X,Y)=P(X,Y)+1
00420 IF S(1)<>S(2) THEN 470
00430 REM SAME/SAME MATRIX
00440 LET P1(X,Y)=P1(X,Y)+1
00450 GOTO 480
00460 REM DIFFERENT MATRIX
00470 LET P3(X,Y)= P3(X,Y)+1
00480 REM SHIFT INPUT STACK
00490 LET Q=1
00500 LET S(Q)= S(Q+1)
00510 LET N(Q) = N(Q+1)
00520 IF Q=L THEN 330
00530 LET Q=Q+1
00540 GOTO 500
00550 PRINT #2, "ANTECEDENT SINGLE ANALYSIS"
00560 PRINT #2
00570 PRINT #2, DATE$
```

```
00580 PRINT #2
00590 PRINT #2,U$
00600 PRINT #2
00610 PRINT #2, "OVERALL"
00620 PRINT #2
00630 PRINT #2
00640 FOR I=1 TO K
00650 PRINT #2, TAB(10*I);I-1;
00660 NEXT I
00670 PRINT #2
00680 PRINT #2
00690 FOR I=1 TO 10
00700 REM ZERO LINE CHECK
00710 LET Y=0
00720 FOR J=1 TO K
00730 LET Y=Y+P(I,J)
00740 NEXT J
00750 IF Y=0 THEN 840
00760 PRINT #2,I-1;
00770 FOR J=1 TO K
00780 PRINT #2, TAB(10*J);P(I,J);
00790 PRINT #3, P(I,J)
00800 NEXT J
00810 PRINT #2, TAB(10+(10*J));Y
00820 PRINT #2
00830 PRINT #2
00840 NEXT I
00850 PRINT #2
00860 PRINT #2
00870 PRINT #2,"SAME SPKR. ANTECEDENT/SUBSEQUENT"
00880 PRINT #2
00890 PRINT #2
00900 FOR I=1 TO K
00910 PRINT #2,TAB(10*I);I-1;
00920 NEXT I
00930 PRINT #2
00940 PRINT #2
00950 FOR I=1 TO 10
00960 REM ZERO LINE CHECK
00970 LET Y=0
00980 FOR J=1 TO K
00990 LET Y=Y+P1(I,J)
01000 NEXT J
01010 IF Y=0 THEN 1100
01020 PRINT #2,I-1;
01030 FOR J=1 TO K
01040 PRINT #2,TAB(10*J);P1(I,J);
01050 PRINT #4, P1(I,J)
01060 NEXT J
01070 PRINT #2, TAB(10+(10*J));Y
01080 PRINT #2
01090 PRINT #2
01100 NEXT I
01110 PRINT #2
01120 PRINT #2
01130 PRINT #2,"DIFFERENT SPKR. ANT/SUB ANALYSIS"
01140 PRINT #2
```

```
01150 PRINT #2
01160 FOR I=1 TO K
01170 PRINT #2,TAB(10*I);I-1;
01180 NEXT I
01190 PRINT #2
01200 PRINT #2
01210 FOR I=1 TO 10
01220 REM ZERO LINE CHECK
01230 LET Y=0
01240 FOR J=1 TO K
01250 LET Y=Y+P3(I,J)
01260 NEXT J
01270 IF Y=0 THEN 1360
01280 PRINT #2,I-1;
01290 FOR J=1 TO K
01300 PRINT #2,TAB(10*J);P3(I,J);
01310 PRINT #5,P3(I,J)
01320 NEXT J
01330 PRINT #2,TAB(10+(10*J));Y
01340 PRINT #2
01350 PRINT #2
01360 NEXT I
01370 STOP
01380 REM RAW DATA TO MATRIX SUB ROUTINE
01390 LET S(M)=N(M)=T(M)=0
01400 IF A1<100 THEN 1440
01410 LET S(M) = S(M)+1
01420 LET A1=A1 -100
01430 GOTO 1400
01440 IF A1<10 THEN 1480
01450 LET N(M) = N(M)+1
01460 LET A1=A1 -10
01470 GOTO 1440
01480 LET T(M)=A1
01490 RETURN
01500 END
```

```
00010 REM PROGRAM 4     SECOND-ORDER ANALYSIS
00020 REM
00030 REM EXTRACTS ANTECEDENT PAIR/SUBSEQUENT UNITS
00040 REM A1/K= RAW DATA/NO. OF CODES USED
00050 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00060 REM U$/W$= INPUT/OUTPUT FILE NAME
00070 REM V$/V1$/V2$= RAW DATA OUTPUT FILES
00080 REM P= OVERALL MATRIX
00090 REM P1= SAME SPKR. ANT/SAME SPKR. SUB MATRIX
00100 REM P2= SAME SPKR. ANT/DIFF SPKR. SUB MATRIX
00110 REM P3= A/B/A MATRIX
00120 REM P4= A/B/B MATRIX
00130 DIM P(100,10),P1(100,10),P2(100,10)
00140 DIM P3(100,10),P4(100,10)
00150 LET V$="PO.DAT"
00160 LET V1$="PW.DAT"
00170 LET V2$="PB.DAT"
00180 PRINT "INPUT FILE NAME"  ;
00190 INPUT U$
00200 PRINT
00210 PRINT "OUTPUT FILE NAME";
00220 INPUT W$
00230 FILE #1,U$;#2,W$;#3,V$;#4,V1$;#5,V2$
00240 MARGIN #2,132
00250 SCRATCH #2,3,4,5
00260 PRINT
00270 PRINT "NUMBER OF CODES USED (MAX 10)"
00280 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00290 PRINT
00300 PRINT "NUMBER OF ACTIVITY CODES"
00310 INPUT K
00320 LET L=2
00330 INPUT #1,A1
00340 LET M=1
00350 GOSUB 2030
00360 INPUT #1,A1
00370 LET M=2
00380 GOSUB 2030
00390 IF END #1 THEN 700
00400 INPUT #1,A1
00410 LET M=L+1
00420 GOSUB 2030
00430 REM ADD TO MATRICES
00440 LET X=N(1)* 10 + N(2) +1
00450 LET Y= N(3)+1
00460 REM OVERALL MATRIX
00470 LET P(X,Y)=P(X,Y)+1
00480 IF S(1)<>S(2) THEN 570
00490 IF S(1)=S(3) THEN 540
00500 REM SAME/DIFFERENT MATRIX
00510 LET P2(X,Y)=P2(X,Y)+1
00520 GOTO 630
00530 REM SAME/SAME MATRIX
00540 LET P1(X,Y)=P1(X,Y)+1
00550 GOTO 630
00560 REM DIFFERENT/DIFFERENT MATRIX
00570 IF S(2)=S(3) THEN 610
```

```
00580 REM A/B/A MATRIX
00590 LET P3(X,Y)=P3(X,Y)+1
00600 GOTO 630
00610 REM A/B/B MATRIX
00620 LET P4(X,Y)=P4(X,Y)+1
00630 REM SHIFT INPUT STACK
00640 LET Q=1
00650 LET S(Q)= S(Q+1)
00660 LET N(Q) = N(Q+1)
00670 IF Q=L THEN 390
00680 LET Q=Q+1
00690 GOTO 650
00700 PRINT #2, "ANTECEDENT PAIR ANALYSIS"
00710 PRINT #2
00720 PRINT #2, DATE$
00730 PRINT #2
00740 PRINT #2,U$
00750 PRINT #2
00760 PRINT #2, "OVERALL"
00770 PRINT #2
00780 PRINT #2
00790 FOR I=1 TO K
00800 PRINT #2, TAB(10*I);I-1;
00810 NEXT I
00820 PRINT #2
00830 PRINT #2
00840 FOR I=1 TO 100
00850 REM ZERO LINE CHECK
00860 LET Y=0
00870 FOR J=1 TO K
00880 LET Y=Y+P(I,J)
00890 NEXT J
00900 IF Y=0 THEN 990
00910 PRINT #2,I-1;
00920 FOR J=1 TO K
00930 PRINT #2, TAB(10*J);P(I,J);
00940 PRINT #3, P(I,J)
00950 NEXT J
00960 PRINT #2, TAB(10+(10*J));Y
00970 PRINT #2
00980 PRINT #2
00990 NEXT I
01000 PRINT #2
01010 PRINT #2
01020 PRINT #2,"SAME SPKR. ANT/SAME SPKR. SUB"
01030 PRINT #2
01040 PRINT #2
01050 FOR I=1 TO K
01060 PRINT #2,TAB(10*I);I-1;
01070 NEXT I
01080 PRINT #2
01090 PRINT #2
01100 FOR I=1 TO 100
01110 REM ZERO LINE CHECK
01120 LET Y=0
01130 FOR J=1 TO K
01140 LET Y=Y+P1(I,J)
```

```
01150 NEXT J
01160 IF Y=0 THEN 1250
01170 PRINT #2,I-1;
01180 FOR J=1 TO K
01190 PRINT #2,TAB(10*J);P1(I,J);
01200 PRINT #4, P1(I,J)
01210 NEXT J
01220 PRINT #2, TAB(10+(10*J));Y
01230 PRINT #2
01240 PRINT #2
01250 NEXT I
01260 PRINT #2
01270 PRINT #2
01280 PRINT #2,"SAME SPKR. ANT/DIFF SPKR. SUB"
01290 PRINT #2
01300 PRINT #2
01310 FOR I=1 TO K
01320 PRINT #2,TAB(10*I);I-1;
01330 NEXT I
01340 PRINT #2
01350 PRINT #2
01360 FOR I= 1 TO 100
01370 REM ZERO LINE CHECK
01380 LET Y=0
01390 FOR J=1 TO K
01400 LET Y=Y+P2(I,J)
01410 NEXT J
01420 IF Y=0 THEN 1510
01430 PRINT #2,I-1;
01440 FOR J=1 TO K
01450 PRINT #2,TAB(10*J);P2(I,J);
01460 PRINT #5, P2(I,J)
01470 NEXT J
01480 PRINT #2,TAB(10+(10*J));Y
01490 PRINT #2
01500 PRINT #2
01510 NEXT I
01520 PRINT #2
01530 PRINT #2
01540 PRINT #2,"A/B/A MATRIX"
01550 PRINT #2
01560 PRINT #2
01570 FOR I=1 TO K
01580 PRINT #2,TAB(10*I);I-1;
01590 NEXT I
01600 PRINT #2
01610 PRINT #2
01620 FOR I=1 TO 100
01630 REM ZERO LINE CHECK
01640 LET Y=0
01650 FOR J=1 TO K
01660 LET Y=Y+P3(I,J)
01670 NEXT J
01680 IF Y=0 THEN 1760
01690 PRINT #2,I-1;
01700 FOR J=1 TO K
01710 PRINT #2,TAB(10*J);P3(I,J);
```

```
01720 NEXT J
01730 PRINT #2,TAB(10+(10*J));Y
01740 PRINT #2
01750 PRINT #2
01760 NEXT I
01770 PRINT #2
01780 PRINT #2
01790 PRINT #2,"A/B/B MATRIX"
01800 PRINT #2
01810 PRINT #2
01820 FOR I=1 TO K
01830 PRINT #2,TAB(10*I);I-1;
01840 NEXT I
01850 PRINT #2
01860 PRINT #2
01870 FOR I=1 TO 100
01880 LET Y=0
01890 FOR J=1 TO K
01900 LET Y=Y+P4(I,J)
01910 NEXT J
01920 IF Y=0 THEN 2000
01930 PRINT #2,I-1;
01940 FOR J=1 TO K
01950 PRINT #2,TAB(10*J);P4(I,J);
01960 NEXT J
01970 PRINT #2,TAB(10+(10*J));Y
01980 PRINT #2
01990 PRINT #2
02000 NEXT I
02010 STOP
02020 REM RAW DATA CONVERSION ROUTINE
02030 LET S(M)=N(M)=T(M)=0
02040 IF A1<100 THEN 2080
02050 LET S(M) = S(M)+1
02060 LET A1=A1 -100
02070 GOTO 2040
02080 IF A1<10 THEN 2120
02090 LET N(M) = N(M)+1
02100 LET A1=A1 -10
02110 GOTO 2080
02120 LET T(M)=A1
02130 RETURN
02140 END
```

```
00010 REM PROGRAM 5     THIRD-ORDER ANALYSIS
00020 REM
00030 REM EXTRACTS ANTECEDENT TRIPLE/SUBSEQUENT UNITS
00040 REM A1/K= RAW DATA/NO. OF CODES USED
00050 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00060 REM U$/W$= INPUT/OUTPUT FILE NAME
00070 REM V$/V1$/V2$= RAW DATA OUTPUT FILES
00080 REM P= OVERALL MATRIX
00090 REM P1= SAME SPKR. ANT/SAME SPKR. SUB MATRIX
00100 REM P2= SAME SPKR. ANT/DIFF SPKR. SUB MATRIX
00110 REM P3= DIFF SPKR. ANTECEDENTS MATRIX
00120 DIM P(1000,10),P1(1000,10),P2(1000,10),P3(1000,10)
00130 LET V$="TO.DAT"
00140 LET V1$="TW.DAT"
00150 LET V2$="TB.DAT"
00160 PRINT "INPUT FILE NAME";
00170 INPUT U$
00180 PRINT
00190 PRINT "OUTPUT FILE NAME";
00200 INPUT W$
00210 FILE #1,U$;#2,W$;#3,V$;#4,V1$;#5,V2$
00220 MARGIN #2,132
00230 SCRATCH #2,3,4,5
00240 PRINT
00250 PRINT "NUMBER OF CODES USED (MAX 10)"
00260 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00270 PRINT
00280 PRINT "NUMBER OF ACTIVITY CODES USED"
00290 INPUT K
00300 LET L=3
00310 INPUT #1,A1
00320 LET M=1
00330 GOSUB 1750
00340 INPUT #1,A1
00350 LET M=2
00360 GOSUB 1750
00370 INPUT #1,A1
00380 LET M=3
00390 GOSUB 1750
00400 IF END #1 THEN 670
00410 INPUT #1,A1
00420 LET M=L+1
00430 GOSUB 1750
00440 REM ADD TO MATRICES
00450 LET X=N(1)*100 + N(2)*10 + N(3)+1
00460 LET Y= N(4)+1
00470 REM OVERALL MATRIX
00480 LET P(X,Y)=P(X,Y)+1
00490 IF S(1)<>S(2) THEN 590
00500 IF S(2)<>S(3) THEN 590
00510 IF S(1)=S(4) THEN 560
00520 REM SAME/DIFFERENT MATRIX
00530 LET P2(X,Y)=P2(X,Y)+1
00540 GOTO 600
00550 REM SAME/SAME MATRIX
00560 LET P1(X,Y)=P1(X,Y)+1
00570 GOTO 600
```

```
00580 REM DIFFERENT/DIFFERENT MATRIX
00590 LET P3(X,Y)=P3(X,Y)+1
00600 REM SHIFT INPUT STACK
00610 LET Q=1
00620 LET S(Q)=S(Q+1)
00630 LET N(Q)=N(Q+1)
00640 IF Q=L THEN 400
00650 LET Q=Q+1
00660 GOTO 620
00670 PRINT #2, "ANTECEDENT TRIPLE ANALYSIS"
00680 PRINT #2
00690 PRINT #2,DATE$
00700 PRINT #2
00710 PRINT #2,U$
00720 PRINT #2
00730 PRINT #2, "OVERALL"
00740 PRINT #2
00750 PRINT #2
00760 FOR I=1 TO K
00770 PRINT #2, TAB(10*I);I-1;
00780 NEXT I
00790 PRINT #2
00800 PRINT #2
00810 FOR I=1 TO 1000
00820 REM ZERO LINE CHECK
00830 LET Y=0
00840 FOR J=1 TO K
00850 LET Y=Y+P(I,J)
00860 NEXT J
00870 IF Y=0 THEN 960
00880 PRINT #2,I-1;
00890 FOR J=1 TO K
00900 PRINT #2, TAB(10*J);P(I,J);
00910 PRINT #3,P(I,J)
00920 NEXT J
00930 PRINT #2, TAB(10+(10*J));Y
00940 PRINT #2
00950 PRINT #2
00960 NEXT I
00970 PRINT #2
00980 PRINT #2
00990 PRINT #2,"SAME SPKR. ANT/SAME SPKR. SUB"
01000 PRINT #2
01010 PRINT #2
01020 FOR I=1 TO K
01030 PRINT #2,TAB(10*I);I-1;
01040 NEXT I
01050 PRINT #2
01060 PRINT #2
01070 FOR I=1 TO 1000
01080 REM ZERO LINE CHECK
01090 LET Y=0
01100 FOR J=1 TO K
01110 LET Y=Y+P1(I,J)
01120 NEXT J
01130 IF Y=0 THEN 1220
01140 PRINT #2,I-1;
```

```
01150 FOR J=1 TO K
01160 PRINT #2,TAB(10*J);P1(I,J);
01170 PRINT #4,P1(I,J)
01180 NEXT J
01190 PRINT #2,TAB(10+(10*J));Y
01200 PRINT #2
01210 PRINT #2
01220 NEXT I
01230 PRINT #2
01240 PRINT #2
01250 PRINT #2,"SAME SPKR. ANT/DIFF SPKR. SUB"
01260 PRINT #2
01270 PRINT #2
01280 FOR I=1 TO K
01290 PRINT #2,TAB(10*I);I-1;
01300 NEXT I
01310 PRINT #2
01320 PRINT #2
01330 FOR I=1 TO 1000
01340 REM ZERO LINE CHECK
01350 LET Y=0
01360 FOR J=1 TO K
01370 LET Y=Y+P2(I,J)
01380 NEXT J
01390 IF Y=0 THEN 1480
01400 PRINT #2,I-1;
01410 FOR J=1 TO K
01420 PRINT #2,TAB(10*J);P2(I,J);
01430 PRINT #5,P2(I,J)
01440 NEXT J
01450 PRINT #2,TAB(10+(10*J));Y
01460 PRINT #2
01470 PRINT #2
01480 NEXT I
01490 PRINT #2
01500 PRINT #2
01510 PRINT #2,"DIFF SPKR. ANT/DIFF SPKR. SUB"
01520 PRINT #2
01530 PRINT #2
01540 FOR I=1 TO K
01550 PRINT #2,TAB(10*I);I-1;
01560 NEXT I
01570 PRINT #2
01580 PRINT #2
01590 FOR I=1 TO 1000
01600 REM ZERO LINE CHECK
01610 LET Y=0
01620 FOR J=1 TO K
01630 LET Y=Y+P3(I,J)
01640 NEXT J
01650 IF Y=0 THEN 1730
01660 PRINT #2,I-1;
01670 FOR J=1 TO K
01680 PRINT #2,TAB(10*J);P3(I,J);
01690 NEXT J
01700 PRINT #2,TAB(10+(10*J));Y
01710 PRINT #2
```

```
01720 PRINT #2
01730 NEXT I
01740 STOP
01750 REM RAW DATA CONVERSION ROUTINE
01760 LET S(M)=N(M)=T(M)=0
01770 IF A1<100 THEN 1810
01780 LET S(M) = S(M)+1
01790 LET A1=A1-100
01800 GOTO 1770
01810 IF A1<10 THEN 1850
01820 LET N(M) = N(M)+1
01830 LET A1=A1-10
01840 GOTO 1810
01850 LET T(M)=A1
01860 RETURN
01870 END
```

```
00010 REM PROGRAM 6     SAMPLE HOMOGENEITY
00020 REM
00030 REM OVERALL/WITHIN-/BETWEEN-SPEAKER TRANSITIONS
00040 REM PROGRAM LIMITED TO 25 SEQUENCES
00050 REM A/K1/N5= RAW DATA/MATRIX SIZE/NO. OF FILES
00060 REM H/H1-U5= OVERALL/INDIVIDUAL ROW PROB.
00070 REM Z9/V1-Z5= TOTAL/INDIVIDUAL L.R.
00080 REM K= OVERALL/WITHIN/BETWEEN MATRIX IDENTIFIER
00090 REM P/B1-D5= OVERALL/INDIVIDUAL MATRICES
00100 REM P1/E1-G5= OVERALL/INDIV. MATRIX ROW SUMS
00110 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00120 REM S1-S9,N1-N9,T1-T8= PROBABILITY MATRICES
00130 REM Z6/Z7/Z8= PROB/FREQ/ROW TOTAL PRINT MATRIX
00140 REM R$/U$/V$= TITLE/INPUT/OUTPUT FILENAME
00150 DIM B1(10,10),B2(10,10),B3(10,10),B4(10,10)
00160 DIM B6(10,10),B7(10,10),B8(10,10),B9(10,10)
00170 DIM C(10,10),C1(10,10),C2(10,10),C3(10,10)
00180 DIM C5(10,10),C6(10,10),C7(10,10),C8(10,10)
00190 DIM D(10,10),D1(10,10),D2(10,10),D3(10,10)
00200 DIM D5(10,10),E1(10),E2(10),E3(10),E4(10)
00210 DIM E6(10),E7(10),E8(10),E9(10),P(10,10),U$(25)
00220 DIM F(10),F1(10),F2(10),F3(10),F4(10),F5(10)
00230 DIM F6(10),F7(10),F8(10),F9(10),Z6(10,10)
00240 DIM G(10),G1(10),G2(10),G3(10),G4(10),G5(10)
00250 DIM S1(10,10),S2(10,10),S3(10,10),S4(10,10)
00260 DIM S6(10,10),S7(10,10),S8(10,10),S9(10,10)
00270 DIM N1(10,10),N2(10,10),N3(10,10),N4(10,10)
00280 DIM N6(10,10),N7(10,10),N8(10,10),N9(10,10)
00290 DIM T1(10,10),T2(10,10),T3(10,10),T4(10,10)
00300 DIM T6(10,10),T7(10,10),T8(10,10),Z8(10)
00310 DIM B5(10,10),C4(10,10),C9(10,10),D4(10,10)
00320 DIM Z7(10,10),S5(10,10),N5(10,10),T5(10,10)
00330 DIM E5(10),P1(10)
00340 REM OPEN FILES
00350 PRINT "RESULTS FILE NAME"  ;
00360 INPUT V$
00370 FILE #2,V$
00380 MARGIN #2,132
00390 SCRATCH #2
00400 PRINT
00410 PRINT "NUMBER OF ACTIVITY CODES"  ;
00420 PRINT "IF 5 CODES ARE USED, CODES 0-4 ASSUMED"
00430 PRINT
00440 PRINT "NUMBER OF CODES USED"  ;
00450 INPUT K1
00460 PRINT
00470 PRINT "TITLE"  ;
00480 INPUT R$
00490 PRINT
00500 PRINT "EACH SEQUENCE TO BE IN SEPARATE FILE"
00510 PRINT
00520 PRINT "NUMBER OF FILES (MAX 25, MIN 2)"  ;
00530 INPUT N5
00540 REM INPUT FILE NAMES
00550 FOR I=1 TO N5
00560 PRINT
00570 PRINT "DATA FILE";I;"="  ;
```

```
00580 INPUT U$(I)
00590 NEXT I
00600 LET K=0
00610 LET K=K+1
00620 FOR I=1 TO K1
00630 FOR J=1 TO K1
00640 LET B1(I,J)=B2(I,J)=B3(I,J)=B4(I,J)=B5(I,J)=0
00650 LET B6(I,J)=B7(I,J)=B8(I,J)=B9(I,J)=0
00660 LET C1(I,J)=C2(I,J)=C3(I,J)=C4(I,J)=C5(I,J)=0
00670 LET C6(I,J)=C7(I,J)=C8(I,J)=C9(I,J)=0
00680 LET D1(I,J)=D2(I,J)=D3(I,J)=D4(I,J)=D5(I,J)=0
00690 LET P(I,J)=C(I,J)=D(I,J)=0
00700 NEXT J
00710 NEXT I
00720 FOR I= 1 TO K1
00730 LET P1(I)=E1(I)=E2(I)=E3(I)=E4(I)=E5(I)=0
00740 LET E6(I)=E7(I)=E8(I)=E9(I)=0
00750 LET F(I)=F1(I)=F2(I)=F3(I)=F4(I)=0
00760 LET F5(I)=F6(I)=F7(I)=F8(I)=F9(I)=0
00770 LET G(I)=G1(I)=G2(I)=G3(I)=G4(I)=G5(I)=0
00780 NEXT I
00790 FOR I=1 TO N5
00800 FILE #1,U$(I)
00810 LET L=1
00820 INPUT #1,A
00830 LET M=1
00840 GOSUB 6300
00850 IF END #1 THEN 1550
00860 INPUT #1,A
00870 LET M=L+1
00880 GOSUB 6300
00890 REM ADD TO MATRICES
00900 REM OVERALL MATRIX
00910 LET X=N(1)+1
00920 LET Y=N(2)+1
00930 ON K GOTO 980 ,940 ,960 ,6290
00940 IF S(1)=S(2) THEN 980
00950 GOTO 1510
00960 IF S(1)<>S(2) THEN 980
00970 GOTO 1510
00980 LET P(X,Y)=P(X,Y)+1
00990 REM INDIVIDUAL MATRICES
01000 ON I GOTO 1010,1030,1050,1070,1090,1110
                1130,1150,1170,1190,1210,1230
                1250,1270,1290,1310,1330,1350
                1370,1390,1410,1430,1450,1470,1490
01010 LET B1(X,Y)=B1(X,Y)+1
01020 GOTO 1510
01030 LET B2(X,Y)=B2(X,Y)+1
01040 GOTO 1510
01050 LET B3(X,Y)=B3(X,Y)+1
01060 GOTO 1510
01070 LET B4(X,Y)=B4(X,Y)+1
01080 GOTO 1510
01090 LET B5(X,Y)=B5(X,Y)+1
01100 GOTO 1510
01110 LET B6(X,Y)=B6(X,Y)+1
```

```
01120 GOTO 1510
01130 LET B7(X,Y)=B7(X,Y)+1
01140 GOTO 1510
01150 LET B8(X,Y)=B8(X,Y)+1
01160 GOTO 1510
01170 LET B9(X,Y)=B9(X,Y)+1
01180 GOTO 1510
01190 LET C(X,Y)=C(X,Y)+1
01200 GOTO 1510
01210 LET C1(X,Y)=C1(X,Y)+1
01220 GOTO 1510
01230 LET C2(X,Y)=C2(X,Y)+1
01240 GOTO 1510
01250 LET C3(X,Y)=C3(X,Y)+1
01260 GOTO 1510
01270 LET C4(X,Y)=C4(X,Y)+1
01280 GOTO 1510
01290 LET C5(X,Y)=C5(X,Y)+1
01300 GOTO 1510
01310 LET C6(X,Y)=C6(X,Y)+1
01320 GOTO 1510
01330 LET C7(X,Y)=C7(X,Y)+1
01340 GOTO 1510
01350 LET C8(X,Y)=C8(X,Y)+1
01360 GOTO 1510
01370 LET C9(X,Y)=C9(X,Y)+1
01380 GOTO 1510
01390 LET D(X,Y)=D(X,Y)+1
01400 GOTO 1510
01410 LET D1(X,Y)=D1(X,Y)+1
01420 GOTO 1510
01430 LET D2(X,Y)=D2(X,Y)+1
01440 GOTO 1510
01450 LET D3(X,Y)=D3(X,Y)+1
01460 GOTO 1510
01470 LET D4(X,Y)=D4(X,Y)+1
01480 GOTO 1510
01490 LET D5(X,Y)=D5(X,Y)+1
01500 REM SHIFT INPUT STACK
01510 LET Q=1
01520 LET S(Q)=S(Q+1)
01530 LET N(Q)=N(Q+1)
01540 GOTO 850
01550 NEXT I
01560 REM ROW SUMS
01570 FOR I=1 TO K1
01580 FOR J=1 TO K1
01590 LET P1(I)=P1(I)+P(I,J)
01600 LET E1(I)=E1(I)+B1(I,J)
01610 LET E2(I)=E2(I)+B2(I,J)
01620 LET E3(I)=E3(I)+B3(I,J)
01630 LET E4(I)=E4(I)+B4(I,J)
01640 LET E5(I)=E5(I)+B5(I,J)
01650 LET E6(I)=E6(I)+B6(I,J)
01660 LET E7(I)=E7(I)+B7(I,J)
01670 LET E8(I)=E8(I)+B8(I,J)
01680 LET E9(I)=E9(I)+B9(I,J)
```

```
01690 LET F(I)=F(I)+C(I,J)
01700 LET F1(I)=F1(I)+C1(I,J)
01710 LET F2(I)=F2(I)+C2(I,J)
01720 LET F3(I)=F3(I)+C3(I,J)
01730 LET F4(I)=F4(I)+C4(I,J)
01740 LET F5(I)=F5(I)+C5(I,J)
01750 LET F6(I)=F6(I)+C6(I,J)
01760 LET F7(I)=F7(I)+C7(I,J)
01770 LET F8(I)=F8(I)+C8(I,J)
01780 LET F9(I)=F9(I)+C9(I,J)
01790 LET G(I)=G(I)+D(I,J)
01800 LET G1(I)=G1(I)+D1(I,J)
01810 LET G2(I)=G2(I)+D2(I,J)
01820 LET G3(I)=G3(I)+D3(I,J)
01830 LET G4(I)=G4(I)+D4(I,J)
01840 LET G5(I)=G5(I)+D5(I,J)
01850 NEXT J
01860 NEXT I
01870 REM LIKELIHOOD RATIO
01880 LET V1=V2=V3=V4=V5=V6=V7=V8=V9=W=W1=W2=W3=0
01890 LET W4=W5=W6=W7=W8=W9=Z=Z1=Z2=Z3=Z4=Z5=0
01900 FOR I=1 TO K1
01910 FOR J=1 TO K1
01920 LET S3(I,J)=S4(I,J)=S5(I,J)=S6(I,J)=0
01930 LET S7(I,J)=S8(I,J)=S9(I,J)=0
01940 LET N1(I,J)=N2(I,J)=N3(I,J)=N4(I,J)=N5(I,J)=0
01950 LET N6(I,J)=N7(I,J)=N8(I,J)=N9(I,J)=0
01960 LET T1(I,J)=T2(I,J)=T3(I,J)=T4(I,J)=T5(I,J)=0
01970 LET T6(I,J)=T7(I,J)=T8(I,J)=0
01980 NEXT J
01990 NEXT I
02000 FOR I=1 TO K1
02010 FOR J=1 TO K1
02020 IF P1(I)=0 THEN 2060
02030 IF P(I,J)=0 THEN 2060
02040 LET H=LOGE(P(I,J)/P1(I))
02050 GOTO 2070
02060 LET H=0
02070 IF E1(I)=0 THEN 2110
02080 IF B1(I,J)=0 THEN 2110
02090 LET H1=LOGE(B1(I,J)/E1(I))
02100 GOTO 2120
02110 LET H1=0
02120 LET V1=V1+(B1(I,J)*(H-H1))
02130 IF P1(I)=0 THEN 2160
02140 LET S1(I,J)=P(I,J)/P1(I)
02150 GOTO 2170
02160 LET S1(I,J)=0
02170 IF E1(I)=0 THEN 2200
02180 LET S2(I,J)=B1(I,J)/E1(I)
02190 GOTO 2210
02200 LET S2(I,J)=0
02210 IF N5=1 THEN 4120
02220 IF E2(I)=0 THEN 2270
02230 IF B2(I,J)=0 THEN 2270
02240 LET S3(I,J)=B2(I,J)/E2(I)
02250 LET H2=LOGE(B2(I,J)/E2(I))
```

```
02260 GOTO 2280
02270 LET H2=0
02280 LET V2=V2+(B2(I,J)*(H-H2))
02290 IF N5=2 THEN 4120
02300 IF E3(I)=0 THEN 2350
02310 IF B3(I,J)=0 THEN 2350
02320 LET S4(I,J)=B3(I,J)/E3(I)
02330 LET H3=LOGE(B3(I,J)/E3(I))
02340 GOTO 2360
02350 LET H3=0
02360 LET V3=V3+(B3(I,J)*(H-H3))
02370 IF N5=3 THEN 4120
02380 IF E4(I)=0 THEN 2430
02390 IF B4(I,J)=0 THEN 2430
02400 LET S5(I,J)=B4(I,J)/E4(I)
02410 LET H4=LOGE(B4(I,J)/E4(I))
02420 GOTO 2440
02430 LET H4=0
02440 LET V4=V4+(B4(I,J)*(H-H4))
02450 IF N5=4 THEN 4120
02460 IF E5(I)=0 THEN 2510
02470 IF B5(I,J)=0 THEN 2510
02480 LET S6(I,J)=B5(I,J)/E5(I)
02490 LET H5=LOGE(B5(I,J)/E5(I))
02500 GOTO 2520
02510 LET H5=0
02520 LET V5=V5+(B5(I,J)*(H-H5))
02530 IF N5=5 THEN 4120
02540 IF E6(I)=0 THEN 2590
02550 IF B6(I,J)=0 THEN 2590
02560 LET S7(I,J)=B6(I,J)/E6(I)
02570 LET H6=LOGE(B6(I,J)/E6(I))
02580 GOTO 2600
02590 LET H6=0
02600 LET V6=V6+(B6(I,J)*(H-H6))
02610 IF N5=6 THEN 4120
02620 IF B7(I,J)=0 THEN 2660
02630 LET S8(I,J)=B7(I,J)/E7(I)
02640 LET H7=LOGE(B7(I,J)/E7(I))
02650 GOTO 2670
02660 LET H7=0
02670 LET V7=V7+(B7(I,J)*(H-H7))
02680 IF N5=7 THEN 4120
02690 IF E8(I)=0 THEN 2740
02700 IF B8(I,J)=0 THEN 2740
02710 LET S9(I,J)=B8(I,J)/E8(I)
02720 LET H8=LOGE(B8(I,J)/E8(I))
02730 GOTO 2750
02740 LET H8=0
02750 LET V8=V8+(B8(I,J)*(H-H8))
02760 IF N5=8 THEN 4120
02770 IF E9(I)=0 THEN 2820
02780 IF B9(I,J)=0 THEN 2820
02790 LET N1(I,J)=B9(I,J)/E9(I)
02800 LET H9=LOGE(B9(I,J)/E9(I))
02810 GOTO 2830
02820 LET H9=0
```

```
02830 LET V9=V9+(B9(I,J)*(H-H9))
02840 IF N5=9 THEN 4120
02850 IF F(I)=0 THEN 2900
02860 IF C(I,J)=0 THEN 2900
02870 LET N2(I,J)=C(I,J)/F(I)
02880 LET R=LOGE(C(I,J)/F(I))
02890 GOTO 2910
02900 LET R=0
02910 LET W=W+(C(I,J)*(H-R))
02920 IF N5=10 THEN 4120
02930 IF F1(I)=0 THEN 2980
02940 IF C1(I,J)=0 THEN 2980
02950 LET N3(I,J)=C1(I,J)/F1(I)
02960 LET R1=LOGE(C1(I,J)/F1(I))
02970 GOTO 2990
02980 LET R1=0
02990 LET W1=W1+(C1(I,J)*(H-R1))
03000 IF N5=11 THEN 4120
03010 IF F2(I)=0 THEN 3060
03020 IF C2(I,J)=0 THEN 3060
03030 LET N4(I,J)=C2(I,J)/F2(I)
03040 LET R2=LOGE(C2(I,J)/F2(I))
03050 GOTO 3070
03060 LET R2=0
03070 LET W2=W2+(C2(I,J)*(H-R2))
03080 IF N5=12 THEN 4120
03090 IF F3(I)=0 THEN 3140
03100 IF C3(I,J)=0 THEN 3140
03110 LET N5(I,J)=C3(I,J)/F3(I)
03120 LET R3=LOGE(C3(I,J)/F3(I))
03130 GOTO 3150
03140 LET R3=0
03150 LET W3=W3+(C3(I,J)*(H-R3))
03160 IF N5=13 THEN 4120
03170 IF F4(I)=0 THEN 3220
03180 IF C4(I,J)=0 THEN 3220
03190 LET N6(I,J)=C4(I,J)/F4(I)
03200 LET R4=LOGE(C4(I,J)/F4(I))
03210 GOTO 3230
03220 LET R4=0
03230 LET W4=W4+(C4(I,J)*(H-R4))
03240 IF N5=14 THEN 4120
03250 IF F5(I)=0 THEN 3300
03260 IF C5(I,J)=0 THEN 3300
03270 LET N7(I,J)=C5(I,J)/F5(I)
03280 LET R5=LOGE(C5(I,J)/F5(I))
03290 GOTO 3310
03300 LET R5=0
03310 LET W5=W5+(C5(I,J)*(H-R5))
03320 IF N5=15 THEN 4120
03330 IF F6(I)=0 THEN 3380
03340 IF C6(I,J)=0 THEN 3380
03350 LET N8(I,J)=C6(I,J)/F6(I)
03360 LET R6=LOGE(C6(I,J)/F6(I))
03370 GOTO 3390
03380 LET R6=0
03390 LET W6=W6+(C6(I,J)*(H-R6))
```

```
03400 IF N5=16 THEN 4120
03410 IF F7(I)=0 THEN 3460
03420 IF C7(I,J)=0 THEN 3460
03430 LET N9(I,J)=C7(I,J)/F7(I)
03440 LET R7=LOGE(C7(I,J)/F7(I))
03450 GOTO 3470
03460 LET R7=0
03470 LET W7=W7+(C7(I,J)*(H-R7))
03480 IF N5=17 THEN 4120
03490 IF F8(I)=0 THEN 3540
03500 IF C8(I,J)=0 THEN 3540
03510 LET T1(I,J)=C8(I,J)/F8(I)
03520 LET R8=LOGE(C8(I,J)/F8(I))
03530 GOTO 3550
03540 LET R8=0
03550 LET W8=W8+(C8(I,J)*(H-R8))
03560 IF N5=18 THEN 4120
03570 IF F9(I)=0 THEN 3620
03580 IF C9(I,J)=0 THEN 3620
03590 LET T2(I,J)=C9(I,J)/F9(I)
03600 LET R9=LOGE(C9(I,J)/F9(I))
03610 GOTO 3630
03620 LET R9=0
03630 LET W9=W9+(C9(I,J)*(H-R9))
03640 IF N5=19 THEN 4120
03650 IF G(I)=0 THEN 3700
03660 IF D(I,J)=0 THEN 3700
03670 LET T3(I,J)=D(I,J)/G(I)
03680 LET U=LOGE(D(I,J)/G(I))
03690 GOTO 3710
03700 LET U=0
03710 LET Z=Z+(D(I,J)*(H-U))
03720 IF N5=20 THEN 4120
03730 IF G1(I)=0 THEN 3780
03740 IF D1(I,J)=0 THEN 3780
03750 LET T4(I,J)=D1(I,J)/G1(I)
03760 LET U1=LOGE(D1(I,J)/G1(I))
03770 GOTO 3790
03780 LET U1=0
03790 LET Z1=Z1+(D1(I,J)*(H-U1))
03800 IF N5=21 THEN 4120
03810 IF G2(I)=0 THEN 3860
03820 IF D2(I,J)=0 THEN 3860
03830 LET T5(I,J)=D2(I,J)/G2(I)
03840 LET U2=LOGE(D2(I,J)/G2(I))
03850 GOTO 3870
03860 LET U2=0
03870 LET Z2=Z2+(D2(I,J)*(H-U2))
03880 IF N=22 THEN 4120
03890 IF G3(I)=0 THEN 3940
03900 IF D3(I,J)=0 THEN 3940
03910 LET T6(I,J)=D3(I,J)/G3(I)
03920 LET U3=LOGE(D3(I,J)/G3(I))
03930 GOTO 3950
03940 LET U3=0
03950 LET Z3=Z3+(D3(I,J)*(H-U3))
03960 IF N=23 THEN 4120
```

```
03970 IF G4(I)=0 THEN 4020
03980 IF D4(I,J)=0 THEN 4020
03990 LET T7(I,J)=D4(I,J)/G4(I)
04000 LET U4=LOGE(D4(I,J)/G4(I))
04010 GOTO 4030
04020 LET U4=0
04030 LET Z4=Z4+(D4(I,J)*(H-U4))
04040 IF N=24 THEN 4120
04050 IF G5(I)=0 THEN 4100
04060 IF D5(I,J)=0 THEN 4100
04070 LET T8(I,J)=D5(I,J)/G5(I)
04080 LET U5=LOGE(D5(I,J)/G5(I))
04090 GOTO 4110
04100 LET U5=0
04110 LET Z5=Z5+(D5(I,J)*(H-U5))
04120 NEXT J
04130 NEXT I
04140 LET A5=(V1+V2+V3+V4+V5+V6+V7+V8+V9+W+W1+W2)
04150 LET A6=(W3+W4+W5+W6+W7+W8+W9+Z+Z1+Z2+Z3+Z4+Z5)
04160 LET Z9= -2*(A5+A6)
04170 FOR I=1 TO K1
04180 FOR J=1 TO K1
04190 LET Z6(I,J)=S1(I,J)
04200 NEXT J
04210 NEXT I
04220 REM PRINT OUTPUT
04230 PRINT #2
04240 PRINT #2
04250 PRINT #2
04260 PRINT #2,DATE$
04270 PRINT #2
04280 PRINT #2,R$
04290 PRINT #2
04300 IF K=1 THEN PRINT #2,"COMPOSITE MATRIX (OVERALL)"
04310 IF K=2 THEN PRINT #2,"COMPOSITE MATRIX (WITHIN)"
04320 IF K=3 THEN PRINT #2,"COMPOSITE MATRIX (BETWEEN)"
04330 PRINT #2
04340 REM COLUMN INDICATOR
04350 FOR J=1 TO K1
04360 PRINT #2,TAB(10*J);J-1;
04370 NEXT J
04380 PRINT #2
04390 PRINT #2
04400 REM PRINT MATRIX
04410 FOR I=1 TO K1
04420 PRINT #2
04430 PRINT #2,I-1;
04440 FOR J=1 TO K1
04450 PRINT #2,TAB(10*J);P(I,J);
04460 NEXT J
04470 PRINT #2,TAB(10+(10*J));P1(I)
04480 FOR J3=1 TO K1
04490 PRINT #2,TAB(10*J3);P(I,J3)/N5;
04500 NEXT J3
04510 PRINT #2
04520 FOR J6=1 TO K1
04530 PRINT #2,TAB(10*J6);Z6(I,J6);
```

```
04540 NEXT J6
04550 PRINT #2
04560 NEXT I
04570 REM PRINT INDIVIDUAL MATRICES
04580 REM INDIVIDUAL MATRIX TO PRINT MATRIX
04590 FOR I=1 TO N5
04600 FOR J=1 TO K1
04610 FOR J1=1 TO K1
04620 ON I GOTO 4630,4660,4690,4720,4750,4780
                4810,4840,4870,4900,4930,4960
                4990,5020,5050,5080,5110,5140
                5170,5200,5230,5260,5290,5320,5350
04630 LET Z7(J,J1)=B1(J,J1)
04640 LET Z6(J,J1)=S2(J,J1)
04650 GOTO 5370
04660 LET Z7(J,J1)=B2(J,J1)
04670 LET Z6(J,J1)=S3(J,J1)
04680 GOTO 5370
04690 LET Z7(J,J1)=B3(J,J1)
04700 LET Z6(J,J1)=S4(J,J1)
04710 GOTO 5370
04720 LET Z7(J,J1)=B4(J,J1)
04730 LET Z6(J,J1)=S5(J,J1)
04740 GOTO 5370
04750 LET Z7(J,J1)=B5(J,J1)
04760 LET Z6(J,J1)=S6(J,J1)
04770 GOTO 5370
04780 LET Z7(J,J1)=B6(J,J1)
04790 LET Z6(J,J1)=S7(J,J1)
04800 GOTO 5370
04810 LET Z7(J,J1)=B7(J,J1)
04820 LET Z6(J,J1)=S8(J,J1)
04830 GOTO 5370
04840 LET Z7(J,J1)=B8(J,J1)
04850 LET Z6(J,J1)=S9(J,J1)
04860 GOTO 5370
04870 LET Z7(J,J1)=B9(J,J1)
04880 LET Z6(J,J1)=N1(J,J1)
04890 GOTO 5370
04900 LET Z7(J,J1)=C(J,J1)
04910 LET Z6(J,J1)=N2(J,J1)
04920 GOTO 5370
04930 LET Z7(J,J1)=C1(J,J1)
04940 LET Z6(J,J1)=N3(J,J1)
04950 GOTO 5370
04960 LET Z7(J,J1)=C2(J,J1)
04970 LET Z6(J,J1)=N4(J,J1)
04980 GOTO 5370
04990 LET Z7(J,J1)=C3(J,J1)
05000 LET Z6(J,J1)=N5(J,J1)
05010 GOTO 5370
05020 LET Z7(J,J1)=C4(J,J1)
05030 LET Z6(J,J1)=N6(J,J1)
05040 GOTO 5370
05050 LET Z7(J,J1)=C5(J,J1)
05060 LET Z6(J,J1)=N7(J,J1)
05070 GOTO 5370
```

```
05080 LET Z7(J,J1)=C6(J,J1)
05090 LET Z6(J,J1)=N8(J,J1)
05100 GOTO 5370
05110 LET Z7(J,J1)=C7(J,J1)
05120 LET Z6(J,J1)=N9(J,J1)
05130 GOTO 5370
05140 LET Z7(J,J1)=C8(J,J1)
05150 LET Z6(J,J1)=T1(J,J1)
05160 GOTO 5370
05170 LET Z7(J,J1)=C9(J,J1)
05180 LET Z6(J,J1)=T2(J,J1)
05190 GOTO 5370
05200 LET Z7(J,J1)=D(J,J1)
05210 LET Z6(J,J1)=T3(J,J1)
05220 GOTO 5370
05230 LET Z7(J,J1)=D1(J,J1)
05240 LET Z6(J,J1)=T4(J,J1)
05250 GOTO 5370
05260 LET Z7(J,J1)=D2(J,J1)
05270 LET Z6(J,J1)=T5(J,J1)
05280 GOTO 5370
05290 LET Z7(J,J1)=D3(J,J1)
05300 LET Z6(J,J1)=T6(J,J1)
05310 GOTO 5370
05320 LET Z7(J,J1)=D4(J,J1)
05330 LET Z6(J,J1)=T7(J,J1)
05340 GOTO 5370
05350 LET Z7(J,J1)=D5(J,J1)
05360 LET Z6(J,J1)=T8(J,J1)
05370 NEXT J1
05380 ON I GOTO 5390,5410,5430,5450,5470,5490
                5510,5530,5550,5570,5590,5610
                5630,5650,5670,5690,5710,5730
                5750,5770,5790,5810,5830,5850,5870
05390 LET Z8(J)=E1(J)
05400 GOTO 5880
05410 LET Z8(J)=E2(J)
05420 GOTO 5880
05430LET Z8(J)=E3(J)
05440 GOTO 5880
05450 LET Z8(J)=E4(J)
05460 GOTO 5880
05470 LET Z8(J)=E5(J)
05480 GOTO 5880
05490 LET Z8(J)=E6(J)
05500 GOTO 5880
05510 LET Z8(J)=E7(J)
05520 GOTO 5880
05530 LET Z8(J)=E8(J)
05540 GOTO 5880
05550 LET Z8(J)=E9(J)
05560 GOTO 5880
05570 LET Z8(J)=F(J)
05580 GOTO 5880
05590 LET Z8(J)=F1(J)
05600 GOTO 5880
05610 LET Z8(J)=F2(J)
```

```
05620 GOTO 5880
05630 LET Z8(J)=F3(J)
05640 GOTO 5880
05650 LET Z8(J)=F4(J)
05660 GOTO 5880
05670 LET Z8(J)=F5(J)
05680 GOTO 5880
05690 LET Z8(J)=F6(J)
05700 GOTO 5880
05710 LET Z8(J)=F7(J)
05720 GOTO 5880
05730 LET Z8(J)=F8(J)
05740 GOTO 5880
05750 LET Z8(J)=F9(J)
05760 GOTO 5880
05770 LET Z8(J)=G(J)
05780 GOTO 5880
05790 LET Z8(J)=G1(J)
05800 GOTO 5880
05810 LET Z8(J)=G2(J)
05820 GOTO 5880
05830 LET Z8(J)=G3(J)
05840 GOTO 5880
05850 LET Z8(J)=G4(J)
05860 GOTO 5880
05870 LET Z8(J)=G5(J)
05880 NEXT J
05890 REM PRINT ROUTINE
05900 PRINT #2
05910 PRINT #2
05920 IF K=1 THEN PRINT #2,"SEQUENCE";I;"(OVERALL)"
05930 IF K=2 THEN PRINT #2,"SEQUENCE";I;"(WITHIN)"
05940 IF K=3 THEN PRINT #2,"SEQUENCE";I;"(BETWEEN)"
05950 PRINT #2
05960 PRINT #2,"SOURCE FILE=";U$(I)
05970 PRINT #2
05980 REM COLUMN INDICATOR
05990 PRINT #2
06000 FOR J5=1 TO K1
06010 PRINT #2,TAB(10*J5);J5-1;
06020 NEXT J5
06030 PRINT #2
06040 PRINT #2
06050 REM PRINT MATRIX
06060 FOR I5=1 TO K1
06070 PRINT #2
06080 PRINT #2,I5-1;
06090 FOR J5=1 TO K1
06100 PRINT #2,TAB(10*J5);Z7(I5,J5);
06110 NEXT J5
06120 PRINT #2,TAB(10+(10*J5));Z8(I5)
06130 FOR J6=1 TO K1
06140 PRINT #2,TAB(10*J6);Z6(I5,J6);
06150 NEXT J6
06160 PRINT #2
06170 NEXT I5
06180 NEXT I
```

```
06190 PRINT #2
06200 PRINT #2
06210 PRINT #2,"SAMPLE HOMOGENEITY, USING THE"
06220 PRINT #2,"ANDERSON-GOODMAN L.R. TEST"
06230 PRINT #2
06240 PRINT #2,"L.R. HAS CHI-SQUARE DISTRIBUTION"
06250 PRINT #2
06260 LET K8=(K1*(K1-1))*(N5-1)
06270 PRINT #2,"LIKELIHOOD RATIO=";Z9;"WITH";K8;"DF"
06280 GOTO 610
06290 STOP
06300 REM RAW DATA CONVERSION ROUTINE
06310 LET S(M)=N(M)=T(M)=0
06320 IF A<100 THEN 6360
06330 LET S(M)=S(M)+1
06340 LET A=A-100
06350 GOTO 6320
06360 IF A<10 THEN 6400
06370 LET N(M)=N(M)+1
06380 LET A=A-10
06390 GOTO 6360
06400 LET T(M)=A
06410 RETURN
06420 END
```

```
00010 REM PROGRAM 7     N-STEP CONTINGENCIES
00020 REM
00030 REM LAGS 0-10 (11 MATRICES) PRINTED ON PAPER
00040 REM LAGS 1-10 (10 MATRICES) TO DATA FILES
00050 REM LAGS TO DATA FILES ARE FREQUENCIES
00060 REM A/B= INPUT/LAG DATA
00070 REM C= SINGLE RUN INCREMENT MATRIX
00080 REM C1= SINGLE RUN FREQUENCY MATRIX
00090 REM D-E1 = LAG 0-10 SPECIFIC MATRICES
00100 REM G= OUTPUT MATRIX COUNTER AND INDICATOR
00110 REM K= LAG DATA LOOP
00120 REM K5/K6= NO. OF CODES/MATRIX SIZE
00130 REM L= LAG (N-STEP CONTINGENCY)
00140 REM M= OBSERVED/GENERATED DATA INDICATOR
00150 REM N/R= SEQUENCE LENGTH/ROW TOTALS
00160 REM Z1/N1/T1= INPUT SPEAKER/ACTIVITY/TYPE CODES
00170 REM Z2/N2/T2= LAG SPEAKER/ACTIVITY/TYPE CODES
00180 REM S-T1= LAG 0-10 FREQUENCY MATRICES
00190 REM X/U$= OBSERVED-GENERATED LOOP/MARKER
00200 REM R$ = TITLE
00210 REM V$/W$= INPUT/OUTPUT FILENAMES
00220 DIM A(700),B(700),C(20,20),D(20,20),D1(20,20)
00230 DIM D2(20,20),D3(20,20),D4(20,20),D5(20,20)
00240 DIM D6(20,20),D7(20,20),D8(20,20),D9(20,20)
00250 DIM R(20),S(20,20),S1(20,20),S2(20,20)
00260 DIM S3(20,20),S4(20,20),S5(20,20),S6(20,20)
00270 DIM S7(20,20),S8(20,20),S9(20,20),T1(20,20)
00280 DIM E1(20,20),C1(20,20)
00290 REM SET STORES =0
00300 FOR I=1 TO 700
00310 LET A(I)=B(I)=0
00320 NEXT I
00330 FOR I=1 TO 10
00340 LET R(I)=0
00350 NEXT I
00360 FOR I=1 TO 20
00370 FOR J=1 TO 20
00380 LET C(I,J)=D(I,J)=D1(I,J)=D2(I,J)=D3(I,J)=0
00390 LET D4(I,J)=D5(I,J)=D6(I,J)=D7(I,J)=D8(I,J)=0
00400 LET D9(I,J)=E1(I,J)=C1(I,J)=T1(I,J)=0
00410 LET S(I,J)=S1(I,J)=S2(I,J)=S3(I,J)=S4(I,J)=0
00420 LET S5(I,J)=S6(I,J)=S7(I,J)=S8(I,J)=S9(I,J)=0
00430 NEXT J
00440 NEXT I
00450 REM OPEN INPUT AND OUTPUT FILES
00460 PRINT "INPUT FILENAME"  ;
00470 INPUT V$
00480 FILE #1,V$
00490 PRINT
00500 PRINT "OUTPUT FILE NAME"  ;
00510 INPUT W$
00520 FILE #2,W$
00530 MARGIN #2,132
00540 SCRATCH #2
00550 REM OBSERVED OR GENERATED DATA?
00560 PRINT
00570 PRINT "DOES FILE HOLD OBS. DATA (Y/N)"  ;
```

```
00580 INPUT U$
00590 IF U$="Y" THEN 610
00600 IF U$="N" THEN 630
00610 LET M=1
00620 GOTO 600
00630 LET M=100
00640 PRINT
00650 REM MATRIX SIZE
00660 PRINT "NUMBER OF CODES USED"
00670 PRINT
00680 PRINT "IF 5 CODES ARE USED, CODES 0-4 ASSUMED"
00690 PRINT
00700 PRINT "NUMBER OF CODES USED" ;
00710 INPUT K5
00720 LET K6=K5+K5
00730 PRINT
00740 PRINT "SEQUENCE LENGTH" ;
00750 INPUT N
00760 PRINT
00770 PRINT "TITLE" ;
00780 INPUT R$
00790 REM INPUT RAW DATA
00800 FOR X=1 TO M
00810 FOR J=1 TO N
00820 INPUT #1,A(J)
00830 NEXT J
00840 REM CREATE LAG DATA ARRAY
00850 LET L=0
00860 FOR K=1 TO 11
00870 LET L=L+1
00880 FOR I=1 TO 700
00890 LET B(I)=0
00900 NEXT I
00910 FOR I=1 TO (N-L)
00920 LET B(I)=A(I+L)
00930 NEXT I
00940 FOR I=1 TO 20
00950 FOR J=1 TO 20
00960 LET C(I,J)=0
00970 NEXT J
00980 NEXT I
00990 REM INCREMENT OUTPUT MATRIX
01000 FOR I=1 TO (N-L)
01010 GOSUB 4860
01020 LET Y=N1+1
01030 GOSUB 5000
01040 LET Z=N2+1
01050 IF Z1=1 THEN 1070
01060 LET Y=Y+K5
01070 IF Z2=1 THEN 1090
01080 LET Z=Z+K5
01090 LET C(Y,Z)=C(Y,Z)+1
01100 NEXT I
01110 FOR I=1 TO K6
01120 REM CALCULATE ROW TOTALS
01130 LET R(I)=0
01140 NEXT I
```

```
01150 FOR I=1 TO K6
01160 FOR J=1 TO K6
01170 LET R(I)=R(I)+C(I,J)
01180 NEXT J
01190 NEXT I
01200 REM CALCULATE MATRIX PROBABILITIES
01210 FOR I=1 TO K6
01220 FOR J=1 TO K6
01230 IF R(I)=0 THEN 1270
01240 LET C1(I,J)=C(I,J)
01250 LET C(I,J)=C(I,J)/R(I)
01260 NEXT J
01270 NEXT I
01280 REM INCREMENT LAG SPECIFIC MATRICES
01290 IF L>1 THEN 1370
01300 FOR I=1 TO K6
01310 FOR J=1 TO K6
01320 LET D(I,J)=D(I,J)+C(I,J)
01330 LET S(I,J)=S(I,J)+C1(I,J)
01340 NEXT J
01350 NEXT I
01360 GOTO 2150
01370 IF L>2 THEN 1450
01380 FOR I=1 TO K6
01390 FOR J=1 TO K6
01400 LET D1(I,J)=D1(I,J)+C(I,J)
01410 LET S1(I,J)=S1(I,J)+C1(I,J)
01420 NEXT J
01430 NEXT I
01440 GOTO 2150
01450 IF L>3 THEN 1530
01460 FOR I=1 TO K6
01470 FOR J=1 TO K6
01480 LET D2(I,J)=D2(I,J)+C(I,J)
01490 LET S2(I,J)=S2(I,J)+C1(I,J)
01500 NEXT J
01510 NEXT I
01520 GOTO 2150
01530 IF L>4 THEN 1610
01540 FOR I=1 TO K6
01550 FOR J=1 TO K6
01560 LET D3(I,J)=D3(I,J)+C(I,J)
01570 LET S3(I,J)=S3(I,J)+C1(I,J)
01580 NEXT J
01590 NEXT I
01600 GOTO 2150
01610 IF L>5 THEN 1690
01620 FOR I=1 TO K6
01630 FOR J=1 TO K6
01640 LET D4(I,J)=D4(I,J)+C(I,J)
01650 LET S4(I,J)=S4(I,J)+C1(I,J)
01660 NEXT J
01670 NEXT I
01680 GOTO 2150
01690 IF L>6 THEN 1770
01700 FOR I=1 TO K6
01710 FOR J=1 TO K6
```

```
01720 LET D5(I,J)=D5(I,J)+C(I,J)
01730 LET S5(I,J)=S5(I,J)+C1(I,J)
01740 NEXT J
01750 NEXT I
01760 GOTO 2150
01770 IF L>7 THEN 1850
01780 FOR I=1 TO K6
01790 FOR J=1 TO K6
01800 LET D6(I,J)=D6(I,J)+C(I,J)
01810 LET S6(I,J)=S6(I,J)+C1(I,J)
01820 NEXT J
01830 NEXT I
01840 GOTO 2150
01850 IF L>8 THEN 1930
01860 FOR I=1 TO K6
01870 FOR J=1 TO K6
01880 LET D7(I,J)=D7(I,J)+C(I,J)
01890 LET S7(I,J)=S7(I,J)+C1(I,J)
01900 NEXT J
01910 NEXT I
01920 GOTO 2150
01930 IF L>9 THEN 2010
01940 FOR I=1 TO K6
01950 FOR J=1 TO K6
01960 LET D8(I,J)=D8(I,J)+C(I,J)
01970 LET S8(I,J)=S8(I,J)+C1(I,J)
01980 NEXT J
01990 NEXT I
02000 GOTO 2150
02010 IF L>10 THEN 2090
02020 FOR I=1 TO K6
02030 FOR J=1 TO K6
02040 LET D9(I,J)=D9(I,J)+C(I,J)
02050 LET S9(I,J)=S9(I,J)+C1(I,J)
02060 NEXT J
02070 NEXT I
02080 GOTO 2150
02090 FOR I=1 TO K6
02100 FOR J=1 TO K6
02110 LET E1(I,J)=E1(I,J)+C(I,J)
02120 LET T1(I,J)=T1(I,J)+C1(I,J)
02130 NEXT J
02140 NEXT I
02150 NEXT K
02160 NEXT X
02170 REM OUTPUT RECORD
02180 PRINT #2,DATE$
02190 PRINT #2
02200 PRINT #2,"SOURCE FILE= ";V$
02210 PRINT #2
02220 PRINT #2,R$
02230 PRINT #2
02240 IF U$="Y" THEN 2270
02250 PRINT #2,"GENERATED SEQUENCES"
02260 GOTO 2280
02270 PRINT #2,"OBSERVED SEQUENCES"
02280 PRINT #2
```

```
02290 PRINT #2,"MEAN FREQUENCIES/ROW PROBABILITIES"
02300 PRINT #2
02310 FOR G=1 TO 11
02320 PRINT #2
02330 PRINT #2
02340 PRINT #2
02350 PRINT #2
02360 PRINT #2,"LAG= ";G-1
02370 PRINT #2
02380 REM COLUMN INDICATOR
02390 FOR J=1 TO K6
02400 IF J>K5 THEN 2430
02410 PRINT #2,TAB(6*J);J-1;
02420 GOTO 2440
02430 PRINT #2,TAB(6*J);J-(K5+1);
02440 NEXT J
02450 PRINT #2
02460 PRINT #2
02470 REM PRINT MATRICES (2 DECIMAL PLACES)
02480 IF G>1 THEN 2650
02490 FOR I=1 TO K6
02500 PRINT #2
02510 IF I>K5 THEN 2540
02520 PRINT #2,I-1;
02530 GOTO 2550
02540 PRINT #2,I-(K5+1);
02550 FOR J=1 TO K6
02560 PRINT #2,TAB(6*J);(INT((D(I,J)/M)*100))/100;
02570 NEXT J
02580 PRINT #2
02590 FOR J=1 TO K6
02600 PRINT #2,TAB(6*J);S(I,J)/M;
02610 NEXT J
02620 PRINT #2
02630 NEXT I
02640 GOTO 4330
02650 IF G>2 THEN 2820
02660 FOR I=1 TO K6
02670 PRINT #2
02680 IF I>K5 THEN 2710
02690 PRINT #2,I-1;
02700 GOTO 2720
02710 PRINT #2,I-(K5+1);
02720 FOR J=1 TO K6
02730 PRINT #2,TAB(6*J);(INT((D1(I,J)/M)*100))/100;
02740 NEXT J
02750 PRINT #2
02760 FOR J=1 TO K6
02770 PRINT #2,TAB(6*J);S1(I,J)/M;
02780 NEXT J
02790 PRINT #2
02800 NEXT I
02810 GOTO 4330
02820 IF G>3 THEN 2990
02830 FOR I=1 TO K6
02840 PRINT #2
02850 IF I>K5 THEN 2880
```

```
02860 PRINT #2,I-1;
02870 GOTO 2890
02880 PRINT #2,I-(K5+1);
02890 FOR J=1 TO K6
02900 PRINT #2,TAB(6*J);(INT((D2(I,J)/M)*100))/100;
02910 NEXT J
02920 PRINT #2
02930 FOR J=1 TO K6
02940 PRINT #2,TAB(6*J);S2(I,J)/M;
02950 NEXT J
02960 PRINT #2
02970 NEXT I
02980 GOTO 4330
02990 IF G>4 THEN 3160
03000 FOR I=1 TO K6
03010 PRINT #2
03020 IF I>K5 THEN 3050
03030 PRINT #2,I-1;
03040 GOTO 3060
03050 PRINT #2,I-(K5+1);
03060 FOR J=1 TO K6
03070 PRINT #2,TAB(6*J);(INT((D3(I,J)/M)*100))/100;
03080 NEXT J
03090 PRINT #2
03100 FOR J=1 TO K6
03110 PRINT #2,TAB(6*J);S3(I,J)/M;
03120 NEXT J
03130 PRINT #2
03140 NEXT I
03150 GOTO 4330
03160 IF G>5 THEN 3330
03170 FOR I=1 TO K6
03180 PRINT #2
03190 IF I>K5 THEN 3220
03200 PRINT #2,I-1;
03210 GOTO 3230
03220 PRINT #2,I-(K5+1);
03230 FOR J=1 TO K6
03240 PRINT #2,TAB(6*J);(INT((D4(I,J)/M)*100))/100;
03250 NEXT J
03260 PRINT #2
03270 FOR J=1 TO K6
03280 PRINT #2,TAB(6*J)S4(I,J)/M;
03290 NEXT J
03300 PRINT #2
03310 NEXT I
03320 GOTO 4330
03330 IF G>6 THEN 3500
03340 FOR I=1 TO K6
03350 PRINT #2
03360 IF I>K5 THEN 3390
03370 PRINT #2,I-1;
03380 GOTO 3400
03390 PRINT #2,I-(K5+1);
03400 FOR J=1 TO K6
03410 PRINT #2,TAB(6*J);(INT((D5(I,J)/M)*100))/100;
03420 NEXT J
```

```
03430 PRINT #2
03440 FOR J=1 TO K6
03450 PRINT #2,TAB(6*J);S5(I,J)/M;
03460 NEXT J
03470 PRINT #2
03480 NEXT I
03490 GOTO 4330
03500 IF G>7 THEN 3670
03510 FOR I=1 TO K6
03520 PRINT #2
03530 IF I>K5 THEN 3560
03540 PRINT #2,I-1;
03550 GOTO 3570
03560 PRINT #2,I-(K5+1);
03570 FOR J=1 TO K6
03580 PRINT #2,TAB(6*J);(INT((D6(I,J)/M)*100))/100;
03590 NEXT J
03600 PRINT #2
03610 FOR J=1 TO K6
03620 PRINT #2,TAB(6*J);S6(I,J)/M;
03630 NEXT J
03640 PRINT #2
03650 NEXT I
03660 GOTO 4330
03670 IF G>8 THEN 3840
03680 FOR I=1 TO K6
03690 PRINT #2
03700 IF I>K5 THEN 3730
03710 PRINT #2,I-1;
03720 GOTO 3740
03730 PRINT #2,I-(K5+1);
03740 FOR J=1 TO K6
03750 PRINT #2,TAB(6*J);(INT((D7(I,J)/M)*100))/100;
03760 NEXT J
03770 PRINT #2
03780 FOR J=1 TO K6
03790 PRINT #2,TAB(6*J);S7(I,J)/M;
03800 NEXT J
03810 PRINT #2
03820 NEXT I
03830 GOTO 4330
03840 IF G>9 THEN 4010
03850 FOR I=1 TO K6
03860 PRINT #2
03870 IF I>K5 THEN 3900
03880 PRINT #2,I-1;
03890 GOTO 3910
03900 PRINT #2,I-(K5+1);
03910 FOR J=1 TO K6
03920 PRINT #2,TAB(6*J);(INT((D8(I,J)/M)*100))/100;
03930 NEXT J
03940 PRINT #2
03950 FOR J=1 TO K6
03960 PRINT #2,TAB(6*J);S8(I,J)/M;
03970 NEXT J
03980 PRINT #2
03990 NEXT I
```

```
04000 GOTO 4330
04010 IF G>10 THEN 4180
04020 FOR I=1 TO K6
04030 PRINT #2
04040 IF I>K5 THEN 4070
04050 PRINT #2,I-1;
04060 GOTO 4080
04070 PRINT #2,I-(K5+1);
04080 FOR J=1 TO K6
04090 PRINT #2,TAB(6*J);(INT((D9(I,J)/M)*100))/100;
04100 NEXT J
04110 PRINT #2
04120 FOR J=1 TO K6
04130 PRINT #2,TAB(6*J);S9(I,J)/M;
04140 NEXT J
04150 PRINT #2
04160 NEXT I
04170 GOTO 4330
04180 FOR I=1 TO K6
04190 PRINT #2
04200 IF I>K5 THEN 4230
04210 PRINT #2,I-1;
04220 GOTO 4240
04230 PRINT #2,I-(K5+1);
04240 FOR J=1 TO K6
04250 PRINT #2,TAB(6*J);(INT((E1(I,J)/M)*100))/100;
04260 NEXT J
04270 PRINT #2
04280 FOR J=1 TO K6
04290 PRINT #2,TAB(6*J);T1(I,J)/M;
04300 NEXT J
04310 PRINT #2
04320 NEXT I
04330 NEXT G
04340 LET F$="STEP1.DAT"
04350 LET F1$="STEP2.DAT"
04360 LET F2$="STEP3.DAT"
04370 LET F3$="STEP4.DAT"
04380 LET F4$="STEP5.DAT"
04390 LET F5$="STEP6.DAT"
04400 LET F6$="STEP7.DAT"
04410 LET F7$="STEP8.DAT"
04420 LET F8$="STEP9.DAT"
04430 LET F9$="STEP10.DAT"
04440 LET H$="GSTEP1.DAT"
04450 LET H1$="GSTEP2.DAT"
04460 LET H2$="GSTEP3.DAT"
04470 LET H3$="GSTEP4.DAT"
04480 LET H4$="GSTEP5.DAT"
04490 LET H5$="GSTEP6.DAT"
04500 LET H6$="GSTEP7.DAT"
04510 LET H7$="GSTEP8.DAT"
04520 LET H8$="GSTEP9.DAT"
04530 LET H9$="GSTP10.DAT"
04540 IF M>1 THEN 4790
04550 FILE #5,F$;#6,F1$;#7,F2$;#8,F3$;#9,F4$
04560 SCRATCH #5,6,7,8,9
```

```
04570 FOR I=1 TO K6
04580 FOR J=1 TO K6
04590 PRINT #5,S1(I,J)/M
04600 PRINT #6,S2(I,J)/M
04610 PRINT #7,S3(I,J)/M
04620 PRINT #8,S4(I,J)/M
04630 PRINT #9,S5(I,J)/M
04640 NEXT J
04650 NEXT I
04660 IF M>1 THEN 4820
04670 FILE #5,F5$;#6,F6$;#7,F7$;#8,F8$;#9,F9$
04680 SCRATCH #5,6,7,8,9
04690 FOR I=1 TO K6
04700 FOR J=1 TO K6
04710 PRINT #5,S6(I,J)/M
04720 PRINT #6,S7(I,J)/M
04730 PRINT #7,S8(I,J)/M
04740 PRINT #8,S9(I,J)/M
04750 PRINT #9,T1(I,J)/M
04760 NEXT J
04770 NEXT I
04780 GOTO 4850
04790 FILE #5,H$;#6,H1$;#7,H2$;#8,H3$;#9,H4$
04800 SCRATCH #5,6,7,8,9
04810 GOTO 4570
04820 FILE #5,H5$;#6,H6$;#7,H7$;#8,H8$;#9,H9$
04830 SCRATCH #5,6,7,8,9
04840 GOTO 4690
04850 STOP
04860 REM INPUT DATA CONVERSION ROUTINE
04870 LET Z1=N1=T1=0
04880 LET A1=A(I)
04890 IF A1<100 THEN 4930
04900 LET Z1=Z1+1
04910 LET A1=A1-100
04920 GOTO 4890
04930 IF A1<10 THEN 4970
04940 LET N1=N1+1
04950 LET A1=A1-10
04960 GOTO 4930
04970 LET T1=A1
04980 RETURN
04990 STOP
05000 REM LAG DATA CONVERSION ROUTINE
05010 LET Z2=N2=T2=0
05020 LET B1=B(I)
05030 IF B1<100 THEN 5070
05040 LET Z2=Z2+1
05050 LET B1=B1-100
05060 GOTO 5030
05070 IF B1<10 THEN 5110
05080 LET N2=N2+1
05090 LET B1=B1-10
05100 GOTO 5070
05110 LET T2=B1
05120 RETURN
05130 STOP
```

05140 END

```
00010 REM PROGRAM 8      MATRIX MULTIPLICATION
00020 REM
00030 REM A/C= RAW DATA/COUNTER
00040 REM E/F/G/H= FREQ/ORIG/POWER/NTH-POWER MATRICES
00050 REM E1/2,F1/2,G1/2,H1/2= ROW/COL INDICATORS
00060 REM M5/N1= MATRIX SIZE/SEQUENCE LENGTH
00070 REM S/N/T= SPEAKER/ACT/TYPE CODES
00080 REM R$= TITLE
00090 REM U$/V$/F$= IN/OUT/DATA FILE NAMES
00100 DIM E(20,20),F(20,20),G(20,20),H(20,20),F$(10)
00110 MAT E=ZER
00120 MAT F=ZER
00130 MAT G=ZER
00140 MAT H=ZER
00150 REM OPEN INPUT AND OUTPUT FILES
00160 PRINT "INPUT FILENAME"  ;
00170 INPUT U$
00180 PRINT
00190 PRINT "OUTPUT FILENAME"  ;
00200 INPUT V$
00210 FILE #1,U$;#2,V$
00220 MARGIN #2,132
00230 SCRATCH #2
00240 FOR I=1 TO 10
00250 READ L$
00260 LET F$(I)=L$
00270 NEXT I
00280 DATA R1.DAT,R2.DAT,R3.DAT,R4.DAT,R5.DAT,
          R6.DAT,R7.DAT,R8.DAT,R9.DAT,R10.DAT
00290 REM PROGRAM SPECIFICATIONS
00300 PRINT
00310 PRINT "MATRIX SIZE (CODES X 2) ="  ;
00320 INPUT M5
00330 PRINT
00340 PRINT "SEQUENCE LENGTH ="  ;
00350 INPUT N1
00360 PRINT
00370 PRINT "TITLE ="  ;
00380 INPUT R$
00390 REM READ DATA
00400 LET L=1
00410 INPUT #1,A
00420 LET C=1
00430 LET M=1
00440 GOSUB 1550
00450 LET C=C+1
00460 IF C>N1 THEN 640
00470 INPUT #1,A
00480 LET M=L+1
00490 GOSUB 1550
00500 REM ADD TO FREQUENCY MATRIX
00510 LET X=N(1)+1
00520 LET Y=N(2)+1
00530 IF S(1)=1 THEN 550
00540 LET X=X+(M5/2)
00550 IF S(2)=1 THEN 570
00560 LET Y=Y+(M5/2)
00570 LET E(X,Y)=E(X,Y)+1
```

```
00580 REM SHIFT INPUT STACK
00590 LET  Q=1
00600 LET S(Q)=S(Q+1)
00610 LET N(Q)=N(Q+1)
00620 GOTO 450
00630 REM ROW TOTALS/INITIAL PROB. MATRIX
00640 FOR I=1 TO M5
00650 LET R=0
00660 FOR J=1 TO M5
00670 LET R=R+E(I,J)
00680 NEXT J
00690 FOR K=1 TO M5
00700 IF R=0 THEN 730
00710 LET F(I,K)=E(I,K)/R
00720 GOTO 740
00730 LET F(I,K)=0
00740 NEXT K
00750 NEXT I
00760 LET P=0
00770 MAT H=F
00780 GOTO 840
00790 REM SET INITIAL=POWER MATRIX
00800 MAT G=F
00810 REM RAISE MATRIX TO N-TH POWER
00820 FOR P=1 TO 10
00830 MAT H=F*G
00840 REM PRINT OUTPUT
00850 PRINT #2
00860 PRINT #2
00870 PRINT #2
00880 PRINT #2,DATE$,TIME$
00890 PRINT #2
00900 PRINT #2,"SOURCE FILE =";U$
00910 PRINT #2
00920 PRINT #2,R$
00930 PRINT #2
00940 PRINT #2,"POWER =",P
00950 PRINT #2
00960 PRINT #2
00970 PRINT #2,"SPEAKER 1"
00980 PRINT #2
00990 PRINT #2
01000 REM COLUMN INDICATOR
01010 FOR J=1 TO (M5/2)
01020 PRINT #2,TAB(10*J);J-1;
01030 NEXT J
01040 PRINT #2
01050 PRINT #2
01060 FOR I=1 TO M5
01070 PRINT #2
01080 IF I>(M5/2) THEN 1110
01090 PRINT #2,I-1;
01100 GOTO 1120
01110 PRINT #2,I-((M5/2)+1);
01120 FOR J=1 TO (M5/2)
01130 PRINT #2,TAB(10*J);H(I,J);
01140 NEXT J
```

```
01150 PRINT #2
01160 PRINT #2
01170 NEXT I
01180 PRINT #2
01190 PRINT #2
01200 PRINT #2
01210 PRINT #2,"SPEAKER 2"
01220 PRINT #2
01230 PRINT #2
01240 REM COLUMN INDICATOR
01250 FOR J=(M5/2)+1 TO M5
01260 PRINT #2,TAB(10*(J-(M5/2)));J-((M5/2)+1);
01270 NEXT J
01280 PRINT #2
01290 PRINT #2
01300 FOR I=1 TO M5
01310 PRINT #2
01320 IF I>(M5/2) THEN 1350
01330 PRINT #2,I-1;
01340 GOTO 1360
01350 PRINT #2,I-((M5/2)+1);
01360 FOR J=((M5/2)+1) TO M5
01370 PRINT #2,TAB(10*(J-(M5/2)));H(I,J);
01380 NEXT J
01390 PRINT #2
01400 PRINT #2
01410 NEXT I
01420 IF P=0 THEN 800
01430 REM PRINT DISK FILES
01440 FILE #3,F$(P)
01450 SCRATCH #3
01460 FOR I=1 TO M5
01470 FOR J=1 TO M5
01480 PRINT #3,H(I,J)
01490 NEXT J
01500 NEXT I
01510 REM SET G MATRIX= H MATRIX
01520 MAT G=H
01530 NEXT P
01540 STOP
01550 DATA CONVERSION ROUTINE
01560 LET S(M)=N(M)=T(M)=0
01570 IF A<100 THEN 1610
01580 LET S(M)=S(M)+1
01590 LET A=A-100
01600 GOTO 1570
01610 IF A<10 THEN 1650
01620 LET N(M)=N(M)+1
01630 LET A=A-10
01640 GOTO 1610
01650 LET T(M)=A
01660 RETURN
01670 END
```

```
00010 REM PROGRAM 9     MATRIX COMPARATOR
00020 REM
00030 REM USING THE ANDERSON-GOODMAN HOMOGENEITY TEST
00040 REM OBSERVED/PREDICTED/TOTAL DATA MATRIX
00050 REM A1/B1/C1= ROW TOTALS
00060 REM A2/B2/C2= PROABILITY MATRICES
00070 REM K/K1= MATRIX SIZE/NO.OF CODES USED
00080 REM R/R$= L.R./TITLE
00090 REM U$/U1$/V$= OBS.IN/PRED.IN/OUTPUT FILES
00100 DIM A(20,20),B(20,20),C(20,20)
00110 DIM A1(20),B1(20),C1(20)
00120 DIM A2(20,20),B2(20,20),C2(20,20)
00130 REM OPEN FILES
00140 PRINT
00150 PRINT "INPUT FILE (PROGRAM 7)"  ;
00160 INPUT U$
00170 PRINT
00180 PRINT "INPUT FILE (PROGRAM 8)"  ;
00190 INPUT U1$
00200 PRINT
00210 PRINT "OUTPUT FILE NAME"  ;
00220 INPUT V$
00230 FILE #1,U$;#2,U1$;#3,V$
00240 MARGIN #3,132
00250 SCRATCH #3
00260 PRINT
00270 PRINT "NUMBER OF CODES USED"
00280 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00290 PRINT
00300 PRINT "NUMBER OF ACTIVITY CODES USED"  ;
00310 INPUT K1
00320 LET K=K1+K1
00330 PRINT
00340 PRINT "TITLE"  ;
00350 INPUT R$
00360 REM INPUT DATA
00370 FOR I=1 TO K
00380 FOR J=1 TO K
00390 INPUT #1,A(I,J)
00400 INPUT #2,B2(I,J)
00410 LET A1(I)=A1(I)+A(I,J)
00420 NEXT J
00430 NEXT I
00440 REM COMPUTE B FREQUENCIES
00450 FOR I=1 TO K
00460 FOR J=1 TO K
00470 IF A1(I)=0 THEN 500
00480 IF B2(I,J)=0 THEN 520
00490 GOTO 540
00500 LET B(I,J)=0
00510 GOTO 560
00520 LET B(I,J)=0
00530 GOTO 550
00540 LET B(I,J)=B2(I,J)*A1(I)
00550 NEXT J
00560 NEXT I
00570 FOR I=1 TO K
```

```
00580 FOR J=1 TO K
00590 LET C(I,J)=A(I,J)+B(I,J)
00600 REM ROW TOTALS
00610 LET B1(I)=B1(I)+B(I,J)
00620 LET C1(I)=C1(I)+C(I,J)
00630 NEXT J
00640 NEXT I
00650 REM COMPUTE MATRIX PROBABILITIES
00660 FOR I=1 TO K
00670 FOR J=1 TO K
00680 IF A1(I)=0 THEN 710
00690 LET A2(I,J)=A(I,J)/A1(I)
00700 GOTO 720
00710 LET A2(I,J)=0
00720 IF C1(I)=0 THEN 750
00730 LET C2(I,J)=C(I,J)/C1(I)
00740 GOTO 760
00750 LET C2(I,J)=0
00760 NEXT J
00770 NEXT I
00780 REM LIKELIHOOD RATIO
00790 LET R=R1=R2=0
00800 FOR I=1 TO K
00810 FOR J=1 TO K
00820 IF C2(I,J)=0 THEN 850
00830 LET E=LOGE(C2(I,J))
00840 GOTO 860
00850 LET E=0
00860 IF A2(I,J)=0 THEN 890
00870 LET F=LOGE(A2(I,J))
00880 GOTO 900
00890 LET F=0
00900 IF B2(I,J)=0 THEN 930
00910 LET G=LOGE(B2(I,J))
00920 GOTO 940
00930 LET G=0
00940 LET R1=R1+(A(I,J)*(E-F))
00950 LET R2=R2+(B(I,J)*(E-G))
00960 NEXT J
00970 NEXT I
00980 LET R= -2*(R1+R2)
00990 REM PRINT OUTPUT
01000 PRINT #3,DATE$
01010 PRINT #3
01020 PRINT #3,R$
01030 PRINT #3
01040 PRINT #3
01050 PRINT #3,"OBSERVED/PREDICTED/TOTAL, SPKR.1"
01060 PRINT #3
01070 PRINT #3
01080 REM COLUMN INDICATOR
01090 FOR J=1 TO K1
01100 PRINT #3,TAB(10*J);J-1;
01110 NEXTJ
01120 PRINT #3
01130 PRINT #3
01140 FOR I=1 TO K
```

```
01150 PRINT #3
01160 IF I>K1 THEN 1190
01170 PRINT #3,I-1;
01180 GOTO 1200
01190 PRINT #3,I-(K1+1);
01200 FOR J=1 TO K1
01210 PRINT #3,TAB(10*J);A(I,J);
01220 NEXT J
01230 PRINT #3
01240 PRINT #3
01250 FOR J=1 TO K1
01260 PRINT #3,TAB(10*J);A2(I,J);
01270 NEXT J
01280 PRINT #3
01290 PRINT #3
01300 FOR J=1 TO K1
01310 PRINT #3,TAB(10*J);B(I,J);
01320 NEXT J
01330 PRINT #3
01340 PRINT #3
01350 FOR J=1 TO K1
01360 PRINT #3,TAB(10*J);B2(I,J);
01370 NEXT J
01380 PRINT #3
01390 PRINT #3
01400 PRINT #3
01410 FOR J=1 TO K1
01420 PRINT #3,TAB(10*J);C(I,J);
01430 NEXT J
01440 PRINT #3
01450 PRINT #3
01460 FOR J=1 TO K1
01470 PRINT #3,TAB(10*J);C2(I,J);
01480 NEXT J
01490 PRINT #3
01500 PRINT #3
01510 NEXT I
01520 PRINT #3
01530 PRINT #3
01540 PRINT #3,"OBSERVED/PREDICTED/TOTAL, SPKR.2"
01550 PRINT #3
01560 PRINT #3
01570 REM COLUMN INDICATOR
01580 FOR J=K1+1 TO K
01590 PRINT #3,TAB(10*(J-K1));J-(K1+1);
01600 NEXT J
01610 PRINT #3
01620 PRINT #3
01630 FOR I=1 TO K
01640 PRINT #3
01650 IF I>K1 THEN 1680
01660 PRINT #3,I-1;
01670 GOTO 1690
01680 PRINT #3,I-(K1+1);
01690 FOR J=K1+1 TO K
01700 PRINT #3,TAB(10*(J-K1));A(I,J);
01710 NEXT J
```

```
01720 PRINT #3,TAB(10+(10*(J-K1)));A1(I)
01730 PRINT #3
01740 FOR J=K1+1 TO K
01750 PRINT #3,TAB(10*(J-K1));A2(I,J);
01760 NEXT J
01770 PRINT #3
01780 PRINT #3
01790 FOR J=K1+1 TO K
01800 PRINT #3,TAB(10*(J-K1));B(I,J);
01810 NEXT J
01820 PRINT #3,TAB(10+(10*(J-K1)));B1(I)
01830 PRINT #3
01840 FOR J=K1+1 TO K
01850 PRINT #3,TAB(10*(J-K1));B2(I,J);
01860 NEXT J
01870 PRINT #3
01880 PRINT #3
01890 FOR J=K1+1 TO K
01900 PRINT #3,TAB(10*(J-K1));C(I,J);
01910 NEXT J
01920 PRINT #3,TAB(10+(10*(J-K1)));C1(I)
01930 PRINT #3
01940 FOR J=K1+1 TO K
01950 PRINT #3,TAB(10*(J-K1));C2(I,J);
01960 NEXT J
01970 PRINT #3
01980 PRINT #3
01990 NEXT I
02000 PRINT #3
02010 PRINT #3
02020 PRINT #3
02030 PRINT #3,"ANDERSON-GOODMAN L.R. TEST"
02040 PRINT #3
02050 PRINT #3,"L.R. HAS CHI-SQUARE DISTRIBUTION"
02060 PRINT #3
02070 LET K8=(K*(K-1))*(2-1)
02080 PRINT #3,"LIKELIHOOD RATIO=";R;"WITH";K8;"DF"
02090 STOP
02100 END
```

```
00010 REM PROGRAM 10     MONTE-CARLO GENERATOR
00020 REM
00030 REM GENERATES 100 CHAINS N ITEMS LONG
00040 REM A/C= PARAMETER/MAX OF RANGE MATRIX
00050 REM R= OUTPUT SEQUENCE MATRIX
00060 REM P$/Q$= PARAMETER/OUTPUT FILE NAME
00070 REM B= ROW PROBABILITY SUM (PARAMETER MATRIX)
00080 REM K/G= GENERATED RANDOM NUMBER/STARTING ROW
00090 REM P/H= ROW/COL INDICATOR (MAX OF RANGE MATRIX)
00100 REM E/F= NO. OF MATRIX ROWS/COLUMNS
00110 DIM A(20,20),C(20,20),R(1000)
00120 PRINT "PARAMETER MATRIX SIZE"
00130 PRINT "MAX 20 ROWS/COLUMNS"
00140 PRINT
00150 PRINT "INPUT NO. OF ROWS"  ;
00160 INPUT E
00170 IF E>20 THEN 190
00180 GOTO 220
00190 PRINT "MATRIX SIZE ERROR"
00200 PRINT
00210 GOTO 150
00220 LET F=E
00230 PRINT
00240 PRINT "SEQUENCE LENGTH"  ;
00250 INPUT N
00260 PRINT
00270 PRINT "WHICH SPEAKER STARTS"  ;
00280 INPUT X
00290 IF X<1 THEN 320
00300 IF X>2 THEN 320
00310 GOTO 340
00320 PRINT "SPEAKER CODE ERROR"
00330 GOTO 260
00340 PRINT
00350 PRINT "START ACTIVITY CODE"  ;
00360 INPUT X1
00370 IF X1<0 THEN 400
00380 IF X1>(E/2)-1 THEN 400
00390 GOTO 420
00400 PRINT "ACTIVITY CODE ERROR"
00410 GOTO 340
00420 LET X2=(X*10)+X1
00430 IF X=2 THEN 460
00440 LET G=X1+1
00450 GOTO 470
00460 LET G=(X1+1)+E/2
00470 PRINT
00480 PRINT "WHICH SPEAKER ENDS"  ;
00490 INPUT Y
00500 IF Y<1 THEN 530
00510 IF Y>2 THEN 530
00520 GOTO 550
00530 PRINT "SPEAKER CODE ERROR"
00540 GOTO 470
00550 PRINT
00560 PRINT "END ACTIVITY CODE"  ;
00570 INPUT Y1
```

```
00580 IF Y1<0 THEN 610
00590 IF Y1>(E/2)-1 THEN 610
00600 GOTO 630
00610 PRINT "ACTIVITY CODE ERROR"
00620 GOTO 550
00630 LET Y=(Y*10)+Y1
00640 REM START AND END POSITIONS
00650 REM ADD TYPE CODE 0 TO START AND END POSITIONS
00660 LET R(1)=X2*10
00670 LET R(N)=Y*10
00680 PRINT
00690 REM PARAMETER FILE
00700 PRINT "PARAMETER FILE NAME"  ;
00710 INPUT P$
00720 FILE #1,P$
00730 PRINT
00740 REM OUTPUT FILE
00750 PRINT "OUTPUT FILE NAME"  ;
00760 INPUT Q$
00770 FILE #2,Q$
00780 SCRATCH #2
00790 REM INPUT PARAMETERS
00800 REM CONSTRUCT MAXIMUM OF RANGE MATRIX
00810 FOR I=1 TO E
00820 LET B=0
00830 FOR J=1 TO F
00840 INPUT #1,A(I,J)
00850 LET B=B+A(I,J)
00860 LET C(I,J)=(1000*B)-1
00870 NEXT J
00880 NEXT I
00890 REM GENERATE RANDOM NUMBER
00900 REM SELECT MATRIX SUBSEQUENT
00910 FOR Z=1 TO 25
00920 RANDOM
00930 LET P=G
00940 LET M=1
00950 LET M=M+1
00960 IF M=N THEN 1070
00970 LET K=INT(1000*RND)
00980 LET I=P
00990 IF K>C(I,F) THEN 970
01000 FOR H=1 TO F
01010 IF K<=C(I,H) THEN 1030
01020 NEXT H
01030 LET R(M)=H
01040 REM SUBSEQUENT (COL)=NEXT ANTECEDENT (ROW)
01050 LET P=H
01060 GOTO 950
01070 REM ADD SPEAKER CODE
01080 FOR I=2 TO N-1
01090 IF R(I)>(E/2) GOTO 1120
01100 LET R(I)=(R(I)-1)+10
01110 GOTO 1130
01120 LET R(I)=((R(I)-E/2)-1)+20
01130 NEXT I
01140 REM ADD TYPE CODE 0, AND OUTPUT TO FILE
```

```
01150 FOR I=2 TO N-1
01160 LET R(I)=R(I)*10
01170 NEXT I
01180 FOR I=1 TO N
01190 PRINT #2,R(I)
01200 NEXT I
01210 NEXT Z
01220 END
```

```
00010 REM PROGRAM 11     INTER-EVENT DISTANCE
00020 REM
00030 REM A= INPUT DATA
00040 REM A1= DATA ITEM IN A(I) TO BE RE-CODED
00050 REM B= MEAN DISTANCE MATRIX
00060 REM B1= STANDARD DEVIATION MATRIX
00070 REM C= DISTANCE COUNT (PER SEARCH)
00080 REM C1= TOTAL DISTANCE COUNT (PER CODE)
00090 REM C2= INDIVIDUAL DISTANCE COUNTS
00100 REM D= NUMBER OF DISTANCE EVENTS (PER CODE)
00110 REM F/G= ANT/SUB CODE
00120 REM I/J= FIRST/SEARCH POINTER
00130 REM K5/K6= NO. OF CODES/MATRIX SIZE
00140 REM M= OBSERVED/MARKOV DATA INDICATOR
00150 REM N= SEQUENCE LENGTH
00160 REM Q= SPEAKER CODE SELECTOR
00170 REM S= STANDARD DEVIATION
00180 REM S1/N1/T1= SPEAKER/ACTIVITY/TYPE CODES
00190 REM S2/N2= SPEAKER/ACTIVTY CODE STORES
00200 REM Y1= FIRST POINTER SPEAKER CODE
00210 REM Y2= SEARCH POINTER SPEAKER CODE
00220 REM Z1= FIRST POINTER ACT CODE
00230 REM Z2= SEARCH PONTER ACT CODE
00240 REM R$= TITLE
00250 REM U$= OBSERVED/MARKOV MARKER
00260 REM V$/W$= IN/OUT FILE NAMES
00270 REM V1$= OBS./GEN. OUTPUT FILE (MEAN)
00280 REM V2$= OBS./GEN OUPUT FILE (S.D.)
00290 DIM A(700),B(20,20),B1(20,20),C2(700)
00300 DIM S2(700),N2(700)
00310 REM SET STORES =0
00320 FOR I=1 TO 700
00330 LET A(I)=C2(I)=0
00340 NEXT I
00350 FOR I=1 TO 20
00360 FOR J=1 TO 20
00370 LET B(I,J)=B1(I,J)=0
00380 NEXT J
00390 NEXT I
00400 REM OPEN INPUT AND OUTPUT FILES
00410 PRINT "INPUT FILE NAME" ;
00420 INPUT V$
00430 PRINT
00440 PRINT "OUTPUT FILE NAME"  ;
00450 INPUT W$
00460 FILE #1,V$;#2,W$
00470 MARGIN #2,132
00480 SCRATCH #2
00490 PRINT
00500 PRINT "DOES FILE HOLD OBS. DATA (Y/N)"  ;
00510 INPUT U$
00520 IF U$="Y" THEN 540
00530 IF U$="N" THEN 580
00540 LET M=1
00550 LET V1$="DISMO.DAT"
00560 LET V2$="DISSDO.DAT"
00570 GOTO 610
```

```
00580 LET M=100
00590 LET V1$="DISMG.DAT"
00600 LET V2$="DISSDG.DAT"
00610 FILE #3,V1$;#4,V2$
00620 SCRATCH #3,4
00630 REM OBSERVED/MARKOV SEQUENCE LENGTH
00640 PRINT
00650 PRINT "SEQUENCE LENGTH"  ;
00660 INPUT N
00670 PRINT
00680 PRINT "NUMBER OF CODES USED"
00690 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00700 PRINT
00710 PRINT "NUMBER OF ACTIVITY CODES USED"  ;
00720 INPUT K5
00730 LET K6=K5+K5
00740 PRINT
00750 PRINT "TITLE"  ;
00760 INPUT R$
00770 REM INPUT RAW DATA
00780 FOR X=1 TO M
00790 FOR J=1 TO N
00800 INPUT #1,A(J)
00810 GOSUB 2710
00820 LET S2(J)=S1
00830 LET N2(J)=N1
00840 NEXT J
00850 REM SELECT SPEAKER CODE
00860 LET Q=1
00870 ON Q  GOTO 880 ,900 ,930 ,960 ,1560
00880 LET Y1=Y2=1
00890 GOTO 990
00900 LET Y1=1
00910 LET Y2=2
00920 GOTO 990
00930 LET Y1=2
00940 LET Y2=1
00950 GOTO 990
00960 LET Y1=2
00970 LET Y2=2
00980 REM SET ACT CODES =0 FOR FIRST RUN
00990 LET Z1=Z2=0
01000 REM SET COUNTERS=0
01010 LET C1=D=I=J=0
01020 FOR K=1 TO 700
01030 LET C2(K)=0
01040 NEXT K
01050 REM SET FIRST POINTER POSITION
01060 LET I=I+1
01070 IF I=N THEN 1260
01080 IF S2(I)<>Y1 THEN 1060
01090 IF N2(I)<>Z1 THEN 1060
01100 REM SET SEARCH POINTER POSITION
01110 LET J=I
01120 LET C=0
01130 LET J=J+1
01140 IF J=N+1 THEN 1060
```

```
01150 REM COUNT DISTANCE
01160 LET C=C+1
01170 IF S2(J)<>Y2 THEN 1130
01180 IF N2(J)<>Z2 THEN 1130
01190 REM COUNT NUMBER OF DISTANCE EVENTS
01200 LET D=D+1
01210 REM COUNT TOTAL DISTANCE
01220 LET C1=C1+C
01230 REM INDIVIDUAL DISTANCE COUNTS
01240 LET C2(D)=C
01250 GOTO 1060
01260 REM ENTER INTO MATRICES
01270 LET F=Z1+1
01280 LET G=Z2+1
01290 IF Y1=1 THEN 1310
01300 LET F=F+K5
01310 IF Y2=1 THEN 1330
01320 LET G=G+K5
01330 IF D=0 THEN 1460
01340 LET B(F,G)=B(F,G)+(C1/D)
01350 REM STANDARD DEVIATION
01360 LET S=0
01370 FOR K=1 TO D
01380 LET S=S+(C2(K)-(C1/D))^2
01390 NEXT K
01400 REM ZERO DIVISION CHECK
01410 IF (D-1)=0 THEN 1440
01420 LET S=SQRT(S/(D-1))
01430 GOTO 1450
01440 LET S=0
01450 LET B1(F,G)=B1(F,G)+S
01460 LET Z1=Z1+1
01470 IF Z1>K5 THEN 1490
01480 GOTO 1010
01490 LET Z1=0
01500 LET Z2=Z2+1
01510 IF Z2>K5 THEN 1530
01520 GOTO 1010
01530 LET Z1=Z2=0
01540 LET Q=Q+1
01550 GOTO 870
01560 NEXT X
01570 REM PRINT MATRICES
01580 PRINT #2,DATE$
01590 PRINT #2
01600 PRINT #2,"SOURCE FILE= ";V$
01610 PRINT #2
01620 PRINT #2,R$
01630 PRINT #2
01640 IF U$="Y" THEN 1670
01650 PRINT #2,"GENERATED SEQUENCES"
01660 GOTO 1680
01670 PRINT #2,"OBSERVED SEQUENCES"
01680 PRINT #2
01690 PRINT #2
01700 PRINT #2,"MEAN DISTANCE MATRIX"
01710 PRINT #2
```

```
01720 PRINT #2,"SPEAKER 1 CONSEQUENT"
01730 PRINT #2
01740 REM COLUMN INDICATOR
01750 FOR J=1 TO K5
01760 PRINT #2,TAB(10*J);J-1;
01770 NEXT J
01780 PRINT #2
01790 PRINT #2
01800 REM PRINT MATRIX
01810 FOR I=1 TO K6
01820 PRINT #2
01830 IF I>K5 THEN 1860
01840 PRINT #2,I-1;
01850 GOTO 1870
01860 PRINT #2,I-(K5+1);
01870 FOR J=1 TO K5
01880 PRINT #2,TAB(10*J);B(I,J)/M;
01890 NEXT J
01900 NEXT I
01910 PRINT #2
01920 PRINT #2
01930 PRINT #2
01940 PRINT #2,"MEAN DISTANCE MATRIX"
01950 PRINT #2
01960 PRINT #2,"SPEAKER 2 CONSEQUENT"
01970 PRINT #2
01980 REM COLUMN INDICATOR
01990 FOR J=K5+1 TO K6
02000 PRINT #2,TAB(10*(J-K5));J-(K5+1);
02010 NEXT J
02020 PRINT #2
02030 PRINT #2
02040 REM PRINT MATRIX
02050 FOR I=1 TO K6
02060 PRINT #2
02070 IF I>K5 THEN 2100
02080 PRINT #2,I-1;
02090 GOTO 2110
02100 PRINT #2,I-(K5+1);
02110 FOR J=K5+1 TO K6
02120 PRINT #2,TAB(10*(J-K5));B(I,J)/M;
02130 NEXT J
02140 NEXT I
02150 PRINT #2
02160 PRINT #2
02170 PRINT #2
02180 PRINT #2,"S.D. DISTANCE MATRIX"
02190 PRINT #2
02200 PRINT #2,"SPEAKER 1 CONSEQUENT"
02210 PRINT #2
02220 REM COLUMN INDICATOR
02230 FOR J=1 TO K5
02240 PRINT #2,TAB(10*J);J-1;
02250 NEXT J
02260 PRINT #2
02270 PRINT #2
02280 REM PRINT MATRIX
```

```
02290 FOR I=1 TO K6
02300 PRINT #2
02310 IF I>K5 THEN 2340
02320 PRINT #2,I-1;
02330 GOTO 2350
02340 PRINT #2,I-(K5+1);
02350 FOR J=1 TO K5
02360 PRINT #2,TAB(10*J);B1(I,J)/M;
02370 NEXT J
02380 NEXT I
02390 PRINT #2
02400 PRINT #2
02410 PRINT #2
02420 PRINT #2,"S.D. DISTANCE MATRIX"
02430 PRINT #2
02440 PRINT #2,"SPEAKER 2 CONSEQUENT"
02450 PRINT #2
02460 REM COLUMN INDICATOR
02470 FOR J=K5+1 TO K6
02480 PRINT #2,TAB(10*(J-K5));J-(K5+1);
02490 NEXT J
02500 PRINT #2
02510 PRINT #2
02520 REM PRINT MATRIX
02530 FOR I=1 TO K6
02540 PRINT #2
02550 IF I>K5 THEN 2580
02560 PRINT #2,I-1;
02570 GOTO 2590
02580 PRINT #2,I-(K5+1);
02590 FOR J=K5+1 TO K6
02600 PRINT #2,TAB(10*(J-K5));B1(I,J)/M;
02610 NEXT J
02620 NEXT I
02630 REM PRINT DATA OUTPUT FILE
02640 FOR I=1 TO K6
02650 FOR J=1 TO K6
02660 PRINT #3,B(I,J)/M
02670 PRINT #4,B1(I,J)/M
02680 NEXT J
02690 NEXT I
02700 STOP
02710 REM RAW DATA CONVERSION ROUTINE
02720 REM FIRST POINTER
02730 LET S1=N1=T1=0
02740 LET A1=A(J)
02750 IF A1<100 THEN 2790
02760 LET S1=S1+1
02770 LET A1=A1-100
02780 GOTO 2750
02790 IF A1<10 THEN 2830
02800 LET N1=N1+1
02810 LET A1=A1-10
02820 GOTO 2790
02830 LET T1=A1
02840 RETURN
02850 STOP
```

02860 END

```
00010 REM PROGRAM 12     PROCESS DATA COMPARATOR
00020 REM
00030 REM USING THE ANDERSON-GOODMAN HOMOGENEITY TEST
00040 REM A/B/C= OBS./GEN./TOTAL MATRIX
00050 REM A1/B1/C1= ROW TOTALS
00060 REM A2/B2/C2= PROABILITY MATRICES
00070 REM K/K1= MATRIX SIZE/NO. OF CODES USED
00080 REM R/R$= L.R./TITLE
00090 REM U$/U1$/V$= OBS.IN/GEN.IN/OUTPUT FILE
00100 DIM A(20,20),B(20,20),C(20,20)
00110 DIM A1(20),B1(20),C1(20)
00120 DIM A2(20,20),B2(20,20),C2(20,20)
00130 REM OPEN FILES
00140 PRINT
00150 PRINT "INPUT FILE (OBSERVED FREQUENCIES)"  ;
00160 INPUT U$
00170 PRINT
00180 PRINT "INPUT FILE (GENERATED FREQUENCIES)"  ;
00190 INPUT U1$
00200 PRINT
00210 PRINT "OUTPUT FILE NAME"  ;
00220 INPUT V$
00230 FILE #1,U$;#2,U1$;#3,V$
00240 MARGIN #3,132
00250 SCRATCH #3
00260 PRINT
00270 PRINT "NUMBER OF ACTIVITY CODES USED"
00280 PRINT "IF 5 CODES USED, CODES 0-4 ASSUMED"
00290 PRINT
00300 PRINT "NUMBER OF CODES USED"  ;
00310 INPUT K1
00320 LET K=K1+K1
00330 PRINT
00340 PRINT "TITLE"  ;
00350 INPUT R$
00360 REM INPUT DATA
00370 FOR I=1 TO K
00380 FOR J=1 TO K
00390 INPUT #1,A(I,J)
00400 INPUT #2,B(I,J)
00410 LET C(I,J)=A(I,J)+B(I,J)
00420 REM ROW TOTALS
00430 LET A1(I)=A1(I)+A(I,J)
00440 LET B1(I)=B1(I)+B(I,J)
00450 LET C1(I)=C1(I)+C(I,J)
00460 NEXT J
00470 NEXT I
00480 REM COMPUTE MATRIX PROBABILITIES
00490 FOR I=1 TO K
00500 FOR J=1 TO K
00510 IF A1(I)=0 THEN 540
00520 LET A2(I,J)=A(I,J)/A1(I)
00530 GOTO 550
00540 LET A2(I,J)=0
00550 IF B1(I)=0 THEN 580
00560 LET B2(I,J)=B(I,J)/B1(I)
00570 GOTO 590
```

```
00580 LET B2(I,J)=0
00590 IF C1(I)=0 THEN 620
00600 LET C2(I,J)=C(I,J)/C1(I)
00610 GOTO 630
00620 LET C2(I,J)=0
00630 NEXT J
00640 NEXT I
00650 REM LIKELIHOOD RATIO
00660 LET R=R1=R2=0
00670 FOR I=1 TO K
00680 FOR J=1 TO K
00690 IF C2(I,J)=0 THEN 720
00700 LET E=LOGE(C2(I,J))
00710 GOTO 730
00720 LET E=0
00730 IF A2(I,J)=0 THEN 760
00740 LET F=LOGE(A2(I,J))
00750 GOTO 770
00760 LET F=0
00770 IF B2(I,J)=0 THEN 800
00780 LET G=LOGE(B2(I,J))
00790 GOTO 810
00800 LET G=0
00810 LET R1=R1+(A(I,J)*(E-F))
00820 LET R2=R2+(B(I,J)*(E-G))
00830 NEXT J
00840 NEXT I
00850 LET R= -2*(R1+R2)
00860 REM PRINT OUTPUT
00870 PRINT #3,DATE$
00880 PRINT #3
00890 PRINT #3,R$
00900 PRINT #3
00910 PRINT #3
00920 PRINT #3,"OBSERVED/GENERATED/TOTAL, SPKR. 1"
00930 PRINT #3
00940 PRINT #3
00950 REM COLUMN INDICATOR
00960 FOR J=1 TO K1
00970 PRINT #3,TAB(10*J);J-1;
00980 NEXTJ
00990 PRINT #3
01000 PRINT #3
01010 FOR I=1 TO K
01020 PRINT #3
01030 IF I>K1 THEN 1060
01040 PRINT #3,I-1;
01050 GOTO 1070
01060 PRINT #3,I-(K1+1);
01070 FOR J=1 TO K1
01080 PRINT #3,TAB(10*J);A(I,J);
01090 NEXT J
01100 PRINT #3
01110 PRINT #3
01120 FOR J=1 TO K1
01130 PRINT #3,TAB(10*J);A2(I,J);
01140 NEXT J
```

```
01150 PRINT #3
01160 PRINT #3
01170 FOR J=1 TO K1
01180 PRINT #3,TAB(10*J);B(I,J);
01190 NEXT J
01200 PRINT #3
01210 PRINT #3
01220 FOR J=1 TO K1
01230 PRINT #3,TAB(10*J);B2(I,J);
01240 NEXT J
01250 PRINT #3
01260 PRINT #3
01270 PRINT #3
01280 FOR J=1 TO K1
01290 PRINT #3,TAB(10*J);C(I,J);
01300 NEXT J
01310 PRINT #3
01320 PRINT #3
01330 FOR J=1 TO K1
01340 PRINT #3,TAB(10*J);C2(I,J);
01350 NEXT J
01360 PRINT #3
01370 PRINT #3
01380 NEXT I
01390 PRINT #3
01400 PRINT #3
01410 PRINT #3,"OBSERVED/GENERATED/TOTAL, SPKR.2"
01420 PRINT #3
01430 PRINT #3
01440 REM COLUMN INDICATOR
01450 FOR J=K1+1 TO K
01460 PRINT #3,TAB(10*(J-K1));J-(K1+1);
01470 NEXT J
01480 PRINT #3
01490 PRINT #3
01500 FOR I=1 TO K
01510 PRINT #3
01520 IF I>K1 THEN 1550
01530 PRINT #3,I-1;
01540 GOTO 1560
01550 PRINT #3,I-(K1+1);
01560 FOR J=K1+1 TO K
01570 PRINT #3,TAB(10*(J-K1));A(I,J);
01580 NEXT J
01590 PRINT #3,TAB(10+(10*(J-K1)));A1(I)
01600 PRINT #3
01610 FOR J=K1+1 TO K
01620 PRINT #3,TAB(10*(J-K1));A2(I,J);
01630 NEXT J
01640 PRINT #3
01650 PRINT #3
01660 FOR J=K1+1 TO K
01670 PRINT #3,TAB(10*(J-K1));B(I,J);
01680 NEXT J
01690 PRINT #3,TAB(10+(10*(J-K1)));B1(I)
01700 PRINT #3
01710 FOR J=K1+1 TO K
```

```
01720 PRINT #3,TAB(10*(J-K1));B2(I,J);
01730 NEXT J
01740 PRINT #3
01750 PRINT #3
01760 FOR J=K1+1 TO K
01770 PRINT #3,TAB(10*(J-K1));C(I,J);
01780 NEXT J
01790 PRINT #3,TAB(10+(10*(J-K1)));C1(I)
01800 PRINT #3
01810 FOR J=K1+1 TO K
01820 PRINT #3,TAB(10*(J-K1));C2(I,J);
01830 NEXT J
01840 PRINT #3
01850 PRINT #3
01860 NEXT I
01870 PRINT #3
01880 PRINT #3
01890 PRINT #3
01900 PRINT #3,"ANDERSON-GOODMAN L.R. TEST"
01910 PRINT #3
01920 PRINT #3,"L.R. HAS CHI-SQUARE DISTRIBUTION"
01930 PRINT #3
01940 LET K8=(K*(K-1))*(2-1)
01950 PRINT #3,"LIKELIHOOD RATIO=";R;"WITH";K8;"DF"
01960 STOP
01970 END
```

```
00010 REM PROGRAM 13     'TYPE' ANALYSIS
00020 REM
00030 REM ACTIVITY+TYPE/TYPE-ONLY ANALYSIS
00040 REM P= ACTIVITY+TYPE/TYPE MATRIX
00050 REM P1= TYPE-ONLY WITHIN MATRIX
00060 REM P2= TYPE-ONLY BETWEEN MATRIX
00070 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00080 REM U$/W$/X$= IN/OUT/DATA FILE NAMES
00090 REM A/T$= RAW DATA/TITLE
00100 DIM P3(800)
00110 PRINT "INPUT FILE NAME"   ;
00120 INPUT U$
00130 PRINT
00140 PRINT "OUTPUT FILE NAME"   ;
00150 INPUT W$
00160 FILE #1,U$;#2,W$
00170 MARGIN #2,132
00180 SCRATCH #2
00190 PRINT
00200 PRINT "DATA RANGES ARE:"
00210 PRINT "ACTIVITY, 0-6; TYPE, 1-28"
00220 PRINT
00230 PRINT "TITLE"    ;
00240 INPUT T$
00250 PRINT
00260 PRINT "TYPE (1) FOR ACTIVITY + TYPE ANALYSIS"
00270 PRINT "OR (2) FOR TYPE-ONLY ANALYSIS"   ;
00280 INPUT C5
00290 PRINT #2,"ACTIVITY/TYPE SINGLE ANALYSIS"
00300 PRINT #2
00310 PRINT #2,DATE$
00320 PRINT #2
00330 PRINT #2,U$
00340 PRINT #2
00350 PRINT #2,T$
00360 PRINT #2
00370 PRINT #2
00380 LET C=0
00390 IF END #1 THEN 440
00400 INPUT #1,A
00410 LET C=C+1
00420 LET P3(C)=A
00430 GOTO 390
00440 IF C5=2 THEN 1220
00450 DIM P(196,28)
00460 FOR J1=1 TO 3
00470 FOR I=0 TO 6
00480 REM SET P MATRIX =0
00490 MAT P=ZER
00500 LET J=0
00510 LET L=1
00520 LET J=J+1
00530 LET M=1
00540 GOSUB 2020
00550 IF J>C THEN 850
00560 LET J=J+1
00570 LET M=L+1
```

```
00580 GOSUB 2020
00590 REM ADD TO MATRICES
00600 IF N(2)<>I THEN 770
00610 LET X=(N(1)*28)+T(1)
00620 LET Y=T(2)
00630 REM INCREMENT MATRICES
00640 REM OVERALL MATRIX
00650 IF J1=2 THEN 690
00660 IF J1=3 THEN 730
00670 LET P(X,Y)=P(X,Y)+1
00680 GOTO 770
00690 IF S(1)<>S(2) THEN 770
00700 REM WITHIN MATRIX
00710 LET P(X,Y)=P(X,Y)+1
00720 GOTO 770
00730 IF S(1)=S(2) THEN 770
00740 REM BETWEEN MATRIX
00750 LET P(X,Y)=P(X,Y)+1
00760 REM SHIFT INPUT STACK
00770 LET Q=1
00780 LET S(Q)=S(Q+1)
00790 LET N(Q)=N(Q+1)
00800 LET T(Q)=T(Q+1)
00810 IF Q=L THEN 550
00820 LET Q=Q+1
00830 GOTO 780
00840 REM PRINT COMPLETED MATRIX SECTION
00850 IF J1=1 THEN PRINT #2,"OVERALL MATRIX"
00860 IF J1=2 THEN PRINT #2,"WITHIN MATRIX"
00870 IF J1=3 THEN PRINT #2,"BETWEEN MATRIX"
00880 PRINT #2
00890 PRINT #2
00900 PRINT #2,"CONSEQUENT ACTIVITY =",I
00910 PRINT #2
00920 PRINT #2
00930 FOR K6= 1 TO 28
00940 PRINT #2,TAB((K6*4)+5);K6;
00950 NEXT K6
00960 PRINT #2
00970 PRINT #2
00980 LET D=0
00990 LET E=0
01000 FOR K5 = 1 TO 196
01010 LET D=D+1
01020 IF D<= 28 THEN 1060
01030 LET E=E+1
01040 LET D=1
01050 REM ZERO LINE CHECK
01060 LET Y=0
01070 FOR K6 = 1 TO 28
01080 LET Y=Y+P(K5;K6)
01090 NEXT K6
01100 IF Y=0  THEN 1070
01110 PRINT #2,TAB(1)E;TAB(3)D;
01120 FOR K6=1 TO 28
01130 PRINT #2,TAB((K6*4)+5);P(K5,K6);
01140 NEXT K6
```

```
01150 PRINT #2,TAB(124);Y
01160 PRINT #2
01170 NEXT K5
01180 NEXT I
01190 NEXT J1
01200 STOP
01210 REM TYPE-ONLY ANALYSIS
01220 DIM P1(28,28),P2(28,28),P4(28,28)
01230 REM SET MATRICES =0
01240 MAT P4=ZER
01250 MAT P1=ZER
01260 MAT P2=ZER
01270 PRINT
01280 PRINT "DATA OUTPUT FILE NAME"   ;
01290 INPUT X$
01300 FILE #3,X$
01310 MARGIN #3,132
01320 SCRATCH #3
01330 LET J=0
01340 LET L=1
01350 LET J=J+1
01360 LET M=1
01370 GOSUB 2020
01380 IF J>C THEN 1610
01390 LET J=J+1
01400 LET M=L+1
01410 GOSUB 2020
01420 REM ADD TO MATRICES
01430 LET X=T(1)
01440 LET Y=T(2)
01450 REM OVERALL MATRIX
01460 LET P4(X,Y)=P4(X,Y)+1
01470 IF S(1)<>S(2) THEN 1520
01480 REM WITHIN MATRIX
01490 LET P1(X,Y)=P1(X,Y)+1
01500 GOTO 1540
01510 REM BETWEEN MATRIX
01520 LET P2(X,Y)=P2(X,Y)+1
01530 REM SHIFT INPUT STACK
01540 LET Q=1
01550 LET S(Q)=S(Q+1)
01560 LET T(Q)=T(Q+1)
01570 IF Q=L THEN 1380
01580 LET Q=Q+1
01590 GOTO 1550
01600 REM PRINT TYPE MATRICES
01610 PRINT #2,"TYPE-ONLY SINGLE ANALYSIS"
01620 PRINT #2
01630 FOR J1= 1 TO 3
01640 ON J1 GOTO 1650 ,1670 ,1700
01650 PRINT #2,"OVERALL MATRIX"
01660 GOTO 1720
01670 PRINT #2,"WITHIN MATRIX"
01680 MAT P4=P1
01690 GOTO 1720
01700 PRINT #2,"BETWEEN MATRIX"
01710 MAT P4=P2
```

```
01720 PRINT #2
01730 PRINT #2
01740 FOR K6= 1 TO 28
01750 PRINT #2,TAB((K6*4)+5);K6;
01760 NEXT K6
01770 PRINT #2
01780 PRINT #2
01790 FOR K5=1 TO 28
01800 REM ZERO LINE CHECK
01810 LET Y=0
01820 FOR K6= 1 TO 28
01830 LET Y=Y+P4(K5,K6)
01840 NEXT K6
01850 IF Y=0 THEN 1920
01860 PRINT #2,K5;
01870 FOR K6= 1 TO 28
01880 PRINT #2,TAB((K6*4)+5);P4(K5,K6);
01890 NEXT K6
01900 PRINT #2,TAB(124);Y
01910 PRINT #2
01920 NEXT K5
01930 REM DATA OUTPUT ROUTINE
01940 FOR I3=1 TO 28
01950 FOR J3=1 TO 28
01960 PRINT #3,P4(I3,J3)
01970 NEXT J3
01980 NEXT I3
01990 NEXT J1
02000 STOP
02010 REM RAW DATA DIVISION SUB-ROUTINE
02020 LET S(M)=N(M)=T(M)=0
02030 LET A1=P3(J)
02040 IF A1<1000 THEN 2080
02050 LET S(M)=S(M)+1
02060 LET A1=A1-1000
02070 GOTO 2040
02080 IF A1<100 THEN 2120
02090 LET N(M)=N(M)+1
02100 LET A1=A1-100
02110 GOTO 2080
02120 LET T(M)=A1
02130 RETURN
02140 STOP
02150 END
```

```
00010 REM PROGRAM 14     CHAIN ANALYSIS
00020 REM
00030 REM A/B= RAW DATA/CHANGED DATA STORE
00040 REM C1= RAW DATA/CHANGED DATA COUNTER
00050 REM D/E= LARGEST CELL MATRIX POSITION
00060 REM F= ACTIVITY/TYPE FLAG + MATRIX SIZE
00070 REM G= FREQUENCY MATRIX
00080 REM K= LARGEST CELL FREQUENCY
00090 REM P/Q= ACT/TYPE
00100 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODE
00110 REM X/Y= MATRIX INDICATORS
00120 REM U$/W$= INPUT/OUTPUT FILE
00130 DIM B(1000),G(100,100)
00140 PRINT "LIMITED TO 1000 DATA POINTS"
00150 PRINT "AND 40 CHAIN PAIRS."
00160 PRINT "PROGRAM ENDS WHEN PAIR FREQUENCY"
00170 PRINT "FALLS BELOW A CRITERION OF 2"
00180 PRINT
00190 PRINT
00200 PRINT "DATA RANGES ARE:"
00210 PRINT "ACTIVITY CODES, 0-9; TYPE CODES, 0-29"
00220 PRINT
00230 PRINT
00240 PRINT "INPUT FILE"   ;
00250 INPUT U$
00260 PRINT
00270 PRINT "OUTPUT FILE"   ;
00280 INPUT W$
00290 FILE #1,U$;#2,W$
00300 SCRATCH #2
00310 MARGIN #2,132
00320 PRINT
00330 PRINT "TITLE"   ;
00340 INPUT T$
00350 PRINT #2,DATE$,"CHAIN ANALYSIS"
00360 PRINT #2
00370 PRINT #2,"ACTIVITY CODES ARE 0-9 (SPKR.1),"
00380 PRINT #2,"10-19 (SPKR.2), 20-100 (NEW)"
00390 PRINT #2
00400 PRINT #2,"TYPE CODES ARE 0-29 (SPKR.1),"
00410 PRINT #2,"30-59 (SPKR.2), 60-100 (NEW)"
00420 PRINT #2
00430 PRINT #2,T$
00440 PRINT #2
00450 PRINT #2
00460 PRINT #2,"ANTECEDENT","CONSEQUENT",
                   "FREQUENCY","NEW CODE","LENGTH"
00470 PRINT #2
00480 INPUT #1,A
00490 REM ACTIVITY OR TYPE ANALYSIS
00500 IF A>299 GOTO 550
00510 LET F=10
00520 LET P=100
00530 LET Q=10
00540 GOTO 580
00550 LET F=30
00560 LET P=1000
```

```
00570 LET Q=100
00580 LET C1=1
00590 LET B(C1)=A
00600 IF END #1 THEN 700
00610 INPUT #1,A
00620 REM DELETE IF SAME AS LAST CODE
00630 IF A=B(C1) THEN 600
00640 LET C1=C1+1
00650 LET B(C1)=A
00660 REM CHANGE FORMAT OF PREVIOUS DATUM
00670 GOSUB 1220
00680 GOTO 600
00690 REM CHANGE LAST DATUM
00700 LET C1=C1+1
00710 GOSUB 1220
00720 LET C1=C1-1
00730 LET F=F+F
00740 REM CONSTRUCT FREQUENCY MATRIX
00750 LET X=Y=0
00760 FOR I=0 TO F
00770 FOR J=0 TO F
00780 LET G(I,J)=0
00790 NEXT J
00800 NEXT I
00810 FOR I=1 TO (C1-1)
00820 LET X=B(I)
00830 LET Y=B(I+1)
00840 LET G(X,Y)=G(X,Y)+1
00850 NEXT I
00860 REM SCAN MATRIX FOR LARGEST CELL
00870 LET K=E=D=0
00880 FOR I=0 TO F
00890 FOR J=0 TO F
00900 IF G(I,J)<=K GOTO 940
00910 LET K=G(I,J)
00920 LET E=I
00930 LET D=J
00940 NEXT J
00950 NEXT I
00960 REM INCREMENT MATRIX SIZE
00970 LET F=F+1
00980 IF K<2 THEN 1200
00990 PRINT #2,E,D,K,F,C1
01000 REM CHANGE SEQUENCE LENGTH
01010 LET C1=C1-K
01020 REM REPLACE OLD CODES WITH NEW
01030 LET L=0
01040 LET I=0
01050 LET I=I+1
01060 IF I>C1 THEN 1190
01070 IF B(I+L)<>E THEN 1160
01080 IF B(I+L+1)<>D THEN 1160
01090 REM CHANGE LENGTH IF 3 OR MORE SAME CODES
01100 IF B(I+L)<>B(I+L+1) THEN 1130
01110 IF B(I+L+1)<>B(I+L+2) THEN 1130
01120 LET C1=C1+1
01130 LET L=L+1
```

```
01140 LET B(I)=F
01150 GOTO 1170
01160 LET B(I)=B(I+L)
01170 GOTO 1050
01180 REM REPEAT WITH MODIFIED SEQUENCE
01190 GOTO 750
01200 STOP
01210 REM DATA DIVISION ROUTINE
01220 LET S=N=T=0
01230 LET X=B(C1-1)
01240 IF X<P THEN 1280
01250 LET S=S+1
01260 LET X=X-P
01270 GOTO 1240
01280 IF X<Q THEN 1320
01290 LET N=N+1
01300 LET X=X-Q
01310 GOTO 1280
01320 LET T=X
01330 REM ACTIVITY/TYPE CHANGE ROUTINE
01340 IF F=30 THEN 1400
01350 IF S=2 THEN 1380
01360 LET Y=N
01370 GOTO 1440
01380 LET Y=N+10
01390 GOTO 1440
01400 IF S=2 THEN 1430
01410 LET Y=T
01420 GOTO 1440
01430 LET Y=T+30
01440 LET B(C1-1)=Y
01450 RETURN
01460 END
```

```
00010 REM PROGRAM 15     UNCERTAINTY ANALYSIS
00020 REM
00030 REM COMPUTES TOTAL, ERROR AND
00040 REM CONTINGENT UNCERTAINTIES FOR A
00050 REM FIRST-ORDER ACTIVITY/TYPE
00060 REM TRANSITION MATRIX
00070 REM P/P1/P2= OVERALL/WITHIN/BETWEEN MATRIX
00080 REM V1/V2= ROW TOTALS MATRICES
00090 REM C5= ACTIVITY/TYPE INDICATOR
00100 REM Z= PROCESSING/PRINTING MATRIX
00110 REM S/N/T= SPEAKER/ACTIVITY/TYPE CODES
00120 REM K= NO. OF CODES  USED
00130 REM F= NO. OF FILES TO BE PROCESED
00140 DIM P(28,28),P1(28,28),P2(28,28)
00150 DIM Z(28,28),V1(28),V2(28)
00160 PRINT "ACTIVITY (1) OR TYPE (2) ANALYSIS"   ;
00170 INPUT C5
00180 IF C5=2 THEN 220
00190 LET V5=100
00200 LET W5=10
00210 GOTO 240
00220 LET V5=1000
00230 LET W5=100
00240 LET V$= "UOVER.RES"
00250 LET W$= "UWITH.RES"
00260 LET X$= "UBETWN.RES"
00270 FILE #2,V$;#3,W$;#4,X$
00280 MARGIN #2,132;#3,132;#4,132
00290 SCRATCH #2,#3,#4
00300 PRINT #2,DATE$
00310 PRINT #2
00320 PRINT #2,"OVERALL UNCERTAINTY"
00330 PRINT #2
00340 PRINT #2
00350 PRINT #2,TAB(5)"LABEL";TAB(30)"TOTAL";TAB(65)
               "ERROR";TAB(100)"CONTINGENT"
00360 PRINT #2
00370 PRINT #2,TAB(30)"U";TAB(45)"%";TAB(65)"U";TAB
               (80)"%";TAB(100)"U";TAB(115)"%"
00380 PRINT #2
00390 PRINT #3,DATE$
00400 PRINT #3
00410 PRINT #3,"WITHIN UNCERTAINTY"
00420 PRINT #3
00430 PRINT #3
00440 PRINT #3,TAB(5)"LABEL";TAB(30)"TOTAL";TAB(65)
               "ERROR";TAB(100)"CONTINGENT"
00450 PRINT #3
00460 PRINT #3,TAB(30)"U";TAB(45)"%";TAB(65)"U";TAB
               (80)"%";TAB(100)"U";TAB(115)"%"
00470 PRINT #3
00480 PRINT #4,DATE$
00490 PRINT #4
00500 PRINT #4,"BETWEEN UNCERTAINTY"
00510 PRINT #4
00520 PRINT #4
00530 PRINT #4,TAB(5)"LABEL";TAB(30)"TOTAL";TAB(65)
```

```
                  "ERROR";TAB(100)"CONTINGENT"
00540 PRINT #4
00550 PRINT #4,TAB(30)"U";TAB(45)"%";TAB(65)"U";TAB
                  (80)"%";TAB(100)"U";TAB(115)"%"
00560 PRINT #4
00570 PRINT
00580 PRINT "NO. OF FILES"  ;
00590 INPUT F
00600 PRINT
00610 PRINT "NUMBER OF CODES USED (MAX 28)"
00620 PRINT "IF 5 CODES USED,CODES 0-4 ASSUMED"
00630 PRINT
00640 PRINT "NUMBER OF CODES USED"    ;
00650 INPUT K
00660 REM LOG2 NO. OF CODES
00670 LET K9=LOG10(K)/LOG10(2)
00680 PRINT
00690 PRINT
00700 REM INPUT DATA
00710 FOR C=1 TO F
00720 PRINT
00730 PRINT "FILE";C;"NAME"    ;
00740 INPUT U$
00750 FILE #1,U$
00760 PRINT
00770 PRINT "LABEL"    ;
00780 INPUT L$
00790 FOR I=0 TO K
00800 FOR J=0 TO K
00810 LET P(I,J)=P1(I,J)=P2(I,J)=0
00820 NEXT J
00830 NEXT I
00840 LET L=1
00850 INPUT #1,A
00860 LET M=1
00870 GOSUB 1860
00880 IF END #1, THEN 1160
00890 INPUT #1,A
00900 LET M=L+1
00910 GOSUB 1860
00920 REM ADD TO MATRICES
00930 IF C5=2 THEN 970
00940 LET X=N(1)
00950 LET Y=N(2)
00960 GOTO 1000
00970 LET X=T(1)
00980 LET Y=T(2)
00990 REM OVERALL MATRIX
01000 LET P(X,Y)=P(X,Y)+1
01010 IF S(1)<>S(2) THEN 1060
01020 REM WITHIN MATRIX
01030 LET P1(X,Y)=P1(X,Y)+1
01040 GOTO 1080
01050 REM BETWEEN MATRIX
01060 LET P2(X,Y)=P2(X,Y)+1
01070 REM SHIFT INPUT STACK
01080 LET Q=1
```

```
01090 LET S(Q)=S(Q+1)
01100 LET N(Q)=N(Q+1)
01110 LET T(Q)=T(Q+1)
01120 IF Q=L THEN 880
01130 LET Q=Q+1
01140 GOTO 1090
01150 REM UNCERTAINTY STATISTICS
01160 FOR R=1 TO 3
01170 ON R GOTO 1180 ,1240 ,1300
01180 FOR I=0 TO K
01190 FOR J=0 TO K
01200 LET Z(I,J)=P(I,J)
01210 NEXT J
01220 NEXT I
01230 GOTO 1370
01240 FOR I=0 TO K
01250 FOR J=0 TO K
01260 LET Z(I,J)=P1(I,J)
01270 NEXT J
01280 NEXT I
01290 GOTO 1370
01300 FOR I=0 TO K
01310 FOR J=0 TO K
01320 LET Z(I,J)=P2(I,J)
01330 NEXT J
01340 NEXT I
01350 REM MATRIX TOTAL FREQUENCY
01360 REM AND ROW TOTALS
01370 LET M5=0
01380 FOR I=0 TO K-1
01390 LET V1(I)=V2(I)=0
01400 NEXT I
01410 FOR D=0 TO K
01420 FOR E=0 TO J
01430 LET M5=M5+Z(D,E)
01440 LET V1(D)=V1(D)+Z(D,E)
01450 NEXT E
01460 NEXT D
01470 REM COLUMN TOTALS
01480 FOR D=0 TO K-1
01490 LET H=0
01500 FOR E=0 TO K-1
01510 LET H=H+Z(E,D)
01520 NEXT E
01530 LET V2(D)=H
01540 NEXT D
01550 REM TOTAL UNCERTAINTY
01560 LET S1=S2=0
01570 FOR D=0 TO K-1
01580 IF V2(D)=0 THEN 1610
01590 LET S1=S1+V2(D)*(LOG10(V2(D))/LOG10(2))
01600 LET S2=S2+V2(D)
01610 NEXT D
01620 LET G= -((S1-((LOG10(M5)/LOG10(2))*S2))/M5)
01630 REM ERROR UNCERTAINTY
01640 LET S2=0
01650 FOR D=0 TO K-1
```

```
01660 LET S1=0
01670 FOR E=0 TO J-1
01680 IF Z(D,E)=0 THEN 1700
01690 LET S1=S1+Z(D,E)*(LOG10(Z(D,E))/LOG10(2))
01700 NEXT E
01710 IF V1(D)=0 THEN 1730
01720 LET S2=S2+(S1-(V1(D)*(LOG10(V1(D))/LOG10(2))))
01730 NEXT D
01740 LET E5= -(S2*(1/M5))
01750 REM PRINT ROUTINE
01760 ON R GOTO 1770 ,1790 ,1810
01770 PRINT #2,TAB(5)L$;TAB(30)G;TAB(45)(G/K9)*100;
               TAB(65)E5;TAB(80)((E5/G)*100);TAB
               (100)G-E5;TAB(115)(100-((E5/G)*100))
01780 GOTO 1820
01790 PRINT #3,TAB(5)L$;TAB(30)G;TAB(45)(G/K9)*100;
               TAB(65)E5;TAB(80)((E5/G)*100);TAB
               (100)G-E5;TAB(115)(100-((E5/G)*100))
01800 GOTO 1820
01810 PRINT #4,TAB(5)L$;TAB(30)G;TAB(45)(G/K9)*100;
               TAB(65)E5;TAB(80)((E5/G)*100);TAB
               (100)G-E5;TAB(115)(100-((E5/G)*100))
01820 NEXT R
01830 NEXT C
01840 STOP
01850 REM RAW DATA CONVERSION ROUTINE
01860 LET S(M)=N(M)=T(M)=0
01870 IF A<V5 THEN 1910
01880 LET S(M)=S(M)+1
01890 LET A=A-V5
01900 GOTO 1870
01910 IF A<W5 THEN 1950
01920 LET N(M)=N(M)+1
01930 LET A=A-W5
01940 GOTO 1910
01950 LET T(M)=A
01960 RETURN
01970 STOP
01980 END
```

REFERENCES

ALLPORT, F. (1955) Theories of perception and the concept of structure. New York and London: John Wiley.

ALTMAN, S.A. (1965) Sociobiology of rhesus monkeys. II: Stochastics of social communication. Journal of Theoretical Biology, 8, 490-522.

ANDERSON, C.W. (1960) The relation between speaking times and decision in the employment interview. Journal of Applied Psychology, 44, 204-268.

ANDERSON, T.W. AND GOODMAN, L.A. (1957) Statistical inference about Markov chains. Annals of Mathematical Statistics, 28, 89-110.

ARGYLE, M. (1969) Social Interaction. London: Methuen.

ARGYLE, M. (1972) The psychology of interpersonal behaviour. Second edition. Harmondsworth: Penguin Books.

ARGYLE, M. (1979) Sequences in social behaviour as a function of the situation. In Ginsburg, G.P. (Ed.), Emerging Strategies in Social Psychological Research. Chichester: John Wiley.

ARGYLE, M. AND KENDON, A. (1967) The experimental analysis of social performance, Advances in Experimental Social Psychology, 3, 55-98.

ARGYLE, M., FURNHAM, A. AND GRAHAM, J.A. (1981) Social Situations. Cambridge: Cambridge University Press.

ASHBY, W.R. (1961) An Introduction to cybernetics. New York: John Wiley.

ASHBY, W.R. (1968) Principles of the self-organising system. In Buckley, W. (Ed.), Modern Systems Research for the Behavioural Scientist. Chicago: Aldine Publishing Company.

ATTNEAVE, F. (1959) Applications of Information Theory to Psychology: A summary of basic concepts, methods and results. London: Holt, Rinehart and Winston.

AULD, F. AND WHITE, A.M. (1956) Rules for dividing interviews into sentences. Journal of Psychology, 42, 273-81.

BALES, R.F. (1950) Interaction Process Analysis: A method for the study of small groups. Reading, Mass.: Addison-Wesley.

BALES, R.F. (1953) The equilibrium problem in small groups. In Parsons, T. et al. (Eds.). Working Papers in the Theory of Action, pp. 111-161. Glencoe, Illinois: The Free Press of Glencoe.

BALES, R.F. (1955) The equilibrium problem in small groups. In Hare, A.P., Borgatta, E.F. and Bales, R.F. Small groups: studies in social interaction. New York: Alfred A. Knopf.

BALES, R.F. (1966) The equilibrium problem in small groups. In Hare, A.P., Borgatta, E.F. and Bales, R.F. (Eds.), Revised edition. Small groups: Studies in social interaction. New York: Alfred A. Knopf.

BALES, R.F., STRODTBECK, F., MILLS, T. AND ROSEBOROUGH, M. (1951) The channels of communication in small groups. _American Sociological Review_, 16, 461-468.

BARTHOLOMEW, D. (1973) _Stochastic models for social processes_. Second edition. New York: John Wiley.

BARTLETT, M.S. (1951) The frequency goodness of fit test for probability chains. _Proceedings of the Cambridge Philosophical Society_, 47, 86-95.

BATESON, G. (1967) Cybernetic Explanation. _American Behavioural Scientist_, 10, 29-32.

BATESON, G. AND JACKSON, D.D. (1964) Some varieties of pathogenic organisation. In Rioch, D.M. (Ed.), _Disorders of communication_. Association for Research in Nervous and Mental Disease, 42, 270-283.

BEATTIE, G.W. (1978) Floor apportionment and gaze in conversational dyads. _British Journal of Social and Clinical Psychology_, 17, 7-15.

BENJAMIN, L.S. (1979) Use of Structural analysis of Social Behaviour (SASB) and Markov chains to study dyadic interaction. _Journal of Abnormal Psychology_, 88, 303-319.

BENNIS, W.G. AND SHEPARD, H.A. (1956) A theory of group development. _Human Relations_, 9, 415-437.

BERGER, C.R. (1977) Interpersonal communication theory and research: An overview. In Ruben, B.D. (Ed.), _Communication Yearbook I_, Brunswick, New Jersey: Transaction Books.

BIRDWHISTELL, R. (1970) _Kinesics and context_. Philadelphia: University of Pennsylvania Press.

BLALOCK, H.M. (1960) Social statistics. New York: McGraw-Hill.

BORGATTA, E.F. AND CROWTHER, B. (1965) A workbook for the study of social interaction processes. Chicago: Rand McNally.

BRINICH, P.M. (1981) Defining the units in observational research. Bulletin of the British Psychological Society, 34, 418-419.

BUGENTAL, J.F.T. (1948) An investigation of the relationships of the conceptual matrix to the self-concept. Unpublished Ph.D. thesis, Ohio State University.

BULL, P.E. AND BROWN, R. (1977) The role of postural change in dyadic conversations. British Journal of Social and Clinical Psychology, 16, 29-33.

CALABRESE, R.J. (1975) The effects of privacy and probability of future interaction on initial interaction patterns. Unpublished Ph.D thesis, Northwestern University.

CAPPELLA, J.N. (1976) Modelling interpersonal communication systems as a pair of machines coupled through feedback. In Miller, G.R. (Ed.). Explorations in Interpersonal Communication. Beverley Hills and London: Sage Publications.

CAPPELLA, J.N. (1979) Talk and silence sequences in informal conversations, I. Human Communication Research, 6, 3-17.

CAPPELLA, J.N. (1980) Talk and silence sequences in informal conversations, II. Human Communication Research, 6, 130-145.

CAPPELLA, J.N. AND PLANAP, S. (1981) Talk and silence sequences in informal conversations, III. Human Communication Research, 7, 117-132.

CARROLL, L. (1871) Through the Looking Glass, in Gardener, M. (Ed.) (1970). The annotated Alice, Harmondsworth: Penguin.

CHATFIELD, D. (1973) Statistical inference regarding Markov chain models. Applied statistics, 22, 7-20.

CHATFIELD, D. AND LEMON, R.E. (1970) Analysing sequences of behavioural events. Journal of Theoretical Biology, 29, 427-445.

CHOMSKY, N. (1957) Syntactic structures. The Hague:Mouton.

CLARKE, D.D. (1975) The use and recognition of sequential structure in dialogue. Br. J. soc. clin. Psychol. 14, 333-339

CLARKE, D.D. (1977) Rules and sequences in conversation. In Collett, P. (Ed.), Social rules and social behaviour, Oxford: Basil Blackwell.

COCHRAN, W.G. (1954) Some methods for strengthening the common χ^2 tests. Biometrics, 10, 417-451.

COHEN, J. (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 10, 1, 37-46.

COLGAN, P.W. AND SMITH, J.T. (1978) Multidimensional contingency table analysis. In: Colgan, P.W. (Ed.), Quantitative Ethology, New York: John Wiley.

COX, D.R. AND MILLER, H.D. (1965) The theory of stochastic processes. London: Methuen.

CROWELL, S. AND SCHEIDEL, T.M. (1961) Categories for analysis of idea development in discussion groups. Journal of Social Psychology, 54, 155-168.

CRYSTAL, D. (1969) _Prosodic systems and intonation in English_. Cambridge: Cambridge University Press.

CUSHMAN, D.P. AND CRAIG, R.T. (1976) Communication systems: Interpersonal implications. In Miller, G.R. (Ed.), _Explorations in interpersonal communication_. Beverly Hills and London: Sage Publications.

CUSHMAN, D.P. AND PEARCE, W.B. (1977) Generality and necessity in three types of theory about human communication, with special attention to rules theory. _Human Communication Research_, 3, 344-353.

DANZIGER, K. (1976) _Interpersonal communication_. Oxford: Pergamon Press.

DANZIGER, K. AND GREENGLASS, E.R. (1970) _Verbal Exchange Analysis_, coding manual. Research Report, York University, Canada.

DAWKINS, R. (1976) Hierarchical organisation: a candidate principle of ethology. In, Bateson, P.P.G. and Hinde, R.A. (Eds.), _Growing points in ethology_, Cambridge: Cambridge University Press.

DAWKINS, R. AND DAWKINS, M. (1976) Hierarchical organisation and postural facilitation: rules for grooming in flies. _Animal behaviour_, 24, 739-755.

DELONG, A.J. (1974) Kinesic signals at utterance boundaries in pre-school children. _Semiotica_, 11, 43-73.

DITTMANN, A.T. AND LLEWELLYN, L.G. (1967) The phonemic clause as a unit of speech decoding. _Journal of Personality and Social Psychology_, 6, 341-349.

DONOHUE, W.A., HAWES, L.C. AND MABEE, T. (1981) Testing a structural-functional model of group decision-making using Markov analysis. Human Communication Research, 7, 133-146.

DORE, J. (1979) Conversational acts and the acquisition of language. In Ochs, E. and Schieffelin, B.B. (Eds.), Developmental Pragmatics, New York: Academic Press.

DUNCAN, S. (1972) Some signals and rules for taking speaking turns in conversations. Journal of Personality and Social Psychology, 23, 283-292.

DUNCAN, S. (1974) On the structure of speaker-auditor interaction during speaking turns. Language in Society, 2, 161-180.

DUNCAN, S. AND NIEDEREHE, G. (1974) On signalling that it's your turn to speak. Journal of Experimental Social Psychology, 10, 234-247.

DUNCAN, S. AND FISKE, D.W. (1977) Face-to-face interaction: Research, Methods and Theory. Potomac, New Jersey: Lawrence Erlbaum.

DUNCAN, S., BRUNNER, L.J. AND FISKE, D.W. (1979) Strategy signals in face-to-face interaction. Journal of Personality and Social Psychology, 37, 301-313.

ELLIS, D.G. AND FISHER, B.A. (1975) Phases of conflict in small group development: A Markov analysis. Human Communication Research, 1, 195-212.

ELLIS, R. AND WHITTINGTON, D. (1981) A guide to social skill training. London: Croom Helm.

EXLINE, R.V. AND WINTERS, L.C. (1965) Affective relations and mutual glances in dyads. In Tomkins, S. and Izard, C. (Eds.), _Affect, Cognition and Personality_, New York: Springer.

FAGEN, R.M. AND YOUNG, D.Y. (1978) Temporal patterns of behaviour: Durations, Intervals, Latencies and Sequences. In Colgan, P.W. (Ed.), _Quantitative Ethology_, New York: John Wiley.

FELDSTEIN, S. AND WELKOWITZ, J. (1978) A chronography of conversation: In defense of an objective approach, In Siegman, A.W. and Feldstein, S. (Eds.), _Nonverbal behaviour and communication_. Hillsdale, New Jersey: Lawrence Erlbaum.

FERGUSON, N. (1977) Simultaineous speech, interruptions and dominance. _British Journal of Social and Clinical Psychology_, 16, 295-302.

FINKBEINER, D.T. (1966) _Introduction to matrices and linear transformations_. San Francisco: W.H. Freeman.

FISHBEIN, M. AND AJZEN, I. (1975) _Belief, Attitude, Intention and Behaviour: An introduction to theory and research_. Reading, Mass.: Addison-Wesley.

FISHER, B.A. (1970) Decision emergence: phases in group decision-making. _Speech Monographs_, 37, 53-66.

FISHER, B.A. (1978a) _Perspectives on human communication_. New York and London: Macmillan.

FISHER, B.A. (1978b) Information systems theory and research: An overview. In Ruben, B.D. (Ed.), _Communication Yearbook II_, New Brunswick, New Jersey: Transaction Books.

FLANDERS, N.A. (1970) Analysing Teaching Behaviour, Reading, Mass.: Addison-Wesley.

FRANK, G.H. AND SWEETLAND, A. (1962) A study of the process of psychotherapy: The verbal interaction. Journal of Consulting Psychology, 26, 135-138.

FRICK, F.C. (1968) The application of information theory in behavioural studies. In Buckley, W. (Ed.) Sociology and modern Systems theory. Englewood-Cliffs, New Jersey: Prentice-Hall.

FRIES, C.C. (1952) The structure of English: An introduction to the construction of English sentences. New York: Harcourt Brace.

GARNER, W.R. (1962) Uncertainty and structure as psychological concepts. New York and London: John Wiley.

GARNER, W.R. AND McGILL, W.J. (1956) The relation between information and variance analyses. Psychometrika, 21, 219-228.

GILES, H. AND ST. CLAIR, R.N. (1979) Language and Social Psychology, Oxford: Basil Blackwell.

GIMSON, A.C. (1962) An introduction to the pronunciation of English. London: Edward Arnold.

GOFFMAN, E. (1972) Relations in public, Harmondsworth: Penguin Books.

GOODMAN, L.A. AND KRUSKAL, W.H. (1954) Measures of association for cross-classifications. Journal of the American Statistics Association, 49, 732-764.

GOTTMAN, J., NOTARIUS, C. MARKMAN, H. AND METTETAL, G. (1977a) Couples Interaction Scoring System (CISS) Research Report, University of Illinois.

GOTTMAN, J.M., MARKMAN, H. AND NOTARIUS, C. (1977b) The topography of marital conflict: A sequential analysis of verbal and non-verbal behaviour. Journal of Marriage and the Family, 39, 461-477.

GOTTMAN, J.M. AND BAKEMAN, R. (1979) The sequential analysis of observational data. In Lamb, M.E., Suomi, S.J. and Stephenson, G.R. (Eds.), Social Interaction Analysis. Wisconsin: University of Wisconsin Press.

GOURAN, D.S. AND BAIRD, J.E. (1972) An analysis of distributional and sequential structure in problem-solving and informal group discussion. Speech Monographs, 39, 16-22.

GUETZKOW, H. (1950) Unitising and categorising problems in coding qualitative data. Journal of Clinical Psychology, 6, 47-58.

HAKE, H.W. AND GARNER, W.R. (1951) The effect of presenting various numbers of discrete steps on scale reading accuracy. Journal of Experimental Psychology, 42, 358-366.

HALL, A.D. AND FAGEN, R.E. (1975) Definition of system. In Ruben, B.D. and Kim, J.Y. (Eds.), General systems theory and human communication theory. Rochelle Park, New Jersey: Hayden Book Company.

HALLIDAY, M.A.K. (1970) Language structure and language function. In Lyons, J. (Ed.), New Horizons in Linguistics. Harmondsworth: Penguin Books.

HARARY, F., NORMAN, R.Z., AND CARTWRIGHT, D. (1965) Structural Models: An Introduction to the Theory of Directed Graphs. New York and London: John Wiley.

HARE, A.P. (1976) Handbook of small group research. Second edition. New York: Free Press.

HARRÉ, R. AND SECORD, P.F. (1972) The explanation of social behaviour. Oxford: Basil Blackwell.

HAWES, L.C. (1972) Development and application of an interview system. The Central States Speech Journal, 23, 92-99.

HAWES, L.C. AND FOLEY, J.M. (1973) A Markov analysis of interview communication. Speech Monographs, 40, 208-219.

HAWES, L.C. AND FOLEY, J.M. (1976) Group decisioning: Testing a finite stochastic model. In Miller, G.R. (Ed.), Explorations in Interpersonal Communication, Beverly Hills and London: Sage Publications.

HERTEL, R.K. (1968) The Markov modelling of experimentally induced marital conflict. Unpublished Ph.D. thesis, University of Michigan.

HERTEL, R.K. (1972) Application of stochastic process analyses to the study of psychotherapeutic processses. Psychological Bulletin, 77, 421-430.

HEWES, D.E. AND EVANS-HEWES, D.E. (1974) Toward a Markov chain model of attitude change. Cited in Ruben, B.D. (Ed.), Communication Yearbook I, p.209. New Brunswick; New Jersey: Transaction Books.

HEWES, D.E., BRAZIL, A.J. AND EVANS, D.E. (1977) A comparative test of two stochastic process models of messages, mediating variables and behavioural expectations. In Ruben, B.D. (Ed.), Communication Yearbook I, New Brunswick, New Jersey: Transaction Books.

HOLSTI, O.R. (1969) Content analysis for the social sciences and humanities. Reading, Mass.: Addison-Wesley.

JACKSON, D.D. (1965) The study of the family, Family Process, 4, 1-20.

JAFFE, J., FELDSTEIN, S. AND CASSOTTA, L. (1967) Markovian models of dialogic time patterns. Nature, 216, 93-94.

JAFFE, J. AND FELDSTEIN, S. (1970) Rhythms of dialogue. New York and London: Academic Press.

JAFFE, J., STERN, D.N. AND PEERY, J.C. (1973) Conversational coupling of gaze behaviour in pre-linguistic human development. Journal of Psycho-linguistic Research, 2, 321-329.

JONES, E.E., AND GERARD, H.B. (1967) Foundations of Social Psychology, New York: John Wiley.

JOURARD, S. (1971) Self-disclosure: An experimental analysis of the transparent self. New York: John Wiley.

KAUFMANN, A. (1972) Points and Arrows: the Theory of Graphs. London: Transworld Publishers.

KELLER, E. (1981) Gambits: Conversational strategy signals. In Coulmas, E. (Ed.), Conversational Routine, The Hague: Mouton.

KEMENY, J.S. AND SNELL, J.L. (1960) Finite Markov chains. New York: Van Nostrand.

KENDON, A. (1967) Some functions of gaze direction in social interaction. Acta Psychologica, 26, 22-63.

KENDON, A. (1972) Some relations between body motion and speech. In Siegman, A.W. and Pope, B. (Eds.), Studies in Dyadic Communication, New York: Pergamon Press.

KENDON, A. (1978) Looking in conversation and the regulation of turns of talk: A comment on the papers of G. Beattie and D.R. Rutter et al. British Journal of Social and Clinical Psychology, 17, 23-24.

KIESLER, C.A., KIESLER, S.G., AND PALLACK, M.S. (1967) The effect of commitment to future interaction on reactions to norm violations. Journal of Personality 35, 585-599.

KRAEMER, H.C. AND JACKLIN, C.N. (1979). Statistical analysis of dyadic social behaviour. Psychological Bulletin, 86, 2, 217-224.

KRUEGER, D.L. (1979) A stochastic analysis of communication development in self-analytic groups. Human Communication Research, 5, 314-324.

KULLBACK, S., KUPPERMAN, M. AND KU, H.H. (1962) Tests for contingency tables and Markov chains. Technometrics, 4, 573-608.

LABOV, W. (1966) The social stratification of English in New York City. Washington, D.C.: Center for Applied Linguistics.

LAING, R.D. AND COOPER, D.G. (1971) Reason and Violence: A decade of Sartre's Philosophy 1950-1960. New York: Vintage Books.

LAKOFF, R. (1973) Language and woman's place. Language and Society, 2, 45-79.

LAZARFIELD, P.S. (1937) Psychological techniques of classification. In _The technique of marketing research_, American Marketing Society, Committee on Marketing Research Techniques. New York: McGraw-Hill.

LAZZERINI, A.J., STEPHENSON, G.M. AND NEAVE, H. (1978) Eye-contact in dyads: A test of the Independence Hypothesis. _British Journal of Social and Clinical Psychology_, 17, 227-229.

LEWIS, G.H. (1970a) Bales' Monte Carlo Model of Small Group Discussions. _Sociometry_, 33, 20-36.

LEWIS, G.H. (1970b) The assumption of stationary parameters in theories of group discussion. _Behavioural Science_, 15, 269-273.

LICHTENBERG, J.W. AND HUMMEL, T.J. (1976) Counselling as a stochastic process: Fitting a Markov chain model to initial counselling interviews. _Journal of Counselling Psychology_, 23, 310-315.

LONGABAUGH, R. (1963) A category system for coding interpersonal behaviour as social exchange. _Sociometry_, 26, 319-344.

LONGABAUGH, R., ELDRED, S.H., BELL, N.W., AND SHERMAN, L.J. (1966) The interactional world of the chronic schizophrenic patient. _Psychiatry_, 29, 78-99.

LYONS, J. (1977) _Semantics_, Vol.2. London: Cambridge University Press.

MARK, R.A. (1971) Coding communication at the relational level. _Journal of Communication_, 21, 221-232.

MATARAZZO, J.D. AND WIENS, A.N. (1972) The interview: Research on its anatomy and structure. Chicago: Aldine-Atherton Publishing Company.

McNEILL, D. (1979) The conceptual basis of language. Hillsdale, New Jersey: Lawrence Erlbaum.

MEHRABIAN, A. (1968) Relationship of attitude to seated posture, orientation and distance. Journal of Personality and Social Psychology, 10, 26-30.

MILLER, A.G. (1972) Role playing: An alternative to deception? A review of the evidence. American Psychologist, 27, 626-636.

MILLER, G.A. AND FRICK, F.C. (1949) Statistical behavioristics and sequences of responses. Psychological Review, 56, 311-324.

MISHLER, E.G. (1975) Studies in dialogue and discourse: II. Types of discourse initiated by and sustained through questioning. Journal of Psycholinguistic Research, 4, 99-121.

MISHLER, E.G. AND WAXLER, N.E. (1968) Interaction in families: An experimental study of family processes and schizophrenia. New York: John Wiley.

MORGAN, B.J.T. (1976) Markov properties of sequences of behaviours. Applied Statistics, 25, 1, 31-36.

MORLEY, I. AND STEPHENSON, G.M. (1977) The social psychology of bargaining. London: George Allen and Unwin.

MYERS, T. (1979) The development of conversation and discourse. Edinburgh: Edinburgh University Press.

NAEGELE, K.D. (1958) Friendship and acquaintance: an exploration of some social distinctions. Harvard Educational Review, 28, 3, 232-252.

NOFSINGER, R.E. (1975) The demand ticket: A conversational device for getting the floor. Speech Monographs, 42, 1-9.

NOFSINGER, R.E. (1976) On answering questions indirectly: Some rules in the grammar of doing conversation. Human Communication Research, 2, 172-181.

NORWINE, A.C. AND MURPHY, O.J. (1938) Characteristic time intervals in telephone conversation. Bell System Technical Journal, 17, 281-291.

ORE, O. (1963) Graphs and their uses. New York: Random House.

PARKS, M.R., FARACE, R.V., AND ROGERS, L.E. (1975) A stochastic description of relational communication systems. Cited in Penman, R. (1980) Communication Processes and Relationships, London: Academic Press.

PARKS, M.R., FARACE, R.V., ROGERS, L.E., ALBRECHT, T. AND ABBOTT, R. (1976) Stochastic process analysis of relational communication in marital dyads. Cited in Penman, R. (1980) Communication Processes and Relationships. London: Academic Press.

PARZEN, E. (1962) Stochastic Processes. San Francisco: Holden-Day.

PEARSON, E.S. AND HARTLEY, H.O. (1970) Eds. Biometrika tables for statisticians, Vol. 1. Third edition. Cambridge: Cambridge University Press.

PECK, A.J. (1981) The alignment of functions and notions and its significance for syllabus construction. The British Journal of Language Teaching, 19, 113-117.

PENMAN, R. (1980) Communication processes and relationships. London: Academic Press.

PRIDE, J.B. (1969) Analysing classroom procedures. In Fraser, H. and O'Donnell, W.R. (Eds.), Applied linguistics and the teaching of English. London: Longman.

QUIRK, R. AND GREENBAUM, S. (1976) A University grammar of English. London: Longman.

RAPPORT, A. AND HORVATH, W.J. (1968) Thoughts on organisation theory. In Buckley, W. (Ed.), Modern systems research for the behavioural scientist. Chicago: Aldine Publishing Company.

RAUSH, H.L. (1972) Process and change - A Markov model for interaction. Family Process, 11, 275-298.

RAUSH, H.L., MARSHALL, K.A. AND FEATHERMAN, J.M. (1970) Relations at three early stages of marriage as reflected by the use of personal pronouns. Family Process, 9, 69-82.

RESTLE, F. (1971) Mathematical models in psychology: An Introduction. Harmondsworth: Penguin Books.

ROBINSON, W.P. (1972) Language and Social Behaviour. Harmondsworth: Penguin Education.

ROGERS, L.E. AND FARACE, R.V. (1975) Analysis of relational communication in dyads: New measurement procedures. Human Communication Research, 1, 222-239.

RUTTER, D.R. AND STEPHENSON, G.M. (1977) The role of visual communication in synchronising conversation. European Journal of Social Psychology, 7, 29-37.

RUTTER, D.R., STEPHENSON, G.M., AYLING, K. AND WHITE, P.A. (1978) The timing of Looks in dyadic conversation. British Journal of Social and Clinical Psychology, 17, 17-21.

RUTTER, D.R. AND STEPHENSON, G.M. (1979) The role of visual communication in social interaction. Current Anthropology, 20, 124-125.

SACKETT, G.P. (1978) The lag sequential analysis of contingency and cyclicity in behavioural interaction research. In Osofsky, J. (Ed.) Handbook of infant development. New York: John Wiley.

SACKS, H. (1972) Unpublished lecture notes (mimeo) UCLA lecture 1, April 4, 1972.

SACKS, H., SCHEGLOFF, E.A. AND JEFFERSON, G. (1974) A simplest systematics for the organisation of turn-taking for conversation. Language, 50, 4, 696-735.

SCHEFLEN, A.E. (1964) The significance of posture in communication systems. Psychiatry, 27, 316-331.

SCHEIDEL, T.M. AND CROWELL, L. (1964) Idea development in small discussion groups. Quarterly Journal of Speech, 50, 140-145.

SEARLE, J.R. (1976) A classification of illocutionary acts. Language in Society, 5, 1-23.

SHANNON, C.E. (1948) A mathematical theory of communication. Bell Systems Technical Journal, 27, 379-423 and 623-656.

SHANNON, C.E. AND WEAVER, W. (1949) The mathematical theory of communication. Urbana, Illinois: University of Illinois, Press.

SIEGEL, S. (1956) Nonparametric statistics for the behavioural sciences. Tokyo: McGraw-Hill Kogakusha.

SINCLAIR, J.McH., AND COULTHARD, R.M. (1975) Towards an analysis of discourse. London: Oxford University Press.

SLUZKI, G.E. AND BEAVIN, J. (1965) Simetriay complementaridad: una definición operacional y una tipología de parejas. Acta Psiquiatrica y Psicología de America Latina, 11, 321-330.

STECH, E.L. (1970) An analysis of interaction structure in the discussion of a ranking task. Speech Monographs, 37, 249-256.

STECH, E.L. (1975) Sequential structure in human social interaction. Human Communication Research, 1, 168-179.

STECH, E.L. (1979) A grammar of conversation with a quantitative empirical test. Human Communication Research, 5, 158-170.

STEPHENSON, G.M., AYLING, K. AND RUTTER, D.R. (1976) The role of visual communication in social exchange. British Journal of Social and Clinical Psychology, 15, 113-120.

SUSTARE, D. (1978) Systems diagrams. In Colgan, P.W. (Ed.), Quantitative Ethology, New York: John Wiley.

SWINTH, R.L. AND TUGGLE, F.D. (1971) A complete dyadic process model of four-man group problem-solving. Organisational Behaviour and Human Performance, 6, 517-549.

THIBAUT, J.W. AND KELLEY, H.H. (1959) The social psychology of groups. New York and London: John Wiley.

THOMAS, A.P. AND BULL, P. (1981) The role of pre-speech posture change in dyadic interaction. British Journal of Social and Clinical Psychology, 20, 105-111.

TROWER, P., BRYANT, B. AND ARGYLE, M. (1978) Social skills and mental health. London: Methuen.

TUCKMAN, B.W. (1965) Developmental sequence in small groups. Psychological Bulletin, 6, 384-399.

VALENTINE, K.B. AND FISHER, B.A. (1974) An interaction analysis of verbal innovative deviance in small groups. Speech Monographs, 41, 413-420.

VAN HOOFF, J.A.R.A.M. (1973) Structural analysis of the social behaviour of a semi-captive group of chimpanzees. In Von Cranach, M. and Vine, I. (Eds.), Social communication and movement. New York and London: Academic Press.

VAN HOOF, J.A.R.A.M. (1979) Methods of Research in Nonverbal Communication. Unpublished mimeo.

VON BERTALANFFY, L. (1962) General systems theory - A critical review. General Systems, 7, 1-20.

VON BERTALANFFY, L. (1971) General systems theory. London: Allen Lane, The Penguin Press.

WATZLAWICK, P., BEAVIN, J.H. AND JACKSON, D.D. (1968) Pragmatics of Human Communication. London: Faber and Faber.

WAXLER, M.H. AND MISHLER, E.G. (1966) Scoring and reliability in Interaction Process Analysis: A methodological note. Sociometry, 29, 28-40.

WIEMANN, J. AND KNAPP, M. (1975) Turn-taking in conversations. _Journal of Communication_, 25, 75-92.

WILDEN, A. (1972) _System and Structure: Essays in communication and exchange_. London: Tavistock Press.

WILSON, C. (1974) _Watergate words: A naturalistic study of media and communication_. Research Report, Communications Studies Group. CSG: E/74240/CW.

YNGVE, V.H. (1970) On getting a word in edgewise. _Papers from the sixth regional meeting of the Chicago Linguistic Society_, pp. 567-578. Chicago: Chicago Linguistic Society.