Semantic and Social-Pragmatic Aspects of Meaning in Task-oriented Dialogue

Anthony Anderson

Thesis submitted in fulfilment of the Degree of Doctor of Philosophy in the Faculty of Science, University of Glasgow, November 1983.

PAGE NUMBERING AS ORIGINAL

DEDICATION

To my wife, Kay, and my parents, Smith and Teresa.

ACKNOWLEDGEMENTS

A considerable debt of gratitude is owed to a number of my colleagues for their help during both the experimental work and the preparation of this thesis. In particular, my supervisors, Professor Robert Farr and Dr. Simon Garrod, provided a great deal of information and guidance during the empirical phase of the work, and their suggestions regarding the organization of the thesis and their detailed comments on earlier drafts have resulted in a greatly improved product.

I would also like to acknowledge the help of Professor Tony Sanford, who also provided much information and encouragement.

Helpful discussions regarding the work were had with Michael Harker, Robert Henderson, Susan Jandoo, Bob Leiser, Angela Lucas, Kathy Moar, Andy Rutherford, and Paddy O'Donnell. Practical help in the running of the experiments was provided by Jim Mullin, Stuart Oliphant and Tony Cullen.

I would also like to express my gratitude to Tom Aitchison, Peter Breeze, Dr. Arthur Shirley and Dr. Ben Torsney for their advice and assistance regarding statistics.

A special debt of gratitude is owed to my wife Kay for her continued support and encouragement throughout both the empirical work and the writing of this thesis.

Finally, I would like to thank Sheena Bryson for typing the various drafts of the thesis.

C O N T E N T S

Page

ABSTRACT

The research reported in the present thesis constitutes an attempt to examine empirically the semantic and social-pragmatic aspects of the meanings of natural language expressions generated by subjects whilst participating in spontaneous task-oriented dialogues. Pairs of subjects were asked to solve a computer game task which involved them in moving their respective position markers through a _maze_. The successful solution of the task required the subjects to cooperate verbally, and the present research focusses on the meanings of subjects' descriptions of locations within the spatial networks of the various mazes.

The semantic analysis undertaken in the present work took the form of a classification study, using hierarchical cluster analysis, of the lexical items observed in the subjects' location descriptions. This analysis indicated that there are (at least) four basic ways of identifying locations within such a spatial network; it is argued that these four ways of describing locations could reflect four different underlying types of _mental model_ of the maze shape.

It was also argued that there is an important social-pragmatic component in the meanings of such spontaneous descriptions. Data were discussed which lend support to the view that (a) each pair of interlocutors constrain one another's choice of referring expression such that both

use similar types of referring expression, and (b) the interlocutors tacitly <u>negotiate</u> the meanings of the descriptions they generate. The final theoretical analysis of the data, therefore, was in terms of negotiated mental models of the maze.

Introductory Overview

The work reported in the present thesis constitutes an attempt to examine empirically certain aspects of the meanings of natural language expressions. The problem of meaning has proven to be a complex and exceedingly difficult one and is currently being tackled by theorists in a variety of disciplines, including linguistics, philosophy, psychology and artificial intelligence.

Like many empirical studies of meaning, the present research focusses upon one limited subproblem within the larger domain of enquiry, namely, the meanings of descriptions of locations within a spatial network generated by subjects during the course of an ongoing task-oriented dialogue. In particular, the present set of studies involved an investigation of the semantic aspects of such location descriptions (that is, the relationship between the location descriptions themselves and the stimuli being described) and their pragmatic aspects (that is, the relationship between the speakers and the types of description that they generate).

The advantages of studying such a narrowly-constrained domain as subjects' descriptions of locations within spatial networks come from the tight empirical control which can be exercised over the characteristics of the referent, and the empirical assessment of critical contextual variables. This high degree of control over relevant extralinguistic variables allows a more detailed analysis of the linguistic data than would otherwise be possible.

The advantage of studying task-oriented dialogues comes from the fact that subjects participating in such studies tend to become quite task-involved and consequently do not reflect upon the language that they produce, devoting their attention instead to the experimental task. The result of this is linguistic data of a truly natural and spontaneously-generated kind. The experimental task used in the present set of studies takes the form of a computer-controlled maze game, in which two subjects, who are located in separate soundproofed rooms, have to cooperate verbally (via a headset and microphone link) to assist one another to move their respective position markers through a maze (visible to each player on a computer terminal screen) from starting positions to goal positions. The verbal cooperation is required because various blockages exist within each player's maze which impede his progress towards his goal, and the mechanism by means of which these blockages are removed involves a given player's (e.g. player A) partner (player B) reporting his position within the maze and subsequently moving into one of player A's <u>switch nodes</u> visible to player A on his screen. This mechanism is discussed more fully in chapter 4; suffice it to say that a given player requires <u>his partner</u> to clear blocked pathways for him, and this partly involves his partner in describing his location within the maze. As a result, the rules of the game require that each player periodically reports his location to his partner, and these spontaneously-generated descriptions of maze location constitute the data base for the present set of studies.

The present thesis consists of two parts. Part I, comprising chapters 1 to 4, is a review of some of the extensive literature on the topic of meaning, and Part II, comprising chapters 5-11, is a report of the empirical work carried out using the game task described briefly above.

Morris's (1938) division of semiotics (i.e. the study of sign systems, of which natural language is an example) into syntactic, semantic and pragmatic aspects is used as a suitable structuring with respect to which the variety of theories and empirical studies of meaning can be discussed in Part I of this thesis. Thus, chapter 1 is a review of what could be described as 'syntactic' theories of meaning (that is, those theories which deal primarily with meaning relations between different signs in the language); chapter 2 is a review of those theories of meaning which could be described as 'semantic' theories of meaning (that is, those theories which focus upon the relation between linguistic signs and their referents); and chapter 3 is a review of theories of meaning which could be described as 'pragmatic' theories of meaning (that is, those theories which focus primarily upon the relations between the particular interlocutors involved and what is understood by what they say).

Morris's distinction thus provides a tripartite division with respect to which a variety of genuinely different approaches to meaning can be classified (although it is rare that a given theory is related exclusively to one of these categories).

This (by no means exhaustive) review of previous theoretical and empirical work on this problem will underline both the complexity of the issues involved and the variety of theoretical approaches which have been adopted.

The final chapter of part I of the present thesis is a detailed exposition of the experimental maze game task employed in the present work to elicit spontaneous dialogue. The maze game task is discussed in the context of previous experimental studies which also used the methodology of seeking to constrain the topic of discourse and eliciting spontaneous dialogue on that topic.

Part II contains details of the various experimental studies conducted and reports the results obtained. Chapters 5-10 present descriptions of the data obtained using the task and the statistical analysis of these data.

Whilst the present work constitutes an attempt to deal with both the semantic and the pragmatic aspects of the meanings of the subjects' descriptions of locations within the mazes, the particular theoretical standpoint from which the data are interpreted is a social-pragmatic one. Thus, whilst four basic ways of describing locations within a maze are described in chapters 6 and 7, further complexities of a purely pragmatic nature arise: different dyads of subjects would appear to place very different interpretations on the same description. Thus, the importance of what individual pairs of subjects understand by a particular expression, and how each dyad selects one sort of interpretation of the expressions they use rather than another, will be emphasized.

In chapter 9, an experiment is described in which subjects were set the task of describing locations within mazes but doing so outwith the context of playing the maze game. In this experiment, a rather different distribution of the four basic ways of describing locations within mazes observed in the computer game studies was obtained, with one of the types of description observed in the computer game studies' data being entirely absent in the 'game-independent' location description task. The discussion in chapter 9 of the possible reasons for this difference in the distributions of the different types of location description in the two tasks highlights some of the subtle and complex problems of studying dialogue.

The thesis concludes, in chapter 11, with a discussion of the experimental results in relation to some of the theories of meaning which were discussed in part I. In chapter 11, both the potential wider significance of the present set of results, and the further issues raised by the results of the studies reported in this thesis, are considered.

Part I

Theoretical Approaches to the Problem of Meaning

CHAPTER 1

'SYNTACTIC' APPROACHES TO MEANING

a) Introduction

The study of the broad field of sign systems (including natural language, letters and numerals, diagrams, pictures, sketches, gestures and facial expressions, and tokens) is called semiotic. Morris (1938), following the philosophy of Charles Pierce, distinguished between three different levels of semiotic, representing different degrees of abstraction, as follows:

i) SYNTACTICS constitutes the study of signs and the relations between signs;

ii) SEMANTICS is the study of the relation between signs and what they designate (i.e. the designata); and

iii) PRAGMATICS, which is the study of the signs in relation to the users of those signs.

This distinction between three sub-areas of study of the sign systems in language has been used down to the present day - see, for instance Cherry (1957), and Lyons (1977). According to Morris, these three levels are not separate from one another, but are inclusive: they overlap one another such that the study of syntactic relations between signs is the highest level of abstraction from actual usage and is directed to the signs themselves and their orderings, studying purely formal aspects, and is wholly subsumed within the area of semantics.

Correspondingly, both syntax and semantics are necessary parts of the study of pragmatics (see diagram below).

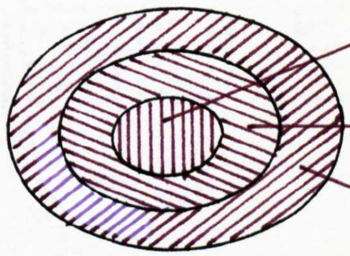


Figure 1.1 The three levels of semiotic (shown schematically as successive abstractions) (From Cherry, 1957, p. 222).

The 'rules' studied at the three levels are inherent in the analysis of language rather than being inherent in the language itself (Cherry, 1957) and are expressed in meta-language (the distinction between object language and meta-language corresponds to the distinction between what is described - the object language - and the meta-language in which the description is couched: see chapter 2). Pragmatics, notes Cherry (1957) is the "real-life" level. Semantics abstracts from all specific communication events and concerns only signs and their designata (qualities, objects, actions etc.). Syntactics abstracts still further and concerns signs only: it treats language as a calculus.

The term 'semantics' has since become used to refer to theories of meaning, an area of great interest to philosophers, psychologists and linguists and an area of central concern to this thesis. Not all theories of semantics are 'semantic' according to the above definition; there have been theories of meaning which have considered 'semantic' phenomena in terms of relations between the signs (or rather, between the hypothesized mental representations of the signs). Such theories are, therefore, more properly termed 'syntactic' theories of meaning. There have, likewise, been 'semantic' theories of meaning which attempt to relate the signs to their designata or mental models of their designata. In recent years, pragmatic approaches to the study of meaning have also been mooted (see, for example, Rommetveit, 1974). These three levels of tackling the problem of meaning will be considered in turn. In chapter 1, theories of an essentially 'syntactic' nature will be considered: in chapter 2, theories of a 'semantic' nature will be considered. Miller and Johnson-Laird (1976) contrast the two positions of 'intensionalism' (which considers the relation between signs) and 'extensionalism' (which considers relations between the signs and their designata) and contend that both approaches pursued independently must fail, and that the failures are related since each approach contains too much of what the other lacks. They advocate an integration of the two approaches, and an exposition of their theoretical framework will be given in chapter 2. In chapter 3, theoretical approaches which explicitly challenge traditional linguistics-based approaches to communication will be considered. These theories emphasize the importance of pragmatic factors in the understanding of language, and contend that an emphasis on linguistic-type analyses effectively ignore phenomena of central importance.

b) Historical Introduction

The meaning of words and larger expressions has fascinated inquirers for many hundreds of years, from the time of the ancient Greeks until the present day, and has been the source of many theoretical disputes. The fundamental question of "what is meaning?" has engendered a great deal of speculation. As Katz (1973) puts it,

> "We find, historically, a variety of direct answers, including Plato's answer that meanings are eternal archetypes, John Locke's answer that meanings are the mental ideas for which words stand as external signs, the answer that meanings are the things in the

world to which the words refer, Wittgenstein's answer that meaning is use, the behaviourist's answer that meanings are mental images associated with verbal behaviour, and so on. But every attempt to give a direct answer has failed. Some, such as the Platonic answer proved too vague and speculative. Others gave the wrong answer".

(Katz, 1973, pp 36-37)

That such a variety of answers has been given to the fundamental question serves to underline the fact that it is a question of interest not only to linguists, but also to philosophers and psychologists. Philosophers, in particular, have always been interested in meaning, since it is necessarily involved in vital and notoriously controversial philosophical issues, including the nature of truth (Lyons, 1968).

When considering what the meaning of a word might be, we soon end up embroiled in philosophical controversy. Lyons (1968) gives the example of the word cow. What is the meaning of the word cow? It cannot be a particular individual animal. Is there a set of properties which distinguishes cows from other objects (for which we have different words)? Immediately we are caught up in the philosophical controversy between nominalism and realism. Nominalism is the position that the set of referents to which we apply the same name have nothing in common other than the name which we have learned to apply to them. In contrast, realism is the position that the things to which we apply the same name have some common 'essential' properties by which we identify them. Thus, as soon as we begin to consider meaning, we become involved with philosophical issues.

The term 'meaning' has itself a number of distinguishable meanings. Lyons (1977) gives some examples: some of the senses of 'mean' can be enumerated as follows:

i) as synonymous with 'intend' (as in 'I did not mean to hurt you').

ii) as equivalent to 'significance' or 'value' (as in 'Life without Faith has no meaning').

iii) a relation of reference (as in 'It was John I meant, not Harry').

Lyons concentrates upon the sense as exemplified by iii) above.

Katz (1972) considers the question "What is meaning?" to be at the same level as the question "What is matter?" and accordingly advocates the theoretical tackling of more manageable subproblems which themselves are related to the larger problem. The breakdown of the general question into more specific questions which he lists results in a listing of phenomena which a theory of meaning would have to be able to explain, namely:

1) What is synonymy (sameness of meaning) and paraphrase?
2) What is semantic similarity (e.g. "aunt", "nun", "cow", "filly" and "actress" all share a common component) and semantic difference?
3) What is antonymy? (i.e. incompatibility of meaning - for example, "whisper" and "shout" are incompatible in meaning).
4) What is superordination? (For example, "human" and "boy").
5) What are meaningfulness and semantic anomaly? (For example, compare "a smelly soap" and "a smelly itch").
6) What is semantic ambiguity? (That is, multiplicity of sense - e.g. "button", "foot").
7) What is semantic redundancy? (An example of this would be where a phrase contained superfluous semantic information, such as the phrase "naked nude").
8) What is semantic truth? (That is, truth by virtue of meaning, or analyticity - such as "Uncles are males").
9) What is semantic falsehood? (That is, contradiction, as in "Kings are females").
10) What is semantically undetermined truth or falsehood? (For example, syntheticity - that is, where the truth or falsity of a sentence is not settled on the basis of its meaning alone but depends upon what is the case in actuality).

11) What is inconsistency? (For example, where "John is alive" and "John is dead" refer to the same individual, one sentence is necessarily true and one is necessarily false; they cannot both be true nor both be false).

12) What is entailment? ('Entailment' is where two sentences are related by virtue of their meaning such that one follows necessarily from the other - for example, "The car is red" entails "The car is coloured").

13) What is presupposition? ('Presupposition' is where one sentence presupposes the truth of another - for example, the sentence "When did you stop beating your wife?" presupposes the truth of a declarative to the effect that the listener had previously indulged in wife-beating).

14) What is a possible answer to a question? (For example, given a question like "When did John arrive?", possible answers include "John arrived at noon" and "John arrived a minute ago" but not "John loves to eat fruit").

15) What is a self-answered question? (For example "What colour is my red wagon?" expresses a question that is answered in the asking).

The above phenomena are all due to the meaning of the expressions in the examples. A semantic theory, Katz argues, must explain why the meaning of a linguistic construction makes it a case of a certain property or relation, or makes it exhibit the phenomena of synonymy or ambiguity, etc.

Katz (1972, p. 10) concludes from the foregoing that " . . . once we construct a theory that can successfully explain a reasonably large portion of . . . semantic phenomena, we can base our answer to "what is meaning?" on what the theory had to assume meaning was in order to provide its explanations".

Clearly, the question of what meaning is requires a complicated answer and a variety of phenomena and meaning relations require to be explained. As noted above, many issues and distinctions are raised by a consideration of meaning. One such distinction is that between the <u>intension</u> of a word or phrase, and its <u>extension</u> (sometimes expressed as a distinction betwen <u>sense</u> and <u>reference</u>). The <u>intension</u> or <u>sense</u> of, for example, a word, is the set of properties which determine the applicability of the word: the <u>extension</u> or <u>reference</u> of the word is the class of things to which the word may be applied. Thus, the intension is the 'intrinsic' meaning of a word and its relation to other words - for example, the word 'chair' has a meaning which is related in some ways to other words like 'table', 'bench', 'wardrobe' etc. The extension of the word is the entire class of entities to which it can refer, in any real or imagined world. So, in the above example, the extension of 'chair' is any real or imaginary chair.

Many theories of meaning (especially those based on linguistic considerations and which will be described in this chapter) have concerned themselves exclusively with the intensional aspect of meaning. Extensional considerations, however, become important when words are considered _in use_ because then the relation of language to the world becomes important. Extensional theories of meaning have been formulated particularly in recent years with the advent of procedural semantics and the postulation of theories of comprehension in terms of mental models of the reference situation (see chapter 2).

An early psychological theory of meaning which retains theoretical appeal today was that advanced by C. K. Ogden and I. A. Richards (1972, originally published 1923). Their fundamental argument was that the relation between words and referents is _indirect_: ". . . we need a theory which connects words with things through the ideas, if any, which they symbolize. We require, that is to say, separate analyses of the relations of words to ideas and of ideas to things . . ." (Ogden and Richards, 1972, p. 7). They attribute a great deal of the confusions and arguments among theorists to their overlooking the indirectness of this relationship. This point is summarized in a diagram taking the form of a triangle, which has become known as the "semiotic triangle" (Lyons, 1968, p. 404).

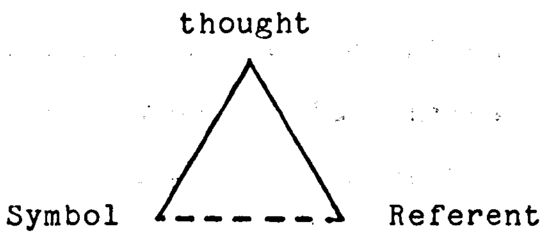


Figure 1.2: The 'semiotic triangle'

(Adapted from Ogden and Richards, 1972, p. 11.

Key: ________ direct links - - - - - indirect link)

Ogden and Richards argue that words mean nothing by themselves: it is only when a thinker makes use of them that they stand for anything or have 'meaning'. Words are instruments. The indirect relation between the word (or symbol, to use the more general term) is illustrated by the above triangular diagram. Between the thought and the symbol is the 'symbolizes' relationship, a causal relationship; between the thought and the referent there is also a direct relation. Between the symbol and the referent, however, there is only an indirect relation which consists in the symbol being used by someone to stand for the referent. The base of the triangle (shown dotted in the above diagram) is an inferred relationship. The symbol and the referent are not connected directly but instead are connected via the other two sides of the triangle. The direct connection represented by the dotted line is an inferred relation, and to treat this inferred relation as being of the same nature as the other two direct relations is the " . . . fundamental and most prolific fallacy . . ." (Ogden and Richards, 1972, p. 15). They point out in a footnote that philosophers, psychologists and logicians are wont to confuse the imputed relation between a sign and that to which it refers either with the referent or with the processes in the mind of the interpreter, when speaking of the meaning of a sign. They comment that it is this sort of confusion which had made so much previous work on the subject of signs and their meanings unfruitful.

Ogden and Richards' characterization of the word-referent relation as an indirect relation via the thoughts of the understander has its modern counterpart in current psychological approaches to meaning; Sanford and Garrod (1981) in their preface to a book on understanding written language, for example, comment that " . . . the squiggles on the page can live only through the agency of the minds of readers" (p. xiv). Likewise, if for "thought" or "reference" at the apex of Ogden and Richards' triangle we substitute instead "conceptual system" or "mental models" we have the basis of many current approaches to meaning.

That a characterization of the operations of the mind of a reader/listener would be fundamental to a plausible theory of meaning may today seem uncontroversial; it was, however, challenged by theorists in the past. In particular, the doctrine of Behaviourism in psychology attempted to rid psychology of such unobservable entities as 'mind'; the Behaviourists, led most notably by B. F. Skinner, argued against the reification of meaning in the mind of the language user and contended intead that words were mere verbal responses which were under the stimulus control of the environment of the speaker. This antimentalistic outlook was shared by a school of linguists led by Bloomfield. For the Bloomfieldian linguists, the meaning of a word was a full 'scientific' description of its referent. Lyons (1968) points out that this approach is more detrimental to progress in semantics than the definition of meaning in terms of 'concepts' which it attempted to supplant, since the Bloomfieldian approach gives preferential treatment to the relatively small set of words which refer to 'things' describable, in principle, by the physical sciences. Furthermore, it rests upon two tacit and unjustified assumptions, namely - (i) that a 'scientific' description of the referents of these words is relevant to the way in which the words are used by speakers on the language; and (ii) that the meaning of all words is ultimately describable in the same terms (Lyons, 1968, p. 408).

Skinner's extreme form of behaviourism was accompanied by a less extreme manifestation of the same doctrine, in the form of Osgood's (Osgood, Suci and Tannenbaum, 1957) mediation theory of meaning, which characterized meanings as symbolic mediation processes. These are unobservable meaning responses to words. They represent only a part of the overt response that would have been made to the word's referent, and in turn stimulate appropriate responses to the word. This type of learning theory approach was popular as a psychological theory of meaning during the 1950's. Osgood et al's approach to meaning also had an impact on the development of social psychology through Osgood and Tannenbaum's congruity theory. Congruity theory linked the mediation theory of meaning, and in particular the use of its associated measuring instrument, the semantic differential, to the study of persuasive communication and attitude change.

A revolution in thinking about language took place following the publication of Chomsky's book 'Syntactic structures' in 1957. In this book, Chomsky argued that the best way of explaining a speaker/listener's ability to understand language was in terms of a system of internalized rules. Chomsky demonstrated mathematically that for natural languages, the number of permissible word strings (that is, grammatical sentences) is infinite. It therefore follows that a descriptive grammar has to be projective, that is, it must be couched in terms of rules which will generate all of the infinite number of grammatical sentences, and no non-sentences. The notion of there being an unlimited number of grammatical sentences undermined the behaviourists' approach to language. As Greene (1972) puts it, "The arguments used by Chomsky and his supporters were designed to show, first, that learning theory is in principle unable to account for the speaker's ability to use language, and second, that in any case, acquisition of stimulus-response probabilities would be a wildly uneconomical explanation of language learning" (Greene, 1972, p. 15). Chomsky's emphasis on the need to characterize a speaker's ability to use language as a system of internalized rules caused something of a revolution in psychological theorizing: the notion that cognitive functions could best be treated as centrally planned, rule-governed mental operations swept away behaviourist antimentalistic theorizing and brought mental operations into the limelight once more.

Chomsky's two-tier theory for the description of language, involving surface syntactic structures, deep syntactic structures and transformational rules for deriving these from one another, was taken up by some psychologists as a possible psychological model of language use. Thus, the psychological reality of transformational grammar was tested, experiments being conducted which attempted to ascertain whether transformationally complex sentences were also psychologically more complex and consequently required more processing time, or occupied more space in memory, etc. (For example, Savin and Perchonok, 1965: Miller and

McKean, 1964). Chomsky also distinguished between _competence_ (the abstract, underlying form of linguistic knowledge possessed by a speaker) and _performance_ (the actual use of the language in concrete situations). Performance reflects competence in only indirect ways: memory span limitations, fatigue, distractability, switching of attention etc. affect linguistic performance in ways which deviate from a competence-based prediction. Slobin (1971) identifies the task for the linguist as that of characterizing competence. The psychologist, on the other hand, has to cut through the maze of performance factors in order to convince himself that competence has "psychological reality", and also examine the psychological factors which cause performance to deviate from competence. Kintsch (1974) rejects this neat division of labour, maintaining that as long as a linguistic theory is strictly a competence theory it is of no interest to the psychologist.

Chomsky's theory emphasized syntactics heavily. It also heavily emphasized the study of isolated sentences _in vacuo_. So, for example, a typical task set to a subject would be to give an intuitive judgement as to the grammaticality or otherwise of a set of word strings, one string at a time. This emphasis on having subjects judge the grammaticality of isolated sentences abstracted from communication settings and having no significance for the subject was subsequently attacked, particularly by

Rommetveit (1974; see chapter 3).

Broadbent (1973) also attacked Chomsky's theory from the standpoint of empiricism (learning theory) and simultaneously defended learning theory accounts of language acquisition. Specifically, Broadbent denied the belief that Chomsky's description of language has any psychological reality, and he also denied the transformationalists' claim that language is innate. He stated his admiration for the elegance of Chomsky's mathematical description of language but, as he put it (p. 193), "Elegance of mathematical description of the effect of an operation does not bear any necessary relation to the empirical facts about the operation itself". Thus, his argument against Chomsky's theory is that, while it describes language by being able to generate all of the grammatical sentences in the language, it is quite possible that the actual mechanism with which people produce language could be very different.

He also attacked the transformationalists' claim that, because of the potentially infinite number of grammatical sentences in a natural language, it could not be acquired using the mechanism of learning but instead had to be considered as an innate capacity. Broadbent countered this claim by pointing out that most learning theorists, including those interested in animal learning, believe that what is learned is not an association between particular concrete events but instead an association between conceptual classes, each of which may have many members. A word or a short sequence of words is treated as a member of a class, and sequences of classes would also be learned. Broadbent's emphasis on language expressions being learned as conceptual classes rather than as concrete entities undermines Chomsky's criticism of learning theories being unable to account for language acquisition.

Chomsky's theory did constitute a revolution at the time it was advanced and the 'psycholinguistic' approach which attempted to test the psychological reality of the theory constituted a "far cry from (the) simpliste pre-Chomskyan attempts to look at verbal associations and so on" (Greene, 1972, p. 196). That the search for an exact one-to-one relationship between grammatical rules and subjects' performance failed does not diminish the revolutionary impact of the theory: it directed to the attention of psychologists the influence of many unexpected factors, and underlined the complexity of language.

At least one influential theory of meaning, that of Katz and Fodor (1963), was developed specifically to dovetail into the transformational grammar framework as the 'semantic component' required by, but not specified in, the theory. Katz and Fodor's approach focussed upon intensional relations between lexical items and so could fairly be described as a 'syntactic' theory of meaning.

Katz and Fodor's theory was in fact subsequently accepted by Chomsky as a suitable theory of the semantic component for incorporation within his general theory. Another two 'intensional' approaches to meaning were also formulated during the 1960's/early 1970's. One approach, based on 'meaning postulates' has its origin in extensional theories of meaning of a type known as model-theoretic semantics (which will be described in detail in chapter 2). The third 'syntactic' approach to meaning to be considered in this chapter is Collins and Quillian's semantic network theory, which owes its origins to attempts at computer modelling of semantic memory (Quillian, 1968; Collins and Quillian, 1972). These three theories will each be reviewed in turn.

c) Three 'intensional' theories of meaning

i) Katz and Fodor's (1963) theory of meaning

The approach outlined by Katz and Fodor has as its origins the componential analyses which were first evolved in anthropological linguistics as a means of studying relations between words within a semantic field (i.e. a collection of semantically related lexical items such as kinship terms). The aim in componential analysis is the reduction of a word's meaning to its ultimate contrastive elements; for example, 'Man' has as its definition the collection of semantic elements + MALE, + ADULT, + HUMAN, whilst 'woman' has as its definitiom the set of semantic elements - MALE, + ADULT, + HUMAN, and so on, each word in the semantic field having a definition chracterized in terms of these elements.

Thus, the central tenet of decompositional theories of meaning is simply that the meaning of an individual word is a composite of 'atoms' of meaning in the form of semantic markers or semantic primitives. The semantic interpretation of a sentence can be obtained by replacing its words with their dictionary definitions and combining them according to the syntactic relations of the sentence. The process of combining word meanings is sensitive to the constraints, or selection restrictions, that a word may place upon the meaning of the other words with which it can be combined. The theory thus attempts to account both for the meanings of words and the meanings of larger linguistic structures.

This basic notion that in the course of understanding, we decompose the linguistic input into the appropriate set of semantic elements or markers has been used in several theories. One of the most famous explications of the idea was that of Katz and Fodor (1963). In their article, "The structure of a semantic theory", they outlined the general characteristics of a theory of meaning rather than advancing one specific theory. Their approach was in terms of semantic markers, and they indicated some of the phenomena which a semantic theory would be required to explain, such as synonymy, syntheticity, etc., outlined earlier.

Such relationships do require explanation by a theory of meaning, and Katz and Fodor explain these relationships by suggesting that, for example, synonymy can be explained in terms of the relevant lexical items sharing one <u>sense</u>. Accordingly, their outline of the form of a semantic theory concentrates on explicating the notion of how a sense could be represented theoretically and they explain the various meaning relationships between words in terms of sense relations of one kind or another.

As noted above, Katz and Fodor's theory is explicitly designed to dovetail into the 'semantic component' theoretical space left by Chomsky's syntactic analysis, or as Katz and Fodor put it, "Synchronic linguistic description minus grammar equals semantics" (p. 172). They take as their starting point the observation that much of the understanding of sentences is left unexplained by grammar, and this is where their theory of semantics fits in. Some examples of ambiguity, such as "The bill is large" cannot be characterized grammatically, but instead are best characterized in terms of meaning. The ambiguity in this sort of example takes the form of there being two different senses (that is, two different interpretations) of the word 'bill'; it can be taken to mean either an invoice for money or as the beak of a bird.

Katz and Fodor give as an example the decomposition of the word "bachelor" into its four senses (a man who has never married, a young knight serving under the standard of another knight, a person possessing the first or lowest academic degree, and a young male fur seal when without a mate during the breeding time). They represent these four senses in the form of a tree diagram. The first entry for 'bachelor' is the syntactic marker 'noun', to indicate that only those senses of a word classified as nouns are being considered (this can be an important distinction for many words, such as run, dog and chase, which can either be nouns or verbs). Proceeding down the tree, the next entries are the semantic markers which represent the meaning of the word in terms which are used to systematically define many words in the dictionary: (human), (male), (animal), etc. Also in the dictionary entry are semantic distinguishers, which distinguish only between senses of the one word, in this case 'bachelor'. These appear in square brackets at the lowest part of the diagram:

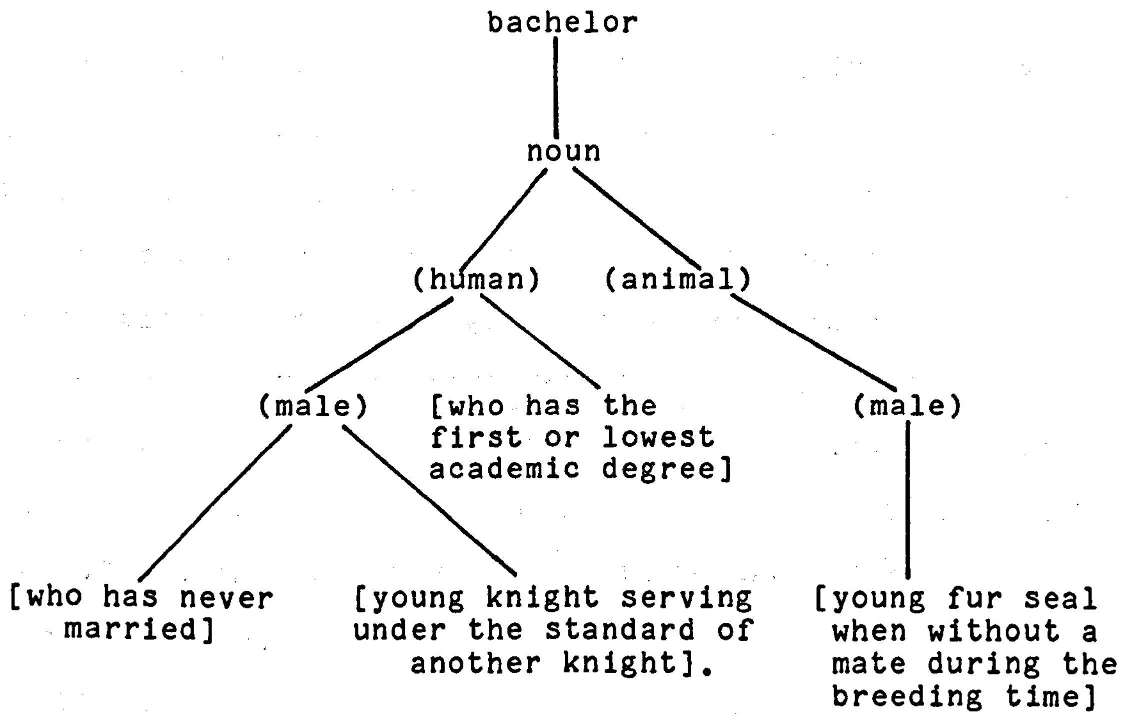


Figure 1-3 Katz and Fodor's tree diagram representing the four senses of 'bachelor'. (From Katz and Fodor, 1963, p. 186).

As Katz and Fodor (1963) p. 185-6) put it "The semantic markers and distinguishers are the means by which we can decompose the meaning of one sense of a lexical item into its atomic concepts, and thus exhibit the semantic structure _in_ a dictionary entry and the semantic relations _between_ dictionary entries. That is, the semantic relations among the various senses of a lexical item and among the various senses of different lexical items are represented by formal relations between markers and distinguishers" (emphasis in original).

Thus, the semantic markers are the atoms of meaning with which to characterize sense relations between different lexical items, whilst the distinguishers serve to relate the different senses of a single lexical item. The semantic markers are the constructs which characterize that part of the meaning of a lexical item which is systematic for the language. The distinguishers, on the other hand, deal with that part of the meaning of a lexical item which do not enter into theoretical relations within a semantic theory - that is, they characterize that part of meaning of which a semantic theory offers no general account.

Where meanings in the distinguishers make more general distinctions than merely between alternative senses of one word, Katz and Fodor suggest (p. 195-196) that the distinguishers can be reanalyzed to yield semantic markers - for example, 'young' could appear as a semantic marker for the second and fourth sense of the word (see diagram, above).

Dictionary entries also contain, in addition to syntactic markers, semantic markers and semantic distinguishers, <u>selection</u> <u>restrictions</u> which limit the possible senses of the other words with which the given lexical items can be combined. Thus, to borrow Katz' (1973) example, a selection restriction would be exemplified by the case where the noun which "waterproof" modifies must contain the semantic marker (Physical Object), otherwise semantic anomalies of the type illustrated by the phrase "waterproof shadows" would result.

As noted earlier, the decompositional theory also attempts to account for the meanings of larger lingustic constructions such as phrases and sentences. It achieves this by means of using Frege's "principle of compositionality" which holds that the meaning of a syntactically complex constituent such as a sentence, is a compositional function of the meanings of its parts. This is realised by using projection rules in addition to the decompositional dictionary. "The projection rule specifies how lexical readings for the syntactic atoms can be combined to form derived readings for a whole expression or sentence". (Katz, 1973, p. 43). Thus, the meaning of a phrase or sentence is a compositional function of the meaning of its parts. Furthermore, the meaning of the parts can in turn be specified by dictionary entries.

This type of dictionary theory employing decomposition into semantic elements can be quite a useful way of characterizing meaning relations between words. Intuitively, such decompositional analyses do seem to capture something of word meanings. However, certain subtleties are ignored by such an analysis. To use Sanford and Garrod's (1981) example, the word 'spinster' might be broken down by componential analysis into (adult), (woman), [who has never married]. However, in common usage, 'spinster' generally is taken to convey the feature 'middle aged', or 'elderly' also. There seems to be no such constraint on 'bachelor', intuitively speaking, so characterizing the difference in meaning between 'bachelor' and 'spinster' in terms of the ±MALE marker fails to capture certain subtleties. Similarly, as Clark and Clark (1977) have pointed out, the word 'girls' can be used sometimes to refer to adult women, as in 'the girls in the typing pool'. Such subtleties fail to be captured by a conventional decompositional theory.

Other criticisms have been levelled at this type of theory, and the criticisms are summarized by Johnson-Laird (1981a) as follows:

1. Kintsch (1974, p. 12) argues that "if one starts to decompose, it is hard to see where to stop". To quote Johnson-Laird (1981a, p. 107) . . . "(Kintsch's) point seems no more cogent here than it would be in opposing the atomic theory of matter".
2. Fodor, Fodor and Garrett (1975) argue that the rapidity with which comprehension occurs is not easily explicable if comprehension involves replacing words with their dictionary entries, combining the dictionary entries in accordance with the selection retrictions, etc. However, this argument carries little weight, since many complicated mental processes occur with great

rapidity.

3. Fodor, Fodor and Garrett claim that the decompositional theory makes erroneous predictions about the comprehension of sentences - for example, predicting that sentences involving 'bachelor' should be harder to understand and occupy more space in memory than sentences using the synonymous phrase 'is married'. Fodor, Fodor and Garrett investigated the times taken to evaluate deductive arguments, some involving explicit negation ('is unmarried') others involving lexical items like 'bachelor' with a putative semantic representation containing a negation. Such items (that is, those like 'bachelor' with a putative semantic representation containing a negation) made an inference reliably easier to evaluate than one containing an explicit negation, from which Fodor, Fodor and Garrett conclude that a word such as 'bachelor' does not seem to contain a negative in its semantic representation although it ought to according to the decompositional theory. Such evidence is not, however, conclusive. Fodor, Fodor and Garrett's conclusion that the representation of 'bachelor' does not include a negative cannot be justified by a difference in latency of response; it is possible that a negation in a dictionary

entry may be responded to faster than a morphological or explicit negation.

4. Fodor, Fodor and Garrett argue that decompositional dictionary entries are unable to capture certain inferences that depend upon the meanings of words. For example, in the case of KILL, if 'kill' is replaced by a decompositional entry of 'cause to die', then the inference from 'x causes y to die' to 'y dies' requires other machinery. Miller and Johnson-Laird (1976, p. 506) present a decomposition of 'cause' from which this inference does follow, however.
5. Kintsch (1974) claims that, if decomposition was invariably required in extracting meaning from linguistic input, it is difficult to see why complex words would have evolved, only to be decomposed again in the process of comprehension. Johnson-Laird (1981a) comments that this is a sensible point, but contends that there may be economies to be obtained by processing one semantically complex word rather than a synonymous string of simple words. Johnson-Laird concludes overall that the case against decompositional theories of meaning is "hardly overwhelming" (p. 109). Decomposition has also been advocated by Miller and Johnson-Laird (1976) and Schank (1975).

ii) The Meaning Postulate Theory of Semantics (Kintsch, (1974), Fodor, Fodor and Garrett (1975))

The central notion underlying this theory is that meaning relations between lexical items is best captured by means of a meaning postulate. The meaning postulate was a device introduced by Carnap (1956) into the (extensional) theory of semantics known as model-theoretic semantics. A meaning postulate is essentially a rule which relates the meanings of two lexical items or expressions. Thus, in 'possible-worlds' semantics (a variant of model-theoretic semantics), a meaning postulate is introduced in order to eliminate some, otherwise possible, worlds. So, for example, a meaning postulate which expresses the synonymity of 'seek' and 'try to find' eliminates those possible worlds in which one may seek something without trying to find it.

An example of a meaning postulate would be that which relates the meanings of 'buy' and 'sell', as follows:

for any x, y, and z, x sells z to y $\equiv$ y buys z from x

Bar-Hillel (1967) argued that decompositional theories cannot explain the semantic relation between 'buy' and 'sell' and that the best way of explicating this meaning relation is in terms of a meaning postulate. This replacing of dictionary entries by meaning postulates has formed the core of a psychologically oriented theory of meaning advocated by Kintsch (1974). The meaning postulate in Kintsch's system characterizes the semantic relations between items in semantic memory. Kintsch argues against lexical decomposition and provides experimental evidence against the decomposition of lexical items. He suggests that lexical concepts exist as unanalyzed wholes, and their logical entailments can be pursued by means of meaning postulates. Thus, Kintsch postulates a semantic memory consisting of lexical concepts, which he represents in a capitalized form, e.g. KILL. These lexical concepts contain sensory information and motor programs, these latter allowing the theory to escape the circularity of exclusively defining lexical items in terms of other lexical items by means of meaning postulates. The sensory and motor parts of the lexical descriptions provide the interface between the real world and the semantic structure.

Fodor, Fodor and Garrett (1975) also advocate a theory couched in terms of meaning postulates. They argue against lexical decomposition and indeed argue that semantic representations, as linguists have conceived of them, do not exist. They cite the failure of correlation between definitional complexity (as defined by decompositional theory) and psychological complexity as being of a considerable enough magnitude to justify the rejection of decompositional theory and the hypothesizing of a rather different theory based on meaning postulates.

Their proposed theory involves the following:

i) To each morpheme of the surface vocabulary (of a natural language) there corresponds a primitive expression in the vocabulary of the representational system. This does not imply, they stress, that the vocabulary of the natural language is identical with the vocabulary of semantic representations; they do not suppose, for example, that the formatives of the semantic level are phonologically interpreted, merely that there is a one-to-one correspondence between the formatives of the natural language and the formatives of the representational system - whatever these latter may turn out to be.

ii) They consider that it is "practically mandatory to assume that meaning postulates mediate whatever entailment relations between sentences turn upon their lexical content" (p. 525-526).

Thus, meaning postulate theories have been advanced in explicit opposition to decompositional, dictionary theories.

iii) Semantic network theories of meaning (e.g. Collins and Quillians, 1972)

As already noted above, semantic network theories owe their origins to attempts at modelling semantic memory in computers, with the wider aim of making a language user

(Collins and Quillian, 1972). The essence of this theory of semantic memory is its characterization of semantic memory as a network of interrelationships between concepts, using a variety of inferential links.

In this type of network, rather than having a list of words as the properties of a concept, the list corresponding to the properties of a concept is a list of pointers to other concepts - thus, a concept would be a set of interrelationships among other concepts. From the outside, such a network appears as a whole set of interrelated lists, with pointers to words found on many of the lists. In such a network there are no primitive or undefined terms: everything is defined in terms of everything else. In this respect, it is like a dictionary, where words are defined in terms of other words. However, they explicitly reject the dictionary as an adequate model of human semantic memory, considering concepts in humans to be much more encyclopaedic: "As a first approximation, it makes sense to assume that the content of a concept is everything that has been heard or read or seen about that concept" (Collins and Quillian, 1972, p. 318). Thus, concepts are not mere word concepts.

There are a variety of possible relationships between concepts in semantic memory - for example, the superset (IS A) relation, similarity, part, proximity, consequence, precedence, parent etc. relationships. They also stress their notion of semantic distance between concepts; concepts can be related more or less distantly, and the semantic distance between adjacent, direcly linked concepts can in fact be greater than the semantic distance between concepts linked via one or more other concepts. This happens when the sum of the accessibilities in the latter case is less than the accessibility of the former case. The semantic distance between concepts is not simply proportional to the number of nodes in the path between the concepts.

According to this theory, comprehension of a sentence or string of words involves an attempt to construct an interpretation based on a configuration of paths in memory between the various concepts referred to by the words in the string. This involves an extensive search to determine how the words can be interrelated within the constraints of syntax and context. Thus, a network type representation of the sentence is set up (for an example, see Johnson-Laird, 1981a, p.110) with pointers into the appropriate nodes of semantic memory, and an extensive search for associated concepts (whilst a parallel syntactic analysis proceeds simultaneously) results in an interpretation of the sentence being arrived at. The entailments of the sentence can be captured by pursuing the links in semantic memory that initiate from the nodes which are initially activated.

Collins and Quillian's account thus stresses the importance of the various types of relationship between concepts in semantic memory, resembling the meaning postulate theory in this regard. However, the nature of the relations between concepts differ in the two theories.

d) <u>Overview of the three intensional theories of meaning</u>

The three theories of meaning outlined above share some common features, most notably an emphasis on intensional semantics. Indeed, some of the authors of the above three theories explicitly acknowledge their commitment to intensional aspects of semantics. This theorizing at the level of intensional semantics is, they argue, a necessary simplification for the purpose of devising theories of manageable complexity: to attempt to devise theories which take into account extensional and pragmatic phenomena would, they suggest, be premature. To quote two of the authors of the above theories, Kintsch (1974, p. 14/15), says:

> "It is quite true that the pragmatic factors are of great significance in psychological research in language behaviour . . . The study of the logical-semantic structure of language and memory appears to be sufficiently rich and promising in itself. It is surely complex enough, and if any simplification can be gained by disregarding further complexities this appears to be a sound research strategy.

> Certainly, an understanding of both semantic and pragmatic factors will eventually be necessary, but for the moment, concentration upon one or the other seems quite appropriate".

Later, Kintsch does concede (p. 15) that " . . . pragmatic considerations can never be neglected for long . . ." Similarly, Katz (1973, p. 38) emphasizes the importance of tackling manageable subproblems, and Katz and Fodor (1963) argue against the possibility of a theory of semantics taking the sociophysical setting of speaker and listener (i.e. pragmatics) into account, because such a theory could not . . . "be completed without systematizing all the knowledge about the world that the speakers share and keeping the systematization up to date as speakers come to share new knowledge". (Katz and Fodor, 1963, p. 181). They further argue that even a limited theory of how sociophysical setting determines the understanding of an utterance blurs the distinction between the speaker's knowledge of his language and his knowledge of the world. They conclude that " . . . since it is unlikely that anything stronger than a theory of semantic interpretation is possible and since such a theory is clearly an essential part of a linguistic description, it is reasonable to fix the upper bound of a semantic theory of a natural language at the point where the requirements upon a theory of semantic interpretation are satisfied"

(Katz and Fodor, 1963, p. 181).

Thus, these theorists are all in agreement upon the necessity to tackle the intensional semantics problem since it is of manageable proportions. This means that their theories are of an essentially 'syntactic' nature, to use Morris' (1938) definition of 'syntactic', that is, they theorize about the relationships between signs in the language system, attempting to account for intensional phenomena and neglecting semantic (the relation of signs to their designata) and pragmatic aspects.

A further similarity shared by the three theories is that, in attempting to theorize about semantics, they (explicitly or implicitly) focus upon the <u>understanding</u> aspect of semantics, rather than the <u>production</u> aspect. These theories give an account of the understanding process which is, to a greater or lesser extent, intuitively plausible; however, the theories then assume that speech production is, as it were, merely the converse of the understanding process. When considering the theories from the point of view of speech production, they lose some of their plausibility. For example, whilst a meaning postulate theory will attempt to give an account of what occcurs when a listener is presented with discourse, intuitively speaking it is difficult to see how meaning postulates could be employed in the selection of lexical items for the speech production process.

Katz and Nagel (1974) compare and contrast meaning postulate theory with Katz and Fodor's (1963) decompositional semantic theory. Katz and Nagel argue that decomposition theory is the logical theoretical successor to a meaning postulate theory, and that decomposition theory provides a fuller account of semantics, since it is capable of accounting for semantic relations between entire sentences by use of the principle of compositionality, whereas meaning postulate theory can only explain semantic relations between individual lexical items. In addition to the above difference between the two types of theory, Katz and Nagel specify three others:

1) Dictionary entries in decompositional theory contain selection restrictions but dictionary entries in meaning postulate theory do not.
2) Semantic analysis in decompositional theory (that is, <u>readings</u>) are sets of semantic markers, but semantic analyses in meaning postulate theory are sets of predicate letters (representing either language - independent abstract concepts, or abbreviations of words in a particular language; Katz and Nagel argue for the former interpretation). Furthermore, the semantic marker of decompositional theory not only names a concept, but also represents the logical structure of the concept in terms of its own formal structure in the same way in which a phrase marker not only designates a sentence but also represents

its constituent structure in terms of its own formal structure. Semantic analyses in meaning postulate theory, on the other hand, yield sets of predicate letters which are not phrase markers.

3) Decomposition theory contains a set of definitions of semantic properties and relations which explicate questions concerning synonymy, redundancy, ambiguity, and so on, while systems of postulates lack such explications.

Katz and Nagel further argue that, in order to convert meaning postulate theory into a satisfactorily complete theory of intensional phenomena, additional machinery is required (in particular, selectional restrictions and some principles for combining the meanings of words in syntactically complicated sentences). They argue that when those supplements are effected, the theory becomes a notational variant of decompositional theory.

Fodor, Fodor and Garrett (1975), however, reject the view that meaning postulate theory is a notational variant of decompositional theory. Among the differences between the two theories pointed out by Fodor et al. are that meaning postulate theory hypothesizes an internal mental language (into which natural language input is translated) that stands in a one-to-one relation to the vocabulary of natural language, whereas decompositional theory is committed to a language of semantic primitives that need have no such correspondence to natural language. Secondly, meaning postulate theory hypothesizes a less abstract type of semantic representation than does decomposition theory. The third major difference existing between the two types of theory is that, whereas decompositional theories hypothesize that inferences such as entailment take place during the course of ordinary comprehension, meaning postulate theory holds that initial comprehension is superficial, with the entailments of a sentence determined at some later time.

In contrasting decompositional theory with meaning postulate theory, Kintsch (1974) points out that, whereas decompositional theories require complex encoding of sentences, a theory based on meaning postulates requires a large number of such rules in semantic memory and there is, therefore, no real advantage of parsimony for meaning postulate theory. Neither way, as Kintsch puts it, is really simpler. Kintsch even goes so far as to concede that decomposition may provide substantial gains in economy for artificial intelligence devices. He maintains, however, that since no-one has written a computer simulation employing meaning postulates, it is impossible to say how such an approach compares in terms of efficiency with the decomposition technique.

Kintsch's main point is, however, that whether decomposition or meaning postulates are used in human comprehension and memory is an empirical psychological question, and his advocacy of a theory based on meaning postulates is based on his own evidence against lexical decomposition as an obligatory process in comprehension. Kintsch's experiments tested the notion that lexically complex words are harder to process than lexically simple words, in various ways. For example, in one experiment, processing difficulty was measured in terms of sentence initiation time, where subjects were instructed to generate a sentence using a given (lexically simple or complex) word. If decomposition takes place, the lexically complex word should take more time to comprehend than the lexically simple word. Kintsch did not, however, find any effect of semantic complexity on the time taken to begin to speak a sentence in the above experiment. Johnson-Laird (1981a) points out that the latency of speaking does not correlate with the difficulty of defining a word, however (Johnson-Laird and Quinn, 1976). Johnson-Laird further notes that Kintsch's other experimental tasks probably do not require subjects to decompose the meanings of lexical items in order to perform satisfactorily in the experiments. Under these conditions, accepting Kintsch's results as hard evidence against lexical decomposition during comprehension would be unwise.

Johnson-Laird concludes that it is best to consider lexical decomposition to be an <u>optional</u> process which can be used if required but which was probably not required by

Kintsch's subjects. He gives the example of someone asking one's name; under such circumstances, one is hardly likely to decompose the word 'name' in order to reply. It is likely that a fact such as one's name will be directly represented in memory. However, if someone asks what is meant by the word 'name', then a process of decomposition is necessary in order to answer. Johnson-Laird concludes that, in the course of ordinary comprehension, the listener may retrieve no direct information from the relevant lexical entry, but instead may merely access it and check that it contains some semantic information. This information may be required subsequently when, for example, attempting to verify the sentence, but mere access is sufficient for normal comprehension.

Semantic network theory has been criticized in some of its aspects. Woods (1975) pointed out a number of problems for semantic network theory. Most notably, three criticisms of importance advanced by Woods are:

1) There is a need to adopt a consistent notation, and also there is a need to ensure that the links between concepts are not treated in an entirely _ad hoc_ manner.
2) There is no means of distinguishing between the intension and eaxtension of expressions in semantic network theory. It is not clear how

this distinction could be captured within a semantic network.

3) The treatment of quantifiers (such as 'some', 'all' etc.) require special care within semantic networks. Quantifiers have been treated as though they were adjectives that modify a noun phrase in semantic networks. Quantifiers and their scope do present problems for psychologists and proponents of semantic networks are not the only theorists who have failed to do full justice to the complexities of quantification (Johnson-Laird, 1981a, p.112).

In summary, all three theories share an emphasis on intensional semantics, and have been directly compared with one another by some theorists. Katz and Nagel argue that meaning postulate theories are a notational variant of decompositional theories, and Hollan (1975) takes the view that semantic network theories are notational variants of decompositional theories. These similarities aside, all three theories share a fundamental weakness, that of being inadequate as psychological models of the comprehension process. Johnson-Laird (1981a) advances several arguments to support this claim, and we shall now consider the points raised by him.

e) Johnson-Laird's (1981a) case against the three 'syntactic' theories of meaning

The first major criticism advanced by Johnson-Laird of the three intensional theories of meaning is that all three theories are based on what he terms the "autonomy of semantics" - that is, the theories assume that the meaning of any sentence can be established independently from what the sentence may refer to. Indeed, all three theories have nothing of note to say about referential matters. Johnson-Laird points out that it is taken to be natural to assume that deriving the intensional meaning of an expression is a precursor to determining its extension. Natural language does not, however, always work in this orderly fashion. Johnson-Laird argues that intensional meaning and reference in fact _interact_ with one another. This interaction between intensional meaning and reference is evident in the case of selection restrictions. He gives the example of the expression "it is pregnant"", where the selection restriction for 'pregnant' constrains the _referent_ of 'it', not its intensional meaning; 'it' must refer to either an idea or a female animal.

Johnson-Laird further argues that sentence interpretation is based upon _factual knowledge of the world_ rather than selection restrictions. So, for example, a selection restriction for the verb 'lift' is a constraint upon the verb's subject; the subject must be human, animal, or machine. This specification, however, fails to allow certain acceptable sentences, such as:-

"The wind lifted the leaves over the fence".

"Hot air lifted the balloon".

and "The rope lifted the weight".

In such cases, Johnson-Laird argues, the listener utilizes factual knowledge (that hot air rises, that ropes can support weights, etc.) rather than selection restrictions.

A further example of the interaction between intensional meaning and reference (and the use of factual knowledge to constrain reference) is given by the sentence:

"He found it difficult to grasp".

The point is nicely illustrated by considering what 'grasp' will be taken to mean if the referent of 'it' is a mathematical theorem, as opposed to what 'grasp' will be taken to mean if the referent of 'it' is a boa constrictor.

Johnson-Laird (1981a) also argues that reference is also a critical importance for certain logical implications that hold between expressions. This is illustrated by considering the transitivity of spatial prepositions. So, for example, the relation <u>in</u> can be specified as being fully transitive. For example, the conclusion of the following inference is valid:

Fred is in his office.

Fred's office is in the University.

Fred is in the University.

The transitivity of _in_ can be specified using a meaning postulate:

If x is in y and y is in z, then x is in z.

However, other spatial prepositions exhibit varying degrees of transitivity, depending upon the nature of the reference situation. The classic example from Johnson-Laird is the relation 'on the right of'. The transitivity of this particular preposition depends upon the reference situation, and it therefore has a deictic component. For example, if all individuals being considered are seated along one side of a straight table, then the transitive inference:

Matthew is on the right of Mark.

Mark is on the right of Luke.

Matthew is on the right of Luke.

is valid, granted the truth of the premises, and 'on the right of' is transitive. However, if the three individuals were equally spaced around a circular table, the above transitive inference would not be valid. If we consider a third case where the three individuals are seated adjacently at a large-radius circular table around which many individuals are seated, then the transitive inference is again valid.

Thus, the transitivity of "on the right of" varies depending upon the nature of the reference situation; how large the table is, how the individuals are spaced around it, and how many individuals are being considered when attempting to make the transitive inference all affect the validity of the resulting inference. This example forcefully emphasizes Johnson-Laird's point that the logical entailments of expressions depend critically upon the reference situation.

Since none of the theories are capable of accounting for the problem of reference, such inferences would have to be embodied in the form of a set of meaning postulates for each spatial preposition. However, the above example illustrates a fundamental difficulty for this way of tackling the problem: the number of meaning postulates required to capture the varying degrees of transitivity in different reference situations would be very large indeed and is in fact potentially infinite. Meaning postulates would not, therefore, constitute a parsimonious solution to this problem.

Johnson-Laird postulates his own theory of comprehension, which overcomes the problems which he raises. His theory will be considered in more detail in the next chapter. Johnson-Laird concludes his criticism of the three theories by noting that the problems of logical inference "destroy any theory based on the assumption that meaning is autonomous and independent of the reference of expressions (Johnson-Laird, 1981a, p. 117)". This comment serves as a prelude to his own theory of comprehension, since it does tackle the reference problem.

f) Conclusions

Three theories of meaning which attempt to account for intensional relations between expressions, but which do not account for reference or extensional phenomena, were considered. Similarities were noted between the theories, and also differences between them - for example, decompositional theory has it that the logical entailments of an expression are recovered in the course of comprehension, whereas meaning postulate theory held that such entailments are determined at a later time, initial comprehension being of a superficial nature. The arguments concerning the relative merits of the three theories of semantic representation are, however, overshadowed by the problems facing all three theories.

Arguments were considered which suggest that a psychologically plausible theory of meaning has to account for the reference situation as well as intensional phenomena. It was also argued that the machinery of 'selection restrictions' would best be replaced by thinking instead of subjects using their factual knowledge in order to constrain reference of expressions appropriately. It was further argued that logical properties such as transitivity also depend critically upon the nature of the reference situation, and that, since these three theories are unable to account for reference, they are also in principle unable to provide a parsimonious explanation of such logical properties.

A further criticism noted in passing was that all three theories provide accounts of intensional meaning which are biased toward explaining the listener's task in comprehension: they do not offer particularly plausible models of speech _production_.

We now turn to consider 'semantic' (according to Morris's definition) theories of meaning, that is, those theories which attempt to take account of the reference of expressions.

CHAPTER 2

'SEMANTIC' APPROACHES TO MEANING

1) Introduction

In the last chapter, a historically-organized introduction was given which indicated the variety of theories of meaning postulated in recent years. Also reviewed in the last chapter were three 'syntactic' theories of meaning, that is, theories which focussed on sense relations between expressions. After noting certain similarities and differences between the three theories, it was concluded that all three theories suffered from the fundamental problem of being unable to account for the relation between the signs and their referents. The reference situation, it was noted, is of critical importance for certain logical inferences (particularly, transitive inferences), and Johnson-Laird (1981a) argues that it is knowledge of referents rather than the use of selection restrictions that constrains the interpretation of the sense of an expression.

Given that the relation between signs and reference situations is of great importance to the interpretation of the meaning of those expressions, we now turn to examine theories which do attempt to account for, or ascribe importance to, the reference of expressions. Such theories can, therefore, be termed 'semantic' theories of meaning, since Morris (1938) defined 'semantics' as the relations between signs and their designata. The theories to be reviewed in this chapter examine (or give precedence to) the relation between signs and their designata.

The first theory to be reviewed, the theory of model-theoretic semantics, has as its origins a background of linguistic and philosophical interest in the logical structure of language. The second theory (or rather, family of theories) to be reviewed is that of procedural semantics, which owes its origins to the relatively new discipline of Artificial Intelligence. The final theory, which will be reviewed in some detail, derives largely from psychological considerations and uses procedural semantics as a proper part; this is the theory of mental models. The theory of mental models is a recent and powerful theory, being considered by its exponents to be capable not only of explaining the problem of meaning, but also of being capable of explaining errors in reasoning (for example, in syllogistic inferences) and also as a possible way of resolving the longstanding controversy regarding imagery (Johnson-Laird, 1980).

The widely differing backgrounds to these three theories (model-theoretic semantics, procedural semantics, and the theory of mental models) serves to emphasize the interdisciplinary nature of the whole question of meaning.

b) '*Extensional*' *linguistic theories of meaning*

An extensional approach to meaning which originates in the areas of linguistics and philosophy is due to Tarski (1956) and is known as "model-theoretic semantics". This theory stems from Tarski's definition of truth, which held that truth is relative to actual or possible states of the Universe. Model-theoretic semantics also embodies Frege's *principle of compositionality*, that is, the intension of a complex expression (such as a sentence) can be built up compositionally from the intensions of its constituents in a way that depends only upon their grammatical mode of combination.

Logicians had shown how to formulate a rigorously compositional semantics for a formal language by providing it with a semantic interpretation with respect to a *model* or *state description*. This model is an abstract construct and consists in elements (corresponding to elements in the real world - the model corresponds to a view of the world) and atomic propositions which combine particular elements with particular predicates. The model has a list of atomic propositions which are taken as a list of true descriptions of the world, and elements in the language are mapped onto elements in the world.

The language which is under interpretation is termed the '*object language*', whilst the semantic interpretation with respect to the model is achieved using an *extensional* metalanguage, in the form of a logical calculus. Thus, the structural rules of the model are couched in a metamathematical calculus.

For example, a state description or model can be constructed for a "love triangle" situation in which both John and James love Mary (and Mary loves neither John nor James) and, as a result, John hits James and vice versa, as follows (This example is adapted from the example in Johnson-Laird, 1982, p. 3):

LEXICAL ITEMS

Nouns - "John" "James" "Mary"

Verbs - "Loves" "Hits"

SYNTACTIC RULES

$N^\frown V^\frown N$ is a well-formed sentence

$S^\frown$"and"$^\frown S$ is a well-formed sentence

(N.B. In more complicated models, other logical operators are permissible).

MODEL STRUCTURE LEXICAL RULES

1) "John" has extension JOHN
2) "James" has extension JAMES
3) "Mary" has extension MARY
4) "Loves" has extension {(JOHN, MARY) (JAMES, MARY)}
5) "Hits" has extension {(JOHN, JAMES) (JAMES, JOHN)}

Predicates are extensionally defined in terms of what can be true in the model - for example "John loves Mary" is true but "John hits Mary" is not. The model structure specifies the extensions, as above, and the extensional relationship is one where, for example, element A in the language maps onto element a in the model, and so on.

STRUCTURAL RULES

1) A sentence of the form $\alpha \gamma \beta$ is true only if the extension of $\langle \alpha , \beta \rangle$ is a member of the set comprising the extension of γ.
2) θ ⁀ "and" ⁀ ψ is true if θ is true and ψ is true in the model. (where ψ and θ are well-formed sentences, e.g. N⁀V⁀N, γ is a verb, and α , β are nouns.

The structural rules are normally much more complicated than in the above (purely illustrative) example. More complicated structural rules allow this method of semantics to be very powerful. A variety of such structural rules would enable semantic interpretations (truth values with respect to the model) to be returned for a wide variety of complex expressions.

That the predicates are extensionally defined is evidenced by the very definition of 'extensional definition' - an extensional definition is a listing of the members of the relevant class, and this is precisely what is done in the model structure lexical rules.

The essence of model-theoretic semantics is that, given the extensions of a list of simple expressions and an appropriate set of structural rules, truth values can be assigned to complex sentences and phrases. The final product is therefore a truth value with respect to the model structure for a given complex expression.

This type of theory has been applied not only to formal languages but also to natural languages (Montague, 1974). This is achieved by introducing sets of model structures, each representing a possible state of affairs - a 'possible world' - at a particular moment. The introduction of 'possible worlds' into model-theoretic semantics allowed the theory to encapsulate Frege's (1952, originally published 1892) distinction between sense (intension) and reference (extension) (see chapter 1, p. 10). This is possible since the theory is then capable of dealing with denotations of expressions not only in the actual world but with <u>rules</u> that govern their denotations in all possible worlds: such rules can be identified with the intensions of expressions. As Johnson-Laird (1981b) puts it, the intension of a sentence, i.e. the proposition it expresses, can be treated as a function from the set of possible worlds onto the set of truth values (true and false); the extension of the sentence is its truth value in the particular possible world under consideration.

Thus, we can distinguish the intension of the following two sentences:

1) The Morning Star is identical to the Morning Star.

and 2) The Morning Star is identical to the Evening Star.

by noting that sentence 1) is true in all possible worlds, whilst 2) is true in only some possible worlds, including the real one. Accordingly, their intensions are different. Thus, a logically necessary proposition is one which is true in all possible worlds, whilst a logically possible proposition is one which is true in at least one possible world.

Montague also used the method of model-theoretic semantics as a basis for developing a treatment of pragmatics; although Montague limits pragmatics to the study of indexical expressions, that is, expressions such as 'I', 'here', etc., whose semantic values depend on contexts of use. He considered that pragmatics should at least initially follow the lead of semantics, that is, pragmatics should, like semantics, concern itself with truth - but with respect not only to a given interpretation but also with respect to a context of use.

Model-theoretic semantics provides a formal and rigorous extensional treatment of the object language using an (extensional) metalanguage for the interpretation of the object language. Montague's chief generalization of model theory was to make semantic assignments relative to various factors.

The notion of truth-under-a-given-interpretation and truth relative to a context of use, and the utilization of a model (state-description or view of the world) to represent the 'given interpretation' in semantic analysis is of interest to psychologists. Even although the model is an abstract formal entity, the general approach of the theory is of interest. Important points of note are: a) that the combinations of particular signs which are admissible is given by the model, and therefore what gets accepted as true or rejected as false is determined by the model structure; b) the model in turn represents a 'possible world', that is, it corresponds to a view of a particular state of affairs at a particular time - i.e. a particular view of the world; and c) extensional considerations are given prominence, in contrast to the theories outlined in the previous chapter.

However, despite its admirable formal rigour, there are problems for model-theoretic semantics; in particular, the treatment of sentences about beliefs, desires and psychological states has not yet been satisfactorily dealt with using this framework (Johnson-Laird, 1981b). Lyons (1977) points out that many sentences of ordinary language express what appear to be complex intensional propositions. For example, "Romeo thinks that Juliet is dead" is not a truth-function of the simple proposition "Juliet is dead", since the truth or falsity of the complex sentence is independent of the truth or falsity of "Juliet is dead". The complex sentence would therefore normally be described as an _intensional_ _proposition_. This is true also of many compound sentences of ordinary language containing 'and' or 'if' since they are taken to imply that some kind of causal, temporal or other connexion holds between the propositions expressed by the constituent clauses, as in "If he did that he is very brave".

A further problem noted by Lyons (1977) is that these formal systems consider the descriptive function of language. It is important to bear in mind, following Austin (1962) that there is no simple one-to-one correspondence in the everyday use of language between the grammatical structure of a sentence and the kind of communicative act that is performed in particular situations by the utterance of that sentence. Thus, declaratives do not always make statements, and conversely, interrogatives, for example, can be used to make statements (e.g. "Do you know that I am 92 years old?"). Such considerations are part of the domain of pragmatics and will be considered in the next chapter.

c) <u>Procedural Semantics</u>

The term "procedural semantics" refers not to a specific theory of semantics, but instead a general approach to the problem, a general way of couching semantic theories. Woods (1981) argues that it is possible to formulate incorrect theories in this vein, as well as (he contends) correct theories.

Winograd (1975) contrasts declarative and procedural knowledge, and points out that this contrast corresponds to the philosophical distinction between 'knowing that' and 'knowing how'. A declarative representation of knowledge would be one where knowledge takes the form of a set of specific facts describing particular knowledge domains, and there exists in addition to the facts a general set of procedures for manipulating these facts. In thinking, the general procedures are used to manipulate the domain-specific data base in order to make deductions. This is similar in principle to axiomatic mathematics, with the facts corresponding to axioms and the thought processes corresponding to proof procedures for drawing conclusions from the axioms.

The proceduralists, on the other hand, assert that our knowledge is primarily a "knowing how": the human information processor is a stored program device, with its knowledge of the world <u>embedded</u> in its programs. Thus, by this account, what we know about language is coextensive with our set of programs for operating with it.

The notion that the meaning of language expressions could be identified with a procedure or set of procedures stems from the area of Artificial Intelligence and has caught the interest of psychologists interested in language (e.g. Miller and Johnson-Laird, 1976; Johnson-Laird, 1977a).

The approach originated from a consideration of the semantics of high level computer programming languages like Fortran and Algol. These high level languages, which are used to commmunicate programs of instructions to the machines, have both a syntax and a semantics. The syntax consists of rules for writing well-formed programs; the semantics consists of the procedures that the computer is instructed to execute. In the computer programming sense, therefore, procedural semantics deals with the meaning of the procedures that the computer is instructed to execute, whether the result returned by the program* is what the programmer would have expected, and so on (Johnson-Laird, 1977a).

There are therefore two stages involved in running a computer program. The first step is to compile the program, that is, translate it into the operational code of the particular machine to be used. The second step is

---

* The American spelling of 'program' will be used to refer to computer programmes in this thesis, and other senses of 'programme' will use the British spelling.

to execute the compiled program, using data which may or may not be supplied with the program. This "compile and execute" strategy has been metaphorically applied to natural language processing as a theory of meaning. As such, it is an interesting and flexible approach to the problem of meaning, since it is possible to explicate both intensional and extensional meaning in procedural terms (Miller and Johnson-Laird, 1976).

The application of the "compile and execute" strategy to human language processing was first clearly formulated by Davies and Isard (1972), and goes as follows: on hearing an utterance, the person must firstly compile it, that is, translate it into his or her internal mental language. The second step is the decision on the part of the person of whether or not to run the compiled program - the choice is usually under the voluntary control of the listener. The two steps differ with regard to the hearer exerting voluntary control; while a person can, for example, choose not to answer a question (that is, choosing not to execute a compiled program), he has no voluntary control over the compiling of the program - this is automatic and involuntary. As Davies and Isard put it, it is difficult to refuse to understand a sentence in a language you know well, but it is often easy to refuse to verify it. A further interesting comment by Davis and Isard is that loss of conscious control over one's compiler may correspond to knowing a language fluently.

In general terms, the main thrust of procedural semantics as applied to human understanding of natural language is to characterize understanding as a process of translation of input utterances into procedures or programs, that is, sets of operations that may or may not be subsequently carried out. That a variety of such operations are admissible (see below) in terms of the theory gives the theory sophistication and plausibility from a psychological point of view.

There are in fact a variety of possible programs which expressions could be translated into. For example, procedures to verify the proposition expressed by a sentence, procedures to take action satisfying the request made by a sentence, and procedures to find information (for example, by instituting a memory search) answering the question posed by a sentence. Some authors have misunderstood procedural semantics as particularly emphasizing the processes of verification (Fodor, 1978); it has been assumed that the meaning of a sentence is the procedure to verify whether it is the case or not. However, as Johnson-Laird (1978) points out in a reply to Fodor, procedural semantics admits many different types of operation, and furthermore points out that understanding is antecedent to verification, not a consequence of it (Miller and Johnson-Laird, 1976, p. 126). The point to be stressed is that verification is but one of many types of procedure which can be compiled and executed on hearing a sentence.

Woods (1981) points out that primitive perceptual procedures could be used to define truth conditions for elementary propositions. This relates to Tarski's model theory. Tarski's definition of truth was as follows, to quote his own famous example -

"Snow is white" if and only if snow is white.

The object language statement whose truth is to be verified appears on the left hand side of the expression and is enclosed by quotation marks. The metalanguage definition of the truth conditions for the object language expression, which is on the right hand side of the above expression and is not enclosed by quotation marks, could be characterized in procedural terms, that is, in terms of the primitive operations of sensory perception. This point is taken up and amplified by Miller and Johnson-Laird (1976). Miller and Johnson-Laird came up with the general formulation that a person learns many rules of the form

P is true if and only if F(x) = 1

where F(x) describes a mental computation to be performed. If the result of the computation is F (x) = 1, then P is true; otherwise it is not true (false or indeterminate). They argue that the psychological problem is to characterize the mental computations that a person performs when he learns and applies such rules, and that the best metaphor currently available with which to formulate such a theory is the theory of computation developed to describe the operation of computers. Therefore, F (x) is characterized as a program of instructions to be executed, containing instructions such as find (in a given search domain, x) and test (at time t, x satisfies the description D (x)). The description D could be a perceptual paradigm composed of perceptual predicates (in Miller and Johnson-Laird's formulation of the idea). Thus, some high-order executor would request the perceptual system to search for, attend to, and make judgements of various combinations of perceptual predicates. The combination of perceptual predicates identify objects or events taking a label, and each label is associated with a particular perceptual paradigm D(x). Thus, in this case, the relevant label would be "white" and the verification of the natural language expression using the truth-rule involves attention to snow and making judgements concerning the presence or absence of particular perceptual predicate(s) given by D (x) corresponding to "white".

Woods (1981) provides an interesting discussion of language understanding from the point of view of procedural semantics. He argues for the existence of an internal language capable of vastly greater discriminative subtlety than the external natural language, and it is in terms of the internal language that the meanings of natural (external) language expressions are defined. Woods argues that human communication requires a receiver to deduce a much more precise understanding of the intended meaning of an utterance than the external language words and syntactic structure manage to convey. This, he argues, is an economic solution, for if the natural language were to be capable of the same level of discriminative subtlety of which the internal language is capable, it would require a much greater vocabulary than it actually has. He argues that most words in English are highly ambiguous, and the speaker's intended sense is selected by context. (This is in fact a problem of pragmatics - that is, the selection of a particular intended sense of an expression from the many possible senses which it has, and will be returned to in chapter 3). The point that is emphasised by Woods is that both the process of translation (corresponding to selection of a particular sense) and the process of execution (the carrying out of the set of operations selected during translation) can best be modelled by means of procedures.

The advantages of a procedural semantics approach to natural language meaning are many and varied. Firstly, the distinction between compiling and executing a program provides a way of disconnecting the understanding of a sentence from any actions it might entail. Secondly, it forces the theorist to consider processes as well as structure. Thirdly, it admits a wider range of extensions of sentences than does model-theoretic semantics: in procedural semantics, the extension of a program is the result the program returns when the procedure has been executed, and the intension of the program is the particular procedure that is executed when the program is run. Accordingly, whilst model-theoretic semantics admits only "true" or "false" as extensions of sentences, procedural semantics admits a variety of possible extensions, including truth values, answers to questions, compliance with requests, additions to knowledge, modification of plans, etc. Fourthly, the distinction between pragmatics and semantics can be explicated in terms of procedural semantics: pragmatic considerations come into effect at the 'compilation' stage, that is, the particular program which gets compiled depends upon the pragmatics of the situation, whereas the semantics corresponds to the particular procedure selected and the execution of it. Lastly, the procedural approach places a diverse range of speech acts (statements, questions, requests etc.) on an equal footing and provides a theoretical language in which to formulate hypotheses about the mental processes involved.

Miller and Johnson-Laird (1976) developed a complicated procedural theory which attempted to characterize linguistic and perceptual functions in computational (procedural) terms. They point out initially that language and perception are related - in their terms, what is seen and what is said are somehow related. Philosophers have been interested in the nature of that relation for centuries. Miller and Johnson-Laird focus primarily upon how perception of the world affects communication about it, and they concentrate their attention on word meanings. To use their example, consider the word "lamp". Such a word must not only relate to other words through grammatical, conceptual or memorial systems, it must also relate to concrete, objective instances of lamps: otherwise the word would be unable to serve the purposes it does serve.

Miller and Johnson-Laird argue (p.7) " . . . the active use of words like "lamp" depends critically on one's ability to identify instances. The word must be associated somehow with a perceptual procedure capable of deciding which objects are and which objects are not instances of an appropriate kind.* Thus, although verification is only one of the many procedures which a hearer may try to execute on hearing a sentence, it illustrates the need to be able to link words and percepts. This is the starting point of Miller and Johnson-Laird's programme of "theoretical revisionism";

they set out to attempt to achieve a theoretical synthesis of external (referential) relations and internal (intensional) relations of the system.

They conclude, after lengthy and detailed examinations of perception and cognition, that words and percepts are not linked directly but instead they are linked via a conceptual system of inscrutable complexity, this conceptual realm itself being the central concern of cognitive psychology. The final analysis was therefore in terms of concepts and procedures for using them. In their theory, every word is associated with a lexical concept. A lexical concept is anything capable of being the meaning of a word, and, for most words, consists of two parts: 1) a <u>definitional</u> part depending on a functional-perceptual schema for recognizing instances (perceptual predicates alone in the definition are insufficient because of problems caused by vagueness, the contribution of noncriterial features, and other problems for a purely perceptual definition, such as the importance of characteristic orientation). The second part of the lexical concept is 2) a <u>connotative</u> part consisting of knowledge associated with the word, including the relation of the word to other words. This particular definition of a lexical concept is thus capable of capturing both sense relations and referential aspects and is tied to procedural rather than propositional interpretations of sentences.

Miller and Johnson-Laird's consideration of the possible intensional relations bewteen lexical concepts is based upon the notion of the concepts being organized into semantic fields. To quote Miller and Johnson-Laird:

> "We assume that a semantic field consists of a lexical field and a conceptual core. A lexical field is organized both by shared conditions determining the denotations of its words and by a conceptual core, by the meanings of what the words denote. A conceptual core is an organized representation of general knowledge and beliefs about whatever objects or events the words denote - about what they are and do, what can be done with them, how they are related, what they relate to. This lexical-conceptual relation is complex. To say that a lexical field covers a conceptual core like a mosaic is suggestive, but it greatly oversimplifies . . ." (Miller and Johnson-Laird, 1976, p. 291).

For example, for furniture in general there is a core of indoor human activities associated with furniture, such as eating, sleeping, working, playing etc., and furniture exists to accommodate peoples' bodies and the objects and instruments they use as they engage in those activities. These core concepts are essentially commonsense theories about the way the world works. The lexical field associated with the conceptual core consists of lexical concepts organized in interlinked decision tables. The schemata of individual lexical concepts specifies a set of conditions that must be satisfied for an appropriate use of the term associated with the lexical concept, and many of these conditions are shared by several schemata. Thus, a decision table based on shared conditions is a parsimonious way of characterizing interrelationships between individual lexical concepts within a semantic field (see example below). If, for example, we consider a 'furniture world' where there are only three kinds of item of furniture, that is, chairs, tables, and beds, we could represent these three concepts in a decision table as follows:

Decision Table

	1	2	3
Conditions			
(i) It has a seat?	Y	N	N
(ii) It has a work top?		Y	N
Actions (call the object a-)			
"chair"	x		
"table"		x	
"bed"			x

(Adapted from Miller and Johnson-Laird, 1976, p. 284).

Each schema (a column in the table) is represented by a pattern of outcomes for these conditions, plus an indication of which words are assertible given that pattern. The table can be entered with a pattern of conditions and the appropriate word found, or it can be entered with a word and the pattern of conditions found. Thus, extremely flexible access is possible both from the speech production and the speech understanding points of view.

It is possible to interlink the above decision table with other decision tables by means of having a fourth column, where in the case of a larger decision table none of the actions has been satisfied by the conditions provided, a 'call decision table # x' option is available if none of the other actions specified by the table were able to be carried out. Recursion is also possible, with a given table calling itself. The decision table is a specific realization of how schemata might be integrated and is one way of characterizing the organization of conceptual information, and could prove a useful basis of organization from the point of view of both production and reception of speech.

Miller and Johnson-Laird's theory has been described at some length, because it illustrates several points, both about procedural semantics in particular and 'semantic' theories (in Morris's sense) in general. Firstly, it illustrates the power of the computational metaphor; both sense and reference are characterizable in procedural terms. Referential matters and verification are <u>processes</u> according to such a theory, and it thus provides a natural and plausible psychological account of the reference problem, which tends to be finessed in some other 'semantic' theories. Model-theoretic semantics is a 'semantic' theory which emphasizes structure rather than process, in contrast to the procedural theory. Sense relations can also be characterized in procedural terms using semantic fields and interlinked decision tables. The fact that it can deal with <u>both</u> sense and reference is necessary according to Morris's definition of 'semantic'; semantic theories include syntactics as a proper part of their study.

Miller and Johnson-Laird's approach concerns itself primarily with the problem of characterizing lexical knowledge; ideally, however, we would like to characterize how whole utterances and even larger units of discourse are understood and how their meaning could best be characterized. The virtue of Miller and Johnson-Laird's theoretical ideas is that they both illustrate the viability of a procedural approach to these problems, and they also raise theoretically interesting questions.

Procedural Semantics in general is a useful and powerful theoretical tool for developing theories of meaning. Perhaps the most important point about this type of theory is that it emphasizes process rather than structure (Johnson-Laird, 1977a). Procedural semantics is not only a theory of semantics in itself, however. It has been incorporated into other theories which attempt to explain meaning and some of the phenomena of human language understanding. The next theory to be considered uses procedural semantics fairly extensively, and attempts to characterize not only the problems of how humans represent meanings, but it also attempts to explain some of the phenomena of syllogistic reasoning and to serve as a general theory of comprehension. We will consider in detail only its treatment of the problem of meaning.

d) The theory of mental models

Earlier in this chapter, an exposition of model-theoretic semantics was given and it was noted that in model-theoretic semantics the "model" is an abstract construct, a set of rules couched in a metamathematical logical calculus. This "model", it was noted, corresponds to a view of a possible world and specifies the admissible combinations of extensionally defined elements to yield interpretations of complex expressions with respect to the model structure. A rather different notion of a 'model' has been suggested in psychology by several different theorists during the past forty years. The psychological notion of a model is that of an internal mental representation which takes the form of a structural analogue of the state of affairs which is being described or thought about.

One of the earliest exponents of such a notion was Kenneth Craik (1943). Craik pointed out the advantages for an organism of having an internal model of the world with which to mentally 'try out' hypothetical courses of action without physically suffering the potentially adverse consequences of those courses of action - specifically, as Woods (1981) puts it, such a modelling system would allow one's theories to die instead of oneself.

The general notion of mental models has been hypothesized by other theorists since Craik, and is now the core of a psychological theory of comprehension by Johnson-Laird (1980, 1981a, 1981b). The motivation for the interest in such internal structural analogues of the referents of a piece of discourse has been the observation that what is remembered of a sentence corresponds to none of its linguistically motivated representations, a point made by Fodor, Fodor and Garrett (1975). Sanford and Garrod (1981, p. 63) comment that "if memory is the product of the comprehension process, then memory experiments may provide clues about the nature of comprehension itself". It therefore follows that none of the linguistically-motivated representations of discourse will suffice as a psychological model of comprehension, and a different type of theory is called for. This theory is based not on linguistic considerations but on the general notion of a mental model.

The basic point which experiments on memory for text demonstrate is that the surface structure (that is, the actual wording of the text) is rapidly forgotten and the enduring memory is for the 'gist' of the text, that is, the 'logical points', main substance or pith of the text. One of the earliest and most striking demonstrations of this phenomenon was the oft-quoted experiment by Frederick Bartlett (1932). Bartlett read his subjects a North American Indian folk story called 'The war of the ghosts', and then had them recall it at later points in time (with varying length of delay between presentation and recall). Firstly, Bartlett noted that recall was very inaccurate, often being only the outline of the story. Secondly, he noted the presence of various systematic distortions in the recall of the story: things which fitted the story but were not actually present in it appeared in the recall protocols. Finally, when the subjects' memory of the story was so poor that only isolated fragments could be remembered, subjects sometimes invented plausible, stereotyped stories around those details which could be recalled. Such reconstruction processes tend to be used less when less curious experimental stimuli are used; furthermore, when reconstruction errors do appear, they generally do so in the case of long delays between presentation and recall. Recall in the case of less exotic material appears to be fairly accurate in the sense of capturing gist. Sanford and Garrod conclude (p. 65) "In part at least, what is remembered is allegedly a product of the comprehension process".

One experiment whose results point to an explanation in terms of models is that of Bransford, Barclay and Franks (1972) who showed that subjects who had been presented with the sentence

1) Three turtles rested on a floating log and a fish swam beneath them.

later erroneously recognized the sentence

2) Three turtles rested on a floating log and a fish swam beneath it,

but they did <u>not</u> erroneously recognize the sentence

3) Three turtles rested beside a floating log and a fish swam beneath it

This result is consistent with the subjects having somehow represented the sentences in a way which captures the spatial relations of log, turtles and fish. Accordingly, Bransford, Barclay and Franks distinguish between a 'constructive' and an 'interpretive' approach to semantics. The result noted above would constitute strong support for a theory based on mental models.

Another line of support comes from an experiment by

R. C. Anderson et al., (1976) in which subjects were presented sentences such as:

The fish attacked the swimmer

Later, subjects were presented with a recall cue. It turned out that a more specific term such as shark was a better recall cue than 'fish', the more general term actually used in the original sentence. Such a result is problematic for conventional theories of meaning using semantic features or semantic networks. Anderson et al. explain their results in terms of an 'instantiation' hypothesis, that is, the word "fish" has in fact a whole family of potential meanings, and a particular sense is instantiated as a function of the context of the sentence. Johnson-Laird (1981a) argues that a more plausible interpretation of the findings would be one in terms of a mental model of the referents and relations described in the sentence.

Similar results were produced by Garnham (1979), but in this case with verbs. Garnham also used a cued recall technique, and subjects read sentences like:

The housewife cooked the chips

Garnham showed that the recall cue fried was more effective in helping the subjects to remember the original sentence than was the cue 'cooked', the original verb which has been present in the sentence. Garnham interpreted his results in terms of the sentence being represented in the form of a knowledge-based model of the situation, and this model is essentially non-linguistic in nature. The word 'fried' maps into this model better than does 'cooked' in this case (since the subject knows that the particular method of preparing chips is by frying), hence the result.

The work of Garnham and of Anderson et al. demonstrate that, in the case of both nouns and verbs, the particular sense of the word which is selected by the subject in the context of a particular piece of discourse is not easily explained in terms of a breakdown of the relevant lexical item into particular sets of semantic primitives, but rather a better explanation would be in terms of the discourse addressing a model-based package of knowledge, and this model constrains the selection of the sense of the lexical item (rather than using selection restrictions or some such machinery).

Several model-based theories of reading comprehension have been proposed. These theories hold that the problem for the understander is not one of breaking words down and relating their various senses together in order to build up a coherent interpretation, but instead the problem is one of addressing appropriate pre-packaged knowledge structures. These pre-packaged knowledge structures (i.e. models), basically contain information about what to expect in a particular situation around which the discourse is based. The characterization of such knowledge structures is usually that of variations on Minsky's (1975) 'frame' hypothesis, that is, as hierarchically-organized data structures.

Minsky put forward the frame as the basic building-block of knowledge and conceived of memory as comprising millions of frames. Data in frames is hierarachical in that the higher levels of the frame 'package' of data contains necessary, fixed elements in the situation; lower down, the data in the frame becomes increasingly arbitrary, that is, it could be violated without necessarily making the use of that frame implausible. So, for example, in a child's 'birthday party' frame, at the top levels a host and guests constitute fixed, necessary features of the situation. Lower down in the data structure, the definition of the lists of features become less well defined, and/or optional. For example, there are presents, games, decor, a birthday cake, ice cream, etc. - these are normal features of the child's birthday party. However, if one of these expectations were violated, the situation would still be describable as a child's birthday party, whereas violation of the expectation of there being a host or guests would invalidate the use of the party frame: without a host or guests there would be no party as most people would understand it.

Thus, the emphasis in these accounts is on how discourse relates to models taking the form of pre-packaged, available data structures. Frames are usable in many ways; for example, there are frames for objects (such as the use of a frame to contain default information to assist the recognition of a cube), temporal or programmatic frames (these contain information about what to do and expect in a restaurant, at a lecture, etc.), mixed frames for situations (such as the child's birthday party example above, etc.), grammar frames (verbs are viewed as frames or information structures which allocate the other parts of speech in a sentence into a relationship with the verb), narrative or text frames (such as what to expect in a folk story), and even scientific paradigms have been considered from the 'frames' point of view.

Particular versions of the frames hypothesis have been mooted in the form of model based theories of understanding; for example Schank and Abelson's (1977) 'script' account is a model based theory using temporal or programmatic frames to provide models of particular situations on which to base understanding. Sanford and Garrod's (1981) 'scenario' account also is a model based theory of comprehension. In both of these theories, the emphasis is more on how discourse addresses knowledge than on how the discourse elements themselves are broken down semantically. That is, these theories are 'top-down' (i.e. concept-driven) rather than 'bottom-up' (i.e. data driven) theories of comprehension. There may be no necessity to break down input language into components on every occasion; what is important according to these accounts is that utterances act as cues or clues for addressing pre-packaged mental models, or as clues to build new models (see below). Lexical decomposition may therefore be an <u>optional</u> process.

There is, therefore, a good case for supplementing the theories of sentence comprehension with a richer form of representation in the form of a model or internal structural analogue of the state of affairs described by the discourse (Johnson-Laird, 1980, 1981a, 1981b). In particular, it was noted in chapter 1 that theories of comprehension which fail to take referential matters into account run into trouble with logical inferences. Mental model theory attempts to circumvent this difficulty by postulating the existence of an internal model of the referent, modelling its major structural features, etc. Mental models also make logical deductions such as transitive inferences possible without having to rely on <u>rules</u> of inference. Features such as transitivity emerge naturally from a model-based account without having to be explicitly specified in advance.

The essence of Johnson-Laird's (1980, 1981a, 1981b) theory is that utterances provide clues for building mental models. Johnson-Laird's emphasis is on how models are constructed to represent novel situations. Other model-based accounts (such as Sanford and Garrod's 'scenario' theory) consider how discourse is related to pre-packed, 'ready-built' models. Models are not always available for every reference situation and in some discourse situations may have to be constructed at the time of encountering the utterance and it is with these processes which Johnson-Laird concerns himself.

Thus, utterances provide clues for building mental models. The phenomenal content of a model is immaterial - it may or may not be accompanied by visual imagery, for example. What does matter is the structure of the mental models and the fact that we possess procedures for constructing, manipulating and interrogating them. Many of these procedures can take for granted a common background of knowledge, including world knowledge and knowledge relating to the language and the conventions governing conversation. Mental models in fact constitute one option in the encoding process: a hearer can choose to either construct a mental model on the basis of what he or she hears, actively drawing as many implications as possible from what is said, or the hearer can choose to merely register the speaker's discourse in a passive way.

Johnson-Laird contends that there are in fact two stages in language understanding. Firstly, the utterances are translated into the internal mental language of the hearer. The formatives of this internal language are in close correspondence with those of the natural language input (as in the theories of Kintsch (1974) and Fodor,

Fodor and Garrett (1975) mentioned in chapter 1). Thus, a given natural language expression is translated into a representation of essentially similar form - in fact, Johnson-Laird theorizes that the mental representation language may take the form of a propositional representation similar to that advocated by Kintsch (1974). The original surface structure of the linguistic input is readily recoverable from this representation. Following this, the listener may choose to construct a mental model from the initial propositions representation. This mental model would utilize a procedural semantics which would build, for example, a spatial array which would mirror the structural layout of what was being described.

Thus, for example, in the case where one reads a description of the spatial layout of a room when reading a novel, one may build a model of the room, and the positions of objects in the model would correspond to the positions of the objects stated in the description. We should note that such descriptions are radically indeterminate, that is, they are usually consistent with a variety (a large number) of different rooms, not merely in terms of the dimensions of the room but also in terms of the actual spatial layout of the furniture and other objects within it. The mental model builder may notice this indeterminacy - indeed, Johnson-Laird argues that it can only be detected upon attempting to build a model - but he may choose one particular interpretation and build the model to correspond to that interpretation. Thereafter, if the listener has decided to plump for one particular instantiation, then all will go well provided that interpretation is not violated in the subsequent discourse - if such a violation occurs, the listener may attempt to rebuild his model appropriately, or else he may abandon the whole enterprise in some confusion.

The points of note are that model construction is an optional mode of encoding, involving more cognitive work than mere translation of the discourse into mental currency, and that, for indeterminate descriptions, the construction of one model which satisfies the discourse is effected rather than there being attempts to construct multiple models. The fact that two stages are involved is necessary to explain the results of some experiments which demonstrate retention of surface structure without detailed understanding having taken place - in such cases, encoding has gone no further than translation of the discourse into a propositional representation (examples of this will be given below in the discussion of experiments by Ehrlich and Johnson-Laird, 1982).

Johnson-Laird (1980, 1981a, 1981b) reports on a program written by him in a high-level list-processing language which simulates his theory of mental models. The program works by interpreting simple spatial assertions by means of building up a spatial model of the relations between entities, and it can combine information from separate assertions in order to achieve a unified composite representation. The procedures used by the program are mainly general procedures which set up an internal representation of two-dimensional space, add items to the array, test for specified relations between items, etc. So, for example, if verification of a statement that one item is behind another is required, the verification procedure would operate by scanning along a line originating from the second of the two items to determine whether the first item is located somewhere along that line. The direction of this scanning search is controlled by two variables, which are the values by which the X and Y coordinates are incremented to successively spell out the locations to be scanned. The actual values of X and Y are determined by the particular relation being tested for. For example, for the relation 'on the right of' ('A is on the right of B'), the instruction to the scanner is as follows:

FUNCTION (%1, 0%)

This instruction takes a general procedure for verifying the relation between two items, assigns it to the variable FUNCTION, and "freezes in" the value of two of its parameters (the "freezing in " being denoted by the decorated parentheses). In this case, the parameters +1 and 0 specify the direction in which to scan - that is, by incrementing the X-coordinate and holding the Y coordinate constant. Thus, the scanner begins at object B and scans successive locations to the right of B to verify whether A is indeed located to the right of B. The array is thus represented as if it were a graph laid upon a table being viewed from above.

This simulation illustrates several interesting points - firstly, the meaning of 'behind', 'to the right of' and other relational terms is represented in an internal meta language which bears no simple, let alone one-to-one, relation to the natural language. (This is a further advantage of the procedural semantics approach; relational terms, conjunctions and other terms of this nature are handled in a very natural way, as opposed to the treatment which they receive in some other theories). The program described illustrates the point that it is possible to define the meanings of relational terms as procedures that work in a way that is utterly remote from meaning postulates and conventional decompositional theories. "The definitions decompose meanings into the primitive components of specific coordinate values that are only interpretable with respect to the spatial models. The meaning of a word is accordingly not a procedure that can do anything by itself, it is a procedure that applies to other procedures" (Johnson-Laird, 1980, p.166). That is, the meanings of words like 'behind' take the form of decompositional procedures which relate to (constrain the operation of) other general procedures which operate on models. The models in turn represent the subject of discourse; this is the basis of the semantics proposed for mental model theory. The simulation demonstrates the use of lexical procedures that interact with the general procedures for constructing, manipulating and interrogating mental models and constraining the general procedures to operate in particular ways.

A further point to note is that such modelling and scanning operations can effect transitive inferences, etc., without such inferences being prespecified in the system in terms of inference schemata; for example, if the program is told that:

A is behind B

B is behind C

this would be represented in the program as:

A
B
C

If the program were asked to verify whether A was behind C, the program would use an identical scanning operation as in the above example. Thus, on finding that both A and C were present in the array, the program would run its verification procedure, testing whether A is indeed behind C; since it would be in this instance, the program would return the value 'true'. The point of note is that the inference schemata associated with relational terms do not have to be specified with those terms, and the vagaries of transitive inference in different spatial arrays (see chapter 1) can therefore be coped with without requiring an extremely large number of postulates. Transitivity is an emergent property of the mental model rather than being specifically 'written in' in advance.

In real life, of course, the procedures involved in scanning would be more complicated, since the locus of points scanned need not be in a straight line and since the actual sizes and shapes of the objects would need to be taken into account. Nevertheless, the simulation demonstrates the essence of how this sort of modelling can work and indeed it is argued by Johnson-Laird that the only way that the factors alluded to in the above caveat could be taken into account would be by constructing an internal model of the sort described in the theory.

There is therefore a twofold advantage of the mental model based theory over conventional theories based on stative decompositions or meaning postulates. Firstly, the theory based on procedures is well capable of accounting for the extensions of expressions, which meaning postulate theory and other 'syntactic' theories of meaning are unable to do, and secondly, the model-based theory is capable of accounting for transitive inferences in a natural and parsimonious way, without multiplying rules and postulates to implausible extents, indeed without necessitating the existence of rules or postulates at all. The operation of procedures on mental models provides a natural and intuitively satisfying way of accounting for such logical inferences.

Having demonstrated the feasibility of a procedural theory utilizing mental models by means of a computer simulation, the question of whether human beings actually do use such models in the course of comprehension is an empirical issue. Johnson-Laird and his colleagues have conducted experiments which test for the existence of mental models and their use in comprehension and deductive inference. For example, Mani and Johnson-Laird (1982) investigated the notion that a listener or reader can construct a mental model of a spatial layout. The subjects heard a series of spatial descriptions involving everyday objects (for example, spoon, cup, knife, etc.) and then judged whether a diagram was consistent with the description they had just heard. For example, subjects would hear a description such as:

The spoon is to the left of the knife.
The plate is to the right of the knife.
The fork is in front of the spoon.
The cup is in front of the knife.

Subjects would then have to judge whether a diagram such as

spoon	knife	plate
fork	cup	

was consistent or inconsistent with the description. If the diagram is considered to depict the arrangement of these objects on a table top (i.e. in plan view) then the diagram is in fact consistent with the description.

The above description was determinate, that is, only one spatial layout is consistent with it, but half of the descriptions which subjects heard were indeterminate, such as:

The spoon is to the left of the knife.
The plate is to the right of the spoon.
The fork is in front of the spoon.
The cup is in front of the knife.

This description is consistent with at least two radically different spatial arrangements:

spoon	knife	plate		spoon	plate	knife
fork	cup			fork		cup

After subjects had evaluated a whole series of pairs of descriptions and diagrams, the subjects were given an unexpected test of their memory for the descriptions. This test took the form of their ranking four alternative descriptions in terms of their resemblance to the description that had actually been presented. The four alternatives consisted of the original description, a description that was inferable from a model of the original description, and two descriptions which corresponded to arrangements different from those in the original description - that is, two confusion items. An example of an inferable description for the above would be:

The fork is to the left of the cup

The results of the study showed that a) subjects ranked the original and inferable descriptions as closer to the original significantly more often in the case of the determinate descriptions. This result indicates that subjects were retaining the meaning of the determinate descriptions (more so than with the indeterminate descriptions) since they were not accepting the confusion items so often in the case of determinate descriptions. The second finding was that b) the percentage of trials on which the original description was ranked higher than the inferable description was significantly greater for the <u>indeterminate</u> descriptions. This suggests that subjects confuse original and inferable items more often in the case of determinate descriptions, which fits in with mental model theory, and it also suggests that in the case of indeterminate descriptions, subjects are retaining verbatim detail of the original description to a greater extent than is the case with determinate descriptions.

Johnson-Laird (1981b) interprets these results in terms of the subjects constructing mental models for the determinate descriptions but using a superficial linguistic representation when they encounter an indeterminacy; reconstructing an existing model in the light of subsequent information places a heavy load on the cognitive system, so subjects use superficial linguistic representations in the case of indeterminate descriptions. In the case of determinate descriptions, the construction of a mental model encodes little or nothing of the form of the original description on which the model is based, and subjects consequently confuse inferable descriptions with those actually given. In summary, these results are consistent with mental model theory; determinate descriptions permit the easy construction of a mental model. Once constructed, however, the verbatim details of the original sentences on which the model is based are discarded, and hence the greater confusion of original with inferable descriptions in the case of determinate descriptions. Indeterminate descriptions, on the other hand, tend not to have models built from them and greater verbatim retention is the result.

Ehrlich and Johnson-Laird (1982) looked at the problem of referential continuity and its effect on the construction of mental models. In their experiments, subjects listened to three sentences about the spatial relations between four common objects, for example:-

The knife is in front of the spoon.
The spoon is on the left of the glass.
The glass is behind the dish.

Subjects then had to draw a diagram of the corresponding layout using the names of the objects. If subjects were continuously building up a mental model as they heard the sentences, the task would be straightforward. If the references were continuous (as in the case of the above example), the model could be built up continuously. Howver, in cases where the premises are in a discontinouous order, for example, where the first two assertions refer to no item in common, it is more difficult to build a mental model. For example:-

The glass is behind the dish.

The knife is in front of the spoon.

The spoon is on the left of the glass.

In such cases as this, the listener must either represent the premises in a superficial linguistic form until the time comes to make the drawing, or else the subject can attempt to firstly build two separate mental models corresponding to each of the first two assertions, and subsequently combine them. The prediction that discontinuous references in the assertions would prove more difficult for the subjects was borne out by the results; the percentage of correct diagrams was 69% for those based on descriptions with referentially continuous assertions, and only 42% in the case of diagrams based on descriptions with referentially discontinuous assertions.

A third type of ordering of the assertions yielded an interesting result; reference can be continuous for the first two assertions but discontinuous for the third assertion. In this case, mental model theory would predict that a single mental model could be continuously built up on the basis of the first two assertions, and the third assertion could be integrated into this model more easily than in the case where the discontinuity occurs early on, in the first two assertions. For example:

The spoon is on the left of the glass.
The glass is behind the dish.
The knife is in front of the spoon.

The third assertion has nothing in common with the second, but it has one referent in common with the first assertion; accordingly, a hypothesis in terms of mental models would be that this sort of discontinuity should be less difficult than in the case of discontinuity in the first two assertions, since referential continuity in the first two assertions allows a consistent single model to be built. The discontinuity at the later point in the discourse is not, therefore, so disruptive to building and retaining one consistent model. In fact, this ordering of the premises yielded 60% correct diagrams, and this result was not reliably different from performance with continuous diagrams.

Johnson-Laird (1981b) argues that a number of discourse phenomena are interpretable by means of postulating the existence of mental models of the discourse; an example would be the problem of <u>reference</u>. This has been something of a problem for philosphers in the past, resulting in such philosophical distinctions as referential vs. attributive usages of definite description. A referential description is one which picks out a particular unique individual, whilst an attributive description is one which merely picks out whoever satisfies the description. This distinction was made by Donellan 1966).

Johnson-Laird (1981b) argues that the question of existence is independent of the referential-attributive distinction; both referential and attributive designations can be used to pick out an individual who does not exist. As Johnson-Laird argues, reference and the mechanisms that underlie it do not depend on whether there is something in the world that an expression picks out. What is crucial is that there is something in a mental model to which the expression refers; it does not matter as far as psychological processes go whether the mental model bears a veridical relation to reality. This consideration has interesting consequences.

Johnson-Laird and Garnham (1980) make points of great interest when they write,

> "Uniqueness in a model rather than in reality is what controls the use and interpretation of definite descriptions. If a speaker is to communicate felicitously, then he must

consider whether an entity will be unique in his listener's model. Utterances need seldom be more than clues about how to change a discourse model; they depend for their interpretation on what a listener knows, but that interpretation in turn modifies or extends the discourse model. A discourse model is in part a surrogate for reality. Indeed, it is sometimes convenient to speak as if language were used to talk abut discourse models rather than the world." (Johnson-Laird and Garnham, 1980, p. 377).

Thus, mental model theory can be used to provide an explanation of the problem of definiteness; the referential problem centres on the problem of picking out entities within models, not necessarily within the world. The above quote also has a direct bearing on pragmatic phenomena (in Morris's definition, pragmatics is the study of the relation between signs and the users of those signs). The choice of a particular type of referent depends upon the speaker's model of what the listener knows. The speaker, as it were, anticipates the nature of the mental model built by the listener on the basis of the discourse and plans his subsequent discourse accordingly. Thus, there is a possibility of attempting to address pragmatic questions using this general approach; the particular listener's knowledge state (or at least, the speaker's notion of the listener's knowledge state) has an influence upon the speaker's choice of referring expression.

To summarize, mental model theory postulates two levels of representation of linguistic input, an initial propositional representation of a form akin to the linguistic input and from which verbatim details of the input are readily recoverable, and a much deeper level of representation in the form of mental models, that is, structural analogues of the state of affairs being described. Mental models constitute an option at the encoding stage; the subject may choose not to construct such an extended representation. Once a mental model has been constructed, the propositional representation on which it was based is discarded, and as a result, verbatim retention of the original discourse is poor but the richness of meaning derived from the discourse is great, and errors in recall stemming from assumptions made at the model-building stage can occur (as in the results of Bransford, Barclay and Franks (1972) and Garnham (1979)). That is, mental models incorporate more than is actually given and this leads to error.

The advantages of a mental model based theory are that mental models provide the 'interface' between natural language and the world. Furthermore, the semantic procedures employed in the construction, manipulation and interrogation of mental models make possible a variety of inferential techniques to enable, for example, transitive inferences to be made without having to explicitly represent rules to make such inference possible. The parsimony of such an account is appealing.

By mirroring the structure of what is being described, the problems of extension and the vagaries of logical inference are overcome, and a variety of subtle referential phenomena, such as the problem of definiteness, can be addressed using this theory. That the problem of extension is dealt with in a natural way justifies the description of this theory as a 'semantic' theory; there is also the possibility of accounting for pragmatic phenomena in these terms.

Although Johnson-Laird emphasizes that function of language which enables us to gain another person's experience of the world by proxy in his exposition of the theory, that is, although once again the theory is couched in terms of what goes on in the mind of the listener or reader when listening to or reading discourse, the theory places plausible constraints upon speech production. In particular, the particular mental model in the mind of the speaker will constrain what he can say, and as Johnson-Laird and Garnham argue, the knowledge state of the listener will also constrain the speaker's choice of phrase. Thus, there is both a plausible set of constraints upon the speaker acccording to this theory (as compared to the theories outlined in chapter 1) but there is also a possiblity of beginning to take stock of the influence of pragmatic phenomena on speech production and understanding.

Although it is rather removed from the mainstream of interest to the work reported in this thesis, it is perhaps worth noting in passing that Johnson-Laird (1980) also argues that mental-model type representations can be used by subjects to represent the premises in syllogistic reasoning tasks. Johnson-Laird and Steedman (1978) produced results which were consistent with an interpretation of the syllogistic reasoning process as one of manipulating and testing mental models. Thus, an advantage of a theory based on mental models is that it can parsimoniously account for a number of cognitive phenomena.

The next chapter will focus upon pragmatics; it will be argued that, although there are no comprehensive formal theories of how pragmatic phenomena constrain speakers' discourse and listeners' understanding of discourse, the phenomena in the domain of pragmatics are of concern to the psychologist, since issues which directly involve social psychology are implicated. It will be argued that the supplementing of a mental model-type theory with considerations stemming from the area of pragmatics (i.e. pragmatic constraints) results in a more complete and intuitively plausible overall picture of language use which does more justice to the complications of the subject than has been possible with earlier theories. Most theorists would agree with the contention that ultimately pragmatics (that is, including syntactics and semantics) will provide a more complete model of communication than has been hitherto formulated. In the next chapter, we will examine some of the possible pragmatic constraints on discourse production and understanding.

CHAPTER 3

'PRAGMATIC' APPROACHES TO MEANING

a) Introduction

In the last chapter, we considered approaches to meaning which were essentially of a 'semantic' nature, that is, they related the signs of the language both to other signs and also to the referents (or models of the referents). We now turn to examine 'pragmatic' approaches to the explication of meaning. Morris's definition of 'pragmatic' was the domain of phemonena pertaining to the relation between the signs of the language system and the users of those signs. As Cherry (1957) put it, "Pragmatics is the most general, inclusive level of study and includes all personal, psychological factors which distinguish one communication event from another, all questions of purpose, practical results, and value to sign users. It is the "real-life" level" (Cherry, 1957, p. 223). As such, the domain of pragmatics is a rather wide-ranging one, encompassing areas of language use which have received attention from linguists and philosphers (such as the problem of 'deictic' or 'indexical' expressions, for example, 'here', 'there', 'left', 'right', 'I', 'you', and so on, whose meaning depends upon the circumstances of usage) and also areas of language use which have received attention from psychologists and sociologists. An example of the latter would be what Clark and Murphy (1982) refer to as "audience design", that is, the way in which particular audiences constrain a speaker or writer's choice of phrase.

A consideration of pragmatics leads naturally into some of the subtle social aspects of communication. For example, Rommetveit's (1974, 1979, 1984) emphasis is upon what particular listeners understand particular speakers to mean when they make an utterance. He therefore discusses what is made known by particular utterances in particular contexts of use. Similarly, Johnson-Laird (1977b) distinguishes between the meaning of a sentence and its <u>significance</u>. Johnson-Laird and Wason (1977) give an example of a sentence whose <u>meaning</u> is perfectly clear to the reader but which has no <u>significance</u> for him:

> The men object to their new rates of pay.

(Johnson-Laird and Wason, 1977, p. 346). This sentence has little or no significance for the reader, except insofar as it is made to illustrate a point about language. If, however, the reader or hearer of such a sentence was a manager or director of a company who had recently changed its employees' pay rates, the utterance would have a great deal of significance. In this example, the social and physical circumstances of the utterance make a great deal of difference to the significance of the message, or in Rommetveit's terms, what is 'made known' by it. (Such terminological differences will not be explored in any detail; the expositions of various theorists' work given in this chapter will be couched in the theorist's own terms, and theoretical similarities will be of greater importance than terminological differences). One of the general points which will emerge clearly from such considerations is that what is understood by an utterance (what significance it has for the listener or reader) is influenced very much by pragmatic factors.

Johnson-Laird and Wason (1977) state that the significance of an utterance is inferred on the basis of five main components:

1. Its meaning.
2. Its linguistic context.
3. Its social and physical circumstances, including a knowledge of the speaker.
4. A knowledge of the conventions governing discourse.
5. General knowledge.

Johnson-Laird and Wason point out that the structure of discourse is presumably determined by an interplay of these five factors; no purely linguistic analysis is going to succeed in elucidating the way in which discourse is structured, a point made by the ethnomethodologists. A further point to emerge from the above listing of factors is that both Rommetveit and Johnson-Laird and Wason are in agreement about the fact that the significance of an utterance (or what is made known by it on a particular occasion) is not adequately characterized by its meaning (in the narrow linguistic sense of 'literal meaning').

The general point that the significance of an utterance can not be characterized solely in linguistic terms has already been touched upon in the earlier discussion of model-theoretic semantics (see Chapter 2), where it was noted that there is no one-to-one correspondence between the grammatical structure of a sentence and the kind of communicative act that is performed what that sentence is uttered on a particular occasion. These subtleties of language use in everyday situations were explored by the philosopher J. L. Austin. Austin (1962) noted that people often perform actions by uttering a sentence, and that these sentences cannot be classified as true or false (as logicians had previously done; the meaning of a sentence for a logician was its truth value). An example would be in making a promise. The sentence may succeed or fail _as an action_ (the speaker may or may not fulfil his promise) but regardless of this the sentence cannot be described as true or false. Such sentences are termed _performatives_.

These ideas led to the view that all utterances could be described as actions called _speech acts_. Austin classified speech acts into three classes:

i) locutionary acts - i.e. the act _of_ saying something. This involves speaking in a way which conforms to vocabulary and grammar.

ii) illocutionary acts - i.e. the actions performed _in_ making an utterance, such as "promise", "request", "warn" and "inform" - these are examples of illocutionary acts.

The _illocutionary force_ of an utterance is the effect that the speaker _intends_ the utterance to have. The intended effect is usually that the hearer knows what illocutionary act was performed. The hearer must correctly identify the illocutionary force of the utterance in order for the intended effect to be successful.

iii) perlocutionary acts - i.e. the acts performed _by_ making the utterance. This is, in effect, the _actual effect_ upon the hearer - whether he has been scared, convinced, or whatever. A perlocutionary act may or may not be intended by the speaker.

When considering language from this point of view, the notion of _indirect speech acts_ came to the attention of theorists. For example,

"Can you pass the salt?"

is not a question about salt-passing prowess; it is a request for the hearer to pass the salt. In such cases, the hearer has to correctly identify the illocutionary force of the utterance - treating the above utterance as a straightforward interrogative and giving the reply "Yes" would not be the effect intended by the speaker. A reply of "Yes" would thus be a failure to grasp the illocutionary force of the utterance. Determining the illocutionary force of a sentence is always a matter of inference beyond what is actually said. No matter what elaborate linguistic construction is employed by the speaker (of the form "I swear to you that I am really asserting that . . ." (and so on)), he could always be playing the fool (Johnson-Laird and Wason, 1979). Thus, correctly identifying the illocutionary force of an utterance is a pragmatic problem of some importance.

Earlier in this section, we noted that one of the factors which plays a part in determining the significance of an utterance is that of the listener (and speaker) possessing a knowledge of the conventions governing discourse. The philosopher Paul Grice has formulated some of the conventions governing discourse in our culture, and formulates them as a set of precepts or conversational maxims, as follows:

> The cooperative principle: Make your conversational contribution such as is required by the accepted purpose or direction of the exchange in which you are engaged.

Specific conversational maxims:-

Quantity:	Make your contribution as informative as is required - neither too much nor too little information.
Quality:	Don't say what you believe to be false. Don't say what you lack evidence for.
Relation:	Be relevant.
Manner:	Avoid obscurity and ambiguity. Be orderly and brief.

Such maxims are normally followed implicitly. Johnson-Laird and Wason illustrate this point by demonstrating what happens when the maxims are violated. In such cases a listener may draw inferences from such violations: e.g.

Question - How old are you?

Answer - Nice weather we're having.

From the reply, an inference, such as the inference that a relevant answer is unspeakable, may be drawn.

That adherence to such maxims is a social-pragmatic problem is exemplified by the conversational maxim of quantity. Clearly, in order to make one's conversational contribution as informative as is required, giving neither too much nor too little information, one needs to anticipate the listener's knowledge state, beliefs and presuppositions. This anticipation is central to what

Clark and Murphy term "audience design", and an exposition of their views will be given later in this chapter.

The foregoing serves to illustrate Cherry's point that pragmatics is the "real-life" level of study; illocutionary force, indirect speech acts, and conversational maxims are used by us every day in ordinary conversation. Pragmatic phenomena are of importance, therefore, and in this chapter several different aspects of pragmatic phenomena will be examined, and the work of several different theorists will be reviewed.

The first aspect of pragmatics to be considered is one of the most basic pragmatic problems, that of deixis. Deictic or indexical expressions are expressions whose reference depends upon contextual factors like the position in space of the speaker, his direction of gaze, and other contextual factors. A speaker and listener have the problem of jointly agreeing upon the denotata of deictic terms like 'here', 'there', and the like; speaker and listener must set up a common deictic space and coordinate their perspectives to ensure a common interpretation of deictic terms. We will consider Karl Buhler's (1982, originally published 1934) theoretical treatment of deixis since it remains an influential treatment of the subject. We will also consider briefly the work of Klein (1982) and Wunderlich and Reinelt (1982) which examine how some deictic terms are used when an individual gives route directions to a destination.

Kaplan's (1977) theory of deixis and Miller's (1982) discussion of Kaplan's theory also raise interesting pragmatic questions. For example, when pointing to an object in order to refer to one of its properties (pointing to grass to refer to the colour green, for instance) the speaker has to utilize his general knowledge about the relation between the referent and the demonstratum, <u>and also</u> his beliefs about the <u>hearer's</u> knowledge of such relations.

Given that a speaker and listener have successfully set up a common deictic space with agreed-upon delimitations of that space, one of their next important pragmatic problems is the question of how they refer to the entities under discussion, whether those entities are physically present or not. This raises the problem of mutual knowledge, that is, knowledge that is common to both parties and, furthermore, knowledge which both parties know that they share. That shared knowledge influences the choice by a speaker of particular referring expressions is illustrated by work which shall be reviewed, for example, the work of Clark and Murphy and the work of Marslen-Wilson et al. The work of a variety of theorists points, it will be argued, to the importance of pragmatic factors in influencing the speaker's choice of referring expression.

This area of pragmatics leads on rather naturally to a further layer of pragmatic complexity, so to speak: that is, the particular social identities of speaker and listener and how this factor enters into a speaker's choice of phrase and a listener's understanding of what has been said. Particular listeners may understand different things from the same utterance depending upon the identity of the speaker. The work of Rommetveit stands out in its advocacy of a social-pragmatic approach to both the above phenomenon and also the more "elementary" ones of deixis. For this reason, the work of Rommetveit, being of a comprehensive and generally thought-provoking nature, will be reviewed in a separate section of its own.

The last problem of pragmatics to be considered will only be raised briefly. This brief treatment is necessary because very little theoretical work has been done in the area. The problem is the social-psychological question of whether one party to a dialogue might exert overall <u>control</u> of it. That is to say, the possibility is raised that one interlocutor may dominate the dialogue in some sense and that this domination may be reflected in terms of linguistic aspects of the dialogue. This intriguing notion, if it were to be supported empirically, would add a further layer of complexity to the area of pragmatics, and should prove to be of great interest to the social psychologist.

Throughout this chapter, the emphasis on the meanings of natural language expressions in natural dialogue will be retained, since it is the question of meaning which constitutes the primary focus of interest in the work reported in part II of this thesis. The importance of pragmatic factors for meaning, in addition to 'syntactic' and 'semantic' aspects, will be emphasized.

b) The pragmatics of deictic expressions

As noted above, perhaps the most basic problem for two persons engaged in a dialogue is the problem of the referents of deictic expressions. Deictic or indexical expressions are those expressions such as 'here', 'there', 'I', 'you', 'this' and 'that', whose reference depends upon the context is which they are used. In other words, these terms have no fixed denotata, the denotata involved varying depending upon who is speaking to whom, where and when. The first problem for a speaker and listener, therefore, is that they have to set up a system of reference relative to which the denotata of deictic terms are agreed.

In examining the problem of deixis we shall firstly consider the theories of Karl Buhler (1982, originally published 1934), whose views on the subject are today still influential. Other interesting notions are raised by Kaplan's (1977) theory and Miller's (1982) discussion of Kaplan's theory.

Buhler (1982) proposes a two-fields theory of language, partitioning language into the symbolic field (which contains the vast majority of lexical items, such as names, whose referential function is less context dependent and which more abstractly reflect symbolic representation of objects, properties and events) and the _deictic_ _field_ (which contains the deictic or indexical expressions whose reference depends upon the particularities of context). Buhler identifies three areas of use of deictic terms: (a) _demonstratio_ _ad_ _oculos_ or 'verbal pointing' in which a deictic term such as 'there' replaces the pointing finger gesture; (b) _anaphorical_ _deixis_, where a deictic term such as 'here' is used anaphorically to refer forward or backward to an abstract 'place' in discourse; and (c) _deixis_ _at_ _phantasma_ (that is, deixis in the imagination, for example, where a novelist leads a reader into the realm of constructive imagination and there uses deictic terms). Buhler treats the 'pointing' function of deictic expressions as primary and treats the other two functions (anaphora and phantasma) as derivative from the primary pointing function. We shall examine each of the three in turn.

(1) _Demonstratio_ _ad_ _oculos_, or 'verbal pointing'.

In their 'pointing' function, Buhler argues, deictic words function in the same way that a physical signpost does, that is, they act as way-indicators or signals (for example, 'here' and 'there'). The deictic terms replace or accompany the pointing forefinger gesture in the case of _demonstratio_ _ad_ _oculos_, and 'point' verbally, hence the analogy with a roadside signpost. The speech event, however, differs from the physical signpost in that it is an _event_ - a complex human action- in which the speaker has both a position in the terrain (as does the signpost) _and_ a _role_ in the action. The speaker's role is that of sender as opposed to receiver. Deictic words therefore include not only positional terms but also personal pronouns such as 'I' and 'you' which _refer to_ the roles of the interlocutors (rather than _denoting_ them in the manner in which names, for example, do).

Buhler therefore argues that the naming words function as _symbols_, while deictic terms function as _signals_. Deictic terms receive their full precision of meaning within the deictic field of language, and can do so there alone. For example, the denotata of 'here' and 'there' vary with the speaker's position precisely as the referents of 'I' and 'you' change with each change of speaker. Deictic words therefore can refer to a vast variety of persons, locations and objects depending upon who is speaking to whom, where and when they are in conversation, and what the conversation is about. This contrasts with the naming words, each of which is used by all speakers of the language to refer to the _same_ object.

Buhler further argues that there are conceptual similarities between first person pronouns and close-to-the speaker local deictic terms, and that there are conceptual similarities between third person pronouns and demonstratives. He backs up this argument with evidence from linguistic studies which suggested that pronouns and certain deictic words originated from common word stems, and argues that the postulated similarities are phenomenologically plausible (Buhler, 1982, pp. 16-17). There are therefore good arguments to sustain the point of view that pronouns should not be separated theoretically from positional deictic terms: both personal pronouns and positional words are nothing but deictic words.

Buhler describes a coordinate system of subjective orientation with respect to which the reference of deictic terms is fixed by the interlocutors. All parties to the verbal exchange are and remain attached to this system of coordinates. The point of origin for the coordinate system is referred to as the <u>origo</u> and is represented by two perpendicular strokes on the page:

Three deictic words must be put at the place of O, namely HERE (the verbal place marker), NOW (the verbal moment marker) and I (the verbal sender marker). These words, Buhler argues, draw our attention to the speech sound out of which they are constructed, and in so doing they achieve their functions as place, time, and sender markers.

It is from the origo that all other spatial positions are identified ('there' as opposed to 'here', which is located at the origo) and it is from the origo that all other time points are identified ('then' as opposed to 'now', which is located at the origo). Thus, all deictic terms' references are fixed relative to the origo.

Clearly, in fixing such references, differences in visual perspective between the speaker and listener may become important. If the speaker and listener are directly facing one another, the speaker's right will be the listener's left, and vice versa. In some cases, this difference in perspective will be negligible, as when the denotatum of 'here' encloses both speaker and listener. However, the difference does become important in many cases, and in these cases, the listener has to 'take over' the orientation or perspective of the speaker, projecting it onto his own (Klein, 1982; Rommetveit, 1974, 1979). In some cases, the speaker may adopt the listener's visual perspective. Buhler gives the following example: "If I stand as a leader facing an aligned formation of gymnasts, I would conventionally choose commands, such as 'forward, backward, to the left, to the right' not in my own system of orientation, but in theirs, and the translation is psychologically so straightforward that every group leader learns to master it . . ." (Buhler, 1982, p. 14). He further points out that this taking of the perspective of another can be achieved " . . . without mental gymnastics . . ." (p. 14). This notion of taking the perspective of a fellow conversant is further elaborated by Rommetveit

(1974; see section d) of this chapter).

Klein (1982) summarizes what is involved for a speaker and listener when they use deictic terms, as follows: (a) Firstly, the two persons must set up a common deictic space - for example, the deictic space relevant to their conversation might be the space of visual perception, or some geographical space such as, for example, the building in which the speaker and listener are located. (b) The two conversants must then fix the position of the origo. (c) The conversants' next problem is to coordinate their perspectives (see above discussion), and (d) what 'here' and 'there' refer to must be delimited. 'Here' is _some space enclosing the origo_, not the origo itself, and the boundaries of 'here' are fixed by context and often by our knowledge of the world. Exactly how this delimitation of local deictic space is achieved (whether by verbal context, components of factual knowledge, or gestures) is an important question for research. (e) The system of oppositions of the language must be used. For example, in English there is a two-term system of opposition which is summarized as 'proximal-distal' (for example, 'here-there'). Again, this is complicated by the fact that what is proximal and what is distal or non-proximal depends on the context. (f) A final potential complication listed by Klein is that of _analogical deixis_, for example, when reference is made to a spot on a map. Two deictic spaces are involved, the map and the geographical space of which the map is an analogue (hence 'analogical deixis'), and the problem here is how the two deictic spaces are related.

Klein's summary underlines the fact that the use of local deictic terms (that is, those terms like 'here' and 'there' which refer to locations) is a complicated business.

(2) Anaphorical deixis

The first of two derivatives of 'pointing' deixis which Buhler lists is anaphorical deixis. This is where deictic terms, particularly 'this' and 'that', are used to refer to 'locations' in an ongoing piece of discourse. This referring forward or backward to different segments of a text presupposes that the speaker and listener (or isolated thinker) have access to the flow of discourse as a whole, where parts may be retaken up and anticipated. Many terms are used in anaphorical deixis, including 'here', 'there', 'this' and 'that', and 'therefore' and 'thereafter' are used anaphorically (these latter two terms stem from 'there' which is itself a positional deictic). The anaphoric reference to places in the organization of discourse involves largely the same deictic words as does demonstratio ad oculos.

(3) Deixis at phantasma

The other derivative of 'pointing' deixis described by Buhler is termed deixis at phantasma: this is where, for example, a narrator takes the listener into the realm of constructive imagination and uses the same deictic terms to refer to imaginary entities as he would to refer to physically present objects. Buhler asks what happens to a reader or listener in such a case who becomes 'lost' in imaginary travels, for example. What about the verbal pointing used by the author (or by the thinker himself in some cases) which the reader (or thinker) follows in his 'phantasma'? The person is not removed in a literal sense from his perceptual situation. Buhler divides deixis at phantasma into three major cases:

(a) The first major case is where, to use Buhler's metaphor, 'the mountain moves to Mohammed' (Buhler, 1982, p. 28), that is, the person mentally 'shifts' what is being imagined into the physically present current environment where it is 'localized'. In this case, the current body feeling representation together with the visual perceptual orientation are maintained and integrated into what is imagined.

(b) The second major case is where 'Mohammed moves to the mountain'. In this case, the person is transported in the imagination to the geographical place which is being imagined, the person seeing what has been seen before in real wandering. In this case, the current body feeling representation is 'taken along' into the imagination.

(c) The third major case is where both Mohammed and the mountain remain where they are, but Mohammed perceives the mountain from his current location. An example would be where a person when wandering in a strange town is asked the location of the railway station from whence he came. Often, he is able to point in its approximate direction. This third major case is usually a labile and transitory experience, and, Buhler argues, is an additive whole or <u>superposition</u> of two localizations, one of them corresponding to the first and the other to the second major case.

Buhler argues that all cases of deixis at phantasma are reducible to one or other of the above three major cases.

One particular case of deixis at phantasma which occurs commonly is involved when we give route directions to another person who has requested them from us. Klein (1982) and Wunderlich and Reinelt (1982) comment on this phenomenon. It involves the speaker and listener firstly setting up a deictic space and coordinating their perspectives with respect to their joint perceptual field. The speaker (i.e. the route-direction generator) activates his cognitive map, localizes his current location and the destination, and then 'takes' the listener on an 'imaginary walk'. The 'imaginary walk' is used to generate a route-description. The speaker initially adopts the listener's perspective, and in the imaginary walk the origo is thought of as shifting. The skeleton of the route is marked by the selection of fixed reference points or points of orientation (landmarks), and the speaker explicitly or implicitly sets up a new perspective at each landmark in the route description. This route description - giving involving an 'imaginary walk' is an example of Buhler's second major case (see above).

Kaplan (1977) also developed an interesting theory of deixis. Kaplan distinguishes _demonstrative indexicals_ which preserve a gestural component in spoken language, from _pure indexicals_ (such as 'I', 'here', 'now'), which are used to define a point of origin for a spatio-temporal coordinate system relative to which demostrative indexicals can be produced and interpreted. Kaplan argues that all deictic terms are _directly referential_ - that is, in order to evaluate the truth of a sentence which uses a deictic word, it is always the _referent_ of the word in context which is relevant to the evaluation of the truth of the sentence. For example, if we speak the sentence 'He (accompanied by the demonstration of a particular person) is a thief', it is only the person demonstrated (the referent) who is relevant to the truth or falsity of the sentence. This contrasts with a situation where a non-deictic phrase is used, such as 'The butler is a thief'. This latter sentence depends for its truth on who the butler is: 'The butler' is not directly referential in the way that 'He' is. It does not directly refer to a particular individual (Furthermore, to define 'He' as 'the male person that the speaker is demonstrating' does not result in an equivalently directly-referring expression. 'He' is directly referential and the longer phrase is not. Hence, indexical terms do not have synonyms).

Kaplan further distinguishes between the _content_ if a sentence and its _character_. This is a distinction which makes good intuitive sense with respect to deictic expressions. There is one sense in which a deictic word has a variable meaning, since it can be used to refer to a variety of individuals and objects depending upon its context of use. This variability in meaning is the _content_. The content of a sentence is the proposition it expresses (which is equivalent to an intension - see chapter 1): content is obtained by replacing all deictic terms by the demonstrata they determine. Content is thus sensitive to context; thus, for example, the term 'this' may be used to demonstrate a variety of objects in different contexts and accordingly the content (intension) of what is said varies.

On the other hand, there is a sense in which a deictic term has a constant meaning: it has a constant meaning in that all speakers of the language use the term in the same way (for example, the term 'there' always refers to a more distal location whereas 'here' always refers to a more proximal location). The term _character_ is used to refer to the constancy of meaning of a deictic term. A character is (partly, at least) a semantic rule that we can use to determine the content in different contexts.

Character is not sensitive to context (for example, 'there' always refers directly to a demonstrated location whenever it is used in a demonstrative sense).

Thus, Kaplan's notion of 'content' captures the idea that deictic words can have a variety of meanings, and his notion of 'character' captures the constancy of usage of a deictic word. However, the semantic rule is only one component of character in Kaplan's theory; a demonstrative can be thought of as assuming a character only when completed by a particular demonstration. A character may, therefore, be likened to a manner of presentation of a content, and the appearance of the demonstratum is part of the completed character.

In broad terms, the distinction between content and character makes good intuitive sense when applied to deictic expressions. The difference between content and character has been generally overlooked because character is fixed in non-deictic language; the same content is evoked in all contexts by a non-deictic word.

In discussing Kaplan's theory of demonstrative reference, Miller (1982) raises some theoretically interesting complications for Kaplan's theory. One such complication is the case where the referent of a deictic expression is not the demonstratum of the demonstration used - this is referred to as <u>deferred ostension</u>. Quine (1971) gives the example of someone pointing to a gas (petrol) guage in order to show that there is petrol in the tank. Such deferred ostension as occurs in this example may intuitively seem unproblematic given the conventional relationship between the state of emptiness of the tank and the reading of the gauge. However, referential convention is not always necessary to support deferred ostension. Deferred ostension can occur in novel situations, as, for example, in the case when a waitress says 'He's sitting at table 20' whilst pointing to a ham sandwich. Nunberg (1977) gives this latter example and others, and argues that a complete account of demonstrative reference will require (at least) two theories. One, like Kaplan's, will explain how a deictic term enables a hearer to identify a physically present demonstratum. The other theory will explain how the hearer's knowledge of the demonstratum enables him to identify the referent of the deictic term. This latter theory will be required to explain how an ostensive gesture to a ham sandwich does in fact refer to the person who ordered the ham sandwich and not the sandwich itself. Nunberg conceptualizes this inference (from demonstratum to referent) in terms of a _referring function_ of a given use of a term which, given a demonstratum, delivers a referent.

Nunberg argues that deferred ostension will work when the property relating the demonstratum to the referent (that is, the property being used as the referring function) has high _cue validity_. A property has high cue validity if it distinguishes one object from a range of alternatives - thus, for example, if there is only one red book on a shelf, then 'red' has a high cue validity for that book. In using deferred ostension, the speaker has to choose a referring function which has the highest cue validity for relating the demonstratum to the intended referent. In other words, the link between the demonstrated object and the referent has to have high validity as a cue which enables the listener to infer the referent from the demonstratum.

Examples of referring functions include the hypostatic function (i.e. 'is a type of'), in which a particular individual object is demonstrated in order to refer to all individuals of that type. For example, 'That (pointing to a chair) is commonly seen in eighteenth century interiors'. The particular chair being pointed to is not the intended referent; instead, the speaker is referring to all chairs of that type. The listener knows this from his general knowledge of eighteenth century interiors and the conventions of chair ownership. The hypostatic interpretation of the deictic reference goes so smoothly that we scarcely notice that the demonstratum of the subject phrase is not the referent. Another referring function is exemplified by the case where we demonstrate an object (such as grass) in order to refer to one of its properties (e.g., the colour green).

Thus, the choice of referring function should be based on the ready inferability of the referent when given the demonstratum, i.e. the property relating the demonstratum to the referent (the referring function) should constitute a readily inferable relationship. This inferability depends very much on the characteristics of the listener - the choice of a referring function therefore has to be made in the light of what the speaker believes that his listener is likely to believe or know. This underlines the fact that the problem of deferred ostension involves a considerable pragmatic component.

Miller concludes that demonstratives are directly referential even though a referring function may be inserted between a demonstratum and the referent of a demonstrative term; this results in an extended form of direct reference.

Deixis thus stands squarely within the domain of pragmatics, both in the space-delimiting referential function and in the more subtle phenomena such as deferred ostension. The relation between deictic dissection of space and proxemics, (our perception of social and personal space (Hall, 1966)) would be an interesting, if complex, problem.

The more subtle phenomena of deferred ostension illustrate the fact that the presumed knowledge state and beliefs of the listener influence the speaker's choice of phrase (specifically, affecting his choice of referring function in the case of deferred ostension). This in turn raises the more general question of the extent to which the sharing of knowledge and the identities of the interlocutors influence their speech. We shall now examine these issues, reviewing in particular the notions of <u>mutual knowledge</u> (Clark and Marshall, 1981) and <u>audience design</u> (Clark and Murphy, 1982). These topics in turn lead naturally to more general questions of how the meanings of non-deictic, symbolic words are affected by such pragmatic and social factors as the identities of the dialogue participants.

c) <u>Mutual knowledge and audience design: pragmatic factors which influence a speaker's choice of referring expression</u>

In the previous section of this chapter, the pragmatics of how deictic expressions are used was examined. It was concluded that particularly in the case of deferred ostension (where a speaker points to one object in order to refer to <u>another</u> object), the assessment by the speaker of the listener's beliefs and knowledge states is necessary. This section of the chapter will examine how a speaker's choice of referring expression in general is influenced by pragmatic factors (that is, evidence will be reviewed which suggests that not only deictic, but also other forms of referential devices' usage are influenced by pragmatic factors). It will be argued that the work of a variety of different theorists supports the view that, when choosing a particular expression with which to refer to something, the speaker is forced to take account of who his listener is and which particular domains of knowledge they can be expected to share.

The first theory of relevance in this sphere is the cognitive theory of semantics (in the sense of "meaning") advanced by David Olson (1970). Olson argued that semantic decisions (about which particular referring expressions to use) by a speaker are based, not on syntactic or semantic selection restrictions, but on the speaker's knowledge of the intended referent. In arguing that semantic decisions are based on knowledge of referents rather than linguistic selection restrictions, Olson was advancing a similar argument to that of Johnson-Laird (1981a; see chapter 2).

Olson's fundamental argument is that semantic decisions about which type of referring expression to use are based on the speakers having to distinguish the intended referent from a (perceived or imagined) set of possible alternative referents: "We shall adopt the point that words do not 'mean' referents or stand for referents, they have a use - they specify perceived events relative to a set of alternatives; they provide information" (Olson, 1970, p. 263).

This thesis is illustrated by his 'paradigm case', in which a gold star is placed under a small, round, white, wooden block. The speaker, who saw the star being placed under the block, is then asked to tell a listener (who did not see the star being placed under the block) where the gold star is. When the description is being given, the target block is placed in an array of alternative blocks. In one case, the alternative block is a square, white, wooden block. In another case, there are three alternative blocks present, a round, black one, a square black one, and a square, white one. In the first case (with one square, white alternative block also present), the speaker would describe the target block as " . . . the <u>round</u> one", whereas in the second case (with three alternative blocks also present), the speaker would describe the target block as " . . . the round, white one". Thus, the particular set of possible alternatives from which the referent must be distinguished affects the speaker's choice of description. The paradigm case therefore illustrates the general thesis that semantic decisions regarding the type of description to be generated are based upon the set of alternatives from which the referent must be distinguished.

This general thesis has some interesting corollaries. Firstly, this point of view implies that words do not name things: " . . . words designate, signal or specify an intended referent relative to the set of alternatives from which it must be differentiated" (Olson, 1970, p. 264). Another consequence of this point of view is that the conception of meaning is altered: the meaning of an utterance is the <u>information</u> provided by the utterance to a listener, where "information" refers to any perceptual or linguistic cue that reduces the number of alternatives to the intended referent. Thus, traditional views to the effect that words symbolize the objects they designate is challenged: " . . . words (or utterances) neither symbolize, stand for, nor represent referents, objects, or events. They serve rather to differentiate some perceived event from some set of alternatives" (p. 265). The meaning of a word is therefore its partitioning of a set of alternatives; meaning is reconceptualized in informational rather than symbolic terms.

Another interesting corollary of this view follows from the fact that all of the information required by the speaker is perceptual and is available to him before he generates the utterance (as is exemplified by the paradigm case). It follows from this that utterances are redundant for the speaker but provide information for the listener. Furthermore, since a word or utterance specifies the perceived referent <u>and also</u> (implicitly) the set of excluded alternatives, it contains <u>more information</u> than the simple perception of the referent itself. A description of an object implicitly specifies the set of alternatives from which it has been distinguished, and explicitly specifies the critical feature which distinguished the object from the alternatives. Perception of the object alone (for example, the target round, white block) does not specify the critical feature which distinguishes it; the critical feature could be its colour, the grain of the wood from which it is made, the way in which it is handled by the experimenter, and so on. Thus, there is more information in the utterance than in the perception of the object alone.

Much of our conversation about objects involves sets of alternatives which are not physically present. Olson accordingly hypothesizes that the speaker makes his semantic decision on the basis of <u>inferred</u> alternatives - he infers what the alternatives being entertained by the listener are. Such an inferential process is a purely pragmatic one; who the listener is, how well the speaker knows him, what they mutually know, and other such pragmatic factors are going to be of great importance in making this inference.

Olson's theory, whilst emphasizing cognition (knowledge of referents) and perception, clearly involves a pragmatic component which centres on what 'common ground' the speaker and listener share, and this in turn relates to Clark and Marshall's notion of 'mutual knowledge' which will be discussed later.

The principal objection to Olson's theory is that people often use a more specific description than Olson's model predicts. For an array with a rock and a dog in it, people say "Look at the dog" rather than "Look at the living thing" as Olson's model would predict. This level of preferred naming, at a middle level of abstraction, has been called the basic level of abstraction (Rosch, 1976). The basic level of abstraction maximizes two factors simultaneously, namely (1) basic level terms are specific and (2) they are dissimlar to other categories. It would appear that speakers design their descriptions to convey as much information as possible yet not to make unnecessary or inappropriate distinctions in the particular situation of speaking (Clark and Murphy, 1982).

Another study which directly tackles the problem of a speaker's choice of referring expressions is that by Marslen-Wilson, Levy and Tyler (1982). The approach of these theorists was methodologically unusual in that they had a single speaker read a single story then retell it from memory to a listener. The story was a comic-book story concerning a battle between two superhuman monsters, 'The Hulk vs The Thing' or the 'battle of the behemoths'. Whilst the speaker had to retell the narrative from memory, the comic book cover was available to him on his lap, and he used this for the purpose of pointing to the (depicted) protagonists on some occasions. The speaker was video- and audio-taped, and the study examined the manner in which the speaker introduced and subsequently referred to the protagonists with respect to the informational context of the descriptions.

The main motivation behind the study was to assess whether the speaker adjusts the nature of the referential devices he is using to the informational conditions under which they occur, and therefore allows the listener to unambiguously resolve the references at the point at which they occur in the speech stream. This makes the basic assumption that "the ways humans produce speech are necessarily closely adapted to the ways that humans can comprehend speech" (Marslen-Wilson et al., 1982, p. 339). Speech production is approached from the perspective of a theory of the listener, and assumes a complementarity between speaking and listening.

The study thus examined whether there was any systematic pattern in the distribution of different referential devices across different informational contexts. The different types of referential device varied on a dimension of lexical specificity. Referential devices which are less lexically specific (such as pronouns) presuppose for their interpretation information that has to be supplied by the listener, whereas referential devices which are more lexically specific (such as names) require less information to be supplied by the listener. The various types of referential devices examined, with the most lexically specific first and the remainder in order of decreasing lexical specificity, were: names plus definite descriptions (e.g. "The Hulk, the green guy"), names alone, non-specific noun-phrases or pro-forms (e.g. "one of these guys"), personal pronouns, and zero anaphors (zero anaphors are cases in which the specification of the referential device becomes lexically empty, such as "he takes the merry-go-round and 0 whips it around at supersonic speed". The 0 marks the location of the lexically empty referential device).

The informational context in which these different devices occurred was assessed in terms of a dimension of "degree of embedding". The 'degree of embedding' reflected the levels of organization of the narrative, with the 'story' level being the least embedded level, and the 'episode' and the 'event' levels representing increasingly more embedded levels. The 'story' level of embedding is necessary to capture the fact that the protagonists in the story cannot be talked about until they have been introduced into the story in the first place. The episode level of embedding reflects the fact that the comic-book itself is segmented into a series of distinct scenes. The 'event' level corresponds to the fact that each episode typically contained a number of distinct sequences of actions, with each sequence marked by a change of actor(s) and/or a change of location within the location of the episode.

Thus, Marslen-Wilson et al. essentially sought to examine whether the use of the different types of referential device was correlated with the level of informational context such that the most referentially specific devices (names plus descriptions) were used initially where there is little or no informational context, and the least lexically specific devices (pronouns and zero anaphors) were used when there was a great deal of informational context (that is, at the most deeply embedded layers of the story). They did in fact find such a regular correspondence between informational contexts and types of referential device: ". . . the actual deployment of referential devices turns out to be precisely constrained by the local environment in which the devices will have to function, and by the extent to which the available intra- and extra-linguistic context can support the requirements of different types of device" (Marslen-Wilson et al. 1982, p. 355).

A closer examination of the on-line resolvability of referential devices revealed some interesting facets of the data which suggest that pragmatic factors are of importance in the resolution of pronominal reference.

Several factors assist in the resolution of pronominal reference. The lexical properties of the pronoun itself constrain the number and gender of the antecedent (for example, the pronoun "He" requires that the antecedent be singular and male). Syntax and prosody also assist the assignment of reference. Many approaches in the psychological and artificial intelligence literature emphasize the importance of these three types of constraint, suggesting that on encountering a pronoun, a search is instituted for possible antecedents which match these lexical properties of the pronoun.

Such a search might, for example, be based on recency; thus, memory is searched for the most recently encountered possible antecedent which matches the pronoun on all three counts, and this is accepted as the pronoun's antecedent.

Marslen-Wilson et al. argue, however, that such search strategies misconstrue the role of inference in pronoun resolution. Such inferences, or 'pragmatic checking' as Marslen-Wilson et. al. put it, assist in resolving the reference by assessing whether the properties predicated of the pronoun in the utterance are plausibly consistent with the properties assigned to the possible antecedents in the previous discourse.

Marslen-Wilson et al. point to cases in their data where such 'pragmatic inferences' alone could resolve the pronominal reference whilst a search strategy based on recency would deliver the <u>wrong</u> antecedent. An example of this is a case where the two protagonists, The Hulk and The Thing, had been fighting on top of a skyscraper building, and The Hulk had just fallen to the ground sixty storeys below. The speaker continues (N.B.: ' . . .' indicates unfilled pauses):

> " . . . so The Thing has to get down to . . .
> the ground level . . . before The Hulk recovers
> enough . . . while he_1's still stunned . . . so
> he_2 rips open the elevator door and just sort
> of slides down the cable . . ."

A recency strategy would assign the correct antecedent for he_1: The Hulk is indeed still stunned, having fallen sixty storeys to the ground, and is currently in progress of recovery from this fall. He_2 would be wrongly assigned to The Hulk by a recency search strategy; he_2 must refer to The Thing, because the action of ripping open elevator doors in order to slide down the elevator cable to ground floor level is inconsistent both with being stunned, and in particular, with being stunned at the bottom of a building, having just fallen from the top. The two pronominal references can only be resolved by recourse to world knowledge. The 'pragmatic checking' is thus necessary for pronoun resolution in this instance (Marslen-Wilson et al., 1982, p. 363).

The checking is 'pragmatic'* in two senses; the sense used by Marslen-Wilson et al. is that what is predicated of the pronoun is <u>plausibly</u> consistent with the antecedent. The 'pragmatic checking' is also pragmatic in the technical sense employed in this volume. It is pragmatic in the more technical sense in that the speaker is relying upon the listener sharing certain world knowledge and using this to resolve the pronominal references. In this case, the speaker relies upon the listener's knowledge of which states (such as being stunned at the bottom of a building) are consistent or

* N.B. 'Pragmatic' will be enclosed in quotes when used in its non-technical sense.

inconsistent with particular actions (such as ripping open elevator doors in order to slide down to the ground level). It is worth remembering that the speaker was previously shown to display great sensitivity in his choice of different referential devices relative to the informational context in which they occurred. It is therefore unlikely that his reliance on mutual knowledge and 'pragmatic checking' on the basis of that mutual knowledge is a lapse into laziness on his part; rather, it suggests that mutual knowledge is an important entity which speakers and listeners exploit with some frequency. Marslen-Wilson et al. note (p.361) that all 35 pronouns in the sample are disambiguated by the 'pragmatic' properties of the utterances in which they occur, when interpreted relative to the discourse properties of the potential antecedents.

Once again, this example points to the speaker and listener using mutual knowledge to guide the description generation and reference resolution problem. An issue which naturally arises from this takes the form of a simple question: how does the speaker know that a particular piece of knowledge is shared by himself and his listener? The problem of mutual knowledge and its influence upon reference generation has been examined by Clark and Marshall (1981), who examined the relation between mutual knowledge and definite reference.

Definite reference (when we refer to an object name preceded by the definite article "The . .", for example) has inspired a vast literature in the disciplines of linguistics, philosophy, artificial intelligence, and psychology, concerning conditions for the felicitous use of definite reference. Clark and Marshall argue that with definite reference (used in the non-generic, referential fashion), speakers refer to individuals - things in particular knowledge (particular knowledge is a partition of mutual knowledge, a partition containing knowledge about particular individual objects, states, events, and processes). The essence of Clark and Marshall's position regarding definite reference is this: for a speaker to use a definite reference to a thing, he must be confident that because of his speech act the identity of that thing will become mutually known to him and his listener. Thus, a definite reference to an object results in the listener inferring mutual knowledge of the identity of that object. The question then arises as to what mutual knowledge _is_.

Mutual knowledge is knowledge that is shared by speaker and listener, _and which is known to be shared by both of them_. The questions of importance concerning such mutually shared knowledge are the following: (1) What type of shared knowledge is needed for language use? and (2) How is that shared knowledge in practice assessed and secured?

Mutual knowledge is potentially something of a paradox from the point of view of information processing. The reason is simply that in order for persons A and B to mutually know some proposition p, it is necessary that:

(1) A knows that p.

(1') B knows that p.

(2) A knows that B knows that p.

(2') B knows that A knows that p.

(3) A knows that B knows that A knows that p.

(3') B knows that A knows that B knows that p.

et cetera, ad infinitum.

There is thus a potentially infinite regress; in order for A to know that A and B possess mutual knowledge of some proposition p (such as the fact that a Scotland vs. England International soccer match will soon take place), A must know that Scotland will play against England, A must know that B knows that Scotland will play against England, A must know that B knows that A knows that Scotland will play against England, and so on, ad infinitum. Thus, in order for A to <u>use</u> this mutual knowledge in order to felicitously generate a definite reference (for example, to inquire whether B intends to go to "The big match"), A has to check through a potentially infinite number of knowledge (or belief) statements. This checking would require a very large amount of time in order to generate one definite reference. However, interlocutors can generate such references in a matter of seconds.

Clark and Marshall argue that instead of performing such an infinite regress of belief checking, interlocutors use certain heuristics to assess whether they have an evidential basis for <u>assuming</u> mutual knowledge. If there are such grounds for assuming mutual knowledge, they use a <u>mutual knowledge induction schema</u> to infer its existence.

The heuristics used by individuals as a means of collecting evidence from which to infer mutual knowledge are termed <u>copresence heuristics</u>. For example, suppose that two people, A and B, are sitting across a table and that there is a single candle on the table between them. from person A's point of view, there is direct evidence for the truth of (1):

(1) A knows that there is a candle on the table.

A also knows that A and B are looking at each other and the candle simultaneously (this is the <u>simultaneity assumption</u>). A also assumes that B is attending to the candle (the <u>attention assumption</u>). Finally, A assumes that B is rational and will draw the same conclusions as A does (this is the <u>rationality assumption</u>). If B is rational and is attending to the candle, he has evidence for (1'):

(1') B knows that there is a candle on the table.

From the above three assumptions, that B is attending to A and the candle simultaneously, and that B is rational, A can conclude (2):

(2) A knows that B knows that there is a candle on the table.

If B is rational, he will be drawing a parallel inference about A, namely (2'):

(2') B knows that A knows that there is a candle on the table.

Once again, A can conclude (2') from the three assumptions, and can have evidence for (3):

(3) A knows that B knows that A knows that there is a candle on the table.

Likewise, the corresponding statement can be inferred by B, and so on to infinity. The above illustrates the operation of the mutual knowledge induction schema, which Clark and Marshall summarize as folows (Clark and Marshall, 1981, p. 33):

A and B mutually know that p if and only if some state of affairs G ('grounds') holds such that:

1. A and B have reason to believe that G holds.
2. G indicates to A and B that each has reason to believe that G holds.
3. G indicates to A and B that p.

In the candle example above, the grounds G is A and

B's evidence of the triple copresence of the candle, A and B, and their auxiliary assumptions. A and B each have reason to believe that G holds. "These grounds G indicate to each of them that the other has reason to believe that they hold. And the grounds G indicate to both of them that there is a candle on the table. By the induction schema . . . "(A and B) "mutually know that there is a candle on the table" (Clark and Marshall, 1981, p. 33).

Thus, once two people have proper grounds which satisfy the three requirements of the induction schema, they can assume mutual knowledge, rather than computing the infinity of belief statements. This is an economical solution in terms of memory load: A does not represent statements (1), (2), (3), . . . and so on, but instead merely represents directly: A and B mutually know that p.

Mutual knowledge is thus inferred from the induction schema given evidence and assumptions, thereby obviating the processing paradox. If mutual knowledge is to be assumed when the evidence on copresence is weak, the assumptions made in inferring the mutual knowledge must be stronger (for example, B may temporarily not be attending to the candle in the example above; if A then mentions 'this candle', this causes B to look at the candle. This is termed _potential physical copresence_). Mutual knowledge can be of different types, depending upon the type of copresence evidence upon which it is based. There are in fact three types of copresence evidence:

physical copresence, linguistic copresence and community membership.

(i) Community membership

If both A and B are members of a particular community and both mutually recognize that they are members of that community, then both A and B can take for granted the body of knowledge in the community. If both A and B recognize that they are members of the community of educated Americans, for example, then both can take it for granted that they both know the things that the community of educated Americans know, such as the fact that American Independence was declared on July 4th 1776, George Washington was the first president of the U.S.A., and so on. This type of mutual knowledge is preserved over long periods of time: once A and B are aware that they are both members of the community of educated Americans, they are likely to bear that fact in mind for use in reference to all sorts of things. Mutual knowledge of the next three types, however, is ordinarily relevant only for short periods of time. It is based on evidence that is in a sense more direct.

(ii) Physical copresence

The strongest evidence for mutual knowledge is physical copresence, exemplified by the 'candle' example discussed earlier. The fact that the three relevant objects or persons are physically present together plus the physical/perceptual evidence that each of A and B have concerning one another's attention to the candle, is strong enough to render the three auxiliary assumptions relatively trivial. In such conditions, mutual knowledge is secure and definite reference to "this candle" is possible. Mutual knowledge can also be based on <u>past</u> shared experiences, such as both person A and person B having seen a cinema film together at some point in the past.

(iii) <u>Linguistic copresence</u>

Listener and speaker can take as part of their common ground things that have been referred to in previous conversation.

(iv) <u>Mixtures</u>

Very often, mutual knowledge is established by use of evidence based on <u>mixtures</u> of the above three sorts of evidence; for example, physical copresence <u>and</u> community membership.

Clark and Marshall argue that the three forms of definite reference, namely, deixis, anaphora, and proper names, correlate clearly with the different types of mutual knowledge. "Deixis corresponds to physical copresence; anaphora corresponds to linguistic copresence; and proper names corresponds to community membership. The fit could

could hardly be more obvious" (Clark and Marshall, 1981, p.42).

Clark and Marshall argue that the assumptions of many popular models of comprehension which assume that memory contains referential indices for separate entities $E_1 \ldots E_n$, is a mistaken one. Instead, they characterize memory as being more like a diary cross-referenced to an encyclopaedia. The 'diary' component of memory would record instances of copresence; the 'encyclopaedia' component would be necessary for mutual knowledge based on community membership, containing all the generic and particular knowledge which the person believes is universal to each community he belongs to (he could belong to several different communities at the same time, being in the community of Californians, the community of social psychologists, and so on).

This organization of memory underlines Clark and Marshall's thesis that memory must be organized around events (hence the diary component), and these events contain evidence of copresence of one kind or another. This organization enables a person to decide whether a given entity is mutually known both to himself and to his fellow conversant. When a definite reference is made, the listener must search in memory for an event during which himself, the speaker and the referent were copresent, and he can use this event (plus auxiliary assumptions) to inductively infer mutual knowledge of the identity of the referent.

This viewpoint has very interesting social-pragmatic consequences. Clark and Marshall argue that speakers prepare themselves to talk or listen to _a particular other person_ every time they enter into conversation. The identity of our fellow conversant makes a great deal of difference to the knowledge which we access from our memory (the 'encyclopaedia' component of memory). We access the relevant information pertaining to the community that both ourself and our fellow conversant belong.

This in turn implies that we carry around detailed models of other people we know, and when we are in conversation with them, we selectively access only the "common ground" knowledge which we share with that particular person. I may know, for example, that both I and my fellow conversant are members of the community of psychology postgraduate students, and so, when entering into conversation with that particular other person, I would selectively access only the body of knowledge common to that particular community when in conversation with that person.

The social-pragmatic emphasis of this conceptualization of conversation is obvious. It implies that we do not produce language within the constraints of grammar and semantics alone; there are also 'social constraints' in that the identity of our fellow conversant constrains which background knowledge we will access. These 'social constraints' affect quite detailed aspects of our linguistic behaviour, even to the level of something as detailed as our choice of a particular type of referential device. We construct our conversation with specific listeners in mind, and this highly specific construction process is based upon the 'common ground' of background knowledge which we share with the particular listener in question.

This social-pragmatic conceptualization of conversation opens up the possibility that what is understood by what is said will be a function of the particular other with whom we are in conversation: we may understand totally different things by the same sentence depending upon the identity of the speaker. Indeed, Clark and Marshall conclude that mutual knowledge ". . . is at the very heart of the notion of linguistic convention and speaker meaning" (p. 58). This issue of different meanings being understood from the same utterance depending upon the identity of the speaker is explored in more detail by Rommelveit (1974) and will be discussed later in this chapter (section d).

To summarize, Clark and Marshall are arguing that we construct our conversation with particular listeners in mind. Clark and Murphy (1982) have also examined this general phenomenon and refer to it as <u>audience design</u>. The phenomenon of audience design (constructing utterances for a particular audience) has been noted in the case of both speaking (Rommetveit, 1974) and writing (Sanford and Garrod, 1981).

Audience design is based on mutual knowledge, which was discussed above.

Speakers' awareness of what knowledge is shared by themselves and their listeners allows them to construct their utterances in a way which is perfectly understandable to the particular addressee in question, but not to other individuals. For example, if a speaker and listener share the knowledge that a particular woman had just been sitting in a particular chair, the speaker may point at the empty chair and say "That woman is named Veronica" and be confident that the listener will understand who is being referred to. Utterances are not intended to be understood by anyone, but instead are designed to be understood by particular listeners with particular momentary thoughts and beliefs.

Clark and Murphy argue that the listener can use this in figuring out what the speaker meant. For example, if a speaker refers to "the man you met yesterday" this description implicitly instructs the listener to reason as follows: the speaker must have good reason to believe that the listener can readily identify the referent on the basis of their mutual knowledge; furthermore, the listener should believe that the speaker believes this. They may only share knowledge of the listener meeting one person the day before (although the listener may have met many people the day before), and so, the tacit reasoning goes, the speaker must be referring to that one person whom they both know the listener met the day before. Effectively, the listener reasons "If she thinks I can readily identify her referent uniquely, she must be referring to the only person that we mutually know that I met yesterday, namely Jacques". This use of the audience design of the utterance by the listener in a tacit reasoning process is termed the design assumption by Clark and Murphy. Thus, audience design works successfully, according to Clark and Murphy, because listeners tacitly reason from the assumption that the utterance was produced for them, reasoning back to the mutual knowledge basis for the utterance and thereby resolving the definite reference.

The effectiveness of the 'design assumption' reasoning by the listener in the resolution of definite references was demonstrated in an experiment by Clark, Schreuder and Buttrick (in preparation). Buttrick approached students on the campus of Stanford University and showed them a photograph of president Ronald Reagan standing next to his director of the budget, David Stockman (an independent survey had shown that people assumed Reagan was known to everyone but Stockman was not). Buttrick asked each student one of the two following questions:

(1) You know who this man is, don't you?

(2) Do you have any idea at all who this man is?

The same definite description, <u>this man</u>, was used in both utterances. However, students replied "Yes, that's Reagan" to question (1) and "Yes, I believe that's Stockman" to question (2). Question (2) presupposes that the speaker doubts that the listener knows who is being referred to, whereas question (1) presupposes that the listener knows who is being referred to. Clark and Murphy conclude that the respondents were using the presuppositions implicit in the questions, and the public knowledge that Reagan is better known, to enable them to use the design assumption in order to establish who the referent was.

In addition to being of importance for definite reference, audience design is also of importance in anaphoric reference and word meaning, according to Clark and Murphy. Novel coinages, such as the following one, illustrate the importance of audience design:

My sister managed <u>to Houdini</u> her way into the office this morning.

Clark and Murphy argue that such innovations have an indefinite number of possible senses. Clark and Clark (1979) called expressions whose senses are not enumerable <u>contextual expressions</u>. Clark and Murphy argue that such contextual expressions are ubiquitous and are interpretable only because of their audience design. A speaker would only use an expression such as <u>to Houdini</u> if he were confident that the relevant background knowledge (in this case the knowledge of Houdini's capability of going through locked doors) were part of the mutual knowledge of himself and his listener. Clark and Murphy argue that contextual expressions are an ordinary part of language and that they are usually understood quickly and without special effort or disruption.

Clark and Murphy conclude that listeners can never know when they hear a word whether it is being used in one of its well established senses, if it has one, or in an innovative sense, and that ultimately listeners must rely on the design assumption in deciding what the speaker meant. Thus, audience design and reasoning from the design assumption is used both for contextual expressions and for the selection of particular senses of words which are not used in this innovative fashion.

Whilst Clark and Murphy's examples and arguments highlight the importance of the phenomenon of speakers designing their utterances for particular listeners, their explanation of the phenomenon (in terms of a reasoning process, outlined above) intuitively seems implausible. It seems intuitively unlikely that such explicit (or tacit) reasoning processes would take place every time a listener heard an utterance designed specifically for him.

It would intuitively seem that 'audience design' would work via some other mechanism, rather than via the somewhat contrived reasoning process advocated by Clark and Murphy.

Another theorist who emphasizes the importance of 'audience design' as a ubiquitous phenomenon in everyday language use is Rommetveit (1974, 1979); Rommetveit, however, explains the phenomenon in rather different terms, and it is to an exposition of Rommetveit's views that we now turn.

d) The 'pragmatic counterrevolution': the theories of Rommetveit (1974, 1979, 1984)

The title of this section of the chapter is based upon the notion (Greene, 1972; see chapter 1) that the work of Noam Chomsky constituted something of a revolution in thinking about language. Rommetveit, on the other hand, emphasizes pragmatics and his whole approach (see, for example, his 1974 book) is set up in explicit opposition to Chomsky's. The phrase 'pragmatic counterrevolution' is Rommetveit's (Rommetveit, 1984), and neatly summarizes the changes in theorizing about language that have taken place in the last thirty or so years; first, a syntactic revolution and an awareness of complexities in language which had not hitherto been adequately appreciated; subsequently, the theories reported in this chapter explicitly challenge the initial 'syntactic revolution' by pointing to significant pragmatic aspects of language use which the transformational grammar-based approach fails to capture. The pragmatic counterrevolution is not, however, yet fully under way, but it is fair to say that an increasingly great amount of attention is currently being paid to the influence of pragmatic factors.

Rommetveit's theoretical position can be adumbrated as follows: what is missing from the psycholinguistic approach based on transformational grammar is actual _use_ of the language to convey messages. We cannot, Rommetveit argues, analyze language in terms of propositional analyses of single sentences _in vacuo_: this is artificial, since in practice, we make utterances _for a purpose_ and the choice of utterance is affected by the pragmatics of the situation, that is, by whom we are talking to. We therefore should be analyzing _what is made known_ by an utterance _in a given context of use_, rather than attempting to assess the 'literal meaning' of an utterance in terms of its propositional content. In order to assess what is made known, we need to examine the semantic _potentialities_ of the utterance, rather than its 'literal meaning', and we have to assess these meaning potentialities in relation to a social context of actual use. This necessitates us having a way of characterizing the social context.

When actually using language, we choose our phraseology to suit our audience, as was illustrated by the work of Clark and Murphy. Conversely, what our audience understands us to mean by what we say is a function of the fact that it is <u>we</u> who speak (as opposed to someone else). When the speaker/listener roles are exchanged, it is our turn to listen on the premises of the speaker, and he must then speak on our premises. Communication is thus based on a reciprocally endorsed and spontaneously fulfilled contract of complementarity - encoding involves anticipatory decoding (speech is listener oriented), and decoding is speaker oriented. In other words, speech is constructed with a specific listener in mind, whilst the listener must try to reconstruct what the speaker intends to make known.

This is the essence of the contractual nature of language use: we tacitly endorse a contract with the other participant which specifies what we are to understand by what is said, that is, a contract concerning shared strategies of categorization: we agree to categorize things in the same way as the other participant. Speech is therefore listener oriented to the extent that the speaker monitors his speech in accordance with his assumptions about the extent of social world and strategies of categorization which are shared by him and his listener. This raises another vital element in speaking and listening: the taking of the perspective or role of the other participant. This notion relates to the philosopher G. H. Mead's (1934) 'symbolic interactionism'. In order for a speaker to indulge in anticipatory decoding, he must take the role of the listener (i.e. decentrate); conversely, attempting to assess what a speaker intended to make known involves the listener adopting the speaker's prespective or role.

These two notions of intersubjectively sharing a strategy of categorization, and the taking of the role of the other participant in speaking and listening, are the cornerstones of Rommetveit's theoretical edifice.

What we understand by a given utterance therefore depends upon who is actually speaking, and the range of semantic potentialities inherent in any utterance is a function of the range of different 'others' whose roles we can take. For example, consider the utterance:

I too was invited
I went to the ball . . .
and it rolled
and rolled away.

If we are under the impression that the utterance originated from a schizophrenic patient, we would most likely consider the utterance to be incoherent. If, however, we are under the impression that the utterance is that of a poet, we are not bewildered by it quite so much and perhaps we may extract, either intuitively and emotionally, or even in a verbally expressible fashion, a feeling that has been conveyed by the poet. The point of the example is that radically different interpretations ensue depending upon who we think the speaker is. The complementarity of speaking and listening is based on a tacit contract; Rommetveit's whole approach is based upon the question of which social contracts are tacitly and reciprocally endorsed by the participants in a conversation, and on how we take the role or perspective of different others with whom we speak. This decentering and taking the role of the other is such a pervasive feature of normal social interaction that we fail to notice it at all, according to Rommetveit.

The pragmatic emphasis of this approach ties in very well with the work of Clark and Murphy discussed above. We will firstly consider Rommetveit's views on why the Chomskyan approach is inadequate as a psychological model of communication, then we will turn to Rommetveit's characterization of the social context of language use.

i) The case against transformational grammar-based analyses of communication

In Rommetveit's (1974) book, he begins by explicitly challenging the Harvard-M.I.T. (Chomskyan) approach to language understanding as inadequate and argues in favour of replacing the Chomskyan approach with a rather more social-psychological approach. Some of Rommetveit's criticisms resemble points raised by Clark and Murphy

(1982).

Initially, the Chomskyan approach advocated the exploration of language essence rather than its use (Chomsky, 1968, p. 62). Later, however, this standpoint was restated as a cautious working hypothesis - by 1972, the Harvard-M.I.T. school had turned its attention to such problems as presupposition. Rommetveit is also interested in what is presupposed by a speaker and listener, but his approach to the problem is rather different from that of the Chomskyan psycholinguists.

The work which stemmed from Chomsky's theory concentrated upon the study of isolated sentences _in vacuo_; in particular, speakers of the language were asked to use their intuitions to judge the grammaticality of word strings presented to them. This emphasis on the sentence resulted in some theorists taking the sentence to be the most natural unit of analysis, especially since larger segments of discourse are less clearly definable (Lyons, 1968, p. 172). This view has been challenged on two counts; firstly, as Sanford and Garrod (1981) point out, exclusively focussing at the level of the sentence results in a lack of appreciation of interesting and important phenomena (such as anaphoric reference) which operate at a suprasentential level. The second challenge to this methodological convenience is an argument advanced by Rommetveit, which is based on the notion that the study of the isolated sentence _in vacuo_ divorces the sentence from actual _use_ and this renders the task artificial.

Rommetveit points out that in such studies involving intuitive judgements about the grammaticality or otherwise or isolated sentences, the instructions are ambiguous and, as a result, subjects instruct themselves. What is being judged therefore varies, depending upon what the informant thinks the researcher wants him to do, and depending upon the particular judgemental context of the sentence. Thus, for example, the sentence "My spinster aunt is an infant" may be judged nonsensical (i.e. illogical) in the context of having judged 'parents are older than their children', or it can be judged as meaningful (i.e. 'reasonable, can certainly be said') in the immediate judgemental context of 'sleepless green ideas sleep furiously'. The judgements, Rommetveit argues, are related to actual and potential use, and lack of specification of different domains of use. A judgement of 'nonsensical' implies that the judged sequence of words cannot be used, and the discovery of ambiguities implies envisaging distinctively different ways in which a particular meaningful sequence can be used. Thus, under conditions of ambiguous instruction, subjects spontaneously resort to potential use as a criterion of acceptability for the experimental materials.

Rommetveit (1974) likens this sort of work to that of Ebbinghaus on the 'nonsense syllable': Ebbinghaus's 'nonsense' ('meaningless') trigrams were in fact made meaningful by subjects relating them to actual words. Likewise, subjects in studies of intuitive judgements of grammaticality of word strings spontaneously resort to potential uses and plausible contextual frames. Rommetveit (1974, p. 14) describes both the nonsense of the 'nonsense syllable' and the sentence detached from a context of use as <u>experimenter-centred illusions</u>.

A further practice of the Harvard-M.I.T. school of which Rommetveit is critical is the assigning of propositional content to sentences <u>in vacuo</u>. Deep sentence structure and propositional content, he argues, represent abstractions in which the language has been deprived of its connection with the pre-established social reality in which communication takes place. Such abstractions thus tell us nothing about the social context in which discourse is normally embedded and, furthermore, do not therefore succeed in explicating the variety of potential messages which can be made known by the utterance in different social contexts. Instead of considering utterances as having a particular propositional content, we should, he argues, consider them as having a set of semantic potentialities. That is to say, a given sentence could make a variety of assertions depending upon the social context in which it is embedded. For example, in a situation where two tenants of an apartment block are engaged in a conversation about the general property of the neighbourhood, in which all flats have been occupied by young couples, and where an elderly man has recently moved into a vacant flat, one of the conversants may say:

"The old man is poor".

The word "poor" has a variety of meanings. In this context, it could mean reliance upon public financial support; in a dialogue about the third world, it could mean living below the subsistence level - notice that <u>both</u> of these meanings are the opposite of 'wealthy' - hence, simple semantic oppositions do not do justice to the meaning potentialities of such a lexical item in actual use. In a different context, 'poor' can convey a sense of pity in the speaker - as in 'he's a poor old soul'; in yet another context, 'The old man is poor' could refer to a dearth of artistic talent in an elderly artist who is under discussion - etc.

Furthermore, even within the "new tenant" context, different assertions could be made by the sentence: this is easily seen by considering intonational variations of the pronunciation of the sentence:

e.g. "The old <u>man</u> is poor" - i.e. it is a man rather than a woman who is poor.

"The old man <u>is</u> poor" - i.e. the old man is currently poor, regardless of his previous state of wealth or poverty.

The Harvard-M.I.T. approach to such a sentence would be to assign propositional content to the sentence by breaking it down, as follows:

a) The man is old.

b) The man is poor.

Proposition b) is treated as the "focus" (i.e. the phrase which should contain the intonation centre and which is therefore what is being asserted by the sentence) and proposition a) is the 'presupposition' on to which the 'focus' proposition is nested. This distinction is a formula for assessing what is novel information as opposed to what is assumed to be the case.

However, the above example clearly demonstrates that this propositional analysis fails to capture the variety of potential messages which could be conveyed by the sentence even within one given context. Such propositional analyses, which Rommetveit describes as 'straightjackets', fail to do justice to the inherent flexibility of natural language, a flexibility which the poet turns to his advantage. The postulation of unequivocal literal meanings is a futile exercise, Rommetveit argues.

A striking example of how propositional analyses fail to do full justice to an expression's potentialities for conveying meaning is a line from a poem by Dylan Thomas, quoted in Rommetveit (1974):

I bought sincerity a grief ago.

This would be regarded as semantically anomalous if analyzed in terms of propositional content, yet it is beautifully meaningful in a poetic context and (quite strikingly) succeeds in conveying a message. Rommetveit argues that the artist's achievement is often evaluated in terms of his capacity to create novel social realities, that is, to make his reader experience aspects of life which are somehow blurred or even concealed by conventional language use. What is conveyed by the line of poetry can only be fully assessed in the light of the setting in which it is embedded, that is, the partly institutionalized poet-to-reader-of-poetry setting, as opposed to other settings in which it could appear. To quote Rommetveit, "In order to find out what is made known when something is read or heard we therefore have to inquire into what kind of contract has been established by the two participants in that particular act of communication. This implies, in turn, an exploration of which aspects of a multifaceted and pluralistic world constitute the temporarily shared social reality" (Rommetveit, 1974, p.25).

Rommetveit's case for a pragmatic approach to meaning relies on the above two points; firstly, the consideration of message potential as opposed to 'literal meaning', and secondly, the relation of what is said to whatever constitutes the temporarily shared social world of the speaker and listener, i.e. the shared 'here-and-now.'

Communication aims at the transcendence of the 'private' world of the participants, setting up what Rommetveit (1974, 1979) refers to as 'states of intersubjectivity'. One of the important elements of the state of intersubjectivity is a degree of overlap in the strategies of categorization of the two participants. Rommetveit's "message structure" approach essentially relates the structure of the message to the architecture of the particular intersubjectivity onto which it is nested. Earlier, we noted that one element of what is intersubjectively shared is a tacit 'contract' to select among the options of perspectives and categorization. Another possibility is that of the two participants tacitly endorsing a <u>meta-contract</u>, that is, a contract concerning how to deal with the contracts of categorization implied by what is said. For example, in reading a novel, we endorse a meta-contract to deal with the abstract novelties of the first few pages as if they were already familiar, on the assumption that what is initially abstract, unknown and undetermined will become familiar as we continue reading.

The main thrust of Rommetveit's argument against the Harvard-M.I.T. school is that in a consideration of the possible meaning of what has been said, we have to relate the utterance to the particular shared social world of the participants in the conversation. By 'shared social world' he means tacit and reciprocally held presuppositions, contracts concerning shared strategies of categorization and attribution, possibly also meta-contracts (as in the poet-to-reader-of-poetry meta-contract), the identification of speaker by listener and vice versa; these are all elements of the shared social world at the moment of speech. It is to Rommetveit's characterization of the intersubjectively established social reality at the moment of speech to which we now turn.

ii) <u>'On the architecture of intersubjectivity'</u>

Of the theoretical influences which Rommetveit explicitly acknowledges in the introduction of his 1974 book, three are of particular interest in the present context, namely: i) George Herbert Mead's 'symbolic interactionism', ii) Jean Piaget's theory of cognitive growth and decentration, and iii) Fritz Heider's work on attribution. This blend of theoretical concepts results in a pragmatic approach to meaning which is of great interest.

As noted above, among the elements of the shared social world are identification of speaker by listener and vice versa, and a host of other factors. Rommetveit considers that, in order to explore all the significant features of the temporarily shared social world of the dialogue, we need at a minimum a system of three coordinates of the act of speech, namely:

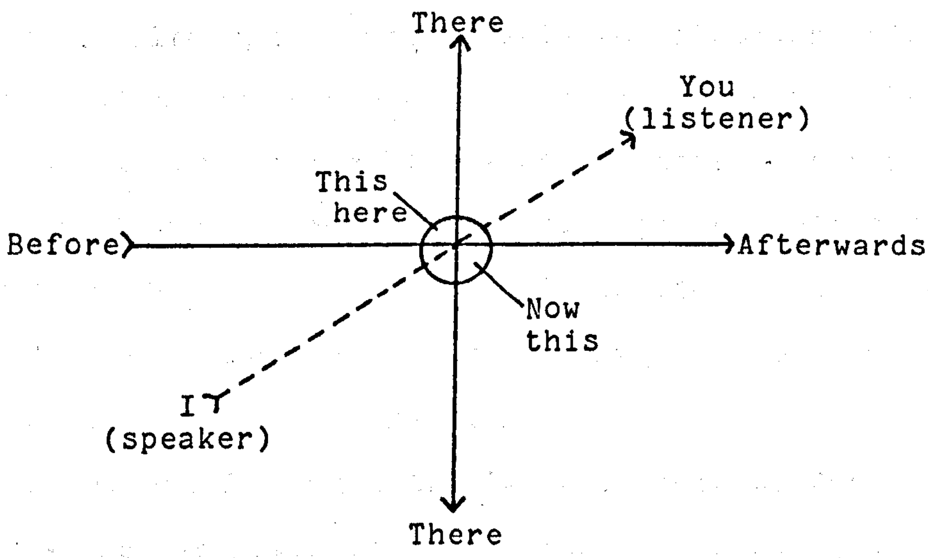


Figure 3.1
The coordinates of the act of speech
(Rommetveit, 1974, p. 36).

The coordinates may be defined in terms of the point in real time at which the act of communication takes place, its location, and (in the case of spoken language) the identification of the listener by the speaker and vice versa.

These coordinates cannot be assessed independently of each other and from the outside. The denotative extension of the intersubjectively established 'now', for example, can be this century in historical discourse, or this very moment in discourse about an ongoing football game. Likewise, 'here' can denote Europe as opposed to Africa in a discussion about geography, or it can denote this room as opposed to the room next door in other discourse. Although we may be able to establish using public indices who is speaking to whom, where and when, we still cannot characterize these coordinates of the act of speech and the relevant intersubjectively established 'values' on the three dimensions in public terms. The 'values' on the coordinate axes can only be known by the participants in the conversation.

Rommetveit describes the I-you coordinate for any utterance as given as a speaking I addressing a listening you, that is, in terms of an unequivocal direction of communication. Two features of the interpersonal coordinate may account for initial commonality and shared presuppositions between speaker and listener, namely:

i) Firstly, some commonality is established by the very fact that the two persons engage in a dialogue - they constitute a temporary established <u>we</u> engaged in an act of communication, as opposed to all others who are not so engaged.

ii) Secondly, there is a complementarity between speaking and listening. This complementarity can be thought of as a reciprocally cognized asymmetry between speaker and listener with respect to what is already known. What is said under optimal conditions of complementarity is monitored in accordance with veridical assumptions concerning what the listener at that moment does not yet know, but wants to know.

The temporarily shared social world at any particular stage of a dialogue can only be assessed in terms of the participants' identification of one another and their mutual knowledge. The identification of speaker by listener and vice versa, their mutual knowledge and their mutual beliefs are central to the tacit presuppositions underlying the dialogue.

An example of nearly perfect complementarity and synchronization of intentions and thoughts is a situation where a middle-aged couple are rather worried about the behaviour of their teenage son. In such a case, very cryptic communicaton may take place, yet be perfectly understood by both participants. The woman may give her husband an interrogating look on seeing his expression of tension and receive the reply 'Pot'. And the wife understands perfectly; what is worrying her husband is the possibility that their son may start smoking pot.

Such cryptic conversations are not unusual. On such occasions, whatever is understood can only be assessed against the background of the intersubjectively established social reality at the moment of speech. In the above example, the uttering of the word 'Pot' by the husband makes known to his wife exactly what she wants to be informed about at that moment - no more, no less. He uses such elliptical expressions because of what _he knows she knows_, namely, that at that particular moment he is anxious about some aspect of their son's conduct, and he identifies in one word precisely which particular aspect it is. The ellipsis is based on a social reality in which much knowledge is mutually shared.

This leads to Rommetveit's interesting approach to ellipsis: he states that ellipsis is "the prototype of verbal communication under ideal conditions of complete complementarity in an intersubjectively established, temporarily shared social world" (Rommetveit, 1974, p. 29). Ellipsis is therefore used under ideal conditions, and full sentences or even larger sequences may be required to make something known under conditions of deficient complementarity and less than perfect sychronization of intentions and thoughts.

The above example relates to states of nearly perfect intersubjectivity, and highly elliptical communication takes place; states of _partial_ intersubjectivity can be established, and yet the intersubjectivity need not be so deficient that nothing can be made known. Indeed, states of partial intersubjectivity are the _norm_, according to Rommetveit: For example, consider a situation where two football fans are watching a game. One is experienced in the rules of the game and has watched football for some time, whilst the other fan is a newcomer to the game. Following a tackle by one of the players, the more experienced fan says: "That was magnificent!". The 'novice' fan may have taken the remark to pertain to only one aspect of the tackle, perhaps the speed and skill with which the final manoeuvre was carried out. However, the more experienced fan was actually praising the preceding tactics _as well as_ the final manoeuvre.

The example illustrates several points of note. Firstly, it demonstrates the point that the social world _is_ multifaceted and strategies of categorization can vary from individual to individual. Secondly, it underlines the openness of language with respect to experientially established social realities. Thirdly, Rommetveit's point (1974, p. 39, emphasis in original) that "_what can be established only as a partially shared social reality will remain only partially shared whenever that social reality is made the topic of a dialogue_" is also brought into focus. Lastly, the example again illustrates the point that what is made known by "That was magnificent!" can only be assessed by examining the nature of the partial intersubjectivity onto which it is nested.

In nesting a message onto a partially shared social reality, that partially shared social reality is itself expanded and/or modified. Further messages may be nested on to the newly expanded social reality, thereby expanding it still further; and so on. Intersubjectivity can also be expanded by means of the participants endorsing a contract concerning shared strategies of categorization, as alluded to earlier. Thus, an initial commonality between speaker and listener by virtue of being participants in an experientially shared here-and-now and participation in a dialogue involving complementarity of roles, and so on, can subsequently be expanded by means of such contracts. These are, according to Rommetveit, some of the outstanding features of natural language in action.

Rommetveit also uses his 'coordinates of the act of speech' as the basis for characterizing the way in which deictic expressions are used. These are the means by which experientially shared objects and events may enter speech. The linguistic devices for introducing such experientially shared objects and events into speech are expressions like this/that, here/there, before/now/afterwards, I/you/they, and suffixes of verbs indicating tense. The origo or point of reference for these expressions is characterized by the temporarily established here-and-now as depicted in the diagram of coordinates of the act of speech (at the intersection of the axes; see Figure 3.1, p.165).

Like Klein (1982), Rommetveit considers that what is not 'here' is 'there', so to speak; in intersubjectivity, it is the 'this, here', that is delimited, and by default everything else is 'that, there'. Thus, whatever is the focus for convergent attention or intention of the two participants may enter their dialogue at that particular time as a 'this, here', whereas what was attended to and/or talked about some time ago, and also what is cognized as being located at a distance by both participants may enter speech as an intersubjectively established that, there. Rommetveit's coordinates of the act of speech are thus one way of characterizing the 'origo' which Buhler also discussed (see section b) of this chapter).

Deictic devices are therefore linguistic ways of 'pointing out' experientially shared social realities. Deictic elements do this by tagging what is experientially common to both participants on to the spatial-temporal coordinates of the act of speech. As Klein (1982) and Wunderlich and Rienelt (1982) have also pointed out, Rommetveit notes that the proper use of deictic expressions habitually involves the hearer in decentration from his own perspective and taking the perspective of the other person. Piagetian egocentrism, which Rommetveit describes as the initial "imprisonment in the immediate <u>here-and-now</u>" (1974, p. 42) is a collapse of the I-you coordinate of the act of speech and an incapacity to decentre and take the role of the other participant in the conversation. Young children display such egocentrism, and decentration is " . . . . a very general and composite developmental process by which an individual acquires the capacity to adopt a variety of interrelated perspectives and strategies of categorization in response to successively more highly differentiated multifaceted objects, events and states of affairs" (Rommetveit, 1974, p. 42). Adult competence in the use of deictic devices (and also successful participation in dialogue in general) is therefore contingent upon a capacity to decentre and adopt the perspective of a variety of other individuals. States of intersubjectivity can thus be explored in terms of potentially shared perspectives and the capacity of the participants to adopt the point of view of one another. The taking of the role of the other in ordinary speech is a persistent theme in Rommetveit's work and owes much to the 'symbolic interactionism' of George Herbert Mead (1934).

This taking the role of different others is the basis of adult semantic competence, according to Rommetveit. Adult semantic competence must be conceived of as a reportoire of interrelated, linguistically mediated modes of cognition and cognitive-emotive perspectives (different ways of 'seeing', that is, categorizing and cognizing, the world), rather than as a set of rules for literal language use. This view is in rather stark contrast to the 'syntactic' theories of meaning outlined in chapter 1 of this thesis.

Before leaving Rommetveit's work, it is worth noting a notion which he briefly introduces in his 1974 book. It is the notion of <u>control</u>, that is, that one of the participants in a dialogue has, in some sense, control (the 'upper hand', or greater influence upon what is talked about and the manner in which it is talked about) of the dialogue. Rommetveit considers control to be in the hands of whoever is currently speaking. And, to some extent, this must be the case: the speaker can select and point out the objects, events and states of affairs which are to enter the field of intersubjectively shared attention. Which of all possible entities of an experientially shared situation will be introduced is in principle determined by the speaker. The listener, according to Rommetveit, has to accept and engage in whatever social reality is being introduced into the formal framework of intersubjectivity whenever the speaker has introduced a topic. This notion of control will be mentioned again in the next section of this chapter, on the views of the philosopher David Lewis. We might note in passing, though, that while the above statement of Rommetveit has a certain amount of truth at one level, at another level it is possible that there is an overall degree of control <u>throughout the dialogue</u> by one of the speakers. For example, consider the case where a speaker is currently replying to a question asked by the other participant. Which of the participants is then in control, the speaker or the questioner? The as yet vaguely formulated notion of control in dialogue may prove to be an issue of great importance for the social psychology of language.

c) The pragmatics of control of a dialogue

David Lewis (1978) raises some interesting social-pragmatic issues in his paper on 'scorekeeping in a language game'. From the point of view of an interest in social-pragmatic factors' influence on various facets of dialogue, two of his notions are of interest. The first is the notion that one interlocutor is in overall control of the dialogue (the master in Lewis's terms) and the other interlocutor, the slave, is under the control of the master. The reason for this relationship arising can be one of a variety of reasons - for example, coercion, deference or common purpose. Control of the slave by the master is exercised verbally.

The other notion of interest is that there exists a conversational score which is analogous to a baseball score in that it has many components. A baseball score is a septuple of numbers denoting the number of runs by the two teams, which half of the game it is, which innings it currently is, and so on. Likewise, Lewis argues, there is a conversational score which is a multiple-component score, and the components of conversational score are abstract set-theoretic entities, such as sets of presupposed propositions, boundaries between permissible and impermissible courses of action for the slave, and the like. As with baseball, what conversation is correct or acceptable at any given time depends on the components of conversational score at the stage of conversation when it is uttered (for example, the acceptability of what is said will depend upon the presuppositions presupposed at the time of utterance. As noted above, presuppositions are a component of conversational score). Like a baseball score, a conversational score evolves in a (more-or-less) rule-governed way.

Lewis lists several possible components of conversational score. These include presuppositions, permissibility (which courses of action are permissible for the slave and which are not), standards of precision (e.g. of descriptions), and the point of reference in a narrative. All such components obey _rules of accommodation_. This is the major difference between baseball and language game scores: conversational score tends to evolve in such a way as is required in order to make whatever occurs count as correct play. There is no such rule of accommodation in baseball.

For example, consider the question of what play (conversation) is permissible for the slave. As noted above, the master exercises control over the slave verbally. The master may verbally shift the boundary of permissibility, by saying to the slave that certain courses of action are permissible (or impermissible). If what the master says would be false were the boundary of permissibility to remain stationary, then immediately the boundary moves appropriately. The range of permissible slave behaviour expands or contracts so that what the master says is true after all. All components of conversational score are governed by such rules of accommodation.

Thus, the master exerts control over the slave by changing the components of conversational score using the rule of accommodation: in saying whatever he says, he automatically shifts the boundary of (for example) permissibility, thus altering the range of permissible utterances of the slave.

Another example would be where the master changed the standards of precision of descriptions, e.g. by saying "Italy is boot-shaped". If this were unacceptable given the current standard of precision in the conversational score, then the standard of precision in the score would change in order to make the utterance by the master acceptable. Thereafter, the slave would be free to say something of an equivalent standard of precision, such as "France is hexagonal".

The master's control over the slave is exercised via the conversational score.

These intriguing notions would seem to be of some importance for the social psychology of language; it would be of interest to the social psychologist to know whether roles, social status, and mood states would have any influence upon who would be the master and who would be the slave in the case of any given pair of dialogue participants. As yet, however, Lewis's master-slave notion, although intuitively appealing, has not yet been subjected to empirical test.

f) <u>Overview</u>

In this chapter, a variety of pragmatic issues of importance were considered, and it is argued that these pragmatic issues should be recognized as being of great importance to a psychological theory of meaning. Such issues as the pragmatics of deictic reference, mutual knowledge and audience design were reviewed and it is concluded that this variety of phenomena points to the importance of social and pragmatic factors in utterance meaning. In all of these phenomena, the identities of the interlocutors, and consequently the mutual knowledge they share, affect the speaker's choice of utterance.

It is perhaps fair to conclude that earlier theories of meaning (particularly those based on purely linguistic, formal aspects of language and which focussed primarily on sense relations) either acknowledged the importance of but did not pay any real attention to, or explicitly rejected as being as-yet-too-complicated-to-study, the pragmatic phenomena considered in this chapter. It is also fair to conclude that any complete theory of meaning which attempts to account for the meaning of natural, spontaneously-produced utterances in context, will have to be able to account for these phenomena.

Some notions considered in this chapter recur in the work of different theorists and one such notion is that of perspective. The notion of perspective is used in the sense of a point in space/time from wich events are viewed, and also in the more subtle sense of 'orientation' or 'approach' towards what is being discussed. Perspective, in the sense of visual perspective, is central to an understanding of deictic phenomena. In using deictic expressions, a speaker and a listener set up a common frame of reference in relation to which the denotational extensions of deictic expressions are agreed. This involves the listener decentering from his own perspective and adopting that of the speaker. Rommetveit stresses the importance of decentering and taking the perspective of the other in the course of non-deictic conversation also.

Perspective can be mentally shifted; listeners adopt the perspectives of speakers during the course of ordinary comprehension, and both listeners and speakers shift the location of their point of reference ('origo') during the course of the 'imaginary walk' embarked upon when giving and receiving route directions.

The notion of listeners adopting the perspective of speakers (and vice versa) has also been stressed in another but related sphere, by Farr and Anderson (1983). Farr and Anderson argue that Jones and Nisbett's (1972) distinction between the attributional tendencies of actors and those of observers was couched in a fashion which failed to take heed of the potentialities of individuals to transcend their own perspective during conversation. Jones and Nisbett had proposed that actors (acting individuals) explain their own behaviour in terms of the characteristics of the situation in which they act (that is, they make situational causal attributions), whilst observers of those same behaving actors explain the actors' behaviour in terms of dispositions of those actors (i.e., they make dispositional causal attributions). Farr and Anderson noted how this hypothesis was very much couched in the language of visual perception, and argued that the divergence in perspective between actor and observer was a divergence in visual perspective. They further argued, following Rommetveit, that such a divergence in visual perspective is ordinarily overcome by means of conversation (which habitually involves us in adopting the perspectives of other people). Thus, the notion of perspective, and of taking the perspective of another person in the course of conversation has implications for attribution theory, and may have other ramifications (see Farr and Anderson, 1983). Attribution theorists have, in general, failed to take into account the importance of language. The pragmatics of natural dialogue may be of some importance to the area of social psychology in general.

The notion that utterances are designed to be understood by particular listeners and that listening in turn relies upon this fact, emerges as also being of some importance. Clark and Murphy (1982) and Rommetveit (1974, 1979) emphasize how approaches to language understanding based around formal linguistic theories are inadequate in this very important respect: what is understood by a given utterance depends upon who is speaking it, and to whom. Thus, standardized analyses in terms of propositional content and 'literal meaning' are inadequate to characterize meaning. In order to assess what is meant by what is said, we must relate what is said to the social situation in which it is said. Thus, the identities of speaker and listener as conceived of by one another, and a characterization of their mutual knowledge, shared beliefs and shared presuppositions, - in short, their 'intersubjectivity' or 'common ground' - are necessarily going to have to be taken into account if we are to adequately assess the meaning that is conveyed by what is said. The inherently social nature of the dialogue is thus emphasized by theorists interested in pragmtics.

Both Rommetveit (1974) and Clark and Murphy (1982) emphasize the importance of this phenomenon of 'audience design', although their explanations of the phenomenon are rather different, Clark and Murphy explaining it in terms of tacit reasoning from the 'design assumption', and

Rommetveit explaining it in terms of contracts of categorization, the complementarity of speaking and listening, and taking the perspective of the other.

Rommetveit's views are a challenge for traditional psycholinguistics; his social-pragmatic emphasis stands in stark contrast to more traditional work. Indeed, Markova (1981) argues that they in fact stem from differing philosphical traditions, with Rommetveit's work being from a Hegelian philosphical tradition, and the transformationalists' views being based on Cartesian presuppositions. Although some theorists described as 'Cartesian' in philosophical orientation by Markova might resist this categorization of themselves, Markova's point does serve to emphasize how radically different the two theoretical approaches are.

The last notion reviewed was Lewis's ideas concerning control of dialogues by one of the participants, and, although very interesting, it has to be concluded that it is an as yet empirically unsupported concept.

The overall conclusion to be stressed is that pragmatic phenomena are of some importance in actual language usage, and pragmatics should therefore be a proper part of any theory of meaning which attempts to account for the meanings of utterances in actual use.

It was concluded at the end of Chapt er 2 that an approach based on mental models of the reference situation is a viable and parsimonious way of characterizing language understanding with respect to the phenomena of semantics. In this chapter, the views of several different theorists were considered, and a general consensus of the importance of pragmatic phenomena, particularly 'audience design', was noted. It could therefore be argued that there is a good case for supplementing a theory based on mental models of referents with machinery of some sort for dealing with pragmatic phenomena.

For example, it is possible that a speaker may nest within his model of the reference situation another model which represents to some extent the knowledge and beliefs of the listener, and use this model-within-a-model to constrain the speech generation process. It is also possible that the model is itself a negotiated product (embodying shared strategies of categorization) and therefore automatically, as it were, builds in "listener constraints". These are questions of some importance.

It was noted that the work of Clark and Murphy (1982) on "audience design" is supported by empirical study. Rommetveit's views, challenging as they are for traditional theories of meaning, are as yet not fully tested empirically. Still less so are the views of Lewis. In order to test such notions empirically, a rather different type of experimental task is required. Instead of assessing how individuals understand carefully prepared experimental materials, what is required is a means of eliciting suitably controlled yet genuinely spontaneous dialogue which is in use for a purpose.

Part II of this thesis reports experimental results obtained using such a speech production task, and attempts to relate these results to the views of some of the theorists whose work has been reviewed in chapter 1 - 3. The next chapter will give an exposition of the task itself and its historical antecedents as a prelude to the work reported in Part II.

CHAPTER 4

THE EXPERIMENTAL PROCEDURE: THE DIALOGUE MAZE GAME

a) Introduction

In the previous three chapters, an indication was given of the variety of approaches there have been to the study of the meanings of natural language expressions. In chapters 1 and 2, approaches to the study of meaning which had their origins in formal linguistic and philosophical approaches were examined, as was also a theory based on psychological considerations (that is, the theory of mental models). In chapter 3, the importance of pragmatic phenomena, and the necessity of studying language under conditions of genuine communicative use was emphasized. In particular, it was concluded that, in order to investigate the pragmatics of meaning, it is necessary to examine speech production and understanding using an experimental task in which speech is produced in a controlled yet natural way. This methodology of investigating speech production in a constrained, experimentally-controlled way, constitutes a rather different methodology from that of traditional experimental approaches investigating language understanding, in which, typically, the subject's task is to respond in some way to isolated sentences or small groups of sentences, these materials having been carefully constructed beforehand by the experimenter.

In this chapter, an exposition will be given of the speech production task used in the work reported here. The task takes the form of a computer-controlled game played by two individuals, and the nature of the game is such that it can only be solved by means of the two individuals cooperating verbally.

The essence of the game task is as follows. The two individuals who are playing the game are located in separate sound-proofed booths, with verbal communication between them being achieved by means of a headset and microphone, this being their only contact with one another. Each player is seated at a computer terminal, and each terminal is linked to a main computer which is under the control of the experimenter. Visible to each subject on his or her screen is a version of a _maze_ presented in plan view (see figures 4-3 and 4-4 for an example of the initial maze configurations presented to two subjects playing a game). The maze consists of a configuration of _nodes_ (small box-like structures placed at equidistant locations, in which players' position markers can be located between moves), and these nodes are connected by _paths_ (which are links between nodes along which players can move: see figures 4-3 and 4-4). The subjects' task is that they must both move their respective position markers through the maze from start positions to goal positions. The locations of start and goal positions are different for the two players; furthermore, in the studies reported here, each player sees only his or her own start position (marked by an X) and goal position (marked by an asterisk), and does not see his or her partner's location or goal position. Players move alternately through the maze.

However, each player's maze also contains obstacles in the form of gates (blockages which cannot be traversed) across some of the pathways, and it is in the overcoming of these obstacles by subjects that verbal cooperation is required. The mechanism by means of which the problem posed by the gates is overcome is located in the nodes which are marked by an S on a given player's screen (see figures 4-3 and 4-4). If a given player's partner moves into one of these switch nodes, then the configuration of gates is reversed for that (given) player. Thus, if player A wishes to traverse a currently-gated path, he must enlist player B's cooperation, find out where player B is located in the maze, and guide B into a switch node near B and which is visible to A on A's screen. If the players achieve this successfully, then the position of player A's gates is reversed such that previously gated paths are now open, and previously open paths are now gated. This does not affect B's gate positions; if B desires to have his own gate positions changed, he likewise has to find out where player A is and thence guide A into a switch node visible to him (B) on his (B's) screen. The positions of barriers and S nodes is also different for both players. Thus, verbal cooperation is required in order to effect appropriate barrier switching and to avoid accidental barrier switching (the switching mechanism operates even if a player moves into one of his partner's S nodes unintentionally). Each game thus takes the form of a number of moves, between each of which there is a certain amount of dialogue which takes the form of discussions of the locations of players' position markers, gates, and S positions, and the formulation of plans. All conversation is taped.

The game task is thus used to elicit genuinely spontaneous and natural dialogue which is produced for a purpose (to play a game), yet which is also appropriately experimentally constrained (the domain of discourse being the game itself). The maze game task employed to constrain the topic of discourse is unique to the work reported here; however, certain of its features are shared by previous studies which have examined spontaneous speech production. Before giving a detailed exposition of the experimental game task used in the work reported here, a brief review will be given of previous experiments in which the investigators have sought to constrain the domain of discourse in tasks involving the generation of spontaneous speech.

The disparate theoretical backgrounds of these precursors to the present task can only be touched upon briefly, since the variety of theoretical backgrounds to these experiments, and consequently the variety of questions which have been addressed using these tasks, is very large. These various tasks are of relevance in the present context in that they all employ the methodology of (to greater or lesser extents) constraining the topic of discourse and examining speech production, and also in that some of them, in addition, share particular features in common with the maze game task used in the present work. It is to a brief exposition of these antecedent speech production experiments that we now turn.

b) Some experimental speech production tasks employing a restricted domain of discourse: a brief review.

As noted above, several studies have investigated naturally produced speech in order to examine a variety of aspects of language production and understanding. These studies have also varied in the extent to which the speech produced by subjects was constrained: some studies (Sacks, Schegloff and Jefferson, 1974; Schegloff, Jefferson and Sacks, 1977) have not constrained the topic of discourse at all, whilst others (e.g. Marslen-Wilson et al. 1982, Levelt, 1982) have constrained the topic of discourse to varying degrees.

The first studies to be reviewed (Sacks, Schegloff and Jefferson, 1974; Schegloff, Jefferson and Sacks, 1977) were conducted in the sociological tradition of research known as ethnomethodology (Garfinkel, 1967), in which one of the primary interests is in uncovering the background workings of the familiar, 'taken-for-granted' social world (Pearson, 1975), for example, the background workings of social interaction in conversation. This study of practical reasoning, of how people make sense of what happens to them in social interactions, manifests itelf in the case of conversation in the form of studies of the organizational features of conversation such as studies of how individuals terminate conversations (Schegloff and Sacks, 1972) and the organization of turn-taking in conversation (Schegloff et al., 1977). These studies indicate that speakers and listeners are sensitive to interactionally consequential features of the conversation, and use quite mundane features of the conversation to control their interaction with one another.

The first such study to be reviewed here is that of Sacks, Schegloff and Jefferson (1974), in which the organization of turn-taking in conversation was examined. These researchers used transcripts of audio-taped naturally-occurring (unconstrained) conversation to investigate turn-taking in conversation, that is, to examine how the participants in the conversation succeed in 'meshing' their contributions such that there is little or no simultaneous talk.

They firstly noted several obvious facts which would have to be accounted for by a model of turn-taking. Their observations included:- (a) the fact that, overwhelmingly, one party talks at a time, though speakers change, and despite variation in the length and sequencing of turns; (b) the fact that transitions are finely coordinated, and that techniques are used for allocating turns; (c) the fact that the length of conversation and what parties say is not specified in advance; (d) the fact that the number of parties to the conversation can vary; and (e) the fact that repair mechanisms exist for dealing with turn-taking errors and violations (for example, if two parties find themselves speaking simultaneously, one of them will stop prematurely, thus repairing the trouble).

Taking the above five observations as their starting-point, Sacks, Schegloff and Jefferson propose a model to account for the organization of turn-taking in conversation. They specify that the model should be both context-free (that is, it should be capable of capturing the most important general properties of conversation) and capable of context-sensitivity (that is, it should be possible to accommodate and be sensitive to various parameters of social reality in a local context). Before turning to their characterization of turn-taking, two other important sources of variation between conversations should be noted:

1) There are various sizes of 'unit' out of which a

speaker's turn may be constructed - such 'turn-constructional units' can be lexical, clausal, or sentential in length. It is possible for a listener to 'project' (that is, anticipate) the location of the completion of a unit type - the first possible completion of a first such unit constitutes a <u>transition-relevance place</u>. It is around such projectable transition-relevance places that transfer of speakership is coordinated. Initially, the speaker is entitled to use one turn-constructional unit when speaking, but he may use more than one unit on any particular occasion.

2) Techniques exist for allocating turns, and these are of two types:

(a) Those in which the next turn is allocated by the current speaker selecting a particular individual to be the next speaker;

and (b) Those in which the next turn is allocated by self-selection.

Sacks, Schegloff and Jefferson characterize the organization of turn-taking in conversation as a set of rules which govern turn construction, provide for the allocation of the next turn to one party, and which coordinate transfer so as to minimize gap and overlap. The rule-set is;

(1) For any turn, at the initial transition-relevance place of an initial turn-constructional unit:

(a) If using a 'current speaker selects next speaker' turn allocational technique, then the party so selected has the right and is obliged to take next turn to speak; no others have such rights or obligations, and transfer occurs at that place.

(b) If <u>not</u> using a 'current speaker selects next' turn allocational technique, then self-selection for next speakership may, but need not, be instituted: first starter acquires rights to a turn, and transfer occurs at that place.

(c) If <u>not</u> using a'current speaker selects next' turn allocational technique, then the current speaker may, but need not, continue, unless another self-selects.

(2) If, at the initial transition-relevance place of an initial turn-constructional unit neither rule 1a (current speaker selects next) nor rule 1b (self-selection by next speaker) has operated, but instead rule 1c's option has been employed (current speaker has continued) then the rule-set a-c reapplies at the next transition-relevance place, and recursively at each next transition-relevance place, until transfer is effected.

Lower-priority rules constrain the use of higher-priority options: for example, the current speaker must select a next speaker before the initial transition-relevance place, otherwise there is a possibility of self-selection by some other taking place at the transition-relevance place. This ordering of the two turn-allocational techniques, such that the 'current speaker selects next speaker' option must be exercised before the 'self-selection by next speaker' option is compatible with the observation that one party talks at a time.

The rules provide for turn-transfers occurring at transition-relevance places - this confines gap and overlap to transition-relevance places and thus eliminates gap and overlap in most single turns. In fact, the rule-set is fully compatible with the observations noted earlier. For example, the rule-set does not specify what is to be said, how long a turn is to be, or the relative distribution of turns, in advance. The existence of a range of possible unit types for the construction of a turn, and the possibility of a current speaker using more than one instance of a unit type at any time provides for the possibility of the observed variation in turn length.

The model is a 'local management' system, that is, all the operations are local in a temporal sense, being directed to 'next turn' and 'next transition' on a turn-by-turn basis. The system is locally managed with respect to both turn size and turn order, and it deals with single transitions at a time, allocating 'next turn'. The model is also a party-administered system, that is, the variability in turn-size and turn-order is under the control of the parties to the conversation. The system fits to conversational interaction - any party's contribution to turn-order determination is contingent on, and oriented to, the contributions of other parties. Turn size is also a product of conversational interaction, since a speaker must speak in such a way that others are able to project possible completion points of his talk, allowing others to use the projected transition places to begin talking. Thus, the turn as a unit is interactionally determined.

The fact that turn-size and turn-order are locally managed, party-administered and interactionally controlled means that turn-size and turn-order can be brought under the jurisdiction of <u>recipient design</u>. Recipient design is the feature of talk being constructed for the particular other(s) who are the co-participants in the conversation (cf. Clark and Murphy's notion of 'audience design' discussed in chapter 3).

Thus, Sacks, Schegloff and Jefferson's model is compatible with both the grossly observable facts noted earlier and other phenomena such as recipient (or audience) design. It should be borne in mind that the phenomena of interaction studied by the ethnomethodologists (such as turn-taking and the closing of conversation) do not rely upon visual contact between the participants; the same phenomena are exhibited in telephone conversations, and indeed some of the ethnomethodologists' data are extracts from telephone conversations (e.g. Schegloff and Sacks, 1972, reprinted in Turner, 1974, examples on pp. 249, 250, and 252). In the experimental task used in the work reported here, subjects are located in separate soundproofed booths with verbal contact being achieved via a headset and microphone, and so in the present experimental task, the situation of speaking and listening is somewhat akin to the situation of a telephone conversation. In such situations, appropriate turn-taking and meshing are of great importance and cannot be mediated by non-verbal cues such as eye contact, and must therefore be mediated by linguistic cues (such as projectability of a turn's completion point) and paralinguistic cues (such as a falling intonation contour at a turn's completion point).

These last points underline a criticism of Sacks et al.'s somewhat atheoretical exposition of the turn-taking phenomena; Sacks et al. do not deal in any detail with the relative importance of visual, paralinguistic and purely verbal cues in the coordination of speaking. Furthermore, they give a primarily empirical exposition of the phenomena, unlike Rommetveit (1974), for example, who would explain the fine coordination of transfers of speakership in terms of intersubjectivity, complementarity and synchronization of intentions and thoughts (see chapter 3, where his example of highly elliptical exchanges being best explained in terms of nearly perfect complementarity and synchronization of intentions and thoughts was discussed).

Schegloff, Jefferson and Sacks (1977) used the same methodology to investigate the organization of 'repair' in conversation. They noted that self-correction and other-correction are related organizationally, with self-correction being preferred to other-correction, and self-correction being vastly more common than other-correction. This differential prevalence of, and preference for, self-correcton rather than other-correction indicates that speaking is indeed listener-oriented, as Rommetveit (1974; see chapter 3 of this volume) has argued. It suggests that speakers are monitoring what they say and are skilled at detecting and correcting errors as they occur in the speech stream. This eases the listener's task inasmuch as he does not often have to infer beyond an erroneous utterance to what he thinks the speaker meant to say.

Schegloff, Jefferson and Sacks distinguish between '<u>repair</u>' and '<u>correction</u>'; 'repair' is the more general case. Repair is sometimes found where there is no hearable error, mistake or fault; and hearable error does not necessarily yield the occurrence of repair/correction. Hence, nothing is in principle excludable from the class 'repairable', and, in view of the possibility of 'repair' taking place where no error has occurred, the term 'repair' is a better one for the general range of occurrences, with 'correction' being a particular type of repair.

Schegloff et al. also distinguish between the <u>initiation</u> and the <u>outcome</u> of a repair, since the person who initiates a repair is not always the same person who completes it.

The noted preference for self-repair rather than other-repair is in practice partly produced by the following factors:

(i) Opportunities for self-initiation of repair come before opportunities for other-initiation. Self-initited repairs can occur in three places - the trouble-source turn, the trouble-source turn's transition space, and in the turn subsequent to that which follows the trouble-source turn. Furthermore, possible other-initiations of repair are often <u>withheld</u> <u>past</u> the possible completion of the trouble-source turn, resulting in an 'expanded' transition-space and giving the speaker more of an opportunity to initiate a repair himself.

(ii) Massively, for those repairables on which repair is initiated, same-turn and transition-space opportunities for self-initiation <u>are</u> <u>taken</u> by speakers of the trouble source.

(iii) Successful self-repair regularly occurs within the same turn as the trouble-source, that is, before the other has had an opportunity to initiate repair.

(iv) Other-initiations lead to self-repair; the other's opportunity to initiate repair is used to locate the trouble, affording the speaker a further chance to effect self-repair, a chance which he usually takes.

A last interesting point noted by Schegloff, Jefferson and Sacks is that even when others do make corrections they are frequently modulated in form, for example, either by downgrading the correction on a 'confidence/uncertainty' scale, or by turning the correction into a joke.

These theorists have thus used unconstrained spontaneous dialogue in order to investigate organizational aspects of conversation. It would, of course, have been possible to investigate turn-taking and repair using a constrained dialogue situation, but it would have been impossible not to have used spontaneous, naturally occurring speech, since the phenomena under investigation are aspects of <u>natural</u> dialogue.

Other theorists have constrained spontaneous dialogue by means of an experimental task which itself acts as the topic of discourse. For example, Deutsch and Krauss (1960, 1962) used a two-person experimental game to examine the phenomenon of how two individuals in an interdependent decision-making situation arrive at bargaining agreements. Deutsch and Krauss' work belongs to a tradition in social psychology in which experimental games are used to examine cooperation and conflict between individuals in situations of interdependent decision-making. There is a vast literature in social psychology on this topic, but it is of little relevance in the present context, since these studies involve individuals making their (simultaneous) choices in the games without any communication occurring between them. Because of this background of studies involving a complete absence of communication between individual decision-makers, Deutsch and Krauss introduced the possibility of communication between the decision-makers as an independent variable. Deutsch and Krauss were thus interested primarily in the question of whether the possibility of communication (either in an optional or a compulsory communication condition) would assist their subjects in reaching an effective agreement which would benefit them both. (In experimental games from the above-mentioned tradition, it is possible for individuals to reach an 'agreement' about their decisions without communication taking place). Like the experimental task used in the present work, Deutsch and Krauss's task takes the form of an experimental game in which the two individuals are in an interdependent relationship insofar as their chosen individual courses of action have conequences for one another.

Deutsch and Krauss's experimental task allows the examination of the effect of two independent variables on subjects' ability to reach a bargaining agreement, the two independent variables being 1) communication and 2) the possibility of (unilateral or bilateral) threat. The experimental task takes the form of a 'trucking game': subjects are asked to imagine that they are in charge of a trucking company that carries merchandise over a road to a destination. For each trip that they complete, they receive 60 cents minus their operating expenses (which mount up at the rate of 1 cent per second). Thus, the longer they take, the higher the expenses deducted from the sixty-cents payout. Each subject is assigned a name, Acme or Bolt, and has to get from a start point to a destination, as follows:

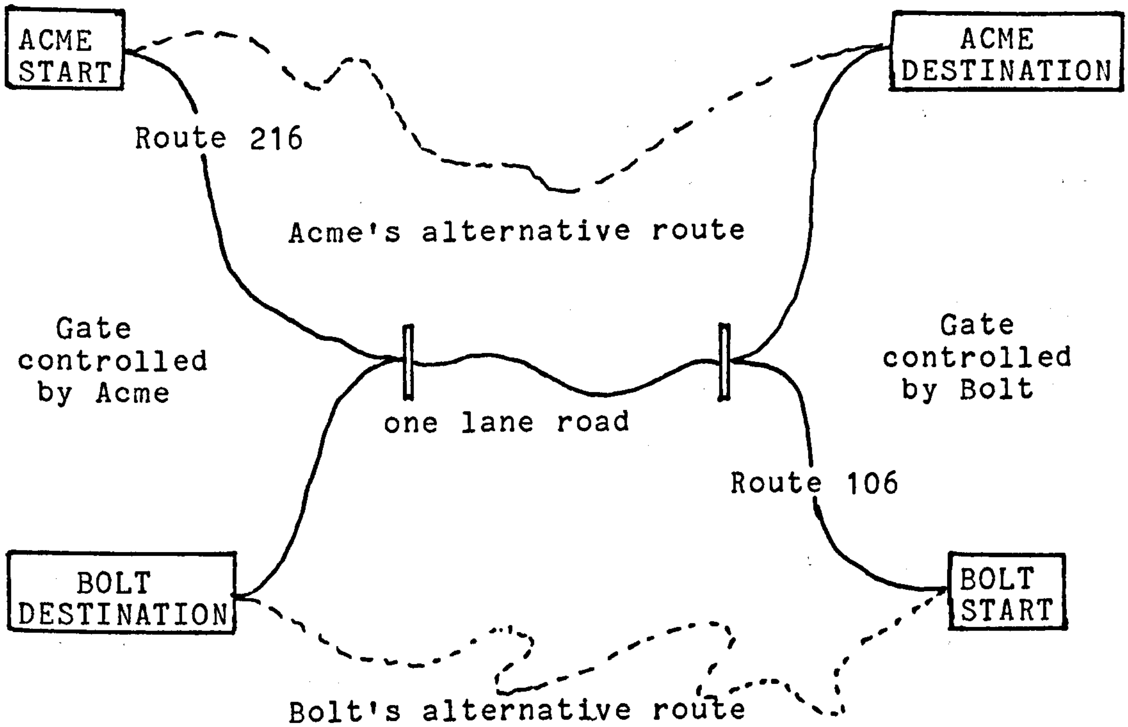


Figure 4.1: The roads available to subjects in Deutsch and Krauss's (1960, 1962) 'trucking game'.

At one point, the paths of the two players coincide - the <u>one-lane road</u>, where two trucks moving in opposite directions cannot pass each other: one has to back up. The alternative route can be taken by the players, but the alternative route is 56% longer than the main route, and subjects are aware that they will lose at least ten cents when using the alternative route. The agreement which requires to be reached is therefore the effective use of the two types of route by both players so that each player may maximize his gains. Subjects were instructed to individually gain as much as possible, regardless of what the other player gains or loses. However, all subjects were aware from the outset that gains and losses would be imaginary. Their respective interests are nevertheless in competition and it is to both their advantages to reach a bargaining agreement.

The potential for threat lies in the gates located at either end of the one-lane road. By closing his gate, a player can prevent his partner from travelling over the one-lane section of the main route. Threat can be bilateral (where both players have gates) or unilateral (where only Acme has a gate).

Players played the game in separate rooms and did not know who their partner was. A panel indicated to each player whether the gates were open or closed, the position of the players' own truck, and also when the two trucks had met head-on in the one way section.

The players were told the two rules of the game:

(1) Routes could only be changed at the start; it was not possible for a player to switch directly from, for example, the main route to the alternative route. If a player had already travelled some distance along a route and desired to change route, she had to reverse back to the start to do so.

(2) Where subjects had control of gates, the gate could only be used (closed) when the subjects were on the main route, not when

they were on the alternative route or at their destinations. Gates could be opened at any point of the game, however.

Subjects were given practice trials to familiarize themselves with the game. There were three conditions with respect to the communication variable (no communication, optional communication and compulsory communication). There were three conditions with respect to the threat variable (no threat, unilateral threat, and bilateral threat). The dependent variable was the median joint payoff, and the value of this reflected how quickly an agreement with respect to the use of the one lane route was reached.

Deutsch and Krauss found that the availability of a threat potential adversely affected players' ability to come to an effective agreement.

Agreement was least difficult in the 'no threat' condition, more difficult in the unilateral threat condition, and exceedingly difficult to arrive at in the bilateral threat condition. The communication variable was manipulated in order to ascertain whether the opportunity (or requirement) to communicate could ameliorate the difficulty that bargainers experience in reaching agreement. In the optional communication condition, the communication variable had no effect upon the players' ability to reach effective agreements. Only a minimum of communication did occur in the optional communication condition, and a compulsory communication condition was therefore run. In the compulsory communication condition (in which subjects were instructed that they had to talk on every trial), subjects' mean joint payoffs were significantly raised in the unilateral threat condition but not in the bilateral threat condition. In the 'no threat' condition, coordination of actions was sufficiently straightforward for players that communication failed to produce any visible effect. The manipulation of the communication variable (particularly, the compulsory communication condition) only had an effect in the unilateral threat condition; in the no threat condition there is a 'ceiling effect' in that payoffs are already high in the 'no communication' condition (since coordination is relatively straightforward) and communication does not therefore have any significant effect. In the bilateral threat condition, on the other hand, the competitive motivation present seems too great to be overcome, even by the compulsory communication treatment. Thus, there is only an effect of communication in the unilateral threat condition.

Although Deutsch and Krauss's study does depart from the tradition of research in which it originated in that it does allow communication in the game situation, communication was treated as an independent variable only and consequently the resulting transcripts of dialogue were not analyzed in any way. Deutsch (1973) says little about the dialogues produced by his experimental task, except that communication in bargaining situations serves several functions, namely a) coordinating efforts (exchanging bids, etc.), b) conveying information (threats, insults, etc.) and c) expediting the development of agreements (see Deutsch, 1973, for a review of these studies).

The Soviet psychologist B. F. Lomov (1979) is another theorist who has examined communication in task-constrained situations. The historical context of Lomov's work is the Soviet tradition of research on the role of language in relation to higher cognitive processes in man, and this tradition goes back via the work of A. R. Luria to the theories of L. S. Vygotsky (1962). Lomov's work is an examination of differences in subjects' performance on tasks under conditions of communication as opposed to solo performance on the tasks. Lomov is interested in the differences between solo- and communication-condition performance of subjects, and also in which aspects of communication are common to all three tasks studied by him (visual search, the drawing of a street map, and the characteristics of text recall).

Lomov begins his paper by criticizing the previous 'abstract functionalism' in investigations of mental functions and processes, which attempted to detach these functions from the whole system of psychic phenomena.

The functions were thus treated as natural, unique qualities in themselves. Subsequent Soviet theory considered psychic functions and processes in the context of a real subject's activity. Lomov argues that communication is an undetachable part of human life and that it is essential to take into account the dependence of dynamics of psychic processes and functions on forms, ways and means of communication.

Lomov presents evidence that "psychic processes" (particularly, visual search, conceptualization and memory processes) "performed under conditions of communication differ from those of individual activity" (Lomov, 1979, p. 212).

For example, in the verse-text reproduction task, subjects were asked to reproduce the first chapter of Pushkin's 'Eugene Onegin', singly and then subsequently in pairs under conditions of direct communication. (The experiment was "natural" in that subjects did not know they were being audio-taped or that they were taking part in an experiment). Recording of individual performance yielded 'classic' memory data, such as primacy and recency effects, substitution of meaningful summaries instead of word-for-word reproduction, and so on. In joint text reproduction under conditions of communication, the extent of word for word recitation was <u>greater than</u> the total sum of both subjects' individual recitations. The material that the two subjects individually remembered forms a 'common memory fund' which acts as a 'skeleton' to help reproduction of other parts of the text. In joint reproduction, 'gaps' (portions of the text which neither subject can remember) are more clearly identified, and become the focus of joint efforts. Having discovered the gaps, the subjects actively propose hypotheses, discuss and correct them. Lomov also notes that in the process of reproduction under conditions of communication, correct recall happens more often than in the case of the individual process, and the process of joint reproduction is more active and emotional than individual text reproduction, which contributes to its greater effectiveness.

Lomov notes the similarities in communication across all three tasks, and these include:

(a) The main function of communication in the three experiments was the exchange of the result of each subject's cognitive activity.

(b) The first stage in the process of communication is the selection of common 'coordinates' of joint activity (that is, 'reference points' or 'foothold images'). It is in relation to these that the whole process of communication is built.

(c) Lomov likens the development of the process of communication to a spiral. There exists an alternation of functions of both communicators, and all communicators have two-way and reversible relations. Significant aspects of communication are synchronization of the communicators' actions, their reciprocal stimulation, regulation, correction and supplement. During the process of communication, common programmes and common strategies of joint activity are being formed. The strategy formed under conditions of communication differs qualitatively from the individual one.

(d) The effectiveness of problem-solving is greater under conditions of communication than under conditions of individual problem-solving in all three experiments. The two subjects taking part in the communication form a kind of interpersonal psychological 'community' with a 'common fund' of concepts, ideas, methods of problem-solving, etc. This increases the effectiveness of the activity.

(e) The role played by speech in communication as opposed to gestures and facial expressions (etc.) varies from task to task. In some cases, speech plays the leading role, whereas in other cases gestures play the leading role. (The latter case is most characteristic of the tasks requiring spatial orientation and reproduction of spatial characteristics of objects). Lomov argues that in the exchange of information regarding emotional states, facial expressions probably play a leading role. Thus, the relative contribution of the various communication channels varies as a function of the task.

Lomov's work is of interest in the present context since he attempts to characterize the general features of communication in task-oriented dialogues. In a discussion of Lomov's paper, Segalowitz suggested that two differences between conditions of communication and individual situations which might partly explain the greater effectiveness of joint problem-solving are:

(1) The level of motivation for subjects may be different.

(2) There is feedback for each individual subject which might improve his performance.

These remarks will be considered again in part II of this thesis (chapter 10), where a comparison will be made between spatial location description under conditions of communication and in a monologue.

Lomov's work involved the tape-recording of natural dialogue in task-constrained situations; this represents a greater level of constraint on the topic of discourse than does the work of Schegloff et al. which was briefly reviewed earlier in this chapter. Another study of interest which used a constrained speech production task was that carried out by Marslen-Wilson, Tyler and Levy (1982; see also chapter 3 of this thesis). Marslen-Wilson et al.'s experimental results were considered earlier, as was also the theoretical background to their study. What is of interest in the present context is their methodology. This methodology was unusual in that it involved a single speaker, a single listener, and a single dialogue. The speaker's task was to re-tell from memory a narrative which he had previously read. The dependent variable studied was the way in which the speaker introduced and subsequently referred to the actors in the story, the speaker being both audio-taped and videotaped. This usage of ony one speaker and one story constitutes a rather unusual methodological strategy. Marslen-Wilson et al. argue that while their results require to be tested against a larger data-base, the particular aspects of this particular speaker's performance that were the most theoretically important are precisely those that are least likely to be peculiar to him as an individual. The general sensitivity of the speaker's use of referential devices to the informational environment in which they occurred (see chapter 3 of this volume for a summary of Marslen-Wilson et al.'s results) would, they argue, be the case for any other normal speaker.

Marslen-Wilson et al.'s results using the constrained speech production task dovetailed into their speech processing theory rather well. They conclude that language has to be studied both as a cognitive process and in terms of its natural communicative function in normal interaction; both these approaches are interdependent. It is, in general, fair to conclude that in the past, the latter aspect of language has been relatively neglected as compared to the former aspect.

Marslen-Wilson et al.'s study again illustrates the general principle of studying constrained speech production, and exemplifies the point that theoretically interesting questions can be addressed using such a task. Indeed, Marslen-Wilson et al. point out that the study of referential devices using isolated sentences is of very limited value compared to their own study using lengthy naturally-produced discourse, since the experiment using speech production involving lengthy, connected discourse allows analysis of the contexts of occurrence of referential devices. They argue that, without an awareness of the entire history of a discourse entity,

> " . . . the proper analysis of any given discourse reference to this entity is simply not possible" (Marslen-Wilson et al. 1982, p. 371).

Another study which utilized a constrained speech production task was reported by W. J. M. Levelt (1982). Levelt concerns himself with the general problem of <u>linearization</u> - that is, the way in which a speaker orders clauses in discourse: speakers have to make decisions on the ordering of clauses. Apart from cases where events are being described, informational structures have no intrinsic linear order, and the speaker has the problem of mapping a two- or multidimensional information structure onto a linear string. For example, if describing the layout of an apartment, the two-dimensional structures in the apartment have to be mapped onto a linear speech string such that the listener can, within certain limits, reconstruct the two-dimensional structure from the description. In the case of more complicated informational structures, like the game of chess, for example, the speaker's task is much more arduous. Levelt's paper is specifically concerned with the problem of linearization in describing spatial networks, and he tests two models of the linearization process using an experimental set-up where subjects have to generate descriptions of the spatial structure of simple two-dimensional arrays.

Previous work on descriptions of spatial layouts had shown that subjects employed certain consistent strategies to aid them in linearization. For example, Linde and Labov (1975) had their subjects describe their apartments and the strategy employed by subjects was to embark on a mental _tour_ of the apartment, starting at the front door and moving through the apartment room by room. At choice points, they take one branch first and after describing it 'jump back' mentally to the last place of choice in order to select a next branch. Ullmer-Ehrich (1978) observed that when describing their living-rooms, subjects use a _gaze-tour_, that is, they mentally position themselves at the living-room door and gaze along the walls, describing the pieces of furniture one by one in the order of the gaze-tour; this strategy led some subjects to forget to mention items of furniture in the centre of the room.

Levelt remarks that speakers can often be captives of their linearization strategy, as is exemplified by the last example (of omitting furniture in the centre of a room when 'gaze-touring' around its walls). A third case where such mental touring operations are used as an aid to linearization in the production of descriptions is that of giving route directions, mentioned briefly in the last chapter. Once again, an 'imaginary walk' or mental tour of the route assists the subject in linearizing a description of route directions appropriately.

Notice the extent to which the above linearization strategies are consistent with an interpretation of cognitive operations in terms of operations on mental models (see Chapter 2). The spatial layout of an apartment would appear to be a prime case for 'mental modelling', so to speak. A gaze-tour would be conceived of in mental model terms as a scanning operation being carried out on a model. Likewise, the 'imaginary walk' employed in generating route-directions would appear to be readily interpretable in terms of internal scanning operations on a mental model in the form of a 'cognitive map' of the locality being described.

Levelt notes that the listener's reactions may become highly important for the speaker's linearization. The speaker has to take into account the listener's presumed foreknowledge and processing capacities when generating descriptions.

Levelt distinguishes between the _ideas_ underlying speech and the processes involved in the choice of linguistic _forms_ for the expression of those ideas, and hypothesizes that linearization takes place during the first, conceptualizing, set of processes. The formulating mechanism operates on a highly specified conceptual input; most work is done by the conceptual preliminaries. Linearization is thus a non-linguistic process. Levelt argues that linearization shows functional properties which may turn out to be fairly general for different types of discourse.

Levelt developed two models of the linearization process. Both take the form of transition networks for generating descriptions. One is more 'speaker-oriented', whilst the other is more 'listener-oriented'. The 'speaker-oriented' ATN model was constructed so as to capture the main features of what Linde and Labov observed in their study of apartment descriptions, that is, that the descriptions took the form of a tour with the moves in the tour preserving maximal spatial connectedness. 'Jumps' only occur back to unfinished choice points, and 'moves' and 'jumps' are as small as possible. The main difference between this and the listener-oriented ATN model is that the listener-oriented ATN model does not '_jump_' back to the last choice point but instead '_moves_' back step by step, making the descriptions 50% longer.

Levelt tested these two models by having speakers describe simple spatial networks consisting of different-coloured nodes connected by horizontal or vertical arcs, for example:

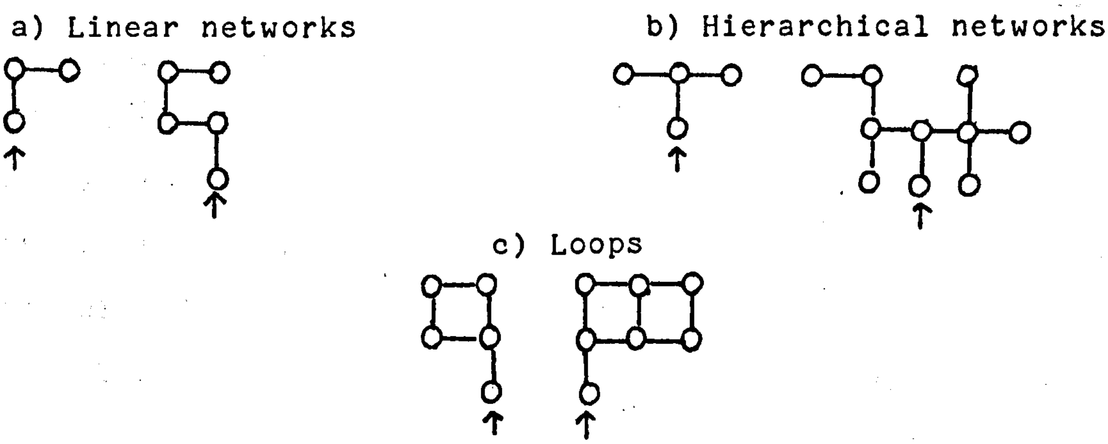


Figure 4.2: Spatial networks to be described by subjects in Levelt's experiments, the description beginning at the arrowed node (from Levelt, 1982).

The subjects were instructed to describe the network, beginning at the arrow, in such a way that a listener would be able to draw the network from the description on the audio-tape. This restricted domain of discourse allows the precise formulation of linearization models yet at the same time it captures some of the important aspects of other spatial domains like city maps, apartment layouts, etc.

Levelt also postulated hypotheses which attempted to predict, probabilistically, which arc would be taken by a subject at a choice-point. These principles of choice are of two types, global constraints on choice (general constraints arising from the nature of the network as a whole and which apply in quite diverse domains of discourse) and local constraints on choice (which consist of two parameters, the probability of 'going' straight, i.e. describing the branch straight ahead, and the probability of describing a right-branching structure. The probability of describing a 'straight ahead' branch and the probability of describing a right-protruding branch are both postulated to be greater than 0.5).

Three global constraints are postulated by Levelt:

(1) The probability of describing a shorter branch before describing a longer branch is greater than 0.5.

(2) The probability of describing branches with fewer embedded choice points before describing branches with more embedded choice points is greater than 0.5.

(3) The probability that, at choice points, loops are described before other branches is greater than 0.5.

Levelt's experimental findings pointed to the existence of two distinct linearization types - 33 subjects out of 53 exclusively 'jumped' in their descriptions back to choice points, 16 subjects exclusively 'moved' in their descriptions back to choice points, and only 4 subjects both jumped and moved. The first global constraint (that shorter branches should be described before longer branches) was strongly confirmed for both 'jumpers' and 'movers'; the second global constraint (that branches with fewer embedded choice points should be described before branches with more embedded choice points) was strongly confirmed for 'jumpers' but its converse (that is, seeking complexity initially) was the case for 'movers'; and the third global constraint (that loops should be described before other branches) was true for 68% of 'jumpers' and 80% of 'movers' (only in the latter case is the figure significantly different from what could be expected by chance).

There is thus confirmation of there being two types of linearization strategy, which correspond to the two ATN models, and Levelt concluded that, up to the level of hierarchical spatial structures, the ATN models are in almost faultless correspondence to the speakers' behaviour. It is only in the case of loop structures that deviations from the ATN models' predictions occur, and the main deviations are incompleteness and the 'cutting' of a loop and treating it as a two-branch hierarchical structure. (It was suggested by one of Levelt's colleagues that the problem with loops may be due to the difficulty of maintaining one's original deictic orientation while turning through a loop).

Levelt argues that the functional properties of linearization which his models embody and which appear to operate in the experimental data (properties such as preservation of connectivity, a general minimal effort principle based upon short term memory economy, etc.) are not limited to the particular domain in question but apply more generally to other discourse also.

Levelt's spatial networks are of interest here in that they share some features in common with the spatial networks embodied in the maze shapes used in the game task employed in the studies reported in this volume (in particular, the mazes used in the present work use arrays of nodes connected by horizontal and vertical pathways in a fashion similar to Levelt's networks but on a larger scale. See appendices I and II). Again, as with the work of Marslen-Wilson et al., the constraining of discourse topic to that of (in the case of Levelt's experiments) a highly controlled domain of discourse does not preclude the study of interesting questions pertaining to natural speech production. Indeed, Levelt argues that the results from his experiments have a generality which extends beyond the original restricted discourse from which they were generated and cites the work of J. Mandler (1978) as an example of the operation of similar principles in the retrieval of stories.

A final experimental task which elicits natural speech in a controlled manner is that described by Blakar (1973), a colleague of Rommetveit. This task has been used to examine a variety of aspects of pathological and normal dyadic communication in a task-oriented situation, and only some of the major findings obtained using the task will be reviewed here (see, for example, Blakar and

Pedersen, 1980; Blakar and Valdimarsdottir, 1981; Mossige, Petterson and Blakar, 1979; Pedersen, 1980; and Solvberg and Blakar, 1975, for examples of studies involving the use of this task).

The task involves one subject explaining a route through a map to another subject, who also has a copy of the map. In one condition, the task is simple in that the maps are identical. However, in the experimental (communication conflict) condition, the maps are different, with the follower's map having one more street than the explainer's. This violates one of the most basic preconditions for successful communication, that of the participants having a "shared social reality", a common here-and-now within which an exchange of messages can take place (Rommetveit, 1974; see also chapter 3 of the present volume).

The difference between the maps induces a situation of conflict in communication which the subjects have to resolve by identifying the error in the maps; if the task is not solved within a fixed time span (forty minutes) the task is terminated. Solvberg and Blakar (1975) demonstrated that parents of schizophrenic patients failed to solve the complex task within the prescribed time limit on four out of five occasions, whilst all of the control dyads (matched parents without pathological offspring) solved the complex task. Solvberg and Blakar also undertook a preliminary analysis of the utterances produced by the dyads (all conversation having been tape recorded and transcribed). Among the observations noted by Solvberg and Blakar were:

(a) A much larger proportion of the comments made by group S (schizophrenic patients' parents) were task-irrelevant as compared to the group N (control) dyads - the figures being 57% and 19%, respectively.

(b) Wives in group S couples gave more active directives, whereas in the group N couples the distribution of active directives was much more equal. This result ties in with Lidz' (1963) notion of "marital skew".

(c) Observations suggested that group S couples were less willing or able to endorse or adhere to contracts which regulate and monitor their communication (for example, contracts concerning the distribution of rôles).

(d) Solvberg and Blakar noted that utterances indicating egocentrism (i.e. where individuals are less able or willing to take the prespective, and speak on the premises of, the other with whom they are in communication) were more frequent in the group S dyads. An example of such egocentric utterances would be the use of deictic linguistic elements (for example "here", "there", "that corner" etc.) <u>without</u> the prior establishment of any intersubjectivity as to what such phrases refer to.

The observations regarding the greater egocentrism and lesser decentration (taking the perspective of the other with whom the individual is in communication) were amplified by a study by Mossige, Petterson and Blakar (1979), who identified from the tapes of dyads performing the task every single utterance that could, on the basis of the ongoing communication process, be assessed as _markedly_ egocentric or decentred. They found that the ratio of egocentric to decentred utterances was much higher (in fact twelve times as high) in the group S families than in the group N families, replicating Solvberg and Blakar's earlier findings.

This experimental task has also been used to examine sex and communication (Pedersen, 1980). In this study, subjects of two different sexes were given two different roles (Explainer _versus_ Follower) when participating in two different kinds of dyads (same-sex _versus_ opposite-sex) in the two different communication situations (_simple_, where the maps are identical, _versus_ _conflict_, where the difference between the maps is important). This experimental design allowed the examination of various 'common sense' hypotheses regarding sex and communication in two different situations (simple _versus_ conflict).

The communication was analysed in terms of efficiency (in terms of time taken to solution, as in the Solvberg and Blakar study) and the qualitative analyses were of measures of self-confidence and control. Among the hypotheses postulated by Pedersen were; (a) that female subjects should make better followers, while male subjects would make better explainers; (b) that the most efficient communication should involve a male leader and a female follower, while the efficiency of same-sex pairs should be intermediate and the least efficient pairing should involve a female leader and a male follower.

The results were of great interest and emphasized the importance of the situational variation (simple <u>versus</u> conflict situation). It was found that, in the simple situation, the sex bound communication pattern fits the common sense assumptions (with a male leader/female follower pairing being most efficient, and the female leader/male follower dyad being least efficient). Female/female same sex dyads were more efficient than male/male same sex dyads in the simple situation. The <u>opposite</u> results were obtained, however, in the conflict situation: the male explainer/female follower dyad was least efficient in the conflict situation, and male/male same-sex dyads were more efficient than were the female/female same-sex dyads. Similarly, the results of the qualitative analyses were complex: sex-linked patterns of communicative behaviour were extremely complex and varied across situations, roles and the sex of the person's partner. Thus, the variation in the situation in which the communication takes place is very important, and Pedersen's conclusion is that we should be cautious in making generalizations about sex-bound patterns of communication across different communication situations.

Pedersen points out that "At our present stage of knowledge, until a far broader selection of social situations has been systematically mapped, we must refrain from any general conclusions about the relationship between language/communication and sex. Through these future tasks a social-psychological perspective will be a necessary guide" (Pedersen, 1980, p. 113).

c) Conclusions

This brief review of previous experimental approaches involving language production tasks has illustrated the variety of methods employed and the variety of questions addressed using such methods. The work of Schegloff et. al. illustrated the use of transcripts of unconstrained natural dialogue to investigate interactional aspects of dialogue, namely, how the mechanism for organizing turn-taking in conversation operates, and how a bias exists whereby self-correction is preferred to other-correction.

Deutsch and Krauss's experimental 'trucking game' task was of interest, although Deutsch and Krauss's interest was primarily in how bargaining agreements were reached rather than in linguistic aspects of communication. The interdependence of the dyads (in decision making terms) in Deutsch and Krauss's work, and the fact that the task took the form of a game, makes their study of relevance in the present context, even although Deutsch and Krauss did not analyze the dialogues generated by subjects in the course of solving the task.

Lomov's interest in how cognitive functions like conceptualizing and memory are affected by communication demonstrated how interesting facets of language use could be explored using such task-oriented situations. Marslen-Wilson et. al.'s experimental task and its results illustrated that fact that important aspects of reference could be studied using a constrained speech production task, and Levelt's research demonstrated also that important general questions (in this case about linearization in speech production) could be addressed using a task where speakers had to describe simple, abstract, experimentally-controlled networks. The last experimental task involving task-constrained speech production reviewed was the 'communication conflict' task devised by R. M. Blakar. Blakar's task has been used to investigate a variety of questions of importance to social psychology, clinical psychology, and the study of communication in general. The two studies reviewed, one being concerned with communication efficiency in couples with and without a schizophrenic offspring, and the other being concerned with sex-bound patterns of communication, serve to illustrate the variety of aspects of language use which could be studied using such a task.

The task which was employed in the work reported in the present volume and which will be described in the next sections of this chapter contains features which are similar to those in some of the above tasks. Like Deutsch and Krauss's 'trucking game' the task used in the present work takes the form of a game in which the courses of action taken by each player affects the other player and in which cooperation is crucial. Like the task set to subjects by Levelt, the present game task involves subjects describing spatial networks (although subjects in this case have to identify specific locations within a spatial nework rather than having to describe the entire network). Like Lomov's subjects, subjects participating in the task used in the present work are engaged in a cooperative effort, <u>jointly</u> solving a problem. Finally, like Blakar's 'communication conflict' task, the present task can be used to examine a variety of interesting questions - semantic aspects of location descriptions generated by subjects, aspects of their strategies in solving the problem, and social psychological questions concerning control of dialogues, in addition to mainstream linguistic questions such as the development of ellipsis, etc., can be studied using the game task described in the next sections of this chapter and used in the present work.

The principal study to be reported in this thesis is of the meanings of the descriptions of locations within a variety of spatial networks. In the meantime, however, an exposition of the details of the task itself will be given.

d) The experimental procedure: the dialogue maze game

As outlined earlier in the present chapter (see section a)), the maze game task essentially involves two subjects alternately moving their respective position markers through a maze from different starting positions to different goal positions. The game task requires the subjects to cooperate with one another by coordinating their moves in such a way that the problems posed by obstacles in the maze in the form of gates are overcome. The mechanism underlying this gate-switching was sketched briefly earlier (in the introduction to the present chapter) and is dealt with in greater detail below. The important feature of the task is that, since subjects require to coordinate their moves through the maze, and since the subjects are located in separate soundproofed booths with only verbal communication being possible via a set of headphones and a microphone for each subject, the successful playing of the game requires that the subjects engage in fairly lengthy dialogues with one another. It is with the analysis of certain detailed aspects of these dialogues that the present work concerns itself.

Before proceeding to a description of the data and analyses, however, it is important that the reader be well acquainted with the nature of the task itself. In order to thoroughly acquaint the reader with the various features of the task, therefore, the following description of the task will be divided into three sections. The first section will reiterate some of the features of the task touched upon briefly in the introduction, specifying the rules of the game, the nature of the mazes, and so on. The second section will describe the opening move from a typical game from the subjects' perspective, in order that the reader may be well versed in what the task involves for the subjects. The last section will provide details of what, from the experimenter's perspective, is important about the task; what can be studied using the task, which variables can be manipulated, and so on. Taken together, the three sections should provide the reader with a suitably deep understanding of the game from the points of view of both experimenter <u>and</u> subject.

(i) <u>The features of the task</u>

As noted earlier, the game task involves two subjects jointly solving a problem by moving their position markers from (different) start positions to (different) goal positions within a <u>maze</u>. A version of the maze in plan view is presented to each subject on the visual display screen of the computer terminal at which each of the subjects is seated. Subjects are seated at separate terminals, each being located in <u>separate</u> soundproofed booths. The subjects' only communication is verbal, and this is achieved by means of a headset and microphone for each subject. (See figures 4-3 and 4-4 for diagrams of the maze display presented to two players at the start of a particular game).

Each maze display consists of:

(a) A collection of paths along which players can move, and nodes, which are located at the ends or at the intersections of these paths. During the course of the game, players move from node to node via available paths.

(b) The players' starting positions and goals (these are marked by an X and an asterisk, respectively). In the case of the games reported in the present work, subjects could only see their own position and goal position, they could not see their partner's location or goal position. (It is, however, possible for the experimenter to allow subjects to see their partner's position throughout the game in one of the optional game types available in the maze game program. This option was not employed in the work reported here - all subjects were unable to see their partner's position).

(c) Gates: these are blockages in some of the paths of the maze. Gate positions are denoted in the diagrams (see figures 4.3 and 4.4) by a small line across the centre of the path and perpendicular to the direction of the path. Gate positions were similarly denoted in the displays on the subjects' screens. If a player attempts to move across a gate by moving towards it, his position marker is automatically rebounded from the gate back to the node from whence it came: this type of move is illegal and impermissible.

Approximately half of the paths in each subject's maze have gates.

(d) Switch nodes Approximately, beween a third and a half of the nodes in a given subject's maze shape are switch nodes, and this is denoted (both in the diagrams used in this chapter and in the actual screen displays on subjects' computer terminals) by their containing an S (see figures 4-3 and 4-4). These switch nodes are the means whereby players can assist one another's passage through the maze. This is achieved as follows: a player who requires to pass along a currently gated path has to i) find out the position of the other player, and ii) guide the other player into one of the S nodes visible to him (the blocked player) on his terminal display, but which is invisible to the player being guided. The positions of gates and switch nodes are different for the two players and each sees only those relevant to himself or herself and not those of his or her partner. If a player does succeed in guiding his or her partner into one of the switch nodes on his or her screen, that player's gate configuration is completely reversed. Thus, if player A successfully guides player B into one of her own (A's) S nodes, then player A's gate configuration changes such that all previously gated paths become open and all previously open paths become gated. Player B's gate configuration is unaffected by this. Player A's gate configuration will change if player B moves into any node marked by an S on player A's screen even if player B does so unintentionally (she might do so unintentionally simply because she is unable to see the locations of her partner's S nodes). Thus, the moves the two players make matter critically to each other, and in the course of playing the game, subjects learn quite quickly how to effect deliberate gate switching and avoid unintentional gate switching.

(e) At the bottom of the display on each player's screen, a <u>tally</u> of the total number of moves made by <u>both</u> players up to the current point in time. When subjects attempt to cross barriers and in so doing make illegal moves, a number of penalty points (arbitrarily set at two in the present studies) registers against them, and the total number of penalty points incurred by <u>both</u> subjects is also displayed in the tally.

Each player moves in turn from his starting node to any other node by an open path. Only one node can be reached on each move and a player's turn terminates at this node. The subjects are exhorted to move to their respective goals in as few moves as possible, and incurring as few penalty points as possible. The game is terminated when both players are simultaneously located in their respective goals - that is, the game solution requires that firstly one player should reach his or her goal, and the other player must then enter his or her goal so that both are simultaneously in their respective goals. When this is achieved, the game is terminated, an appropriate message is displayed on the screen (for example "Task successfully completed on move 30 with 18 penalty points") in place of the tally, and players are allowed to make no further moves.

An additional optional feature of the maze game program was the possibility of a maze 'monster' being present in the maze. This monster, denoted by a capital M when present (visible to both players and occupying the same node on their respective mazes), is a semi-intelligent third player, whose task it is to occupy the same node as one of the players. If this is achieved, that player is considered to be 'eaten' and the game is thenceforth terminated, an appropriate message being displayed on both players' screens in place of the tally. The monster's presence makes the task more difficult for the players, as they then have to not only solve the task by cooperating with each other, but also avoid being 'eaten'. The monster moves once for every third move of both players, pursues one player at a time, and obeys the gate configuration of that player, that is, the monster will not traverse a closed gate in the pursued player's maze.

Prior to playing the game, subjects were fully informed of the rules of the game, and it was emphasized to subjects that they should cooperate with each other. They were informed that the solution of the game requires both subjects to be in their respective goals simultaneously, and the necessity to cooperate in order to achieve this aim was emphasized to the subjects.

(ii) <u>The maze game task: the subject's perspective</u>

An appreciation on the reader's part of what the game task involves for a pair of subjects playing it is best achieved by considering the opening moves in a typical (hypothetical) game: this will better acquaint the reader with the operation of the gate-switching mechanism. Consider figure 4-3: this illustrates what player A sees of the maze at the start of the game. Player A's start position is represented by an X, and her goal position is represented by an asterisk. There is no indication on player A's maze of where her partner (player B) is located or where she is headed for. Consider now figure 4.4: this illustrates what player B sees of the maze at the start of the game; her own (B's) goal and start positions are marked on the maze, but, like player A, she has no idea where her partner's start and goal positions are. The players are thus initially in mutual ignorance of each other's position and destination.

Comparison of figures 4-3 and 4-4 reveal several points of note:

a) The overall shape of the maze (the configuration of nodes and paths) is identical for both players; there are no extra paths or missing paths for one of the players as compared to the other (cf. Blakar's 'communication conflict' task maps);

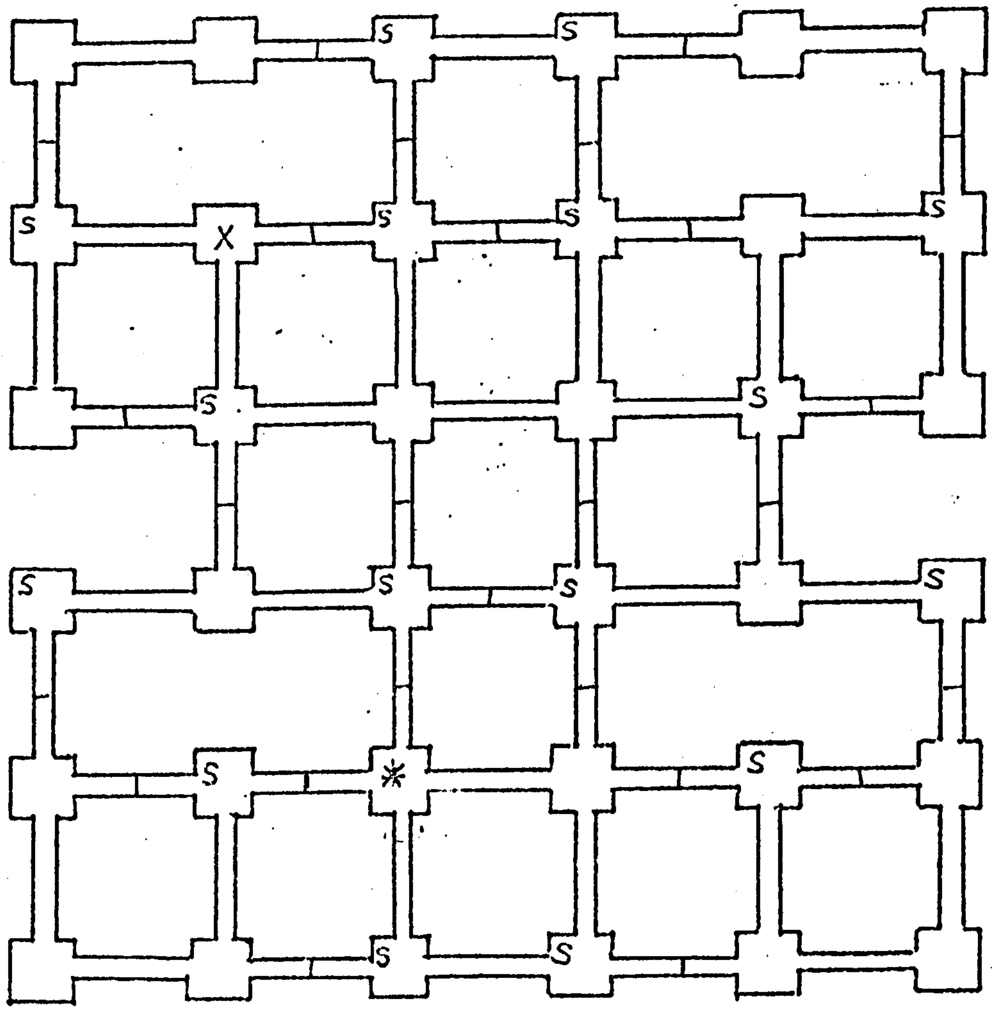


Figure 4-3: Typical Game: State of Game at Game Start for Player A

Key: X = Start Position
Asterisk = Goal Position
Small Black Barriers in some Paths Represent Initial Blockage Positions (i.e. Gates)
S = Switch Node

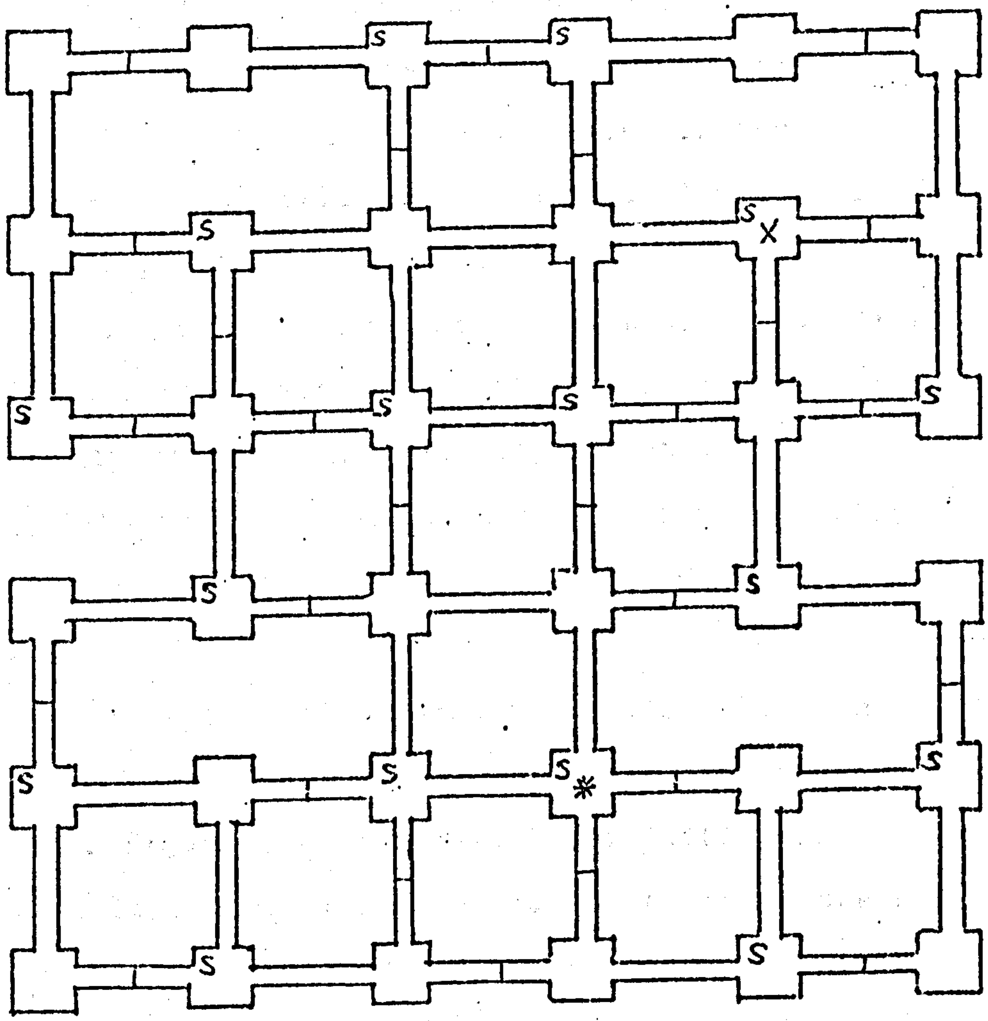


Figure 4-4: Typical Game: State of Maze at Game Start for Player B.

Key: X = Start Position
Asterisk = Goal Position
Small Black Barriers in some Paths
Represent Initially Gated Paths
S = Switch Node

b) Both mazes are absolutely symmetrical about a vertical centre line (in this particular case; asymmetrical mazes were also used. See the later section in this chapter). This symmetry is true for both the maze structure itself (paths and nodes) and also for S node and gate positions;

c) The two players are set the task of moving in a parallel fashion: their start and goal positions are symmetrically placed, although they do not know this initially (and they may not realize that this symmetry exists even when they are aware of one another's position).

d) The locations of S nodes and gates, although symmetrical for each player, are <u>different</u> for the two players (cf. figures 4-3 and 4-4). This difference is non-systematic in that A's gate and switch nodes are not simply the 'converse' of B's (i.e. located at the nodes where B's gate and switch nodes are absent). This means, in effect, that sometimes S nodes coincide for both players (for example, the two topmost S nodes coincide for both players) and sometimes they do not coincide (for example, consider the bottom-most horizontal line of nodes in the two versions of the maze and compare the locations of S nodes). The same is true for barriers. This effective non-predictability of the locations of the other's switches and barriers forces the subjects to discuss their positions, and this in turn guarantees the generation of dialogue.

The first task is for both players to describe their locations and goal positions. Player A might do this as follows:

> "I'm on the second row from the top, and the second box from the left, and my goal is on the second row from the bottom, third from the left".

The main emphasis of this thesis will be an analysis of such descriptions. Note, for the moment, that this description involves treating the maze shape as if it were an array of parallel horizontal lines (rows) of nodes and generating descriptions by first identifying the relevant line of nodes in which the node to be described is located, and then identifying the particular node involved. (A fuller discussion of these descriptions will be presented in chapters 6, 7 and 8 of this thesis). Suppose that player B describes her location in a similar manner:

> "Well, I'm on the second top row, fifth from the left, and my goal is on the second bottom row, fourth from the left".

Having thus described their locations and goals, the players now start planning to move. In all games, player A was first to move (this was indicated to her by her position marker flashing in and out on the computer terminal screen. B's position marker flashed when it was _her_ turn to move, and stayed on continuously on the screen when it was not. Obviously, it is impossible to represent this on these diagrams).

Suppose, then, player A is first to move. Inspection of figure 4.3 reveals that, in effect, only two options are available. There is no path above her position, so no movement is possible at all in that direction; she could move one node to the left or one node downward, where paths are available and there are no gates blocking her way. There _is_ a path available to her right but it is gated, so if she were to move to the right, she would rebound on the barrier and incur a penalty move, ending up at the node from which she started, no further forward and with two penalty points against her. Thus, she really only has two options of movement. That to the left takes her further from her goal, and that downward is in the general direction of her goal.

The node directly beneath player A's initial position is one of her own S nodes. Moving into this will not affect either player, (provided of course that it is not an S node for player B - see figures 4.3 and 4.4). Consultation with player B establishes that the node directly beneath player A is not an S node for player B, and player A makes her first move by pressing one of four computer terminal keys marked with arrows to signal movement direction relative to the plan view presented on the screen. Suppose she moves down, which is in the correct general direction. She moves into her own S node, which affects neither player, since that particular node is not one of player B's S nodes. The state of A's maze is now given in figure 4-5 (see figure 4-5). The only thing that has changed is A's position. B's maze is unaffected and would be as depicted in figure 4-4. As things stand at the moment, player A will be able to move to her right, or back up, on her next move _if_ her maze is unchanged by player B's next move. It is now player B's turn to move.

Consider figure 4-4. B has available paths to her right, to her left, and downwards from her. However, those to her right and downwards contain gates, and so moving in either of those directions would not advance her passage through the maze. Her best option would be to move to her left, that is, towards the centre of the maze, _provided_, of course, that this move did not actually worsen the situation for player A. A consultation with player A reveals (see figures 4.3 and 4.4) that the node located one move to the left of B is a switch gate for A. If B moved to her left, this would change player A's entire barrier configuration. This would, in fact, be to their mutual benefit (see figures 4.4 and 4.5): player B would be slightly nearer her goal, and player A's gates would change such that the paths below her and to her left would become open, and those to her right and above her would

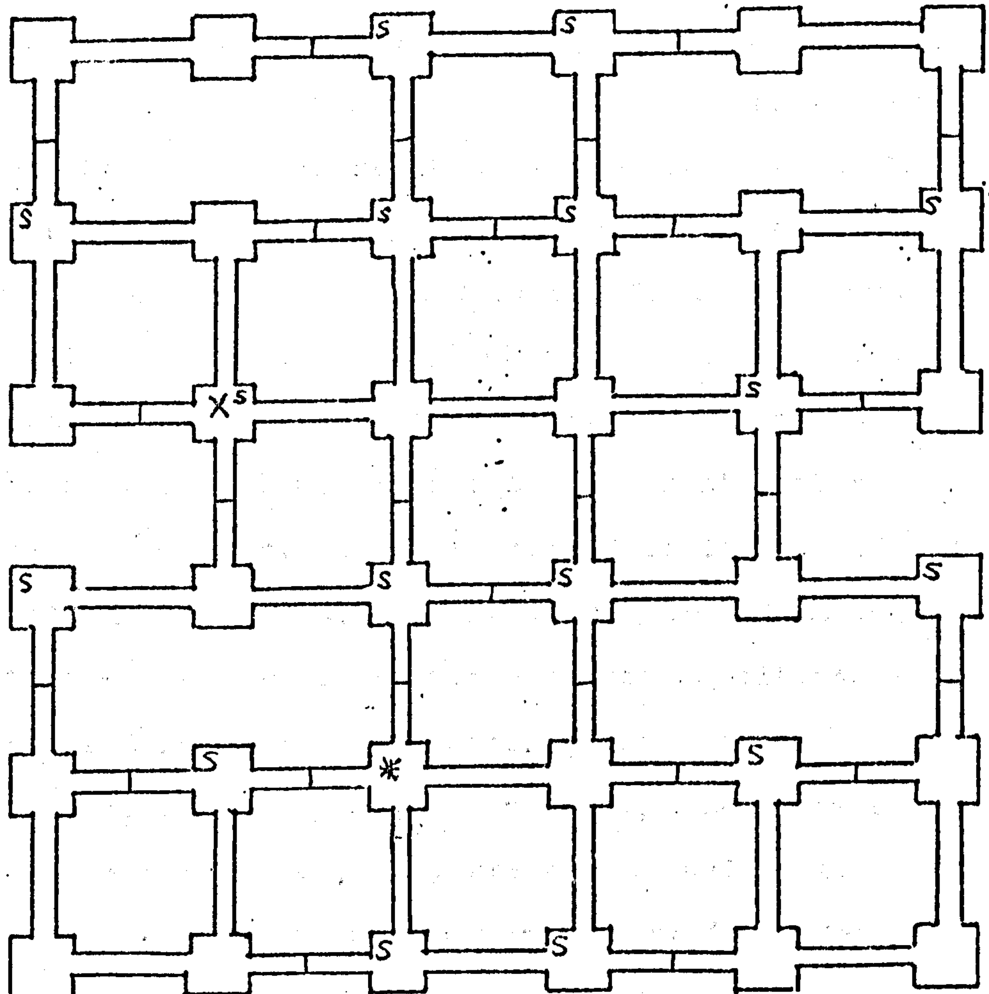


Figure 4-5: Typical Game: State of Game after one Move for Player A.

Key: X = Current Position
Asterisk = Goal Position
Small Black Barriers insome Paths Represent Currently Gated (Blocked) Paths.
S = Switch Node

become blocked. This would allow player A to move down on <u>her</u> next move, taking her in the direction of her goal.

Suppose, then, that player B <u>does</u> move to her left. figure 4.6 shows the state of player B's maze following her move. She is now nearer the maze centre and in direct vertical line with her goal, which is three movements below her in a direct path (see figure 4.6). Her own configuration of gates does not change; only player A's moving can achieve that. Player B has, however, moved into an S node for player A, and so player A's entire configuration of barriers has changed (see figure 4.7 and compare it with figure 4.5 in terms of gate positions). Notice that the change is in terms of gate positions - the locations of S nodes do <u>not</u> change. Where gates previously existed, they no longer exist, and where there were previously no gates, they are now in evidence. Player A is now free to move to her left or down, two courses of action that were impossible before player B made <u>her</u> move.

This discussion of the opening moves in a typical game thus illustrates the essence of the task. The interdependence of the players on one another, and the necessity for them to cooperate and coordinate their patterns of moving is clear. The successive removal of gates is necessary for a solution to the game to be attained. Notice that several changes of gates are necessary for the solution of any one game - examination

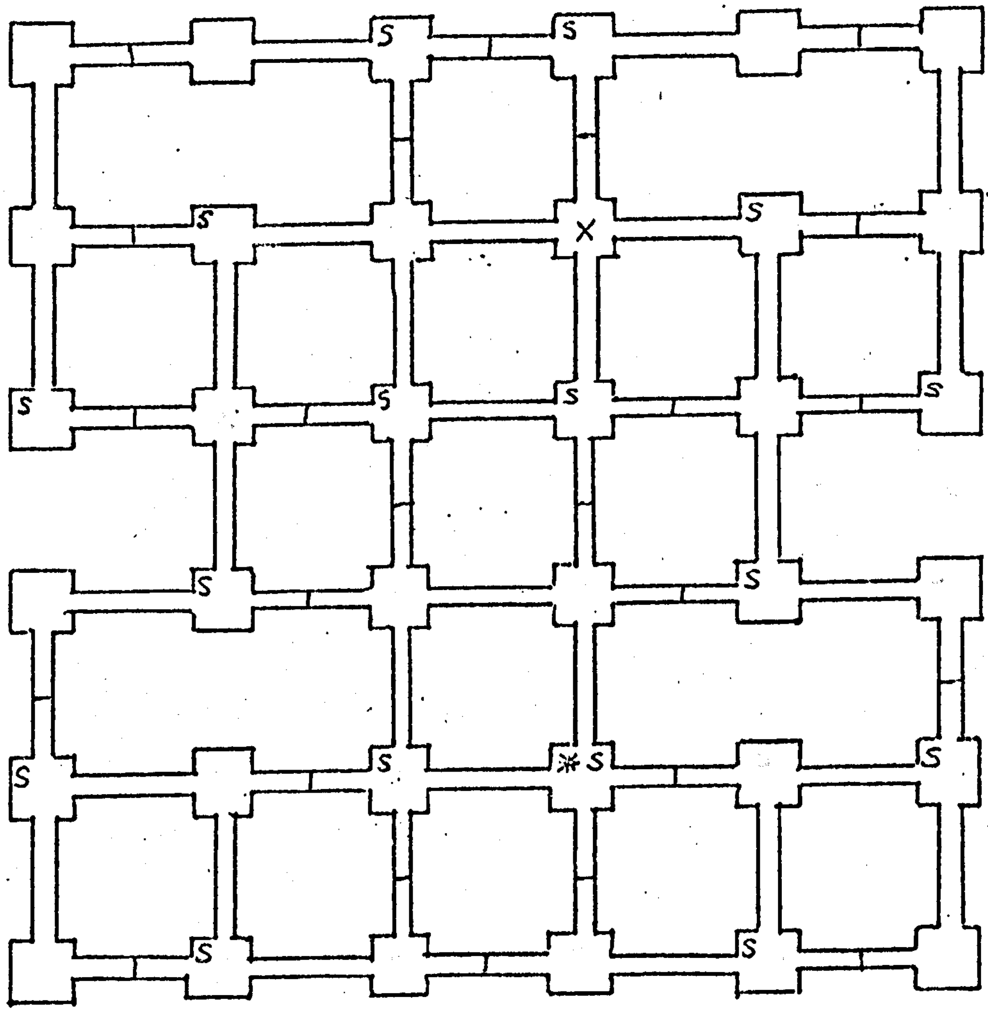


Figure 4-6: Typical Game: State of Player B's Maze following 2 Moves

Key: X = Start Position
Asterisk = Goal Positions
Small Black Barriers insome Paths Represent Initially Gated Paths
S = Switch Node

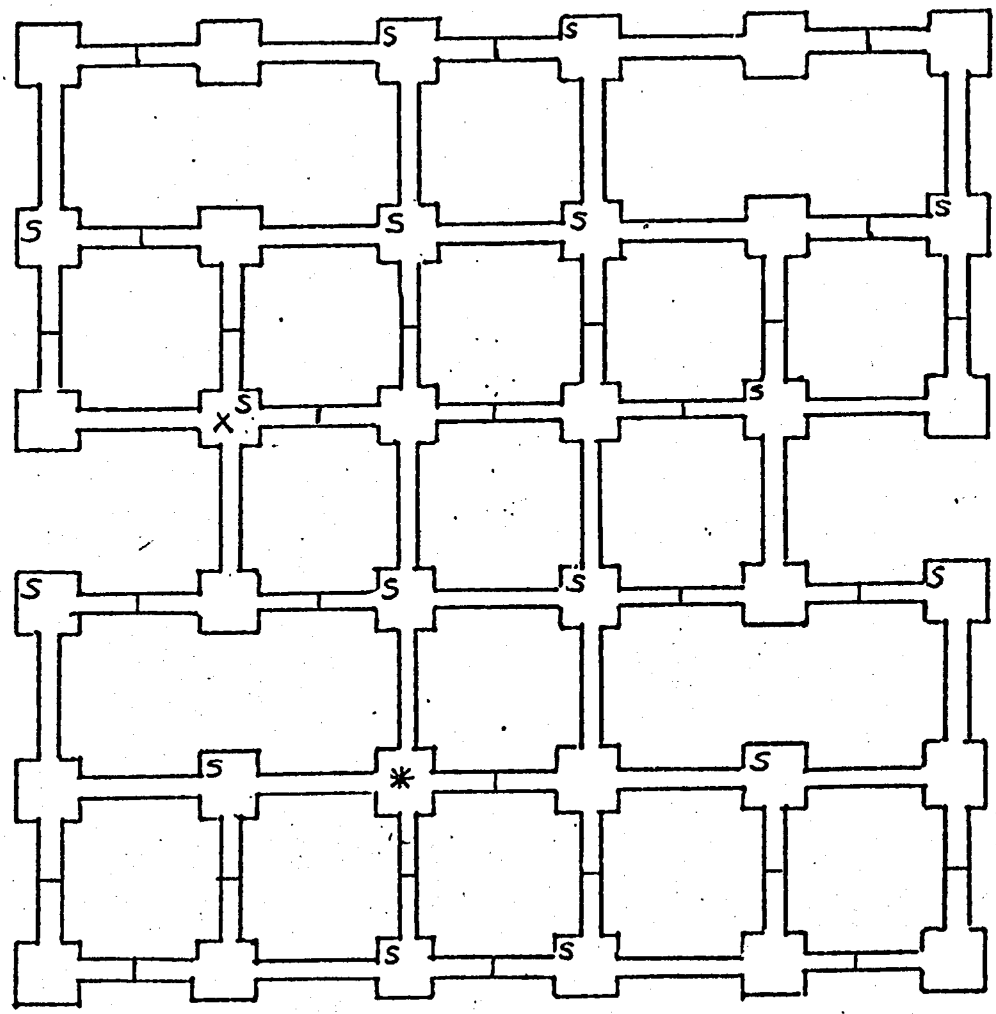


Figure 4-7: Typical Game: State of Player A's Maze after 2 Moves

Key: X = Start Position
Asterisk = Goal Position
Small Black Barriers in some Paths Represent Initially Gated Paths
S = Switch Node

of figure 4.7 will reveal that player A will soon require to enlist player B's cooperation again, since, as things stand, A can now move down one node , but after she has made that move, she can go no further towards her goal without having her gates changed again.

The typical game would continue with players A and B discussing possible moves, node positions, and the positions of one another's location markers. It should be borne in mind that neither player can see her partner's position, and so the repeated reporting of positions is necessary, particularly since players' positions change with each move.

The game would proceed, sometimes smoothly, sometimes with difficulty, until a solution is attained, at which point several minutes of dialogue have been spoken by both players and recorded by the experimenter.

The foregoing serves to illustrate the subjects' perspective on the task. Now we turn to examine the experimenter's perspective.

(iii) The maze game task: the experimenter's perspective

The game has been discussed from two points of view in the previous sections, namely, its properties as a game (the rules involved, and so on), and its properties from the point of view of a subject playing the game (illustrating what the game involves for the subject, and so on). We shall now consider the game from the experimenter's point of view, that is, in terms of how the task is realized in practice (with what equipment, and so forth) and also in terms of what variables can be manipulated and what sorts of aspects of dialogues can be studied using the task. Each of these aspects will be dealt with in turn.

a) Apparatus

The mazes were presented to subjects on Imlac corporation PDS1-G and 'Intel' (Intelligent systems corporation) Intecolor 8001 terminals, driven from a Data General Nova 210 computer using hard disks. (See figure 4.8 below).

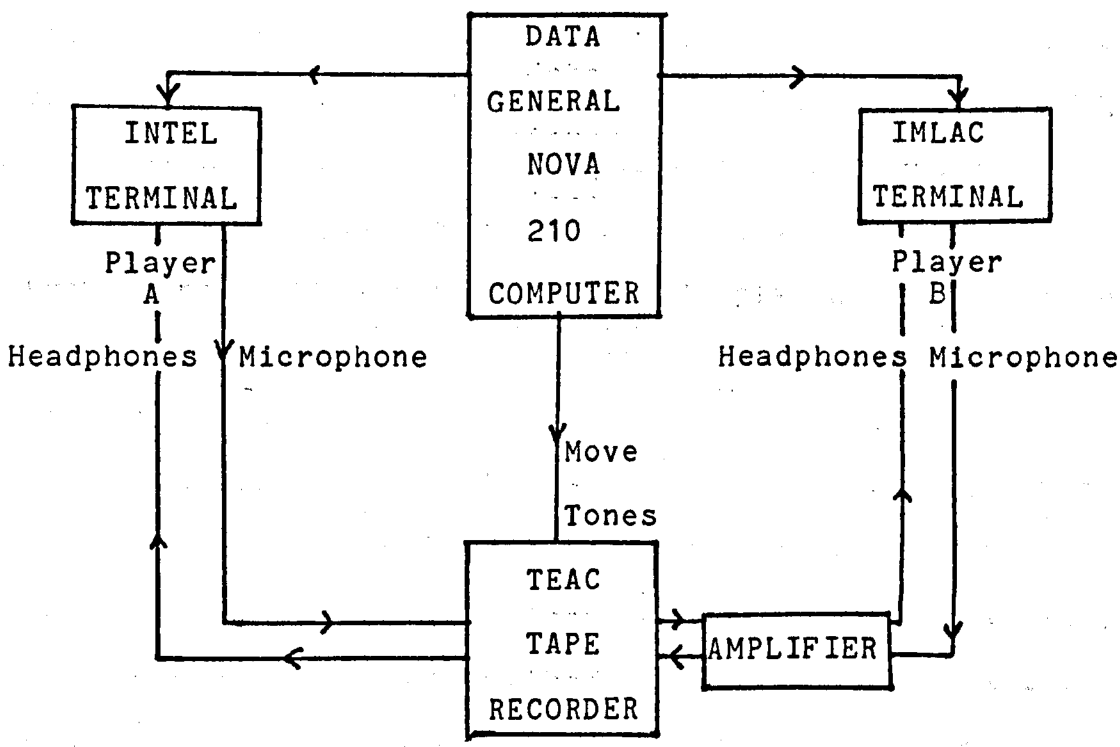


Figure 4.8

The experimental set-up

Each terminal was located in a different soundproofed room and subjects were seated in front of either terminal. Subjects communicated with each other through microphones and headphones, this system being linked up via a Ferrograph '20 + 20' amplifier to a TEAC A-3340 4-channel tape recorder. Each player's speech was recorded on a different track of the tape, while moves by the players were signalled on the other two tracks by means of a tone 'blip' generated by a pulse supplied by the Nova. The tone was of a different frequency for the two players. All dialogues were recorded from start to finish on the TEAC recorder, using EMI and BASF 6 inch reel tapes.

As noted earlier, the terminals displayed an indication of which player's turn it was to move by means of that player's position marker (X) flashing on and off. A tally was kept of the number of moves of both players and of the total number of penalty points incurred up to that point. This tally was also displayed on the screen.

Players made their moves by pressing keys located on the keyboards of the terminals, these keys being appropriately marked with arrows indicating the direction of the chosen move relative to the orientation of the display on the screen, that is, ↑ is up, → is right, etc. If a button was pressed where no movement was possible, i.e. when the subject attempted to move where there was no path whatsoever (neither open nor gated), the subject was free to move again. All moves and times of moves were recorded by the Nova and subsequently printed out.

The maze game program was written in the appropriate machine codes by Mr. J. T. Mullin, computer manager in the department of psychology of the University of Glasgow.*

b) <u>The study of task-oriented dialogues using the maze game task</u>

As noted in section (ii) above, the dialogues produced when using the task are usually several minutes in length each, and contain location descriptions, planning of moves, and so on. As such, a variety of facets of spontaneous task-oriented dialogue can be studied using this type of task. In the present work, a detailed analysis of the semantics and pragmatics of the location descriptions generated by subjects in the course of playing the game is undertaken. As such, only games where players could not see their partner's position were analyzed here, since such 'unseen partner' condition games force subjects to generate location descriptions. Where his partner's position is visible to a subject, description of that position by his partner is superfluous. Thus, the option in the program allowing subjects to see their partner's location was not employed in the games used in the present work.

The general advantage of studying task-oriented

---

* See also Appendix VI for further details of the maze program.

dialogues hinge on the fact that, when subjects are busy attempting to solve a task, they are not self-conscious about the dialogue they produce, and their speech has a spontaneous and natural quality. As Grosz (1977) puts it, "Task-oriented dialogues are a good source of unbiased data on discourse. Concentration on the performance of a task keeps the participants from becoming self-conscious about their language. The resulting dialogues are spontaneous and unrehearsed". (Grosz, 1977, p. 11). In social psychological terms, subjects are in the state of awareness termed "awareness-of-self-as-subject" (or agent, or problem solver) rather than being in the state of awareness known as "awareness-of-self-as-object" (i.e. as an object being studied by someone else).

Thus, although the topic of discourse is constrained, it can be argued that the data produced is quite natural due to the spontaneous nature of its production. It should be emphasized that, although the topic of discourse is constrained, the subjects' choice of phrase is entirely unconstrained. This is true of dialogue in general; the topic of discourse constrains what can be said, but <u>how</u> it is said is entirely up to the participants.

A casual inspection of the transcripts suggests that, as was the case with the various studies employing Blakar's 'communication conflict' task, subjects in the maze game experiments become quite task-involved. This task-involvement further enhances the spontaneity of the dialogue produced. Furthermore, post-experimental comments by subjects indicates that they were often quite highly motivated to perform the task successfully. All of these factors suggest that the dialogues produced were natural rather than artificial in nature.

Subjects were recruited from the first- and second-year (undergraduate) classes of psychology students in the University of Glasgow, and, in some cases, from among students found playing computer games in the University of Glasgow refectory at lunchtime. This selection of subjects ensured that the subjects involved were familiar with computers in general and computer games in particular, ensuring that the task was not an unnatural one for the subjects. Pairs of dyads were recruited at a time such that subjects within a dyad were already acquainted with one another. This obviated any difficulties which might have arisen for subjects playing a complex computer game with a person unknown to them. All subjects were paid for participating in the studies at the rate of £1.00 per hour.

It can be argued, therefore, that the dialogues produced by subjects were of a genuinely spontaneous and natural type.

Only a brief summary of the variables manipulated in the various studies will be reported here, since greater detail concerning the studies undertaken and the results obtained are given in the following chapter.

It is in fact possible to manipulate three types of variable, namely; i) the type of game played by the subjects (i.e. whether the game involved the 'monster' or not), ii) the nature of the maze shape on which subjects played the game - the structural configuration of the paths could be varied in any way desired to produce symmetrically- or asymmetrically-shaped mazes (see appendix I), and iii) subject variables (i.e. deliberately assigning subjects to dyads on the basis of how they played the game previously).

Four studies were in fact run. Experiment 1 was a 'baseline' study in which subjects played games on mazes which were symmetrically shaped (see appendix I). Experiment 2 involved the subjects playing on the same symmetrically-shaped mazes as were used in experiment 1 but in this case the 'monster' was present in the maze.

Experiment 3 involved a manipulation of the maze shape variable:subjects played games on vertically symmetrical and markedly asymmetrical pairs of mazes (see appendix I; see also next chapter). Experiment 4 involved manipulating subject variables. In this experiment, subjects were paired up with one another to form dyads on the basis of their previously having used different ways of describing maze locations.

Experiment 1 acted as baseline study, against which the results from experiments 2, 3 and 4 were compared.

The foregoing serves to introduce the reader to the task employed in the present work and its properties. The present task does share certain of its features with previous studies which have used a similar methodological strategy. The maze game task is, however, unique and has not been employed in studies previous to those reported here.

The next, empirical, chapters will present analyses and results from the above studies using the maze game task, and the relation of these results to the theories reviewed in part I of this thesis will be an interesting matter which will be taken up in the later chapters of the present volume.

Part II

The Semantics and Pragmatics of Spatial Location Description in Spontaneous Dialogue:

The Empirical Studies

CHAPTER 5

PRELIMINARY ANALYSES: THE EFFICIENCY OF PROBLEM SOLVING IN THE MAZE GAME

a) Introduction

In the last chapter, a number of speech production tasks were reviewed, and an exposition was given of the task employed in the work reported in this volume. The speech production tasks antecedent to the present maze game were reviewed in order to indicate the range and variety of issues which can be addressed using such tasks. As noted in the last chapter, all these tasks have the common advantage that the language produced by the subjects in performing the tasks is produced <u>for a purpose</u>, thereby alleviating the problem of the artificiality of both materials and data which some language understanding experiments run the risk of.

A further advantage which was noted previously is that the task-involvement of subjects helps to ensure a true spontaneity in the language produced. In addition to these twin advantages (of speech being produced for a purpose, and of subjects being task-involved) which ensure genuine spontaneity and naturalness, the maze game task used in the work reported here possesses one further advantage: there is an independent record of what is being talked about by the subjects in the form of a computer printout of successive game states.

This makes possible an analysis of the semantics and of some of the pragmatics of spatial location description, since we can relate any spatial location description to a particular game state. In this respect, the present task differs from most of its predecessors which were reviewed in the last chapter: only Levelt possessed a similar independent record of what was being described in his experiments, as well as constraining tightly the topic of discourse. Levelt's task, however, involved monologues as opposed to the present dialogue task. Thus, the present task permits an examination of a range of pragmatic issues, since we can examine whether particular dialogue partners constrain the other participant's choice of phrase, for example.

The major focus of the work reported here will be a detailed analysis of the location descriptions generated by subjects in the course of playing the game. In the next chapter, sample data will be given to exemplify the nature of the location description task given to the subjects. The next chapter will also summarize the statistical analysis of the data, and the results obtained will be presented in chapter 7.

In the meantime, however, some preliminary data analyses are presented in terms of the efficiency with which subjects solve the game task in the different experimental conditions. Since the primary interest lies in the location description data, this preliminary exposition will be brief.

b) Preliminary analyses: the efficiency of problem solving in the various types of game.

As alluded to in the previous chapter, a total of four studies were conducted:

1) a baseline study using symmetrical mazes,
2) a study involving the 'maze monster',
3) a study utilizing specially-constructed (asymmetrical and 'column biased') mazes, and
4) a reassignment study.

Details of the studies, hypotheses and variables manipulated in the different comparisons are given below. The measures employed to summarize the efficiency of subjects at solving the game task under these different conditions were:

a) The number of moves required by players in order to achieve solution.
b) The amount of real time required by players in order to achieve solution.

Data from these two measures are included in the data summary table (table 5-1, below). However, statistical comparisons of different conditions employing these two measures were not carried out. The reason for this is that, although, for example, monster games and standard games employed identical mazes and similar start and goal positions (see appendix II (b) and (d)), the mere presence of the monster, in addition to complicating the game strategically, may alter the optimal solution to this type of game. This would render the two types of game incomparable on this particular measure. The same is also true for analyses of the amount of real time taken by subjects to achieve a solution in the task. If the number of moves required for an optimal solution is changed by the presence of the monster, then the amount of time required until solution is achieved is also likely to change. (We would expect these two measures to be correlated to some extent, with games requiring more moves also taking subjects longer).

The non-comparability of different game types on these measures is even more marked if we consider the comparison between baseline games and games played on specially-constructed mazes. The specially-constructed mazes contain fewer paths, allowing subjects fewer options in their moves. This would probably result in the optimal solution to these games on specially-constructed mazes being different to that in games using symmetrical mazes, and again the two types of game will not be comparable on the 'number of moves' and 'time taken' measures.

For these reasons, summary statistics on these two measures will be presented for the different game types, but no statistical comparisons will be made between different game types (for example, in the comparison between monster games and baseline games).

However, study 2, the baseline study, involved two unseen-partner games on symmetrical mazes, and this allows an examination for a possible effect of practice on subjects' efficiency at solving the task to be undertaken. In this case, the two game types being compared (earlier-played games as opposed to later-played games) involve the same mazes and similar problems being set to subjects (i.e. similar start and goal positions, similar distances to be moved through the maze structure, and so on) and therefore they _are_ comparable on the number of moves/time taken to solution measures.

The other two measures employed to assess efficiency at problem-solving circumvent the above-mentioned problem of different game types not having comparable optimal solutions. These measures avoid this problem because they are measures of _rate_, and are not absolute measures. They are:

c) The rate of moving (the number of moves made by both players per minute).
d) The mean number of utterances spoken by both subjects per move.

The rate of moving measure controls for differing numbers of moves in different game types by measuring how much time is required by subjects, on average, between successive moves. It could be expected that this measure will correlate with the amount of difficulty experienced by subjects in the different game types: games which are more difficult for subjects (whether for strategic or other reasons) would require more careful planning of moves, and this would result in subjects spending more time, on average, planning each move. For example, it could be expected that monster games would involve a slower rate of moving than would baseline games, since subjects would spend greater amounts of time planning each move, taking greater care in order to evade the hazard presented by the monster. The presence of the monster would also constrain freedom of movement within the maze, since moving along some of the paths might be inadvisable in that it would take a subject's position marker closer to the monster, even though these paths were free to be moved along. Thus, subjects in baseline games would effectively have greater freedom of movement within the maze, and have less complicated planning, and thus would be expected to move at a quicker rate.

The other measure, d), the mean number of utterances spoken by both subjects between moves, reflects the average amount of speech generated by subjects between successive moves and again this obviates the comparability problem by measuring the average number of utterances between successive moves (and is therefore effectively a 'rate' measure). Absolute values of moves or utterances are not involved. Again, it would be expected that, in the case of games where subjects are experiencing greater difficulty (whether strategic or communicative difficulty), more utterances would be required on average between successive moves. This greater number of utterances between moves might be needed to formulate more complicated and lengthier plans, for example.

Statistical comparisons between different game types were made on these two measures (rate of moving and mean number of utterances between successive moves - see results section below).

The details of the four studies are set out below.

Study 1. Baseline study

In this study, sixteen subjects (eight males and eight females, comprising four pairs of males and four pairs of females) took part in an exploratory study. Subjects firstly played a practice game on a small, symmetrical maze based on a 4 x 4 matrix of nodes (see Appendix I for diagrams of all the maze shapes used). This was followed by four games, one on each of four different larger mazes based on 6 x 6 matrices of nodes (mazes 5B, 6B, 7A and 8A: these mazes were horizontally symmetrical, that is, symmetrical about a vertical line passing through the centre of the maze. See Appendix I). Two of the games involved a 'seen partner' condition, and two involved an 'unseen partner' condition. A counterbalanced design was used to balance the orders in which the different mazes were used. One of each of the seen- and unseen-partner games involved subjects traversing symmetrical distances of identical length in the same direction (e.g. from the bottom part to the top part of the maze) and the other games involved subjects traversing symmetrical distances of identical length but in opposite directions (one subject proceeding from the bottom part of the maze to the top part, and the other proceeding from the top to the bottom). Appendix II (b) and (c) illustrates these two types of game. Again, their order of occurrence was counterbalanced.

Only the unseen-partner games have been analyzed here. The data from study 1 constitute a 'baseline' against which data from the other studies were compared in an 'independent samples' design.

Since subjects in study 1 played two unseen-partner games, an analysis for the possible effects of practice on efficiency at solving the task was possible by comparing data (number of moves, time taken, rate of moving, and the mean number of utterances between moves) from earlier games with those from later games. This repeated-measures comparison would allow us to ascertain whether or not subsequent games were solved more efficiently than were earlier games.

Study 2. Manipulation of game type (Monster game vs. baseline game).

In study 2, ten pairs of subjects (four pairs of males and six pairs of females) took part in a total of eleven practice games on the small practice maze (maze 1A, as in study 1), three in a seen-partner condition and eight in an unseen-partner condition. These practice games were recorded but not analyzed.

These were followed by a total of twelve unseen-partner games based on the same larger, horizontally symmetrical mazes as were used in study 1 (mazes 5B, 6B, 7A, and 8A) but in which the monster was involved. Only the twelve non-practice games involving the larger mazes and the monster are analyzed here. In all cases, subjects were set the task of proceeding from the bottom end of the maze to the top end of the maze (starting and goal positions being symmetrically placed, see appendix II (d), and the monster's initial position was at the top of the maze, i.e. the opposite end of the maze from the region in which the players' starting positions were located. Start and goal positions for players were similar to those involved in study 1. The data from study 2 were compared with those yielded in study 1, constituting an 'independent samples' design.

It was anticipated that the presence of the monster would complicate the game for the players from the point of view of solving the task, for the reasons outlined earlier. When the monster is present, players not only have to cooperate with each other, make plans, develop strategies and so on, but they also have to avoid occupying the same node as the monster. Furthermore, the monster moves once for every three moves made by both players and actively pursues one player at at time.

This strategic complication, it could be hypothesized, would force the subjects to consider more carefully their patterns of moving in order to avoid being 'eaten'. Thus, we would anticipate that subjects in 'monster' games would move at a significantly slower rate and require significantly more utterances per move on average than would subjects in baseline games: this would reflect the greater care required in planning moves in the monster games.

From the point of view of the location description aspect of the subjects' task, the monster constitutes a moving reference point (visible to both players at identical locations within their respective maze displays) relative to which location descriptions can be generated. This feature was sometimes used by subjects playing 'monster' games, players' position marker and goal positions being described in terms of distances (in terms of numbers of nodes in the pathway) between the location being described and the monster's position.

Study 3. Manipulation of maze shapes

In this study, eight pairs of subjects (three pairs of males and five pairs of females) took part in a total of fourteen games using mazes 9B (the 'column biased' maze) and 10A (the asymmetrical maze: see appendix I), both of which are based on the 6 x 6 matrix of nodes. All games were in the unseen-partner condition and all were analyzed. Data from this study were also compared with data from study 1, constituting an 'independent samples' design.

The principal interest in the comparison between data from study 1 (baseline study) and study 3 (using asymmetrical and 'column biased' mazes) lies in the manipulation of maze shape and its possible effect upon the nature of the location descriptions generated. Mazes 9A and 10B were constructed in a manner such that they would contrast with the symmetrical mazes 5B, 6B, 7A and 8A used in study 1. Hence, maze 10B was constructed such that it possessed a markedly asymmetrical shape in both planes of orientation. The 'column biased' maze (maze 9A) was constructed such that it retained horizontal symmetry, but with the continuity of horizontal lines destroyed, resulting in a maze shape featuring complete columns and incomplete rows, hence the term 'column biased' used to describe it.

As indicated above, therefore, study 3 was designed to manipulate maze shape, and this enables us to assess whether this manipulation affects the nature of the location descriptions generated. The result of this aspect of study 3 will be given in a later chapter.

It is possible that the specially-constructed mazes used in study 3 would, since they contain fewer paths than the symmetrical mazes used in study 1, allow fewer options of movement within the maze and thereby present greater difficulty for subjects than do the baseline games. The greater difficulty inherent in games which allow fewer options of movement would arise because fewer options of movement would allow fewer possible alternative routes towards goal for each of the players. When a player's path towards goal is blocked by barriers, he might be unable to take an alternative route towards goal and instead would be forced to plan with his partner how to remove the barriers from his path. This planning can be quite complicated, involving the blocked player taking penalty moves until his partner can get to one of the blocked player's switch gates. Furthermore, the fact that there are fewer paths available increases the possibility of one or other of the players being unable to move at all on some occasions. This might occur because the few available pathways radiating from the node occupied by him all contain barriers simultaneously. (For example, his partner may have unwittingly moved into a switch gate and hemmed the player in). In such cases, he must plan with his partner how to effect a switching of his barriers and take penalty moves in the meantime. Games involving fewer options of movement are thus more difficult for subjects, since the possibility of one of the players being completely blocked arises, and quite complicated planning and sequences of moves may be required to circumvent this problem.

Thus, an independent samples comparison on the two 'efficiency' measures (rate of moving and mean number of utterances per move) was drawn between data from study 3 and data from study 1, with the hypothesis that subjects in study 3 would show a slower rate of moving and a greater number of utterances per move being postulated.

Study 4. Manipulation of subject variables

In study 4, subjects who had previously taken part in either study 1 or study 2 were contacted and each allocated a new partner with which to play the maze game, such that each member of a dyad was paired with a subject known to him or her but with whom he or she had not played th game previously. One pair of male subjects and two pairs of female subjects played a total of six unseen-partner condition games, each pair playing two games. The first-played game was played on one of the symmetrical mazes used in study 1 (mazes 5B/6B/7A/8A), and this was followed by a game on the asymmetrical maze (maze 10A). Again, data from study 4 used the data from study 1 as a control comparison, constituting an 'independent samples' design. As with the comparison between study 3 and study 1, the main interest in this manipulation of subject variables was in the possible effect upon the type of location descriptions generated.

Subjects had been paired together on the basis of their having used different types of location descriptions, and the main interest of study 4 was an attempt to ascertain how subjects would overcome this initial difference. An exposition of the different description types found in the data is given in the next chapter, and an examination of the effect of the manipulation of subject variables upon the descriptions generated will be undertaken in a subsequent chapter.

It is possible that such a pairing of individuals who had previously used different description types could affect subjects' efficiency of performance in the task adversely, and so an analysis in terms of rate of moving and mean number of utterances per move was undertaken, comparing study 4 data with data from study 1.

In all cases, the rules of the game were explained to subjects beforehand and subjects were instructed to attempt to complete the game in as few moves as possible, in as little time as possible, and incurring as few penalty points as possible. In all, 48 games were analyzed across the four studies.

Exclusions

Some games were excluded from the efficiency analyses for four main reasons:

(a) In some games, both of the subjects failed to understand the instructions properly at first, particularly, failing to understand switch box functioning. As a result, they moved around in an uncoordinated fashion, thereby rendering the games in which they participated spuriously long. Three games out of 48 were excluded on this basis.

(b) On two occasions, one of the subjects playing a 'monster' game was 'eaten' by the monster in an extremely short time span, resulting in two games which were spuriously short. The two games excluded from the analysis on this basis were only 6 and 12 moves in length, corresponding to 1.35 and 5.1 minutes of dialogue, respectively.

(c) There were cases where the tape on the tape-recorder had to be changed during the course of a game, thus spuriously lengthening the time taken by the players, but not affecting the number of moves taken. Accordingly, such games (4 in all, 3 being from study 1 and 1 being from study 2) were excluded from the 'rate of moving' comparisons, but not excluded from the 'mean number of utterances per move' comparison. In such cases, the subjects were not allowed to discuss the game during the period in which the tape was being changed.

(d) There were cases where players were set an insoluble task due to experimenter error. This occurred on two occasions with games from study 3, involving asymmetrical mazes.

In all the tables of data in this chapter, the numbers of subjects (N) shown in each table is the number of subjects <u>after</u> exclusions had been made.

c) <u>Results</u>

In drawing the various statistical comparisons (between data from first-played as opposed to data from later-played games in study 1, and between data from studies 2, 3, and 4 as opposed to data from study 1), appropriate non-parametric statistical tests were used. This was because variances in the measures employed were not equal in the various conditions, and this violates one of the assumptions on which the use of parametric statistical tests are based. The Mann-Whitney U test for independent samples was thus used for comparisons between studies (i.e. when data from studies 2, 3 and 4 were being compared with data from study 1) and the Wilcoxon matched-pairs signed-ranks test was used for comparisons involving repeated measures (that is, in the comparison between data from first-played as opposed to later-played games in study 1).

The first comparison of interest is the comparison between first-played games and later-played games in study 1. This comparison allows an assessment of the possible effects of practise on subjects' efficiency at solving the game task. The relevant summary statistics are shown in table 5-1, below:

Efficiency Measure	First-played Game $\bar{X}$	S.D	Later $\bar{X}$	Game S.D	N
1. Number of moves to game solution	36.57	8.25	29.33	10.307	7
2. Time taken to game solution (mins)	9.8198	3.373	8.05	3.4289	6
3. Rate of moving (Number of moves per min.)	3.6824	0.4162	3.243	0.8013	6
4. Mean number of utterances by both players between successive moves.	6.81	1.212	8.377	3.33	7

Table 5-1: Subjects' efficiency at solving the game task in first-played as opposed to later-played games in study 1, comparing the two types of game on each of four measures of problem solving efficiency.

Examination of the data table reveals that the mean number of moves and the mean time taken to solution both differ between the two types of game in a direction consistent with a practice effect. That is, the mean number of moves required to solution is lower in the later game (the mean being 29.33 moves to solution) than in the first-played game (the mean in this case being 36.57 moves to solution). Likewise, the amount of time required before a solution is reached is slightly lower in the later game (with a mean of 8.05 minutes) than in the case of the first-played game (with a mean of 9.82 minutes). There is a slightly <u>slower</u> mean rate of moving in later-played games (3.24 moves per minute) as compared to first-played games (which have a mean of 3.68 moves per minute), but the difference is slight. Lastly, there is a slightly <u>higher</u> mean number of utterances per move required in the later played games (a mean of 8.38 utterances per move) as compared to earlier-played games (with a mean of 6.81 utterances per move). Both of these later differences are in a direction opposite to that which one might expect on the basis of a hypothesized practice effect; one might expect subjects to require fewer utterances on average between moves in later games. However, the differences are small ones.

The significance of these differences was tested using a Wilcoxon matched-pairs signed-ranks test. This demonstrated that the number of moves required for a solution to be achieved by subjects was significantly lower in the later game as compared to the earlier game ($T = 1$, $N = 7$, critical $T = 2$, $p < 0.025$ (1 tail)). The difference in time taken to solution, however, was not significant ($N = 6$, $T = 6$, critical $T = 0$, NS). The difference in rate of moving proved also to be non-significant ($N = 6$, $T = 6$, critical $T = 0$, NS), as did the difference in mean number of utterances per move ($T = 15$, $N = 7$, critical $T = 2$, NS).

The only significant difference between first-played and later games is, therefore, in terms of the number of moves required to solution, with later games requiring significantly fewer moves before a solution is reached. This suggests that subjects are becoming more efficient in problem-solving terms with increasing practice at playing the game. They do not, however, appear to be becoming verbally more efficient: they are in fact speaking slightly, but not significantly, more between successive moves on average, and this could reflect the greater planning and strategic care required in order that a significant reduction can be made in the number of moves required to solution.

That is, players would appear to be learning how to coordinate their moving, learning which strategies of moving are more efficient in problem solving terms (for example, learning that sometimes the deliberate making of illegal moves against barriers - see chapter 4 - is strategically advantageous rather than the converse), and generally improving in the problem-solving aspects of the task.

The next comparisons of interest are those between the data from the different game types ('monster' games _versus_ baseline games, games on specially constructed mazes _versus_ baseline games and reassignment study games _versus_ baseline games). In all cases, the comparison is between the later studies' games and the _first-played_ games of study 1, since a practice effect has been shown to be in operation, and later-played games from study 1, may not, therefore, be strictly comparable to the games from studies 2 and 3. That is, it was felt desirable that the games from study 1 which were being compared with the monster games should be as equivalent as possible in terms of the subjects' experience and amount of practice at playing the game. Since the monster games were played after one practice game, it was decided to compare the data from the monster games with the data from the earlier-played games in study 1, which were also played following one practice game.

Efficiency Measure	Study 1 $\bar{X}$	S.D.	Study 2 $\bar{X}$	S.D.	Study 3 $\bar{X}$	S.D.	Study 4 $\bar{X}$	S.D.
1. Number of moves to game solution	36.57	8.25	42.0	11.07	34.7	12.64	31.16	5.8
	N = 7		N = 10		N = 10		N = 6	
2. Time taken to game solution	9.82	3.37	17.69	5.43	14.74	7.39	12.15	3.9
	N = 6		N = 9		N = 10		N = 6	
3. Rate of moving (No. of Moves/ Min).	3.68	0.4	2.40	0.57	2.45	0.63	2.63	0.6
	N = 6		N = 9		N = 10		N = 6	
4. Mean No. of Utterances between moves	6.81	1.21	10.49	10.24	10.27	4.55	9.42	2.16
	N = 7		N = 10		N = 10		N = 6	

Table 5-2 Summary statistics of games showing means, standard deviations, and numbers of subjects in study 1 (baseline study) study 2 (monster games study) study 3 (specially constructed mazes study) and study 4 (reassignment study) on four measures of efficiency at solving the task.

Comparing data from study 2 ('monster' games) with data from study 1 (baseline games), the differences on measures 3) and 4) (see table) are in the directions predicted; monster games involve a slower rate of moving (a mean of 2.40 moves per minute as against a mean of 3.68 moves per minute in the baseline games) and a greater mean number of utterances between successive moves (a mean of 10.49 utterances between successive moves, as against a mean of 6.81 utterances between successive moves in the baseline games), as compared to the baseline games.

The differences in rate of moving and mean number of utterances per move both proved to be significant - i.e. monster games involved a significantly slower rate of moving (Mann-Whitney $U = 3$, $n_1 = 6$, $n_2 = 9$, $p < .01$ (1 tail)), and a significantly greater average number of utterances between moves ($U = 2$, $n_1 = 7$, $n_2 = 10$, $p < .001$ (1 tail)) than did the baseline games.

Both of these findings uphold the hypotheses outlined earlier. Monster games would appear to be strategically more complex and this is reflected in their requiring a greater number of utterances between successive moves on average. This greater mean number of utterances between moves significantly slows the rate of moving.

Comparing the data from study 3 (from games played on specially-constructed mazes) with data from study 1 (from baseline games), differences on the appropriate two efficiency measures (rate of moving and mean number of utterances between successive moves) also appear. The rate of moving is slower in the games played on specially constructed mazes (with a mean of 2.45 moves per minute as compared with a mean of 3.68 moves per minute in the baseline games). The mean number of utterances between successive moves is higher in the games played on specially constructed mazes (the mean being 10.27

utterances per move as compared with a mean of 6.81 utterances per move in the baseline games). However, only the difference in the rate of moving is a significant one. Games played on specially-constructed mazes involve a significantly slower rate of moving than do baseline games (Mann-Whitney $U = 4$, $n_1 = 6$, $n_2 = 10$, $p < .01$, (1 tail)). The higher mean number of utterances between moves in the games played on specially constructed mazes turned out to be non-significant (Mann-Whitney $U = 21$, $n_1 = 7$, $n_2 = 10$, N.S.). The size of the difference between the mean number of utterances per move in the two studies is of an order such that it could be expected to be significant. However, the variance in the case of the games played on specially-constructed mazes is very large, approximately triple that of the baseline games, and this is why the difference is a non-significant one.

The manipulation of maze shape did have the effect of slowing down subjects' rate of moving, as noted above. This could be because the games involving specially-constructed mazes afford subjects fewer options of moving than the baseline games, and are therefore more complex strategically in that they require more careful planning of moves.

Another factor which could contribute to the apparently greater difficulty presented by the specially constructed mazes is that, being of a less regular shape, they are more difficult to describe locations within.

Comparing data from study 4 (the reassignment study) with data from study 1, differences on the two efficiency measures of greatest interest are again apparent. Reassignment games show a slower rate of moving than do the baseline games (with mean values of 2.63 moves per minute and 3.68 moves per minute, respectively). Reassignment games also involve a higher mean number of utterances between moves as compared to the baseline games (9.42 utterances per move on average and 6.81 utterances per move on average, respectively). Both of these differences are significant. Thus, the rate of moving was significantly slower in reassignment games (Mann-Whitney $U = 5$, $n_1 = 6$, $n_2 = 7$, $p < 0.04$, (2 tail)), and the mean number of utterances between moves was significantly higher in the reassignment games, as compared to the baseline games (Mann-Whitney $U = 6$, $n_1 = 6$, $n_2 = 7$, $p < 0.017$ (2 tail)).

These last two findings are of the greatest interest in the present context. Here we have evidence that manipulation of subject variables (that is, directly manipulating pragmatic factors) results in a decline in subject's performance in the game task. Specifically, pairing subjects who had previously employed differing types of location description had a significant effect in that their efficiency as a dyad in solving the task was reduced; subjects required significantly more time to make moves, and they made significantly more utterances between moves.

It is possible that, since each pair of subjects in the reassignment study played one game on a symmetrical maze and one game on an asymmetrical maze, the significant difference between reassignment- and baseline games (in terms of rate of moving and mean number of utterances between successive moves) could be due entirely to the games played on the asymmetrical mazes being more difficult. A comparison was therefore drawn between games played on symmetrical mazes in the reassignment study and in the baseline study on both measures. In this comparison, the games (reassignment and baseline) are played on the same mazes, and the result is unequivocal; reassignment gamess involving symmetrical mazes require significantly more utterances between moves on average than do baseline games played on the same mazes (Mann-Whitney $U = 1$, $n_1 = 3$, $n_2 = 7$, $p < 0.017$ (2 tail)). The rate of moving in the reassignment games played on symmetrical mazes is significantly lower than baseline games played on the same mazes (Mann-Whitney $u = 1$, $n_1 = 3$, $n_2 = 7$, $p < 0.017$ (2 tail)).

Thus, the significant difference between reassignment and baseline games in terms of the rate of moving and the mean number of utterances between successive moves cannot be explained in terms of some of the reassignment games being played on the asymmetrical maze. The same significant difference holds true for comparisons between rassignment and baseline games played on the same symmetrical mazes. This supports the interpretation that the subjects' increased difficulty in the reassignment games is not due to the type of maze involved but is instead due to pragmatic factors.

d) <u>General discussion</u>

As was expected, the results from study 1 indicate that there is a practice effect in operation such that subjects tend to solve later games in significantly fewer moves, without significantly lengthening the total time taken, or altering the rate of moving or the mean number of utterances required per move. Furthermore, it was noted that 'monster' games are significantly more difficult for subjects than are baseline games, requiring significantly more utterances between moves and involving a significantly slower rate of moving. Games played on specially-constructed mazes also proved significantly more difficult for subjects in that the subjects' rate of moving was significantly slower; however, the mean number of utterances between successive moves, although higher in the 'specially-constructed mazes' games, was not significantly so. Both of these results were explained in terms of the monster games and the games played on specially constructed mazes being more difficult in problem-solving terms than were the baseline games.

The interesting findings were those which indicated that the reassignment manipulation had the effect of slowing down subjects' rate of moving and resulting in subjects using significantly more utterances between moves. This indicates that reassignment games were significantly more difficult for subjects, and occurs despite the fact that there is an effect of practice on subjects' performance of the task. The effect of practice is to significantly reduce the number of moves required to game solution, and this occurs with no significant change in the rate of moving or the mean number of utterances spoken between moves. The present finding that the reassignment games are significantly more diffiult for subjects than the baseline games is counterintuitive in that, if anything, we would expect subjects in the reassignment games (and therefore with more practice) to play the game at least as well as subjects in the baseline games, and one would certainly not expect them to deteriorate significantly in efficiency.

The fact that subjects in reassignment games show a significant deterioration on both efficiency measures indicates that manipulation of subject variables, that is, the bringing into play of pragmatic factors, has an important effect upon subjects' efficiency at solving the task.

The above results and conclusions, whilst dealing with only basic features of the data in the transcripts, raise many interesting questions. For example, the fact that pairing subjects up on the basis of their previously having used different description types significantly impairs their efficiency at solving the task, despite their having had more practice in the game, is very interesting. It suggests that the description of spatial location within the maze is indeed an important aspect of playing the game, important enough for the manipulation of it as an independent variable to significantly affect subjects' performance. Thus, despite the fact that, as noted earlier, description-giving is only used instrumentally, the findings from the reassignment study indicate that it is important to the successful playing of the game and a detailed analysis of the location descriptions may therefore raise interesting further questions. It can be argued, therefore, that by focussing primarily on the location description aspect of the transcripts, we are addressing ourselves to a significant problem.

The next chapter will begin by giving examples of different types of location description generated by subjects. This will immediately raise interesting points, some of which relate back to the theoretical exposition in part I of this thesis. Later in the next chapter, details will be given of the analyses employed in the attempt to describe statistically the different location description types and to relate them to one another. The following chapter will then present the results of this analysis in detail, and later in the present thesis, some of the issues raised above can be re-viewed in a new light - for example, why the reassignment games caused difficulty for the subjects.

In the meantime, an exposition of the different description types generated by subjects will illustrate the nature of the data and will itself raise further theoretically interesting questions.

CHAPTER 6

THE SEMANTICS AND PRAGMATICS OF SPATIAL LOCATION DESCRIPTION - I: INTRODUCTORY

a) Introduction

In the last chapter, preliminary analyses of subjects' efficiency in achieving a solution to the maze game task in different types of game were undertaken. The results demonstrated certain differences between the various types of game such that some experimental manipulations appeared to cause difficulty for the subjects, since subjects' rate of moving, for example, was significantly reduced in particular game types of game (for example, in the reassignment games as opposed to the 'baseline' games). These results raise interesting questions to which we shall return in a later chapter.

In this and in the next three chapters, the problem of the semantics of location description by subjects will be examined. This chapter will contain an introduction to the location description problem, and provide details of the analysis undertaken. This analysis was designed to produce a statistical description of the various ways in which subjects described their locations within the maze structure.

The descriptions generated by subjects were extracted from the 48 transcripts of data yielded in the various studies. Since each transcript contained several descriptions, this constitutes a substantial body of data. The semantic and pragmatic aspects of these spontaneously generated descriptions constitute an interesting set of problems for analysis, particularly in relation to some of the theories of meaning summarized in part I of this thesis (see chapters 1 - 3). It was decided to concentrate upon the meanings of the descriptions generated, and other problem-solving aspects of the transcripts were only analyzed insofar as they had any bearing on the description problem. The comparatively crude measures of problem-solving used, and the results from the analyses using these measures, were detailed in the previous chapter.

We now turn to an exposition of the location-description problem.

b) The location-description problem

As described already in chapter 4, the maze game requires, for its successful solution, that the subject describe their locations within the spatial array of the maze to one another from time to time during the course of playing the game. The subject therefore has the problem of uniquely identifying different locations within a complicated spatial array. An example is shown in figure 6-1, which depicts a typical starting position and goal position for a subject in the baseline study (study 1).

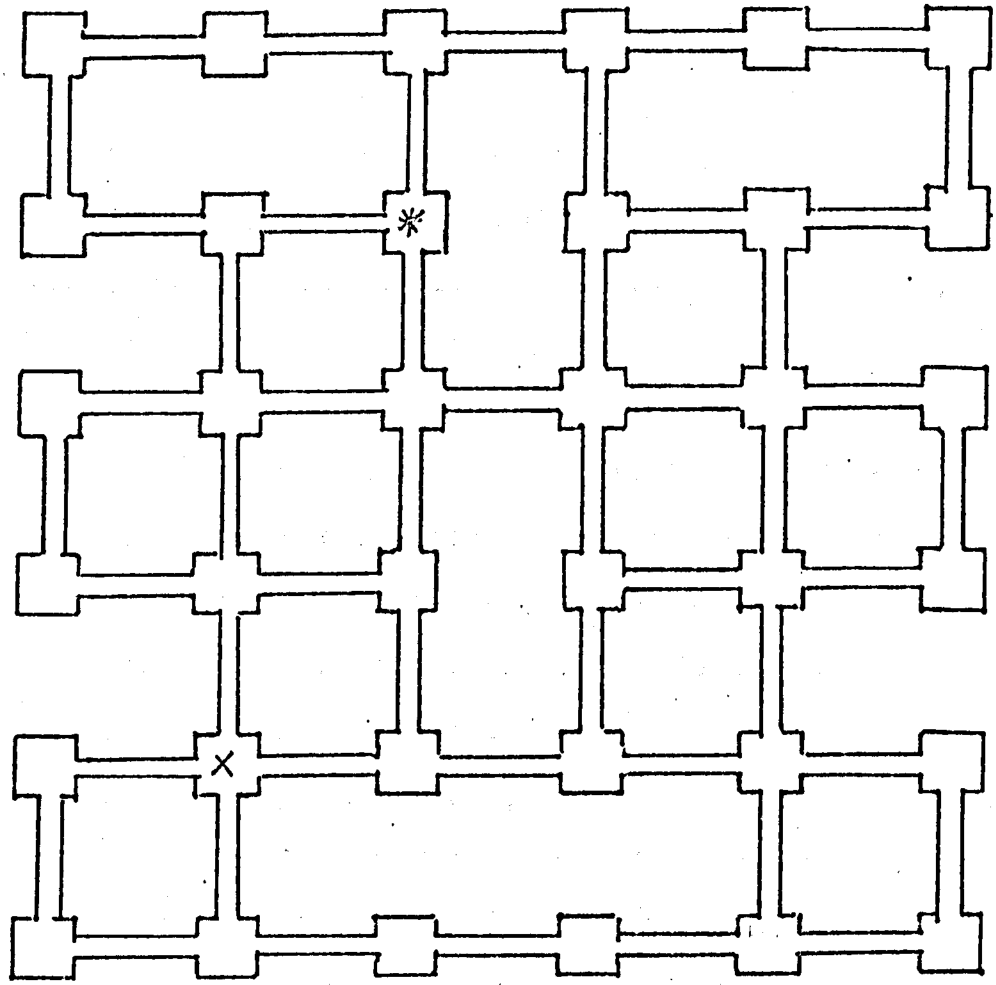


Figure 6-1: diagram showing typical description problem set to subject at the outset of the game. ('baseline' game)

Key: X = start position
* = goal position
(Switch node and barrier positions omitted for clarity)

The subject has the problem of uniquely identifying both of these locations.

Correct identification of one's location is crucial for the proper coordination of moves if barrier switching is to be effected appropriately. Therefore, the generation of unambiguous descriptions is an important aspect of playing the game, even though it serves only as a means to the end of effecting appropriate barrier switching. There are, in fact, several different strategies which a subject could use in order to achieve this identification goal, and a convenient introduction to the data would be in terms of several examples which illustrate the various types of description generated. Therefore, before dealing with the formal statistical description of the data, some sample descriptions will be considered.*

In the examples quoted below, details will be given of the type of game involved and the sex of the subjects. However, descriptions from all 48 transcripts were analyzed as one large data matrix (see parts d) - f) of this chapter, and also the following chapter).

A casual inspection of the transcripts suggests that there are four major types of description generated by

---

*N.B. All diagrams in this chapter will show only the basic framework of the maze shape with the subject's position, or the position being described, being marked in. Barriers and switch positions will be omitted in the interests of clarity. It should, however, be borne in mind that the positions of barriers and switches were always indicated on the subjects' mazes, making the maze shapes more complex as visual stimuli compared to the outlines shown in this chapter. Examples which show barriers and switch node configurations of different game types at the start of each game are set out in appendix II).

subjects. These four generic categories of description could themselves be further subdivided; however, the generic description types will be illustrated here, with differences in detail being discussed later. The four generic description types are as follows:-

i) <u>Path type descriptions</u>

In this type of description, as the name suggests, subjects describe their locations in terms of path movements from a specified start point. This is exemplified by the following extract from a transcript (transcript AALOG15C, standard, 'baseline' game, female dyad):

B I am: if you go to the top right corner,
and go down one and along two to the left,
that's where I am.

(The location being described is depicted in figure 6-2)*

In this example, player B describes the location by means of specifying a starting position (the top right corner of the maze), then gives details of path movements away from that starting position until the required point is reached. An interesting feature of this description is that it can be regarded as being <u>procedural</u> in nature; the first part of the description directs the listener to

* See Appendix III for a summary of symbols used in the transcripts of subjects' speech.

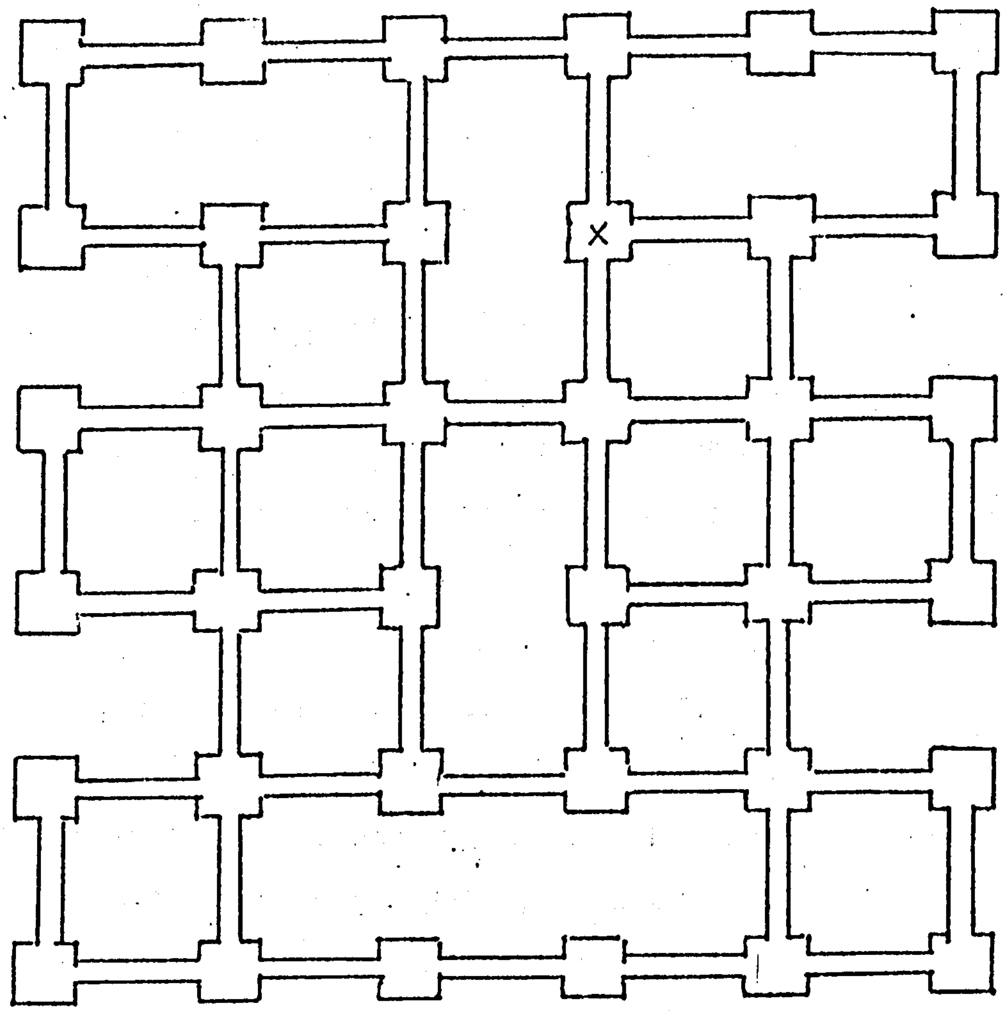


Figure 6-2: The point described by player B in transcript AALOG15C as .. . "go to the top right corner and go down one and along two to the left, that's where I am".

locate the starting point for the path description ("the top right corner"), and the second part of the description instructs the listener to scan from that starting position in specified directions (" . . . go down one and along two to the left . . ."). That is, we could consider the two parts of the description as constituting a set of instructions to the listener as to how to visually scan the maze shape; first, scan for a particular prominent point, then scan in specified directions from that point. This interpretation of the meaning of the descriptions ties in well with the theory of meaning known as procedural semantics, which was discussed in chapter 2.

The above example illustrates the basic principle behind the 'path' type of description. The speaker first locates a visually prominent point, such as a corner of the maze, and identifies it for his listener, and then locates the position to be described relative to that prominent point by means of specifying path movements along available paths from the start to the desired location.

Other sample descriptions from the same pair of female subjects illustrate further properties of path-type descriptions. For example, the following description again illustrates the general principle behind path-type descriptions (transcript AALOG15E, female dyad, 'standard' or 'baseline' game); see figure 6-3):

A Well that- where a-wait a minute, where are you?

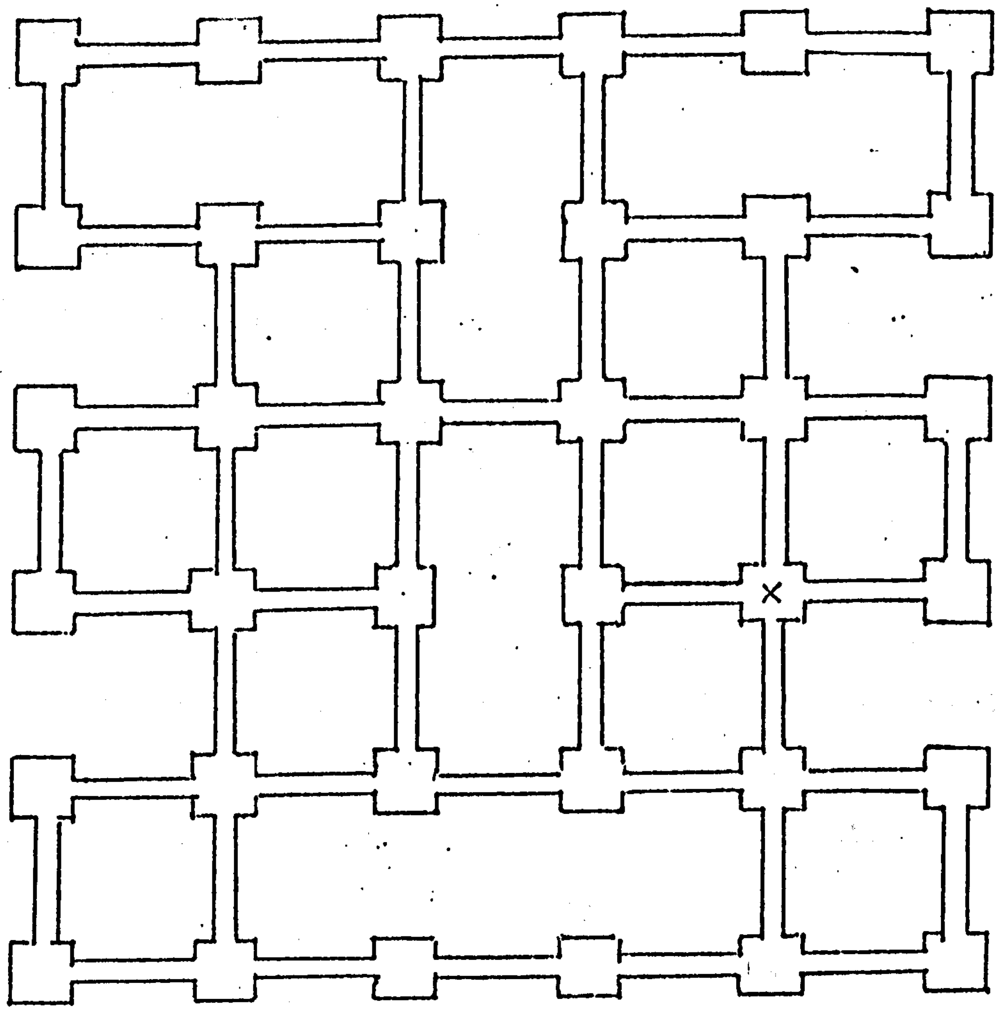


Figure 6-3: The point (X) described by player B in transcript AALOG15E as " . . . if you go to the bottom right corner, one along, two up".

B I'm: if you go to the bottom right corner
{one along, two up}

A {one along } two up.

The start point for the path movement is different in this example as compared to the last example. Notice also that the direction of horizontal movement is left implicit (it can only be to the left; it is not specified in this example, but was specified in the last one).

An interesting characteristic of the true path-type description is exemplified by the following goal position description, also from transcript AALOG15E, where subject B, having described her location (shown as an X on figure 6-4) is asked to describe her goal position (shown as an asterisk on figure 6-4). Her goal is in fact located directly above her, two nodes away, but with a gap intervening (see figure 6-4). Her description is as follows:

A And where are you wanting to go?

B Ehm: sort-I want to go two above myself,
but there's sort of - an empty space, I've
got to go a long way round.

Player B pauses initially; she then describes the location of her goal relative to her own location (which she has just described). There is, however, no directly available path between herself and her goal position. She describes her goal using the direct straight-line, two-node

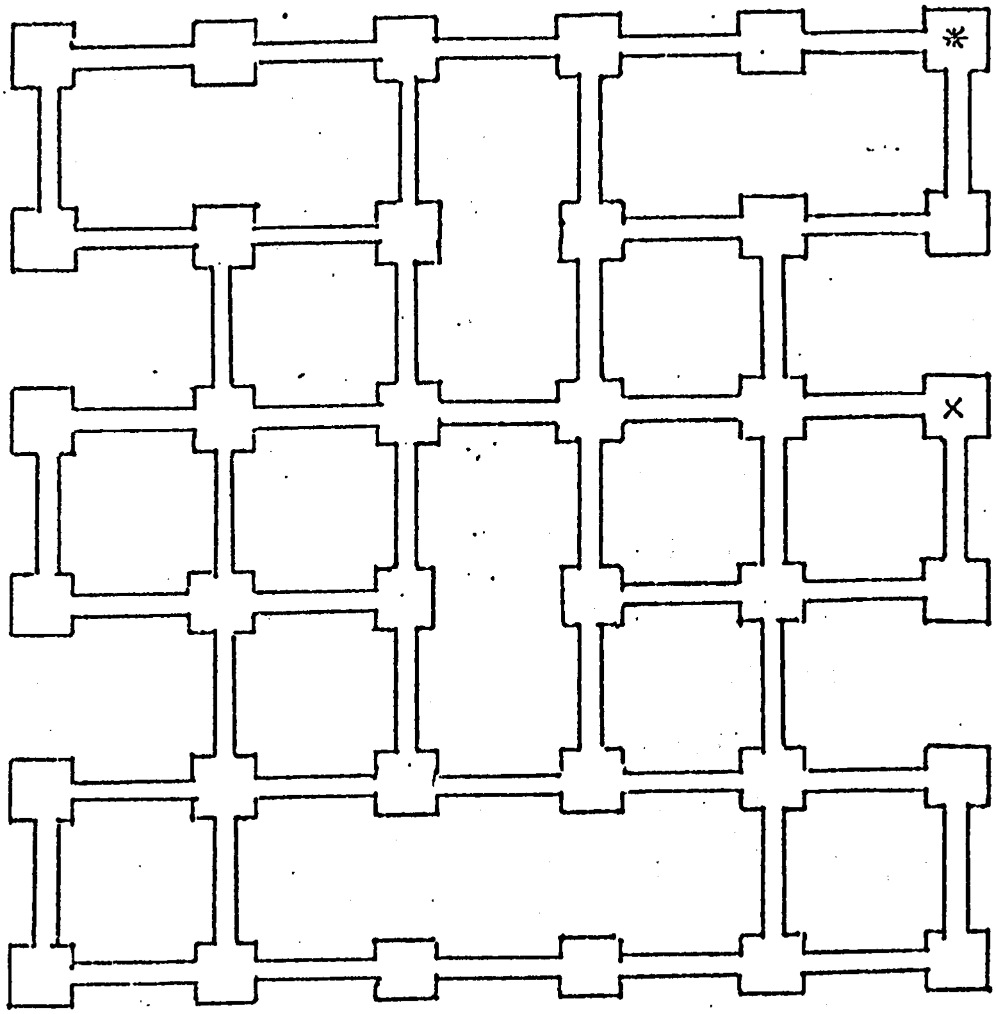


Figure 6-4: The goal position (*) described (by player B in game AALOG15E) relative to the position marker location (X) in the following manner:
" . . I want togo two above myself, but theres sort of: an empty space, I've got togo a long way round".

distance between herself and her goal, but acknowledges that there is a gap which has been mentally 'traversed' in her description. It is possible that the initial filled pause and short hesitation at the beginning of the description was due to her awareness of the gap and being reluctant to 'talk across it', so to speak; however, she decides on the shorter description which violates a true path-type description but which has the virtue of being less complex. This apparent unwillingness to 'talk across gaps' is a hallmark of the true path-type description, and it would be of interest to know whether pausing and self-correction are significantly more prevalent under conditions where a subject's preferred description type is inapplicable or violated in some way.

The apparent unwillingness to 'talk across gaps' is again illustrated by a description from transcript AALOG27B, a 'monster' game with female subjects; see figure 6-5 for an illustration of the position being described. Player A is asked to describe her position:

B Oh, wait a minute, where are you? Where are you?

A Right. See the bottom left hand corner?

B The bottom left.

A There's a box and then there's a gap:

B Uh-huh.

A And then there's a box: and then there's

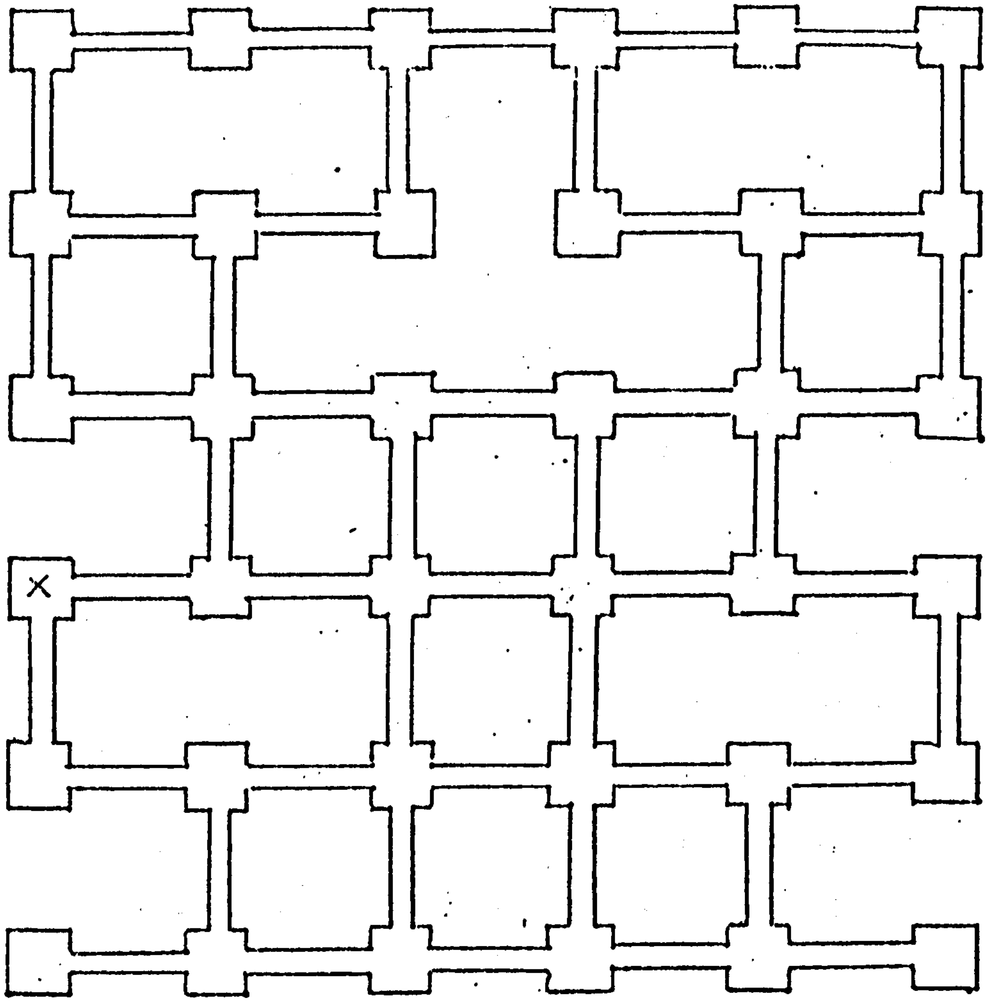


Figure 6-5: The point described by player A in transcript AALOG27B as "See the bottom left hand corner? . . . theres a box and then there's a gap . . . and then theres a box and then theres another box . . . right, I'm there".

another box:

B Uh-huh

A Right I'm there.

Once again, gaps are not 'talked across' without explicit acknowledgement of their existence. Also of note is the fact that the direction of scanning (upward) is left implicit.

In sum, path type descriptions are generated by first directing the listener to a suitable perceptually prominent point (usually a corner of the maze) and instructing them (explicitly or implicitly) to scan along available paths until the relevant point is reached. A corollary of this general strategy is that gaps (that is, spaces where no paths are available to be traversed) are explicitly acknowledged before being 'talked across'. This implies that path type descriptions are, to some extent, dependent upon the particular maze structure being described: this is in contrast to two of the other description types, which are next described in turn.

ii) Line type descriptions

The second major type of description generated by subjects could be called the line type of description. In this type of description, the maze shape is treated as if it were an array of parallel lines. A particular line is firstly identified, then the speaker describes a particular location on that line. The first example comes from a 'baseline' game with female subjects. The locations being decribed are the players' starting positions (A and B, respectively, for the appropriate players on fig. 6-6). The descriptions are as follows:

B Where are you for a start, where are you?

A Oh- oh yes. Second row.

B Second row.

A Second from the left.

B Second from the left, O.K.

A Second row going up the way, that is.

Three utterances later, player A asks player B to describe her (B's) location:

A Where are you-? (B interrupts - 2 utterances)

B Oh: (laughs) ehm: I am: on the top line:

A Uh-huh:

B One, two, third box from the right.

(AALOGI3A)

Both subjects describe their location by treating the maze as if it were an array of parallel horizontal lines (rows). Descriptions are generated by first specifying the relevant row containing the location to be described, and subsequently specifying the relevant place within that row. Once again, we can regard the meaning of this type of description as being procedural in nature; the

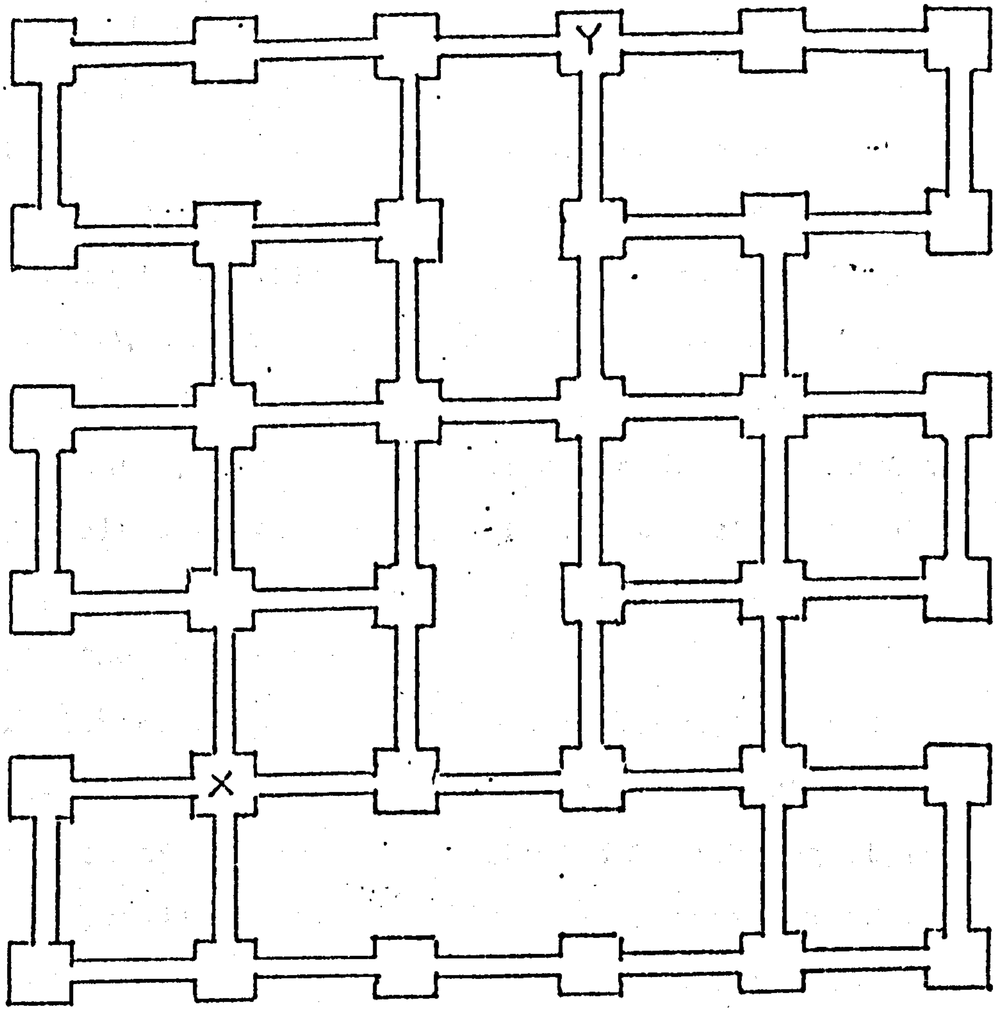


Figure 6-6: The point (X) described by player A in AALOG13A as "Second row, second from the left", and the point (Y) described by player B as ". . . on the top line. . . one, two, third box from the right".

listener, firstly, has to find the appropriate row, then to find a (given) point on that row. Player B is heard to count node positions aloud when generating the second description. We can imagine the listener executing scanning procedures firstly to scan for the correct row from an implicit or explicit start point, and then to scan along that row, using a node-counting procedure to 'mark' the different nodes, from an implicit or explicit start point until the appropriate node is reached. Player A in the example cited above initially leaves the start point for the 'row-finding procedure' implicit but subsequently supplies it for her listener.

It is interesting to note how the descriptive and interpretive task is simplified by using this sort of simplification of the maze shape (that is, treating it as an array of parallel 'rows'). This simplification of the maze shape in row-type descriptions is in contrast to the 'path' type descriptions outlined earlier.

Whilst 'row' type descriptions are used in study 1 and study 2 by subjects, the 'column biased' maze was deliberately designed to contain incomplete rows and complete columns and thus encourage the generation of line-type descriptions using columns rather than rows. An example where subjects did adopt this strategy comes from transcript AALOG41B, where subjects were playing the game on the 'column biased' maze. The (female) subjects are describing their initial starting positions and goals (see figure 6-7 (i)).

A Where- where are you?

B Ehm: (I sec) Oh: oh hell (laughs)

A Do it by vertical lines:

B Right: I'm (laughs)

A (laughs)

B (laughs) Hell's teeth: second (2.5 sec): vertical line: from the: right:

A Uh-huh

B O.K.? And: second from the bottom.

Player A then describes her own position in an identical fashion:-

A Ah, well I'm in the second vertical line from the left and second from the bottom.

B Ha! {Where's your asterisk?}

A {And wh-: } My asterisk is the: the third vertical line: a-the square at the top- where's yours?: (1 sec)

B The third a-: (1 sec) line from the right:

A The square at the top?

B Yes. (See also fig. 6-7 (ii))

The principle behind these descriptions is identical to that underlying the row descriptions of the two previous examples, namely, the identification of a line (in this case, a vertical line) followed by the identification of

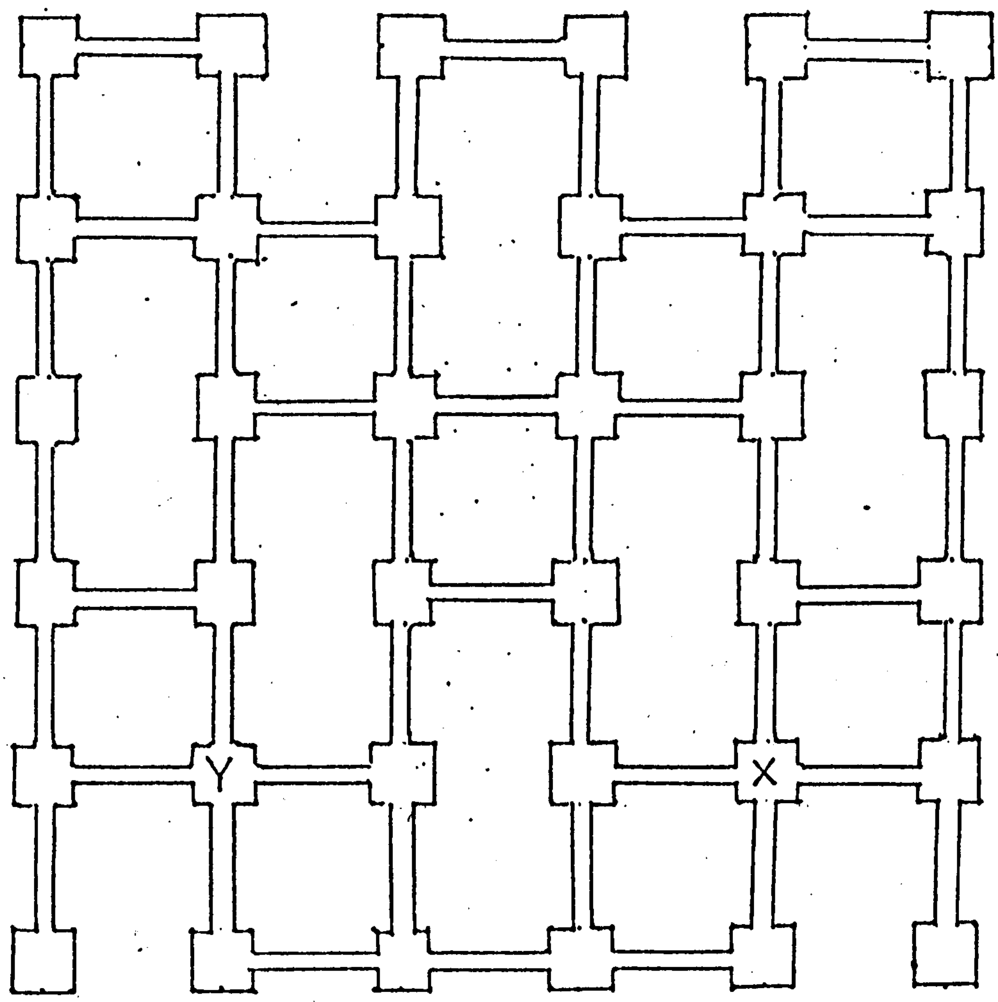


Figure 6-7 (i) the point (X) described by subject B in transcript AALOG41B as " . . second vertical line from the right . . . second from the bottom".
The diagram also shows the point (Y) described by player A as vertical line from the left and second from the bottom".

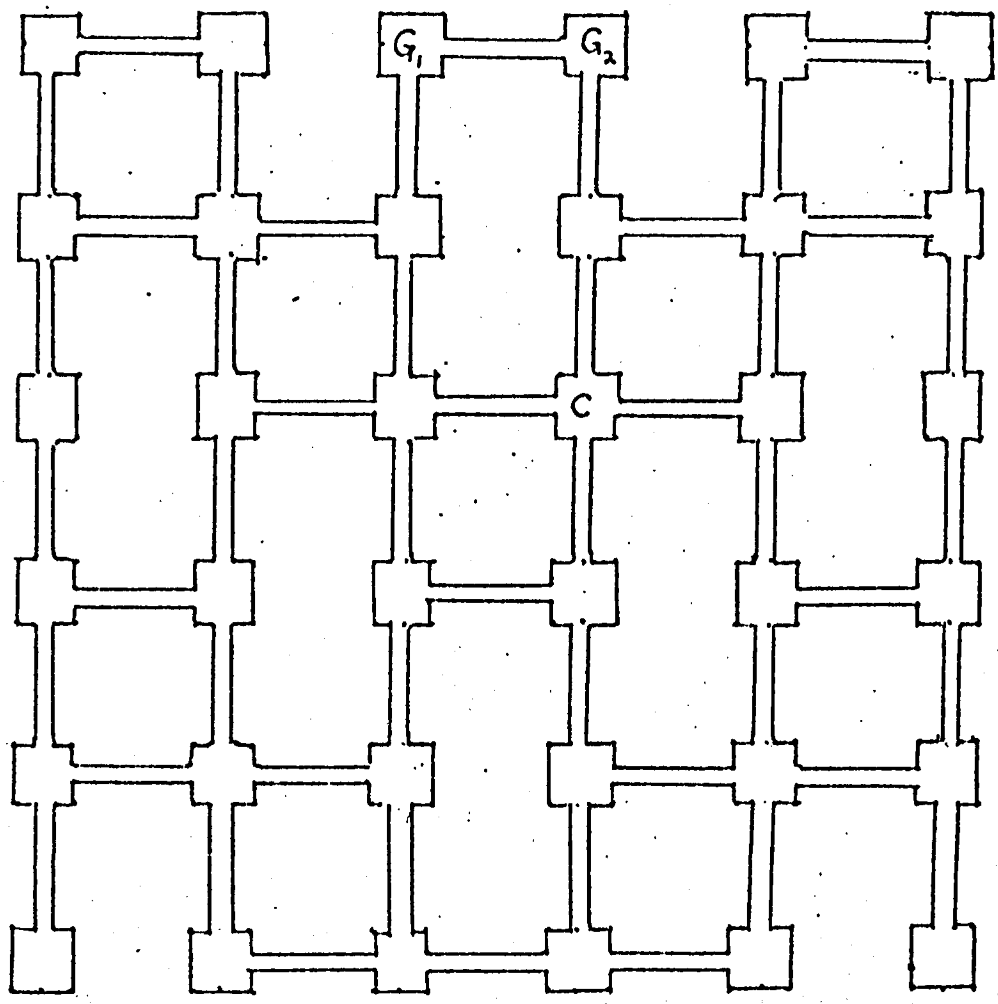


Figure 6-7 : (ii) the point (G_1) described as "the third vertical line: the square at the top" and the point (G_2) described as the third vertical line from the right: the square at the top". Diagram also shows point (C) described as "Three rows in and four up". (All taken from transcript AALOG41B) a position located within that line. Thus, the term line type will be used to refer to the generic type comprising both row- and column-type descriptions.

The possible reason for the hesitations apparent in the above example is that the above descriptions are the very first in the whole game; later descriptions from the same pair of subjects, whilst still involving some pauses (presumably to 'encode' the description) give the impression of, as it were, 'flowing' better, and also illustrate an interesting quirk (see figure 6-7 (ii); point C is being described):

A . . . where are you now?

B Oh I'm: (1 sec)

A Three rows in?

B Three rows in and four up (1 sec).

Two things are notable; firstly, the two pauses. The first pause occurs while speaker B is encoding her description (she is, in fact, 'beaten to it' by her partner), whilst the second pause occurs after B has generated her description, a gap in the dialogue which is presumably because it is taking player A a certain amount of time to 'decode' the message and to execute the relevant procedures. (Whilst this interpretation is necessarily speculative at this stage, it does lead to hypotheses which are testable). The second interesting point is both speakers' use of the term 'row' to refer to columns. Notice that the description given ('three rows in and four up') is literally nonsensical, since rows are ordered in the vertical plane, not the horizontal plane. It makes no sense, therefore, to talk of 'rows in (from one side)'. These two subjects, however, are using the term 'row' to refer to both horizontal and vertical lines; this reflects a certain looseness of usage of the term 'row' by any formal criterion, but plainly it causes no problem since both speakers use it in the same way, and there appears to be no evidence of misunderstanding at any point in the transcript. Whilst this may appear to be something of an idiosyncracy, it has occurred in more than one dialogue and underlines Rommetveit's point about the inadequacy of formal criteria based on 'literal meaning'; provided two speakers tacitly agree to use a particular term in an unconventional way and do so consistently, no problems arise with regard to meaning.

These last examples illustrate the properties of line-type descriptions; the next generic type of description we shall refer to as the matrix or coordinate type.

iii) Matrix type

This type of description is very similar to the Cartesian coordinates of mathematics. In this case, the speakers agree upon a notation with which to identify individual rows and columns, and thereafter generate descriptions by quoting two numbers or a letter and a number, each of which denotes a particular row or column. Thus, the descriptions generated are two-item coordinates, which correspond to count procedures operating vertically and horizontally from an agreed-upon origin point (always either the top or the bottom left hand corner of the maze).

Some examples are given below to illustrate the properties of this type of description. The first example (see figure 6-8) comes from a 'monster' game with a female dyad (transcript AALOG28B). Player A describes her position:

A Right, will I try and tell you with - as if: it was just S: all squares? (1 sec)

B Uh-huh

A Right, and that'll ma-: right, I'm like: (1, 1) (1 sec).

B Right, well I'm er: (1.5 sec) (5, 1):

A Wait a minute: hhh 1, 2, 3, 4, F- (5, 1) 5-: (B interrupts)

B No: no, no, no, no. (1 sec)

A {4, 1)?

B {1, 2, 3, (4, 1)

A (4, 1)

B Yeah.

In the above example, the columns are numbered zero

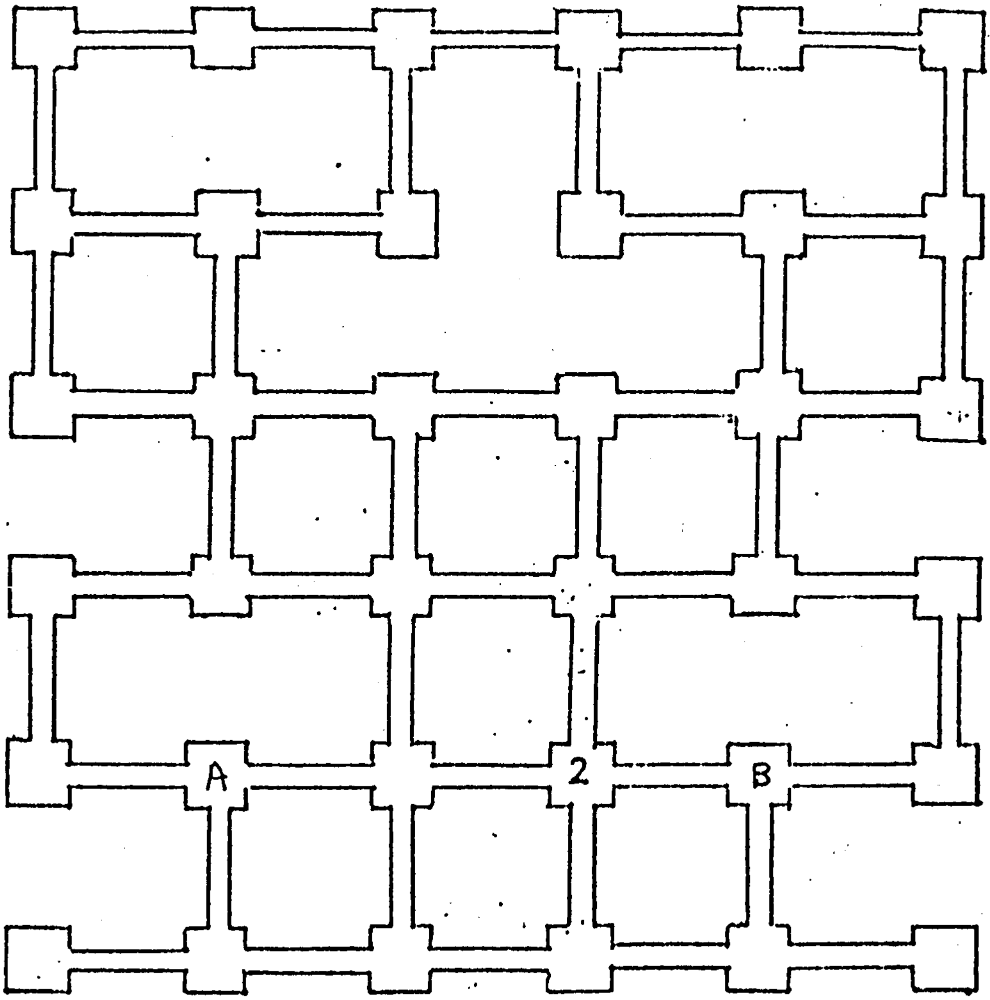


Figure 6-8: The point (A) described as "(1, 1)" and the point (B) described as "(4, 1)". Also illustrated is the point (2) described as "(3, 1)" (all three descriptions from transcript AALOG28B)

to five from left to right, and the rows are numbered zero to five from bottom to top. Stating the column followed by the row numbers (as in the ordering of X and Y axis positions in Cartesian coordinates) is implicitly procedural. The listener's task would be to scan from the origin point to find the relevant column, then scan to find the relevant row, and thence to seek their intersection, thus yielding the point in question.

This type of description could be regarded as highly elliptical and rather abstract. We noted earlier that path descriptions relate directly to available paths in the maze structure and so 'depend' on the maze structure. Line-type descriptions on the other hand, abstract parallel lines from the maze structure even though on the actual maze some of the, for example, rows, are incomplete. Line-type descriptions thus tend to ignore gaps in the path structure of the maze. Matrix or coordinate descriptions abstract from the maze structure still further, and ignore missing paths in both horizontal and vertical planes, and even, to some extent, <u>overspecify</u> the maze locations since they implicitly include a certain number of 'imaginary' points, for example, in matrix descriptions of the asymmetrical maze. Thus, matrix descriptions represent a further level of abstraction than do line type descriptions.

Matrix descriptions are highly elliptical since they specify only a number for each of the row and column combinations corresponding to the particular location being described: the listener has rather a lot of 'decoding' to do. We would therefore expect even more in the way of hesitating and pausing in generating and understanding matrix type descriptions than was the case with row- and column-type descriptions, a hypothesis which will be tested later.

Casual inspection of the transcripts does tend to support the impression of pausing and hesitations being prevalent whilst subjects are producing and understanding matrix descriptions.

Also prevalent is a good deal of counting aloud whilst the listener or speaker executes the scanning procedures. Both pausing and counting feature in this example (from transcript AALOG28B, see figure 6-8, point 2; this example comes from the same game as the previous example).

A Right, where are you now?

B I'm now: (1 sec) at (3, 1) (2.5 sec) And . . .

A (3, 1) Wait 1, 2, th-: (1 sec) One (3 sec) And . .

The pausing by both players to encode and decode the description is readily apparent, as is the counting out (i.e. the execution of the count procedures marking the row and column positions).

A final example, from a 'monster' game with a male dyad (transcript AALOG25B, see figure 6-9) again

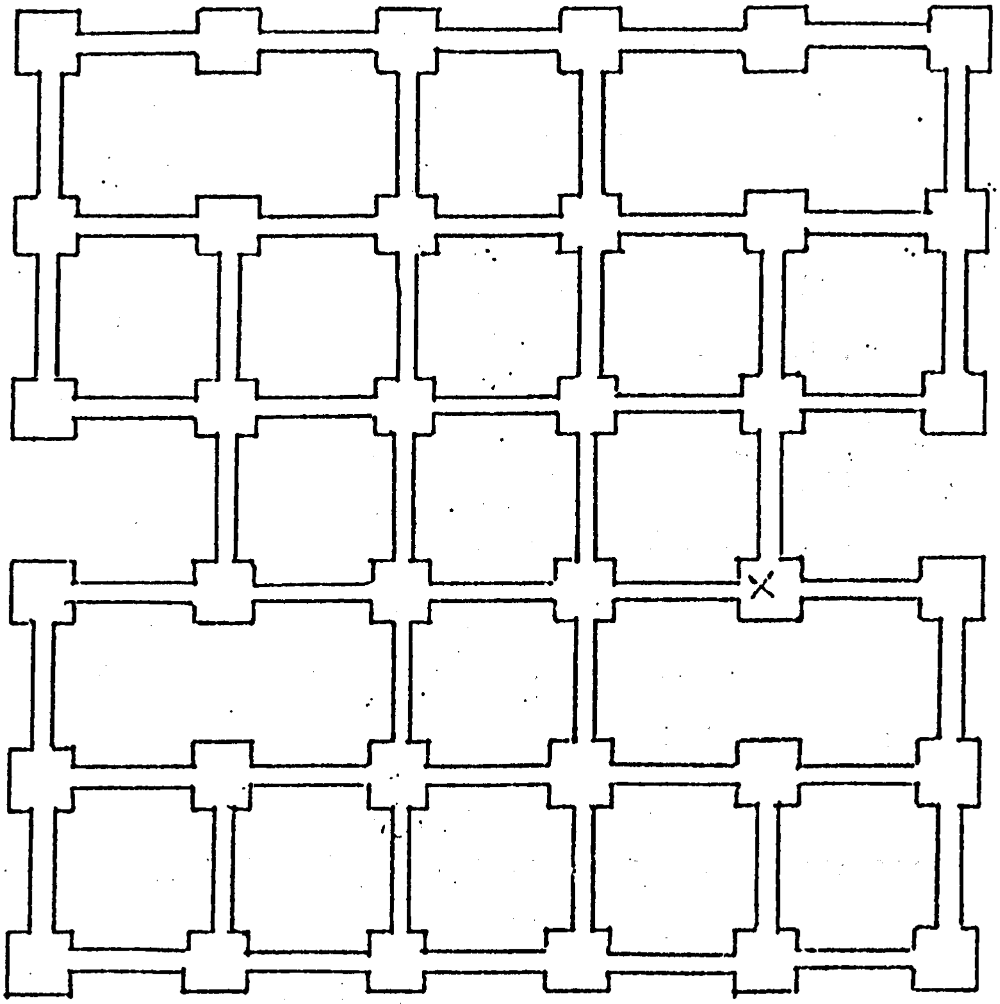


Figure 6-9: The point (X) described as "(4, 5)" by subject B in AALOG25B. The origin point for the coordinate counting in this example is the top left hand corner, and "Y" coordinates are here specified before the "X" value. This is different to the situation illustrated in figure 6-8 where the origin was the bottom left hand corner, and "X" values were quoted before "Y" values.

illustrates these two features noted above, that is, pausing and the counting:

B Er: I'm in (4, 5), where do you want me?

A (4, 5): Er: 4: 1,2,3,4, 1,2,3,4,5: (2 sec). Could you move. . . .

Thus, matrix descriptions function in much the way that Cartesian coordinates do, and are highly elliptical, being short in actual utterance length but implicitly packaging a great deal of information into each pair of numbers, and thus there is a good deal of pausing and counting out apparent when reading transcripts of games involving this type of description.

The final type of description is much less frequent in occurrence than the other three types, but does constitute a genuinely different category of description in its own right. This type shall be referred to as the _figural type_ of description.

iv) _Figural type descriptions_

In this type of description, subjects identify particular locations in the maze by making a metaphorical reference to a _figure_ which constitutes part of the maze structure, such as a rectangle, T or L shape. This serves to identify a particular region in the maze structure, and the particular point to be described is located relative to the named figure. Examples (below) will help to clarify this general strategy. For example, from transcript AALOG38A, involving a male dyad and the asymmetrical maze (see figure 6-10) came the following description (of point X in figure 6-10):

> B I can only go backwards, I'm: quite near a corner, there's a sort of: T shape: I'm in one corner of it: and no matter what way I go I'll be stuck.

Another example of a figural description is one which refers to a point also shown on figure 6-10, and the description comes from transcript AALOG35B (female dyad, asymmetrical maze). Player A is asked to describe the location of her goal box (i.e. point Y in figure 6-10):

> B Where's the asterisk? (2 utterances)
>
> A Er:mine is: if you take the uppermost left hand side: right?
>
> B Right
>
> A See where that L is?
>
> B Yes
>
> A Right: ehm: my asterisk is at the: (1 sec) sort of: (1 sec) last part of the L: see what I mean? In the middle square of that line:
>
> B O.K.: ehm:

Once again, this type of description could be interpreted to be procedural in nature: the listener must scan for

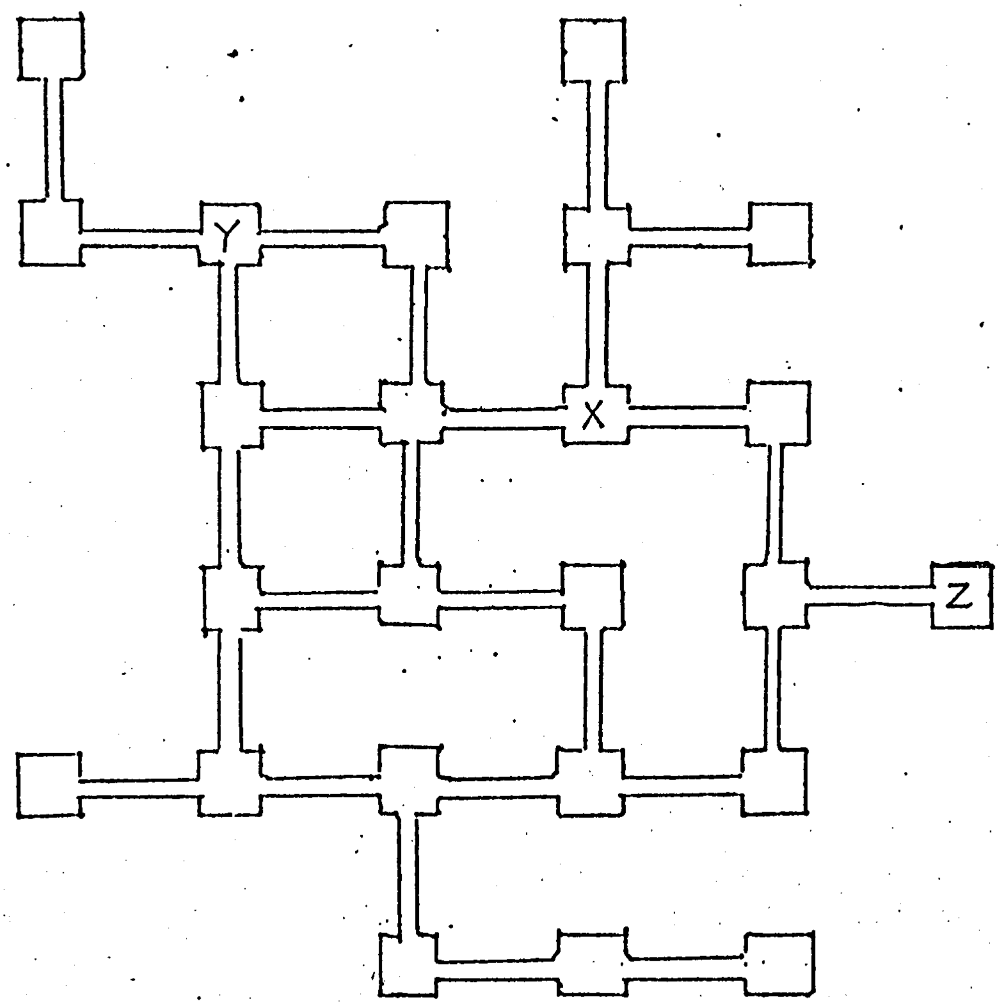


Figure 6-10: The point (X) described as "one corner" of a "T shape" (AALOG38A)
The point (Y) described as "last part of" an "L" (AALOG35)
The point (Z) described as "sticking out like a sore thumb . . like a right indicator" (AALOG40B)

the appropriate figure and then scrutinize the region in which the figure is located, in a manner described by the speaker, to find the relevant point.

Another example of such a description comes from transcript AALOG40B, again from a game involving the asymmetrical maze (see point Z in figure 6-10):

B O.K. Stan, right, let's- let's talk about
  this, whereabouts a-whereabouts are you?
A Right: er: I'm: I'm: extreme right:
B Extreme right
A Yes, the o-the-there's one er: what d'you
  call it, there's just one box:
  {to the extreme right:
B {W- away out on the right?
A O-on my right, I don't know which way you're
  facing the screen but I:I suppose it should
  be the same.

B {Yeah
A {You know, the extreme, there's one box.
B Yeah, right, the extreme right, st-sticking
  out like a sore thumb?
A Yes
B Yeah O.K.
A That's where I am.
B It's like a right indicator?

A Yes.

This example is an interesting one. The metaphorical references to 'sticking out like a sore thumb' and 'it's like a right indicator' serves to emphasize the physical isolation of the point being described. The term 'right indicator' is a metaphorical reference to a bicyclist's hand signal for turning right (the protruding of the right arm perpendicularly from the vertical line of body symmetry). Interestingly, this term 'right indicator' was adopted by these two players to refer to all three right-protruding parts of the maze (that is, the bottom-most horizontal line, the protruding node just described, and the topmost right-protruding node which the first description on fig. 6-10 described as part of a 'T shape').

Another point worthy of note is that, at one point, player A questions player B's, deictic orientation relative to the screen. This is not unique to this particular pair of players. In normal face-to-face conversation, of course, A's right would be B's left, and vice versa, and as Buhler pointed out in his example of a physical training instructor, A would try to adopt B's perspective and use his (B's) orientation to generate discourse. This player A tries to do; however, the orientation of both players' bodies relative to the maze on the screen is identical and A's effort is unnecessary. This attempt by player A to anticipate the other's spatial orientation is interesting in relation to the discussion in chapter 3 of Rommetveit's notion, based on Mead's philosophy, of speakers and listeners taking one another's perspective.

The reference to analogues in the maze stucture of certain figures is the main characteristic of figural-type descriptions. This helps the listener to narrow the search zone by indicating the region of the maze within which, or near to which, the point to be described is located.

Figural descriptions have also been used by subjects to describe locations within symmetrical mazes; an example is given below (see also figure 6-11) from transcript AALOG36A (female dyad, study 3). Player B describes her position:

B O.K.: My-I'm now: see-see the three columns, there's three rows: (1 sec) tak- taking it downwards: there's three groups:
A One: (1 sec) two: (1 sec) three?
B Three groups:
A Right:
B Well I'm on: the second group at the left hand side:
A Second group up at the top.
B Uh-huh:

The three vertically oriented groups described by

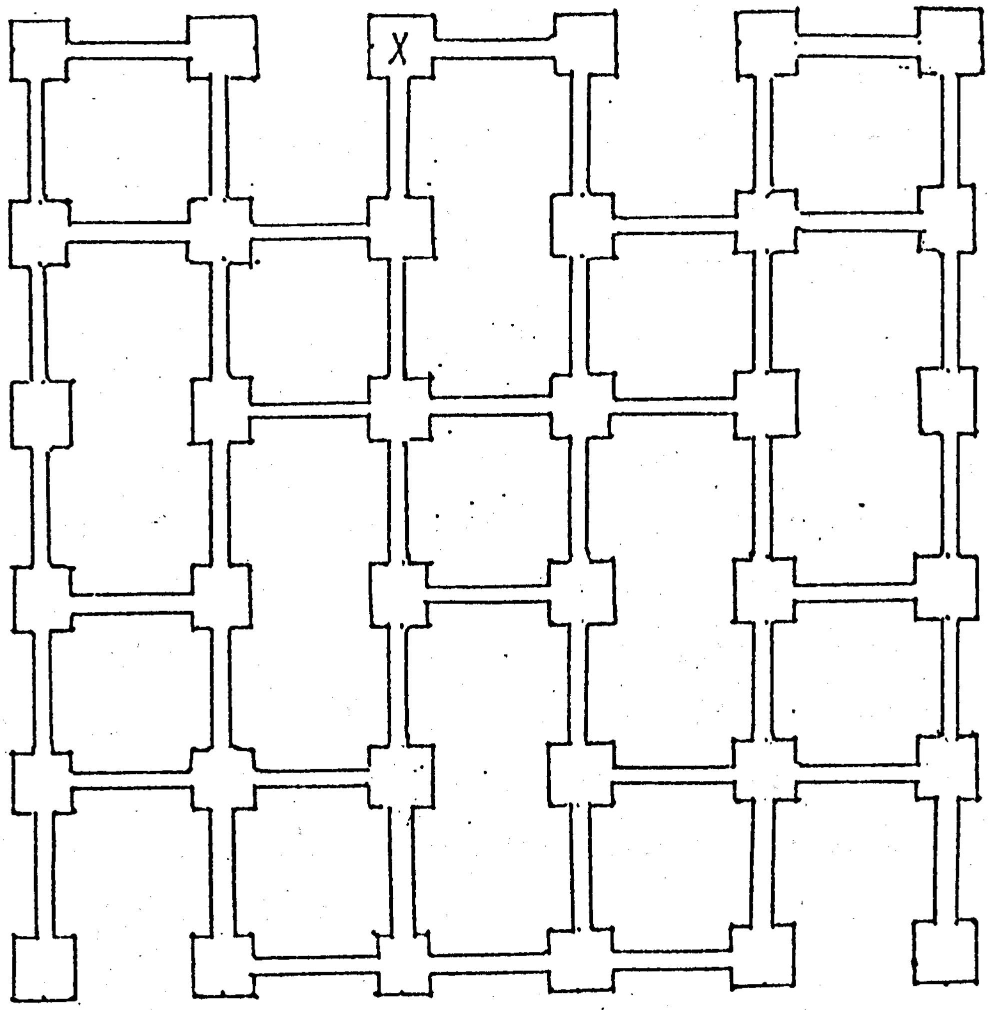


Figure 6-11: The point (X) described as "the second group at the left hand side . . up at the top" (AALOG36A)

player B are the three pairs of columns in the maze shape (see figure 6-11); the groupings are apparent because of the discontinuity of the top horizontal line of the maze. Once again, the use of figures (in this case, 'groups' of columns) and the identification of the second (centre) group narrows the search zone. As it turns out, the listener already has an idea of where player B's position marker is located and guesses correctly.

These four examples serve to illustrate the general properties of the figural type of description. It should be emphasized in passing that figural-type descriptions are far less frequently observed than are path, line and matrix descriptions, and tend to be accompanied by path-type descriptions, resulting in a mixture of path- and figural type descriptions. The figural type of description does, however, constitute a genuinely distinguishable type of description in its own right.

c) Conclusions

The examples given illustrate the four generic types of description used by subjects, which have been termed the path, line, figural and matrix types. Common to all four, it has been argued, is an implicit or explicit procedural component: the descriptions take the form of tacit instructions regarding visual scanning operations. Indeed, in a study to be described subsequently in which the experimenter was present in the same room as both subjects whilst subjects generated and listened to such descriptions, subjects were observed to execute the procedures hypothesized above using ostensive gestures with their fingers or a pen in a manner consistent with the above interpretation during the course of understanding and generating descriptions. The procedural interpretation of the descriptions is therefore backed up by informal observation.

Within the generic categories outlined above, certain detail differences in the extensional semantics of the procedures implicit in the descriptions arise, which will be discussed more fully later. In the next section, the methodological rationale behind the analysis will be introduced. The analysis was an attempt to describe statistically the place-descriptions by means of a categorization of lexical item types, and to ascertain which groupings of lexical item types consistently occurred.

For the moment, we can conclude that the descriptions generated comprise four broad classes which we have described as path, line, matrix and figural. These four types of description exhibit certain commonalities, notably their use of procedures for generating and interpreting descriptions. It was argued that some description types involve conceptual simplifications of the complex stimulus of the maze structure: this is particularly true of line- and matrix-type descriptions.

(Paradoxically, these simplifications result in implicit overdescription of the maze, as happens, for example, when line- or matrix-type descriptions are used to describe locations on the asymmetrical maze). The classes of description differ in the particular manner in which they conceptually "dissect" the maze shape in order to simplify the description task.

Having thus indicated some of the semantics of the location descriptions in informal terms, we now turn to examine the background to the statistical analyses undertaken.

d) The statistical description of the data: methodological considerations

In the first part of this chapter, an exposition was given of some typical descriptions of spatial locations generated by subjects in the course of playing the game. In particular, four broad classes of description type were informally identified, corresponding (it was argued) to four different ways of conceptualizing the maze shapes and visually scanning them. The classes identified were: a) line-type descriptions, in which the maze is treated as if it were an array of parallel lines, and the subject firstly identifies lines followed by positions on those lines; b) matrix-type descriptions, in which the array of nodes is treated as a two-dimensional space and particular nodes are identified by means of Cartesian-type two-place coordinates; c) path-type descriptions, in which descriptions are generated by means of firstly identifying a visually prominent point in the maze structure, and the desired location is then described relative to that visually prominent point by means of path movements from the visually prominent point to the location to be described; and d) figural-type descriptions, in which the subjects generate descriptions by referring metaphorically to a figure (a shape of some kind) which forms part of the structure of the maze's network of pathways, and the relevant position is then described relative to that figure. These four types actually constitute <u>classes</u> of description, within which detail differences occur. An example of such a detail difference aleady encountered is the contrast between line-type descriptions which use vertical as opposed to horizontal lines.

Another detail difference is whether in counting node positions, subjects count <u>movements</u> from node to node or instead count the nodes <u>themselves</u>. (See figure 6-12 for an exposition of this difference). The difference is exemplified by comparing figure 6-3 (illustrating a path-type description) and figure 6-7 (i) (illustrating a line-type description). In the line-type descriptions shown in figure 6-7 (i) (i.e. the <u>nodes</u> themselves are counted - the right-most vertical line of nodes is line 1, the next line is the 'second line from the right', and so on. Thus, <u>node positions</u> are counted. This contrasts with

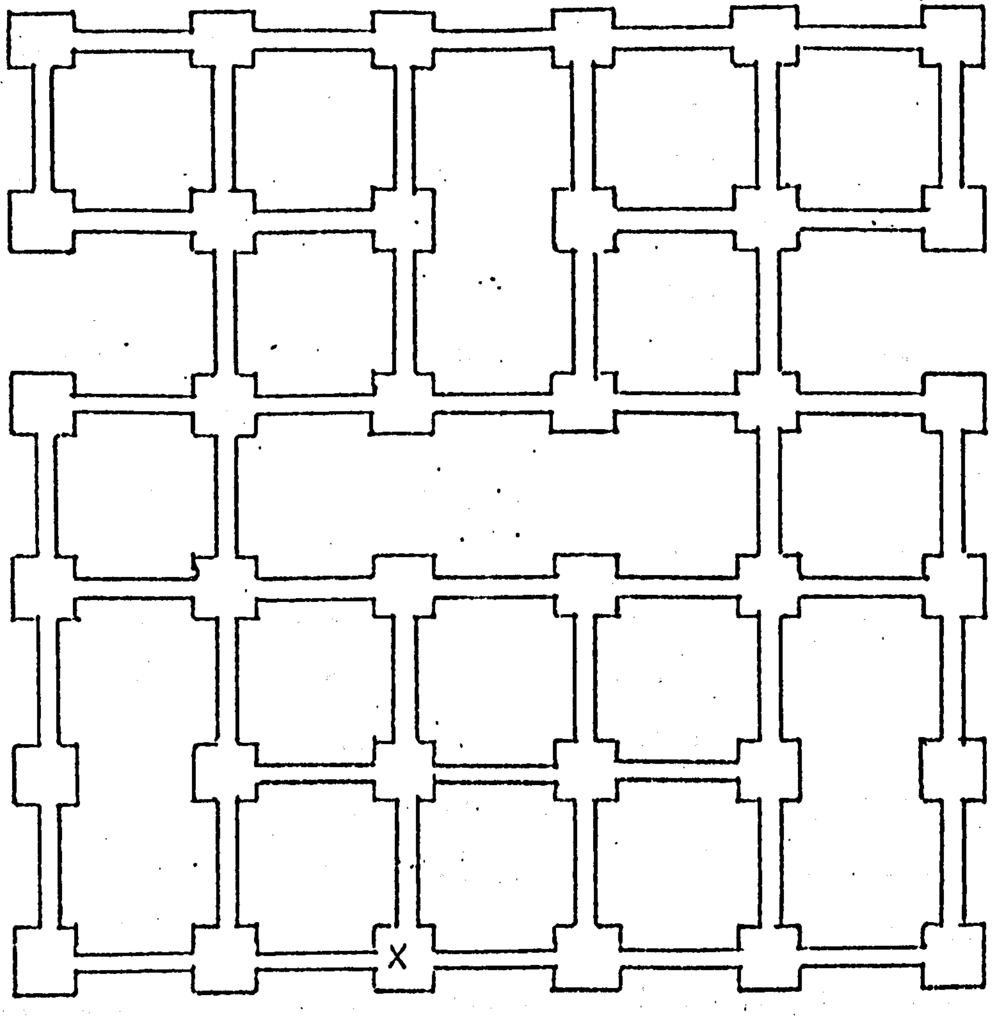


Figure 6-12: The difference between the extensional semantics of node-counting, as opposed to path-movement-counting, scanning procedures. A node - counting scanning procedure would designate X in the above diagram as "third on/from the left" (numbering the nodes from 1 to 6 from left to right). A path-movement-counting scanning procedure would designate X in the above diagram as "two from the left" or "two along" (numbering the nodes from zero to five from left to right).

the path description in figure 6-3, in which 'one along' counts a path movement, as does 'two up' - 'two up' as a path movement brings the visual scanning procedure to the third node from the bottom (cf. figures 6-7 (i) and 6-3; see also figure 6-12). Such a detail difference could cause confusion if one subject in a dyad counts nodes whilst his partner counts path movements between nodes when executing scanning procedures, and so this detail difference is of some importance.

The point to be emphasized here is that, although the four-category classification given above is a reasonable summary of the major categories of description type, the individual categories are by no means uniform classes.

As is readily apparent from a casual inspection of the various sample descriptions presented earlier in this chapter (see figures 6-2 to 6-11), there is a great deal of variation between descriptions in terms of the lengths of descriptions and how explicit or elliptical they are. These, quite considerable, variations in the length and explicitness of descriptions occur both within and between description classes. For example, compare the descriptions of the points illustrated in figures 6-6 ("second row . . . second from the left") and 6-7 (i) ("second (2.5 sec) second: vertical line: from the: right: . . . O.K.? And: second from the bottom"). Both of these descriptions are line-type descriptions, yet one is approximately twice the length of the other in terms of the number of words employed. There is thus a considerable variation in the amount of speech used to produce one description.

A more formal statistical classification of subjects' location descriptions than the earlier-presented informal four-fold classification was thought to be desirable. This formal classification would have to be achieved despite the above-noted variation in the length and explicitness of different location descriptions. It was hoped that the formal classification would produce results which would provide more detail than, and at least partially confirm, the informal four-fold classification presented earlier. Details of how the formal analysis was conducted are set out below.

The formal analysis of the subjects' location descriptions involved two aspects: i) a descriptive aspect, and (ii) a classificatory aspect. These are each dealt with in turn.

i) The descriptive aspect of the formal analysis

The first aspect of the formal analysis employed was a _descriptive_ aspect: the raw data (that is, the actual location descriptions generated by subjects, with all their variations in length and explicitness) had themselves to be described. This was achieved by means of the development of a set of thirty-eight _lexical categories_, to which individual lexical items in subjects' raw data could be assigned. The description of the raw data thus took the form of a categorization: each description generated by each subject in each game was examined, and the key lexical items employed by the subjects were assigned to, that is, counted as instances of the use of, particular categories in the set of thirty-eight lexical categories. (The rationale behind the choice of the particular lexical categories employed is set out below, and a listing of the lexical categories is given in table 6-1).

Thus, all descriptions generated by each subject in the course of playing each game were examined for their usage of lexical items corresponding to each of the thirty-eight lexical categories. The subject's usages of each of the thirty-eight lexical categories were summed across all descriptions generated by him or her in that particular game. This resulted in a profile for each subject in each particular game of the number of occasions on which lexical items corresponding to each of the thirty-eight lexical categories were used. When this is done for all subjects in all 48 games, it is possible to utilize the resulting matrix (of usage of lexical categories by all 96 subjects) as the data matrix for a classification study in which the lexical categories are classifed into groups.

ii) The classificatory aspect of the formal analysis

The second aspect of the formal analysis was the classification of the thirty-eight lexical categories. That is, a classificatory analysis of the thirty-eight lexical categories was undertaken in order to ascertain whether clearly identifiable clusters of lexical categories existed such that the categories within a cluster tended to occur together in the subjects' location descriptions. Such clusters would thus correspond to sets of lexical categories related in such a way that all lexical categories within a set tended to be used together by subjects - for example, it would accord with our earlier informal classification if a cluster of lexical categories corresponding to all the types of lexical item employed in generating 'path'-type descriptions were to be identified.

A suitable descriptive statistical procedure capable of generating such a classification is that of Hierarchical Cluster analysis. A summary of the functioning of this technique is provided in the technical appendix (appendix IV; see also Gordon, 1981, for further details). Essentially, hierarchical cluster analysis operates by treating each of the thirty-eight lexical categories as variables, generating a measure of similarity between each pair of variables (for example, using the correlation between the variables as the measure of similarity between them), and it then imposes a best-fit hierarchical structuring upon the variables. This result in a kind of tree diagram known as a <u>dendrogram</u> which displays groupings (clusters) of related variables nested in a hierarchical fashion (see appendix IV). In the case of the present analysis, the interest centres on how individual clusters are made up rather than with higher level relationships between clusters.

Thus, the formal analysis involves the two aspects of (i) initially describing the data in terms of thirty-eight lexical categories and then (ii) subjecting the summarized data in the thirty-eight lexical categories to hierarchical cluster analysis in order to ascertain which groups of lexical categories tend to be used together by subjects.

In the remainder of the present chapter, a discussion of the methodological rationale behind the selection of the particular lexical categories employed in the descriptive aspect of the analysis is undertaken, and the final section of the chapter will briefly discuss the uses to which hierarchical cluster analysis was put in the present set of studies.

e) The selection of lexical categories

As noted above, the lexical categories used for the purposes of describing the raw data numbered thirty-eight in total. The selection of the thirty-eight lexical categories was based upon a detailed inspection of the raw data. That is, the actual location descriptions themselves were used to assist the process of devising a categorization; the lexical categories were not devised in advance in an arbitrary fashion. In fact, the set of thirty-eight lexical categories was the successor to an earlier set of twenty-seven lexical categories which was initially devised but which had been felt to be insufficiently "fine-grained" in its categorization of the lexical items in the subjects' location descriptions. The thirty-eight category categorization was developed in an attempt to provide a suitably detailed categorization of the lexical items in the location descriptions which would be sufficiently sensitive to enable the categorization of <u>all</u> the descriptions in the data.

The selection of lexical categories was thus 'data driven' in the sense that it was based on the actual raw data; the categories were selected in an attempt to capture the broad semantic distinctions which the subjects appeared to be making. For example, a feature which was common to all of the different types of description was the counting of node positions (or path movements). This counting was of two types, <u>ordinal counting</u> (that is, 'first', 'second, 'third' . . . 'last') and <u>cardinal counting</u> (that is, 'one', 'two', 'three', . . . etc.) Each of these two types of counting was treated as a separate category of lexical item. This treatment of ordinal- and cardinal counting separately was undertaken in case there was a systematic difference in their patterns of occurrence; it is quite possible that ordinal counting is used by subjects under one set of circumstances and cardinal counting is used by subjects under a different set of circumstances. Treating them as separate categories, I hoped, would capture this, with the two categories showing different patterns of clustering, such that they belonged to different clusters of lexical categories. The sheer prevalence of counting in the descriptions indicated that the inclusion of categories which recorded instances of this feature was necessary.

The same rationale was employed in order to select the other lexical categories. Each of the description types (see the earlier informal four-fold classification) was examined in order to ascertain which sorts of semantic distinctions were being made, and these suggested further lexical categories for inclusion in the categorization. Some examples will illustrate the operation of this principle for devising suitable lexical categories.

Consider, firstly, the row-type descriptions, such as the following description (which was discussed earlier, and the location described was illustrated in figure 6.6):

"Second row . . . second from the left"

In this example, two ordinal counts are involved, one for marking the different rows and the other for marking the positions of nodes within the rows. The count denoting the position of the individual node within the specified row is expressed as a distance from an edge (". . . . second from the left"). Notice that in the above example, the specification of the particular row is elliptical in that the relevant edge with respect to which the counting of rows takes place is not specified. In other line-type descriptions, however, the specification of lines and node positions on those lines is less elliptical. Consider, for example, the following description which was also discussed earlier and whose referent is illustrated in figure 6-7 (i):

> " . . .second: vertical line: from the right: . . O.K.? And second from the bottom".

In this example, both the line position and the node position are specified with respect to a relevant edge (the right edge for specifying the relevant vertical line, and the bottom for specifying the relevant node position). This specification of line and node positions with respect to edges is quite common in line-type descriptions, and so a category for recording instances of usage of each of the edge names (top, bottom, left and right) was included in the categorization, as was one category for recording instances of usage of the prepositional phrases which were sometimes used in these constructions (for example, " . . third on the right", " . . third from the left"). The reason that the use of such prepositional phrases was recorded in one category was that in some line-type descriptions the prepositional phrases were absent altogether (as in examples where a subject might say "I'm on the second bottom row"). Thus, a lexical category which recorded instances of "preposition (plus edge name)" was used, in order that the relationship of such prepositional phrases' usage to the usage of other lexical categories could be ascertained. (Specifically, if consistent groupings of lexical categories were revealed by the cluster analysis, it was felt desirable to ascertain which grouping the prepositions would cluster with).

Inspection of line-type location descriptions reveals certain other lexical items used by subjects quite frequently. These were also judged to reflect relevant semantic distinctions and consequently were included as separate lexical categories - these were 'row', column', 'line', 'horizontal' and 'vertical'.

An inspection of path-type location descriptions suggested further lexical categories for inclusion in the descriptive categorization. Consider, for example, the following description (see also figure 6-3 and section b) i) of the present chapter):

> "If you go to the bottom right corner, one along, two up".

Two cardinal counts are involved here. Key terms suggested as important for inclusion in the descriptive categorization are '<u>corner</u>', '<u>along</u>', '<u>up</u>' and '<u>move/go</u>'. These terms are of importance in path descriptions; path movements are often specified with respect to corners, and movements in specified directions are important semantic elements of path descriptions. Thus, a lexical category for recording instances of usage of the term 'corner' was included in the descriptive categorization, as were separate lexical categories for recording instances of usage of terms denoting movement in (either vaguely or precisely) specified directions, namely: '<u>move/go</u>' (one lexical category), '<u>along</u>', '<u>up</u>', '<u>down</u>', '<u>left</u>' and '<u>right</u>'. These last two categories were distinguished carefully from the edge names 'left' and 'right'; each term 'left' and 'right' has two senses, a) a name of an edge, and b) a direction of movement. These two senses were kept separate: separate lexical categories were involved for each of the two senses (see table 6-1).

Other lexical categories were suggested by an inspection of the path descriptions, notably '<u>half/hand side</u>', this category being used to record instances where subjects indicated that a particular half of the maze about a vertical centre line was under discussion. '<u>Gap</u>' was also included in the categorization as a separate lexical category, since mention of gaps or spaces without paths did seem to be a feature of true path-type descriptions. '<u>Start</u>', '<u>own/other</u>' and '<u>monster</u>' were included as separate lexical categories, since subjects often used their own start position , a previously specified position of themselves or their partner, or the monster's position, respectively, as reference points with respect to which path descriptions were generated (for example, "I'm three to the left of the monster", "I'm just two above where I started from", etc.).

An inspection of the figural-type descriptions revealed several terms used by subjects which required to be accounted for in the descriptive categorization. The generic category of '<u>shape</u>' was used to count instances of usage of a variety of terms of this <u>genre</u> - T shapes, L shapes, rectangular shapes, etc. Whenever the term 'shape' was used to refer to a figure in the maze structure, this was recorded as an instance of usage of the lexical category 'shape'. Other key terms suggested by an inspection of figural descriptions as meriting inclusion in the descriptive categorization as separate lexical categories were '<u>middle</u>' (as in " . . . the middle of that line") and '<u>limb</u>' (as in, for example, "there's a limb sticking out to the right").

An inspection of the coordinate descriptions (see section b) iii) of the present chapter and figures 6-8 and 6-9) suggested two further lexical categories of importance: one for recording instances of usage of <u>coordinate descriptions</u> themselves (e.g. "<u>A5</u>", "<u>C3</u>", etc.), and a lexical category for recording instances of what might be termed '_alphabetic counting_', (for example, where a subject might say "I'm in _A, B, C, D,_: D5").

The above categories, which total 29 in all, were devised in the manner indicated above: careful attention to the raw data suggested that certain categories of lexical item were distinguishable, these categories corresponding to semantic distinctions which the subjects appeared to be making.

The remaining nine lexical categories of the total of thirty-eight were included in the categorization in an attempt to describe more idiosyncratic terminology on the part of the subjects.

The above description exemplifies how the descriptive categorization of the lexical items used by the subjects was devised, and simultaneously illustrates that _the basis for the categorization was the distinctions made by the subjects themselves: their_ choice of lexical items guided the descriptive categorization.

It is of importance that a full discussion of the rationale behind the selection of lexical categories should be given here, since the selection of descriptors of the data is a crucial part of the use of the technique of cluster analysis. As Gordon (1981, p. 8) remarks, it is quite possible (and even desirable) that there are several different ways of classifying the same set of objects. It follows that careful thought should be given to the selection of variables (in this case lexical categories) with which to describe each object (in this case, the set of place descriptions generated by a subject). That the content of the subjects' location descriptions was the basis of the categorization employed to describe the data rather than any theory about how the descriptions were generated is an important principle in this regard.

A complete listing of the 38 lexical categories is given in table 6-1. It was noted earlier that some features corresponding to particular lexical categories were used frequently by many subjects (for example, the counting of node positions), whilst other features were used by far fewer subjects. In table 6-1, an exposition will be given of which particular lexical items would qualify as exemplars of each of the lexical categories in addition to comments on those lexical categories of notably high or low frequency of usage. A further point of note is that some lexical categories were used for recording instances of usage of single lexical items, whilst other lexical categories were used for recording instances of usage of any of a group of synonymous lexical items. For example, the name of an edge like 'top' or 'left' has no synonymous expression and so each of these lexical categories were used to record instances of usage of single lexical items. Other lexical categories, on the other hand, like 'middle', were used to record instances of any of a group of synonymous expressions (in this case, two: 'middle' and 'centre'). Table 6-1 is situated at the end of the present chapter.

The thirty-eight lexical categories were devised in an attempt to account both for descriptions of a common, frequently occurring nature and also for those of a more idiosyncratic nature.

Each description generated by each individual subject in each game was examined and instances of usage of lexical items corresponding to each lexical category were recorded and summed across descriptions. This count was not conducted in terms of the <u>absolute</u> frequency of usage of lexical items corresponding to a given category in a given description but on a <u>presence/absence</u> basis. Thus, each description was examined in order to ascertain which features were present (corresponding to different lexical categories.) If an exemplar of a given lexical category was present more than once within one description, this was recorded as <u>one</u> instance of usage. The count was conducted in this fashion in order to circumvent the problem of variation in the length of the descriptions. Long descriptions might repeat the use of particular lexical items, and an absolute frequency count could thus misleadingly inflate the value of the count for those lexical categories corresponding to lexical items used in long descriptions. Such lexical items would have associated with them a high value of frequency of occurrence were the count conducted on an absolute frequency basis, despite the fact that these particular lexical items may have been used in only a small number of descriptions.

Thus, what was counted in the initial description of the data was the presence or absence of a given feature within a given description, the count being incremented by 1 if a feature was present in a given description, regardless of whether it was used more than once in that description. The counts for each category were summed across descriptions for each subject, yielding a profile for each subject of <u>the numbers of descriptions in which he used each lexical category</u>.

Examples are given below to illustrate how sample descriptions were classified in terms of the lexical categories. The underlined terms are those being counted and the relevant lexical category numbers and names of which they are an example are given below each term (terms used repeatedly are underlined repeatedly but the associated lexical category is only named once):

1) " . . . <u>Second</u> <u>row</u> . . . <u>second</u> <u>from</u> <u>the</u>

(13, <u>order</u>) (23, <u>row</u>) (21, <u>to edge</u>)

<u>left"</u> (from transcript AALOG13A).

(3, <u>left</u>

2) "<u>Second:</u> <u>vertical</u> <u>line</u>:

(13, <u>order</u>) (26, <u>vertical</u>) (22, <u>line</u>)

from the: right: . . . and

(21, to edge) (4, right)

second from the bottom ." (from transcript

(2, bottom)

AALOG41B)

3) "I am: if you go to the top right corner, and go

(15, move/go) (1, top) (6, corner)

down one and along two to the left, that's where

(18, down) (16, along) (19, left).

I am". (from transcript AALOG15C)

4) "I can only go backwards, I'm: quite near a corner,

(6, corner)

there's a sort of: T shape: I'm in one corner

(7, shape)

of it . . ." (from transcript AALOG38A).

5) "Well, I'm in 1, 2, 3, 4,: (1s) Er: (1s) . . (4, 4)"

(12, count) (33, coords)

(from transcript AALOG25B).

When the numbers of descriptions in which exemplars of each lexical category were employed were summed up for all subjects, the result was a data matrix of 38 lexical categories X96 subjects (representing 48 dyads) and this matrix served as the data for a cluster analysis of lexical categories. The uses to which cluster analysis was put in the present set of studies are outlined below.

f) The analyses employed in the present set of studies*

As noted above, the data description phase of the analysis yields a matrix of the scores of 96 subjects in 48 games on a set of thirty-eight lexical categories. As has been indicated already, the first analysis to be undertaken was the cluster analysis of lexical categories. This was carried out in order to ascertain which groupings of lexical categories consistently resulted across different measures of similarity and different clustering criteria (see appendix IV for the rationale behind comparisons of results from different cluster analyses). It was hoped that the groupings of lexical categories yielded would be readily interpretable in terms of the earlier-presented informal four-fold classification of the location descriptions. The results yielded in this analysis will be examined in detail in the next chapter.

The second analysis undertaken was also an analysis involving the same data matrix. It is possible to mathematically transform the data matrix, rotating it through 90 degrees. Having done this, one then has a 96 subjects X 38 lexical categories matrix (having started with a 38 lexical categories X 96 subjects matrix).

* The reader is urged to consult appendix IV which presents a summary of the cluster analysis technique and the method of comparing the results from different cluster analyses of the same data.

It is then possible to perform a <u>cluster analysis of individual subjects</u>. This analysis would be of interest for several reasons. Firstly, an informal inspection of the transcripts suggests that dyads are consistent with respect to their usages of particular types of description: the casual impression that each dyad uses one form of description (for example, using line-type descriptions) throughout their game, rather than changing the type of description used from time to time within a game, is difficult to resist. It would be of great interest to ascertain whether consistent groupings of subjects could be detected across cluster analyses such that all individuals within a grouping used similar types of description. It would also be of interest to find out whether two individuals within a dyad playing the game were more likely to be located within the same cluster of individuals rather than in different clusters. If this were the case, it would point to the possibility that subjects within a dyad were entraining one another linguistically - that is, subjects within a pair playing a game were constraining one another's choice of referring expressions, such that both used the same type of description rather than using different types of description, and this would suggest that pragmatic factors are indeed of importance.

Thirdly, it would be interesting to relate the cluster analysis of individual subjects to the cluster analysis of lexical categories in order to ascertain whether the groupings of subjects were sampling from the groupings of lexical categories in a consistent way. If this were the case, it would suggest that not only are clusters of lexical categories detectable and clusters of individual subjects are detectable, but also that the individual subjects are clustering together in the particular way that they do because <u>everyone within a given cluster of subjects selects among the lexical categories in a similar way</u>. There is, unfortunately, no formal statistical methodology with which to carry out such an analysis, other than by simple inspection. More details of this analysis will be provided in the next chapter.

The next chapter will report the results obtained using the three analyses outlined above.

Table 6-1

The thirty-eight lexical categories for the description of the raw data

Category Number	Category Name	Explanation of the Category
1	Top }	The categories numbered 1 to 4 were used for recording instances of usage of edge names. For example, where a subject said something like "I'm third <u>top</u> line, three in from the <u>left</u>" an increment to the counts recorded in categories 1 and 3 would be made.
2	Bottom}	
3	Left }	
4	Right}	
5	Handside	This lexical category was used to record instances where subjects referred to a particular half of the maze (on either side of an imaginary central vertical line); for example, "I'm away over on the far left <u>hand side</u>". This lexical category was a multiple item category: the expressions "half" and "handside" or "hand" were used synonymously by subjects and instances of usage of any of these

Category Number	Category Name	Explanation of the Category
		three expressions were counted in this lexical category.
6	Corner	This single-item category was used to record instances where subjects referred to a corner of the maze (or of part of the maze structure), and its count would be incremented in the case of expressions like "I'm at the bottom left hand <u>corner</u>" (categories 2, 3 and 5 would, of course, also have an increment made to their counts in this example) or ". . . there's a wee T shape: I'm in one <u>corner</u> of it".
7	Shape	This lexical category is a multiple-item category, and its count would be incremented in the case of any of a variety of expressions being used, particularly those encountered in figural-type descriptions, such as "I'm in a the corner of that T <u>shape</u>", "There's an L <u>shape</u> up at the top left", and "I'm in the middle <u>grouping</u>, the column on the left at the top".

<u>Category Number</u>	<u>Category Name</u>	<u>Explanation of the Category</u>
7	Shape (continued)	The idiosyncratic expression "right turn indicator" (see the earlier discussion of figural descriptions in the present chapter) was also counted as an instance of his category, since it is being used by the subjects as a description of a shape. In all cases, examination of what was being described by subjects when they used an expression (that is, examination of the computer printout of the relevant game state) was carried out for clarification.
8	Middle	This lexical category's count was incremented whenever the terms "<u>middle</u>" or "<u>centre</u>" were used by subjects.
9	Limb	This was a multiple-item lexical category used for recording instances of reference to a protruding part of the maze structure in terms of a "<u>limb</u>" or a path "<u>jutting out</u>" or "<u>sticking out</u> on its own".

<u>Category Number</u>	<u>Category Name</u>	<u>Explanation of the Category</u>
10	No path	This multiple-item category was used for recording instances of reference to a gap in the network of paths in the maze. A variety of terms were used by subjects to refer to such gaps, notably "<u>space</u>" "<u>gap</u>" and "<u>no path</u>".
11	Symmetry	This lexical category was used for recording instances of reference to a figural resemblance between parts of the maze structure, such as "<u>symmetry</u>", "<u>opposite</u>" and "<u>parallel</u>".
12	Count	This was a highly frequently-used lexical category which was used to record instances of cardinal counting; an example given earlier was the description " . . . <u>1, 2, 3, 4,</u> F-(5, 1) s-: "Any separate instance of a cardinal count being used would result in the current count in this category being incremented. The exception to this was two-place coordinates, which were recorded in a separate lexical category (number 33). Other

Category Number	Category Name	Explanation of the Category
12	Count (continued)	examples where an increment to the value of the count in this category would be made are examples such as "I'm <u>two</u> along from the left on the top row".
13	Order	This was another heavily-used lexical category, and its count was incremented whenever a subject used an ordinal term (namely, first, second, third, fourth, fifth, sixth, or last). For example, "I'm <u>second</u> top row, three from the right".
14	Node	This multiple-item lexical category was used for recording instances of usage of expressions referring to nodes in the maze structure. Expressions used to refer to nodes included <u>"node",</u> <u>"box",</u> and, more infrequently, <u>"place"</u> or "<u>one</u>".
15	Move	This multiple-item lexical category was used to record instances of usage of instructions regarding visual scanning of the form "<u>Go</u> to the bottom left hand corner. <u>Move</u>

Category Number	Category Name	Explanation of the Category
	Move (continued)	one along to the right . . .". 'Move' and 'go' were the principal lexical items counted as belonging to this category.
16	Along	This multiple-item lexical category was used to record instances where a subject specified a horizontal movement, as in, for example, "I'm three <u>in</u> from the right . . ." "Three <u>across</u> . . .". "<u>In</u>" "<u>along</u>" and "<u>across</u>" were the principal lexical items counted here.
17 18 19 20	Up Down Left Right	Lexical categories 17, 18, 19 and 20 were used for recording instances of specification of path movements in descriptions, for example, "I'm two <u>to the left of</u> you", "I'm three <u>up from</u> the monster", and so on. Such cases were quite distinct from cases where the edge names 'left' and 'right' were specified, and could readily be disambiguated from them by reference to the computer printout showing the particular game state in question.

Category Number	Category Name	Explanation of the Category
21	To Edge (continued)	This multiple-item lexical category was used for recording instances where distances were specified relative to an edge- for example, "I'm third _from_ the right". The lexical items "to" , "on" and "from" were the principal exemplars of this category.
22	Line	This was a single-item lexical category and was used for recording instances of usage of the lexical item _"line"_.
23	Row	This single-item lexical category was used for recording instances of usage of the word "_row_".
24	Column	This single-item lexical category was used for recording instances of usage of the word "_column_".
25	Horizontal	These three single-item lexical categories were used for recording instances of usage of the lexical items in their respective category names, for example, "I'm _diagonally_ one away from my goal . . ." or "second _vertical_ line from the left
26	Vertical	
27	Diagonal	

Category Number	Category Name	Explanation of the Category
		. . ."
28	End	This multiple-item lexical category was used for recording instances of reference being made by subjects to the end of a line of nodes, as in, for example, "I'm in the last box on that line", or " . . . the very end node". The terms 'last' and 'end' were the principal lexical items which were counted in this category, but the term 'outside' was also used by subjects in this fashion.
29	Extreme	This multiple-item lexical category was used to record instances of usage of lexical items qualifying the 'end' of a line of nodes (see category 28). An example would be "I'm at the very end", or " . . . the extreme edge".
30	Groundfl	This multiple-item lexical category was used very infrequently, in fact being used by only one pair of subjects to metaphorically describe row positions and was of the form "I'm on the first floor line, say, rather than the ground floor line".

Category Number	Category Name	Explanation of the Category
31	Rowpcol	This is another infrequently-used lexical category and was used to record instances where subjects specified their position in terms of both row and column (row plus column) as in "I'm at row five column two"). This type of description is similar in nature to matrix- or coordinate-type descriptions but was infrequently observed.
32	Dfedges	This is another infrequently-used category and was used for recording instances where subjects described their positions in terms of distances away from two edges, for example, "I'm three from the top, two from the right". Again, there is a similarity to matrix-type descriptions and again this is an infrequently-used strategy for generating location descriptions.
33	Coords	This frequently-used lexical category was used to record instances of coordinate descriptions

Category Number	Category Name	Explanation of the Category
		being generated by subjects, for example" (2, 1)", "A5", "C3," etc.
34	ABCDEF	This lexical category was used to record instances of cases where subjects 'counted' out the letters of their coordinate system, for example: "Let's see I'm in: A,B,C,D, D5."
35 36 37	Start Ownother Monster	These three categories were used to record instances where subjects described their location relative to a previously specified position, which might be a start or goal position (category 35), one's own or one's partner's position (category 36) or the monster's position (category 37). For example, "I'm three to the left of you "(category 36), "I'm one above the monster" (category 37) or "I'm right back to where I started" (category 35).
38	Junction	This very infrequently-used single-item lexical category was used for recording instances of usage of the term 'junction' as in, for example, "I'm at the junction of the T".

CHAPTER 7

THE SEMANTICS AND PRAGMATICS OF LOCATION DESCRIPTION - II: DATA ANALYSES

a) Introduction

In the previous chapter, an exposition was given of the nature of subjects' descriptions of locations within the maze. The formal statistical treatment of the subjects' location descriptions was also outlined. This formal treatment consisted of two aspects, i) a descriptive aspect (in which subjects' location descriptions were themselves described in terms of their usage of each of thirty-eight lexical categories) and ii) a classificatory aspect (in which the lexical categories were classified into groups (clusters) by means of hierarchical cluster analysis). The hierarchical cluster analysis was conducted in order to ascertain which clusters or groupings of lexical categories could be identified. Such groupings could then be examined in the light of the earlier informal four-fold classification of location descriptions.

Since the technique of hierarchical cluster analysis to some extent imposes a hierarchical partitioning on the data, it is necessary to compare the results from several analyses in order to assess which particular clusters of variables (in this case, lexical categories) occur consistently across different cluster analyses. One could be more confident about the reality of these consistently-occurring groups: they would be less likely to be an artefact of the technique of cluster analysis and would be more likely to reflect the actual structure of the data.

Therefore, several different cluster analyses of the lexical categories were conducted and compared in order to ascertain which clusters of lexical categories were consistently revealed by different clustering techniques. It is also possible to conduct cluster analyses of individual subjects, the object of these analyses being the assessment of whether consistently-occurring groupings of subjects could be identified on the basis of their having used location descriptions of a similar nature. Again, several different cluster analyses were conducted and compared in an attempt to ascertain which groupings of individual subjects occurred consistently across different cluster analyses.

Finally, it was also noted that it would be of interest to relate the results from the cluster analyses of subjects to the results from the cluster analyses of lexical categories. The purpose of this analysis would be the assessment of whether particular identifiable groupings of subjects were generating location descriptions which were themselves readily explicable in terms of the clusters of lexical categories revealed by the first set of analyses. That is, it would be of interest to know whether the groupings of subjects revealed by the cluster analyses of individual subjects [illegible] sampled from the set of lexical categories in a manner consistent with the groupings obtained in the cluster analysis of lexical categories (e.g. cluster X of subjects using only those lexical categories present in cluster F and G of lexical categories and no others, rather than sampling one or two lexical categories from each of the clusters of lexical categories).

In the present chapter, we will consider the results obtained from (a) the cluster analyses of lexical categories and (b) the cluster analyses of subjects, <u>after</u> a cluster comparison technique (Gordon, 1981, see appendix IV) had been employed. That is, only those groupings which are consistently revealed by more than one cluster analysis will be presented. Following this, an attempt will be made to fit the two sets of results ((a) and (b) above) together.

b) Cluster analysis of lexical categories

As noted in the last chapter, the Biomedical computer programs (BMDP) statistical software package was used to analyze the data, and the particular program involved was program P1M (Hartigan, 1981, in Dixon et al. 1981), which cluster analyzes a set of variables. In the case of the present data, the thirty-eight lexical categories (in terms of which subjects' location descriptions were themselves described) were treated by the BMDP P1M program as thirty-eight variables and were analyzed to yield clusters of lexical categories. In program P1M, it is possible to manipulate both the measure of similarity or distance between variables on which the clustering is based, and also the criterion or linkage rule according to which clusters are formed (see appendix IV of this thesis and Gordon (1981) for an explanation of the principles behind cluster analysis). Two measures of similarity, and two measures of distance between variables are available in program P1M. The measures of similarity are (i) the correlation between the variables and (ii) the absolute value of the correlation between the variables, and the measures of distance available are (iii) the angle between the two variables (i.e. the arccosine of the correlation between the variables) and (iv) the acute angle corresponding to the arccosine of the absolute value of the correlation between the variables.

Three criteria for combining clusters are available in program P1M. These are (i) the minimum distance or maximum similarity over all pairings of the variables between the two clusters (single linkage), (ii) the maximum distance or minimum similarity over all pairings of the variables between the two clusters (complete linkage) and (iii) the average distance or average similarity over all pairings of the variables between the two clusters (average linkage).

In the first analysis of the groupings of lexical categories, four cluster analyses were conducted, all using the average linkage clustering criterion and each using one of the four measures of similarity or distance between variables noted above (see table 7-1 below).

Analysis (a)	Chosen Clustering Criterion	Chosen Measure of Similarity or Distance
(i)	Average linkage	Correlation
(ii)	Average linkage	Absolute correlation
(iii)	Average linkage	Angle of correlation
(iv)	Average linkage	Absolute angle of correlation

Table 7-1. The four cluster analyses conducted in analysis (a), varying the measure of similarity/distance between variables for a given clustering criterion.

Thus, there were four cluster analyses, involving the average linkage clustering criterion coupled with each of the four measures of similarity/distance between variables. These four cluster analyses were compared using Gordon's (1981; see appendix IV of this volume) cluster comparison technique.

The second analysis of the groupings of lexical categories involved a further three cluster analyses, the results from which were again compared with each other. In this case, the measure of similarity between variables was held constant (the correlation measure of similarity between the variables being used in all three analyses) and the criterion on which clustering was based was varied. This resulted in three cluster analyses (see table 7-2, below):

Analysis (b)	Chosen Clustering Criterion	Chosen Measure of Similarity or Distance
(i)	Single linkage	Correlation
(ii)	Average linkage	Correlation
(iii)	Complete linkage	Correlation

Table 7-2. The three cluster analyses conducted in analysis (b), varying the clustering criterion for a given measure of similarity.

All four cluster analyses in analysis (a) above were compared with one another, and the resulting consistently-occurring clusters appear in the left-hand column of table 7-3 (overleaf). Likewise, the results from the three analyses of procedure (b) above were compared. The results of this comparison are shown in the right-hand column of table 7-3. The results of the two sets of analyses (a) and (b) are shown side by side for easy comparison.

The analyses which vary the measure of similarity on which clustering is based (i.e. analysis (a)) yield six consistently-occurring clusters of lexical categories, whilst the analyses which vary the clustering criterion and hold constant the measure of similarity (analysis (b)) yield nine consistently-occurring clusters. What is striking is the similarity between the results from these two sets of analyses. Examination of table 7-3 reveals the following (see table 7-3):

(i) Clusters (1) and (2) in analysis (b) taken together are identical to cluster (1) in analysis (a).

(ii) Cluster (8) in analysis (b) is identical to cluster (2) in analysis (a).

(iii) Cluster (4) in analysis (b) is identical to cluster (3) in analysis (a).

(iv) Cluster (5) in analysis (b) is identical to cluster (4) in analysis (a), except that the lexical category named 'start' is included in cluster (5) of analysis (b) in two analyses out of three. However, this lexical category is also included in cluster (6) of analysis (b) on two cluster analyses out of three and is therefore likely to be intermediate between clusters (5) and (6) of analysis (b).

(v) Cluster (3) in analysis (b) is identical to cluster (5) in analysis (a).

(iv) Cluster (6) in analysis (b) is identical to cluster (6) in analysis (a), except for the lexical category 'start' discussed above.

(vii) Two clusters are yielded on at least two analyses out of three in analysis (b), these being the clusterings (7) and (9) of analysis (b), which do not have counterparts in analysis (a).

The striking similarity between the overall results from these two sets of analyses points to the existence of a definite structure in the data, and claims can be made for the veridicality of the clusters numbered (1) to (6) in analysis (a) of table 7-3, since they are identical to clusters (1), (2), (3), (4), (5), (6) and (8) of analysis (b). Furthermore, each of these groupings of lexical categories is readily explicable in terms of the intuitive analysis of place descriptions presented in the last chapter. Each of these clusters will be discussed in the light of this analysis.

Considering cluster (1) of analysis (a) in table 7-3, which corresponds to clusters (1) and (2) of analysis (b) in table 7-3, the grouping of lexical categories is as follows: the categories, TOP, ORDER, ROW, BOTTOM, LEFT, TO EDGE, RIGHT and LINE cluster together. This cluster therefore contains lexical categories corresponding to all of the edge names, ordinal counting, prepositions, and the terms 'row' and 'line'. This cluster of lexical categories thus contains all those lexical categories involved in the generation of row-type descriptions. For example:

"I'm on the	<u>third</u>	<u>row</u>	<u>from</u>	<u>the</u>	(<u>Bottom</u> or <u>Top</u>)
	↓	↓	↓		↓
	(ORDER	(ROW)	(TO EDGE)		(BOTTOM or TOP)

<u>second</u>	<u>from</u>	<u>the</u>	(<u>Left</u> or <u>Right</u>)
↓	↓		↓
(ORDER)	(TO EDGE)		(LEFT or RIGHT)

Thus, a prototypical row-type description could be generated using the lexical categories clustered together in cluster (1) without the use of any other lexical categories from other clusters being necessary. Cluster (1) of analysis (a) in table 7-3, therefore, does appear to agree well with our earlier intuitive analysis, appearing to correspond to that grouping of lexical categories used in row- or line-type descriptions.

Considering cluster (2) of analysis (a) in table 7-3, which corresponds to cluster (8) of analysis (b) in table 7-3, the clustering of lexical categories is as follows: categories COORDS and ABCDEF group together. This grouping makes good intuitive sense in terms of our earlier informal classification of description types, and there are numerous coordinate-type descriptions in the data which use these two lexical categories and no others. For example:

"I'm at A, B, C, D, E,:F-:E5"
(ABCDEF) (COORDS)

The third major cluster in analysis (a) in table 7-3, which corresponds to the fourth major cluster in analysis (b), consists of the lexical categories COUNT, UP, ALONG, DIAGONAL, and DOWN. This clustering of lexical categories could correspond to the lexical items used in the more elliptical path-type descriptions (see chapter 6). Elliptical path descriptions often omit the starting point for the path movement which comprises the description. For example:

"I'm three along, three up".
(COUNT) (ALONG) (COUNT) (UP)

The lexical category DIAGONAL included in this cluster is used by subjects in a variety of contexts, one of which is in the form of a 'short cut' of such path descriptions as the example above, as in the following example:

"Diagonally I'm one up from the bottom left corner".
(DIAGONAL)

Cluster (4) of analysis (a) in table 7-3, corresponding to cluster (5) in analysis (b) in table 7-3, contains lexical categories which are associated with figural-type descriptions. This cluster also contains lexical categories corresponding to lexical items used in path-type descriptions. The cluster is as follows: SHAPE, HORIZONTAL, MIDDLE, MOVE, NODE and LIMB. It has already been noted (see chapter 6) that it is difficult to describe an entire maze in terms of figures or shapes and so path descriptions are employed in addition to figural descriptions. A sample description which uses lexical items corresponding to the above lexical categories in cluster (4) of analysis (a) would be:

"I'm in the middle of that T"
(MIDDLE) (SHAPE)

or

"You see that limb? Move horizontally along one, I'm there".
(LIMB) (MOVE) (HORIZONTAL)

The small cluster (5) in analysis (a) in table 7-3 appears as cluster (3) in analysis (b), and consists of two lexical categories, HANDSIDE and CORNER. Lexical items corresponding to these two categories are often used by subjects as parts of larger descriptions, and are used in a variety of path- and figural-type descriptions, such as the following example:

"From the bottom left-hand corner I'm one along, one up".
(HANDSIDE) (CORNER)

or

"At the top right hand corner of the maze there's a wee T shape".

(HANDSIDE)(CORNER)

The last cluster common to both analysis (a) and analysis (b) of table 7-3 consists of four lexical categories and appears as cluster (6) of analysis (a) and cluster (6) of analysis (b). It consists of the lexical categories LEFT, OWNOTHER, START and RIGHT. This cluster would seem to reflect the groupings of lexical items found in the often-used strategy of describing one's location in terms of a previously specified position, as in the following example:

"I'm to the left of you"

(LEFT) (OWNOTHER)

or

"I'm just to the right of where I started".

(RIGHT) (START)

These six clusters occur consistently across two sets of analyses. Intuitively speaking, each cluster consists either of a group of semantically related lexical categories, or of a group of lexical categories which grouping is readily explicable in terms of the manner in which the subjects generate place descriptions.

Thus, the clusters yielded are both consistent in occurrence and do appear to correspond to the sorts of phrases which subjects do in fact employ in the data.

The next analysis conducted was the cluster analysis of individual subjects on the basis of their usages of the lexical categories. In this analysis, the aim was to ascertain whether any clearly identifiable groupings of individuals could be discerned such that all of the individuals within a cluster employed similar types of description. It is to a summary of the results of this analysis that we now turn.

c) Cluster analysis of individual subjects

As noted in chapter 6, the cluster analysis of individual subjects was conducted on the same data matrix as was used for the cluster analysis of lexical categories: the 38 lexical categories X96 subjects data matrix was mathematically transformed through 90^{o} to yield a 96 subjects X38 lexical categories data matrix. This transformed matrix was again analyzed using BMDP program P1M, which treated each subject as a variable and performed the standard cluster analysis of variables.

As was the case with the analysis of lexical categories, two sets of analyses (holding the clustering criterion constant whilst varying the measure of similarity, and holding the measure of similarity constant whilst varying the clustering criterion, (a) and (b) respectively), were conducted. These are summarized in table 7-4:

Analysis (a)	Chosen clustering criterion	Chosen measure of similarity of distance
(i)	Average linkage	Correlation
(ii)	Average linkage	Absolute correlation
(iii)	Average linkage	Angle of correlation
(iv)	Average linkage	Absolute angle of correlation
Analysis (b)		
(i)	Single linkage	Correlation
(ii)	Average linkage	Correlation
(iii)	Complete linkage	Correlation

Table 7-4 The two sets of cluster analyses employed in the cluster analysis of individual subjects. Analysis (a) holds the linkage criterion constant and varies the measure of similarity on which clustering is based and analysis (b) varies the linkage criterion and holds the measure of similarity constant.

Thus, analysis (a) involved four cluster analyses, the results from which were compared with one another in order to ascertain which groupings of subjects occurred consistently, and analysis (b) involved three cluster analyses, the results from which were also compared with one another. The consistently-occurring groupings of subjects found in both sets of analyses are set out in table 7-5 (over). The consistently-occurring groupings of individual subjects found in analysis (a) are set out in the left-hand column of table 7-5, and the consistently-occurring groupings of individual subjects found in analysis (b) are set out in the right-hand column of table 7-5.

The following convention is employed for denoting individual subjects. Each dyad of subjects who played a game or games has a number assigned to it, for example 11 or 43. Following this number is a letter which denotes an individual game (A, B, C, D, E, or F). Thus, the subjects in dyad number 11 played two games, games 11D and 11F. The second letter after the dyad number denotes the individual player, A or B; player A always played the game using the 'Intel' terminal and player B always played the game using the 'Imlac' terminal. Thus, subject 11DB is player B in dyad 11 playing game D; subject 43AA is player A in dyad 43 playing game A, whilst subject 43AB is the partner of subject 43AA in game A of dyad 43; and so on (see table 7-5, overleaf).

As can be ascertained from an inspection of table 7-5, there is a striking similarity in the clusters consistently yielded across both sets of analyses (for ease of comparison, the order of items within the clusters of analysis (b) has been altered slightly in some cases to conform with the ordering of the corresponding items within the clusters of analysis (a): this has been done to facilitate comparison of the two sets of analyses, since the fact that a particular set of items cluster together is more important from the present point of view than is the internal structure of each cluster).

There is a close correspondence between the result from the two sets of analyses in the case of all clusters. Where differences do exist, they are small ones, usually being the inclusion of one or two subjects in a particular cluster in one analysis but not in the corresponding cluster of the other analysis (for example, subject 43AB is included in cluster (1) of analysis (b) but is not included in cluster (1) of analysis (a)). Thus, the two clusters numbered (1) in both analyses are identical except for the fact that cluster (1) of analysis (b) contains the 'extra' subject 43AB.

The two clusters numbered (2) in both analyses agree in the case of 15 of the subjects comprising them, the difference between the two clusters being that cluster (2) of analysis (a) contains one subject (40AB) who is not included in cluster (2) of analysis (b) (see table 7-5).

The cluster numbered (3) in both analyses agree on eight out of the nine subjects in each cluster, the solitary 'extra' subjects not common to cluster (3) in both analyses being subject 11FB in cluster (3) of analysis (b) and subject 43AB in cluster (3) of analysis (a) (see table 7-5).

The two clusters numbered (4) in both analyses (see table 7-5) are identical except that cluster (4) in analysis (b) includes subject 29BB, whereas the cluster numbered (4) in analysis (a) does not.

The two clusters numbered (5) in both analyses show the greatest number of differences, agreeing on the clustering together of eight subjects (24BB, 41AB, 29BA, 33AA, 35BB, 40BA, 33AB, and 35AB), and differing from one another on a total of 5 subjects; analysis (a) includes subjects 40BB, 31BA, and 35BA whereas analysis (b) does not, and analysis (b) includes subjects 11FA and 38AA where analysis (a) does not.

Cluster (6) in analysis (a) is absolutely identical with cluster (6) in analysis (b); agreement on the inclusion of all 26 subjects within one cluster is perfect.

Since the two sets of six clusters in the two sets of analyses agree so closely, there is a good basis for concluding that there is a definite structure in the data. There is a substantial core of the 96 subjects who are consistent in their choice of description type to such an extent that they are consistently classified into six groups by several different clustering criteria. An investigation of the basis for this consistency is reported in the next section.

Before leaving the cluster analyses of individual subjects, however, there are one or two further points of note. Notice, for example, that in cluster (1) of both analyses in table 7-5 that both subjects in both games played by dyad number 10 cluster together. This suggests that both subjects within this dyad are repeatedly using the same type of description as one another. Other cases where both members of a dyad playing a particular game cluster together are games 13A and 13C (cluster 2), 17D and 17F (cluster 2), 18D and 18F (cluster 2) 11D, 43B and 38B (cluster 3) 15E, 27B, 37A and 37B (cluster 4) 33A (cluster 5), and games 20B, 20D, 23B, 26B, 27D, 28B, 28C, 19F, 42A, 44A, 44B, 25B and 42B (cluster 6). In the case of all of these 29 games, both of the subjects within a dyad and playing a particular game consistently cluster together on the basis of the descriptions they generate. This consistency is striking and suggests that subjects are linguistically entraining one another to some extent, that is, both subjects within a dyad playing a particular game are using the same type rather than different types of location description as one another.

Thus, the cluster analyses of individual subjects reveal consistently-occurring clusters of individual subjects such that all individuals within a cluster use the same type of description. Furthermore, the above observation suggests that there is a strong tendency for both members of a dyad to group together. This supports the impression one gets on reading transcripts of games that each individual subject was consistent with respect to the type of description he or she used, and that the two members of a dyad were broadly consistent with respect to one another's choice of description type.

We now turn to examine the basis for the above-revealed consistency in the composition of the clusters of individual subjects.

d) The lexical content of subjects' descriptions: the relation between the cluster analysis of individual subjects and the cluster analysis of lexical categories

In section (b) of this chapter, the results of several cluster analyses, when compared with one another, revealed a consistently-occurring set of clusters of lexical categories. An examination of the content of each of these clusters of lexical categories suggested that the particular groupings of lexical categories which were revealed by the analyses partly reflected subjects' usages of lexical items corresponding to the lexical categories. That is, each cluster seemed to comprise a grouping of lexical categories which made sense, intuitively speaking, in relation to our earlier informal classification of subjects' location descriptions (see chapter 6). A cluster of lexical categories might, for example, group together those lexical categories corresponding to the types of lexical item found in line-type descriptions.

In section (c) of the present chapter, the results of two sets of cluster analyses of individual subjects revealed a set of consistently-occurring clusters comprising groupings of subjects. Subjects within a cluster were grouped together on the basis of their having used similar description types.

In the present section, we turn to examine exactly how each of the clusters of subjects (see section (c) of this chapter) samples from (i.e. uses lexical items corresponding to) each of the lexical categories in the different clusters of lexical categories (see table 7-3). That is, we wish to ascertain whether the lexical categories used by the clusters of subjects relate in a systematic way to the clusters of lexical categories: do each group of subjects use lexical items corresponding to lexical categories which are all grouped together in the cluster analyses of lexical categories, or do the subjects sample more randomly?

Reference to table 7-5 reveals six groupings of subjects. The data for each subject in each group was examined in order to ascertain which categories in the classification of lexical categories were used by that subject. The results of this examination of each subject's data (the frequencies with which the subject used lexical items corresponding to each of the lexical categories) were summed across all subjects in a given cluster of subjects (see table 7-5) in order to ascertain which lexical categories were used by the group as a whole. The results of this analysis are shown in table 7-6 (over).

In table 7-6, the left-hand column shows the consistent clusters of lexical categories already tabulated in table 7-3 and discussed in section (b) of this chapter. These clusters of lexical categories are lettered (a)-(l). The numbered columns ((1) - (6)) in the main body of table 7-6 correspond to the six clusters of subjects revealed by the analyses presented in section (c) of the present chapter. The ticks (✓) located alongside the different lexical categories represent those categories used at least once per game, on average, by each subject in the group.

The first cluster of subjects, (cluster (1) in both analysis (a) and analysis (b) in table 7-5 and comprising subjects 10BA, 10BB, 10GA, 10GB and 41AA) use the following lexical categories (see table 7-6); TOP, ORDER, ROW, BOTTOM, LEFT, TO EDGE, RIGHT, LINE, COUNT, ALONG, UP and START. That is, they use all of the lexical categories in clusters (a) and (b) of the lexical category classification (see table 7-6) and selected items from cluster (e). The other item they use is 'start' from cluster (g) of the lexical category classification. Thus, these subjects appear to be using row-descriptions, using some path movement terms, and occasionally describing their positions in terms of path movements relative to previously specified goal or start positions.

Example descriptions from individual subjects within this cluster back up the above interpretation:

e.g. "I'll go up to the third box on the: fifth row".

(MOVE)(UP) (ORDER)(NODE) (ORDER)(ROW)

(Subject 10BA)

"I'm on the top line: fourth box along" (10GB)

(TOP)(LINE)(ORDER)(NODE)(ALONG)

As can be seen from table 7-6, the description of one's position relative to a previously specified start or goal position is quite common, since the subjects in clusters 2, 4 and 5 of tables 7-5 and 7-6 do this also.

Cluster (2) of subjects (see table 7-6) use the following lexical categories: TOP, ORDER, ROW (cluster (a) in the cluster analysis of lexical categories in table 7-6), COUNT, ALONG, DOWN (from cluster (e) of the cluster analysis of lexical categories in table 7-6), plus the category START from cluster (g). The particular lexical categories used by this group of subjects would suggest that they are also using row-type descriptions, but row-type descriptions of a more elliptical nature than those used by the subjects comprising cluster (1) of table 7-6. The lexical categories used would suggest descriptions being couched in terms of 'row plus counts along that row'. Examination of the data of subjects in cluster (2) bears this out:

e.g. "I'm on the top row, third left". (17FA)
(TOP) (ROW) (ORDER) (LEFT (EDGE))

"Second row, second from the left" (13AA)
(ORDER)(ROW) (ORDER)(TO EDGE)(LEFT EDGE))

Thus, according to the present analysis, the subjects in clusters (1) and (2) of the cluster analysis of subjects both use row-type descriptions but of slightly differing types, those used by cluster (2) of subjects being slightly more elliptical in nature. The subjects in cluster (2) of the cluster analysis of subjects would seem to specify the edges relative to which the row descriptions are being generated less often than the subjects in cluster (1), according to the present analysis.

Cluster (3) of subjects (see table 7-6) use the following lexical categories; ORDER, ROW (cluster (a) in table 7-6), LINE (cluster (b)) and ALONG (cluster (e)). This combination of categories would again suggest row-type descriptions, in this instance of a highly elliptical nature. An example from a subject in this group bears this out:

". . . the third: row: third along. . ." (43AA)
(ORDER)(ROW) (ORDER)(ALONG)

Thus, the subjects in the first three clusters of the cluster analysis of individual subjects all use row-type descriptions, but of slightly differing types.

The subjects in cluster (4) of the cluster analysis of individual subjects (see table 7-5) would appear to use the following lexical categories (table 7-6); TOP, ORDER (cluster (a) of the cluster analysis of lexical categories) BOTTOM, LEFT, TO EDGE, RIGHT (cluster (b) of lexical categories) COUNT, ALONG, UP (cluster (e)), HANDSIDE (cluster (c)) and START and OWNOTHER (cluster (g)). This collection of lexical categories would appear to suggest that the subjects in this group generate path-type descriptions; both types of counting, edge names and the movement terms 'along' and 'up' are included among the categories used by this cluster of subjects, as are previously-specified-position-denoting categories (start/goal and own/other's previous position). The term 'row', furthermore, is not used by this group of subjects. Typical descriptions generated by some of the subjects in cluster (4) lend support to the conclusion that this group of subjects generate path-type descriptions:

e.g. (37BB) "I'm just one along from you: to the:

(COUNT)(ALONG) (OWNOTHER)

to the left: (3.5 sec)".

(LEFT (direction))

(37BB) "I'm four places up from the bottom and

(COUNT)(NODE)(UP)(TO EDGE)(BOTTOM)

I'm three columns in from the right".

(COUNT)(COLUMN) (TO EDGE) (RIGHT)

(15EA) ". . . One along and one up from the

(COUNT)(ALONG) (COUNT)(UP)(TO EDGE)

left-bottom left"

(LEFT)(BOTTOM)(LEFT (EDGE))

(15CA) "Top left . . go along two . . to the

(TOP)(LEFT) (MOVE)(ALONG)(COUNT)

(EDGE)

right and down one".

(RIGHT (direction)) (DOWN) (COUNT)

These typical descriptions from some of the subjects who group together in cluster (4) of the cluster analysis of individual subjects confirm that these subjects do indeed use path descriptions.

Cluster (5) of individual subjects (see tables 7-5 and 7-6) use the following lexical categories: TOP (from cluster (g) of the cluster analysis of lexical categories), BOTTOM, LEFT, RIGHT, LINE (cluster (b)) HANDSIDE (cluster (c)) COUNT, ALONG, UP, DOWN (cluster (e)), SHAPE, MIDDLE (cluster (f)) and LEFT, START, OWNOTHER and RIGHT (cluster (g)). This is quite a varied selection of lexical categories from the different clusters. Noticeable for their absence from the set of lexical categories used by cluster (5) of subjects are the lexical categories 'row' and 'order'. the lexical categories used by this group of subjects include path terms (along and up) edge names, and the figural term 'shape'. This would seem to indicate that these subjects are describing their locations in terms of a mixture of path- and figural-type descriptions (see chapter 6).

Sample descriptions taken from protocols from the subjects in this group would seem to bear this impression out. For example:

(35BA) " . . My asterisk is at the: (1s) sort of: (1s) last part the <u>L</u>: see what I mean? In the <u>middle</u> <u>square</u>

(SHAPE) (MIDDLE) (NODE)

of that <u>line</u>".

(LINE)

Furthermore, the above analysis suggests that subjects extensively use prespecified positions to generate descriptions (that is, describing locations relative to a previously specified start, goal, the other player's or one's own previous, position). This is also confirmed by an examination of the data:

(34BA) "See where I started? . . I'm one along to the

(START/GOAL) (COUNT)(ALONG)

left from that.

(LEFT (direction))

(35BB) "I'm one below where I started from".

(START)

Thus, the subjects comprising cluster (5) of the cluster analysis of individual subjects do appear to be using the earlier-noted (see chapter 6) mixture of path- and figural-type descriptions.

The sixth cluster of subjects revealed by the cluster analysis of subjects use only four categories of lexical item - COORDS and ABCDEF from cluster (d) of the cluster analysis of lexical categories (see table 7-6), COUNT (cluster (e)) and OWNOTHER (cluster (g)). In other words, this would appear to be the group of subjects who use matrix- or coordinate-type descriptions. As with other types of description, these subjects appear to sometimes describe their positions relative to a previously specified position (of the self or of one's partner).

This is also done by the subjects in clusters 1, 2, 4 and 5 of the cluster analysis of subjects.

The impression that cluster (6) of subjects use coordinate-type descriptions is again borne out by an inspection of their data:

e.g. (20BA) "I am (1,5sec) presently at: C:5 (4 sec) O.K.?

(COORDS)

(44BB) "O.K. I'm now then on three: 1, 2, 3:

(COUNT)

4, (3, 4) O.K.?"

(COORDS)

(26BA) ". . . Ehm: A, B, C, D, E,: 3 (6 sec)"

(ABCDEF) (COORDS)

Thus, this group of subjects are indeed using coordinate-type descriptions, confirming what our analysis had suggested.

e) Conclusions

In conclusion, the results obtained from the cluster analyses substantially confirm our earlier, informal, observations. Firstly, the lexical item categories group together in a fashion which is consistent with the types of grouping one would expect from our earlier informal analysis. For example, all of the types of lexical categories corresponding to the lexical items used in row-type descriptions (edge names, prepositions, 'row', 'line' and ordinal counting) consistently group together. Secondly, the earlier observation that subjects are consistent with respect to the type of description they use, with one type of description (e.g. path-type descriptions) being used by a given dyad of subjects throughout a given game (rather than switching from description type to description type within a game) was confirmed by the two latter analyses in the present chapter. These analyses demonstrated that consistent groupings of subjects are revealed across different cluster analyses (the groupings being on the basis of the descriptions used by individual subjects) and each of these groups of subjects consistently use particular selections of lexical categories. The selection of lexical categories used by all of the members of each group relates in a systematic way to the earlier cluster analysis of lexical categories. Thus, for example, the subjects in group (1) use the lexical categories in clusters (a) and (b), plus three lexical categories from cluster (c) of the lexical category classification (see table 7-6) and also one lexical category from cluster (g). They do not sample from all of the clusters of categories.

One group of subjects (group (5) in tables 7-5 and 7-6) do sample from several of the clusters of lexical categories (see table 7-6). They do, however, sample the individual categories from the different clusters of lexical categories in a fashion consistent with our informal four-fold classification of descriptions presented in chapter 6: they use a variety of lexical categories associated with figural- and path-type descriptions but they <u>never</u> use terms like 'row' nor do they use the 'coordinates' category. Thus, although they do sample lexical categories from a large number of different clusters of categories (see table 7-6), they do not do so in an entirely random fashion, but instead in a fashion which is consistent with our earlier informal classification of the location descriptions.

The analyses presented in the present chapter thus indicate that the following conclusions can be drawn:

(1) The lexical categories, in terms of which subject's descriptions of locations were described, cluster consistently across several analyses to form groupings which are readily explicable in terms of our earlier, informal, four-fold classification of the location descriptions.

(2) When the subjects are treated as variables and cluster analyzed, consistently-occurring

groupings of subjects emerge across different cluster analyses. This suggests that there are clearly identifiable groups of subjects which are homogeneous with respect to the type of location description used by individual subjects within each cluster of subjects.

(3) The groups of subjects revealed in the cluster analysis of individual subjects use lexical items corresponding to the lexical categories in a systematic way which is consistent with the four-fold classification of descriptions into path-,line-,figural- and matrix-type descriptions presented in chapter 6.

Thus, one group of subjects (cluster (6) in table 7-6) appear to use coordinate- or matrix-type descriptions; another group of subjects (cluster (5) in table 7-6) appear to use the mixture of path- and figural descriptions discussed in chaper 6; a third group of subjects (cluster (4) in table 7-6) appear to use path descriptions exclusively; and a further three groups of subjects (clusters (1), (2) and (3) in table 7-6) appear to use row-type descriptions exclusively. The row-type descriptions would not, however, appear to be a unitary class of descriptions, since the analyses presented in this chapter distinguish three distinct classes of row description, the three classes varying in explicitness.

Overall, the analyses presented in this chapter entirely confirm the earlier informal classification of the location descriptions into four broad categories: line-, matrix-, path- and figural-type descriptions.

The next chapter will be devoted to a further examination of the four types of description. Firstly, some hypotheses which were first raised in chapter 6 will be tested empirically: in particular, the notion that matrix-type descriptions involve greater amounts of hesitating and explicit counting than do the other three types of description will be subjected to empirical scrutiny.

The final section of the next chapter will be devoted to a reconsideration of the four generic description types. The possibility that they each represent a class of _mental models_ of the maze shape ('models' of a type similar to those described by Johnson-Laird (1980, 1981a, 1981b)) will be considered and this interpretation of the data will raise some theoretically interesting issues which will be taken up again in chapters 10 and 11.

Table 7-3 Clusters arising consistently in different analyses

(a) Across variation in measures of similarity/distance		(b) Across variation in cluster linkage	
(1)	TOP	(1)	TOP
	ORDER		ORDER
	ROW		ROW
	BOTTOM	(2)	BOTTOM
	LEFT (edge)		LEFT (edge)
	TO EDGE		TO EDGE
	RIGHT (edge)		RIGHT (EDGE)
	LINE		LINE
(2)	COORDS	(3)	HANDSIDE
	ABCDEF		CORNER
(3)	COUNT	(4)	COUNT
	ALONG		UP
	UP		(ALONG)
	DIAGONAL		(DIAGONAL)
	DOWN		(DOWN)
(4)	SHAPE	(5)	SHAPE
	HORIZONTAL		HORIZONTAL
	MIDDLE		(MIDDLE)
	MOVE		MOVE
	NODE		(NODE)

LIMB

(5) HANDSIDE
CORNER

(6) LEFT (direction)
START
OWNOTHER
RIGHT (direction)

(LIMB
(START)

(6) LEFT (direction)
OWNOTHER
(START)
RIGHT (direction)

(7) EXTREME
ROWPCOL

(8) COORDS
ABCDEF

(9) (GROUNDFL)
(COLUMN)

N.B. In analysis (a) (left-hand column) groups (1) and (2) cluster together to form one large cluster in three analyses out of four. In analysis (b) (right-hand column) those lexical categories in parentheses group with the others on two analyses out of three, whilst all others group together on all three analyses.

Table 7-5 Clusterings of subjects consistently occurring across different cluster analyses)

(a) Across variation in measure of similarity or distance between variables	(b) Across variation in cluster linkage criteria
(1) 10BA,10GA,10BB,10GB,41AA,	(1) 10BA,10GA,10BB,10GB,41AA,43AB,
(2) 13AA,17FA,13CA,17DB,13AB, 13CB,31CA,17FB,17DA,18DB, 18FB,18FA,41BB,18DA,31BB, 40AB	(2) 13AA,17FA,13CA,17DB,13AB, 13CB,31CA,17FB,17DA,18DB, 18FB,18FA,41BB,18DA,31BB
(3) 11DA,11DB,41BA,43BB,43AA, 43BA,38BA,38BB,43AB	(3) 11DA,11DB,41BA,43BB,43AA, 11FB,43BA,38BA,38BB
(4) 15EA,32BA,37AA,27BA,15EB, 39BA,27BB,37AB,37BB,39BB, 37BA	(4) 15EA,32BA,37AA,27BA,15EB, 39BA,27BB,37AB,37BB,39BB, 37BA,29BB
(5) 24BB,41AB,29BA,33AA,35BB, 31BA,40BA,40BB,35BA,33AB, 35AB	(5) 11FA,24BB,41AB,29BA,33AA, 35BB,40BA,33AB,35AB,38AA,
(6) 20BA,27DA,25BA,28BB,28CB, 27DB,19FA,23BA,20DA,42AA, 26BB,23BB,28CA,44BB,28BA, 44BA,44AA,44AB,25BB,26BA, 42AB,20BB,42BA,20DB,19FB, 42BB	(6) 20BA,27DA,25BA,28BB,28CB, 27DB,19FA,23BA,20DA,42AA, 26BB,23BB,28CA,44BB,28BA, 44BA,44AA,44AB,25BB,26BA, 42AB,20BB,42BA,20DB,19FB, 42BB

KEY - The number refers to a particular dyad; the first letter after the number refers to a particular game; the second letter after the number refers to a particular individual player (A or B).

Table 7-6 Usages of lexical categories by subjects in clusters 1-6 of the cluster analysis of subjects

Grouping of lexical categories		Cluster of subjects (1)	(2)	(3)	(4)	(5)	(6)
(a)	TOP	✓	✓		✓	✓	
	ORDER	✓	✓	✓	✓		
	ROW	✓	✓	✓			
(b)	BOTTOM	✓			✓	✓	
	LEFT	✓			✓	✓	
	TO EDGE	✓			✓		
	RIGHT	✓			✓	✓	
	LINE	✓		✓		✓	
(c)	HANDSIDE				✓	✓	
	CORNER						
(d)	COORDS						✓
	ABCDEF						✓
(e)	COUNT	✓	✓		✓	✓	✓
	ALONG	✓	✓	✓	✓	✓	
	UP	✓			✓	✓	
	DIAGONAL						
	DOWN		✓			✓	
(f)	SHAPE					✓	
	MIDDLE					✓	
	MOVE						
	NODE						
	LIMB						

Table 7-6 Cont.

Grouping of lexical categories		Cluster of subjects (1)	(2)	(3)	(4)	(5)	(6)
(g)	LEFT					✓	
	START	✓	✓		✓	✓	
	OWNOTHER				✓	✓	✓
	RIGHT					✓	
(h)	EXTREME						
	ROWPCOL						

Ticked lexical categories are those used at least once per game by each subject within each cluster of subjects.

CHAPTER 8

SEMANTICS, MENTAL MODELS, AND THE DESCRIPTION OF MAZE LOCATIONS

a) Introduction

In the last chapter, the results from three sets of analyses were presented. These results were interpreted to formally confirm the earlier informal classification (see chapter 6) of subjects' descriptions of locations within mazes into four broad classes: path-type, line-type, figural-type, and matrix-type descriptions. The results presented in the last chapter indicate that:

(i) Consistently occurring clusters of lexical categories are found in several different cluster analyses, and these clusters comprise groupings of lexical categories such that the composition of each cluster is readily explicable in terms of the above four-fold typology.

(ii) Consistently occurring clusters of individual subjects are obtained in several different cluster analyses.

(iii) The groups of subjects revealed in analysis (ii) above can be shown to use lexical items corresponding to the lexical categories classified in analysis (i) above in a manner which is consistent with the four-fold typology presented above. Thus, for example, one group of subjects use lexical items corresponding to several different lexical categories in such a fashion that only those lexical categories which correspond to the terms used in path-type descriptions are in fact used. The different groups of subjects revealed by analysis (ii) sample from (that is, use lexical items corresponding to) the lexical categories in a highly consistent manner rather than a random manner, and the manner in which the different groups of subjects do sample from the lexical categories is readily explicable in terms of their using one of the four types of description listed above (i.e. either path-, line-, figural-, or matrix-type descriptions).

The analyses presented in the last chapter thus lend support to the view that there are four broad categories of descriptions of locations within mazes, representing four distinguishably different ways of describing locations within the maze.

In the present chapter, some further properties of the four types of location description will be explored. This empirical exploration will involve the testing of some hypotheses which were first postulated in chapter 6.

In short, the hypotheses postulate that matrix-type location descriptions should be <u>shorter</u> (in terms of the lengths of the individual utterances which comprise them, and/or in terms of the number of utterances comprising each description) and <u>less fluent</u> (in terms of involving a greater number of pauses, that is, speechless gaps during the production of a description) than path- and figural-type descriptions. (The rationale underlying these hypotheses is stated below). In testing these hypotheses, a comparatively crude analysis was conducted: the length of location description was measured in terms of the average number of words in each utterance and the average numer of utterances in each description, and the pauses during speech production were classified in a binary fashion into <u>hesitations</u> (these being speechless gaps of less than 1 second duration) and <u>pauses</u> (these being speechless gaps which lasted longer than 1 second). Pause lengths were measured to the nearest half second. The relative imprecision of the 'hesitation' and 'pause' measures is justified by the fact that no 'microscopic' hypotheses were being postulated (that is, hypotheses concerning the relative difficulties of psychological processing to the level of the individual word in the different types of description). The hypotheses postulated were rather more macroscopic in nature (involving comparison of the different types of description at the level of the utterance), and for this reason, a 'fine-grained' analysis with pauses being measured in a rather more precise manner than to the nearest half second was not necessary: the comparatively crude analysis used was sufficient.

In chapter 6, differences were noted between the various types of description in terms of the extent to which the actual configuration of paths within the maze constrained the description being generated. In particular, it was suggested that path- and figural-type descriptions involve the speaker and listener in attending to the particular configuration of paths present in the maze, whereas line- and matrix-type descriptions are more abstract in nature, involving the subjects in ignoring such features as gaps in the network of pathways in the maze.

A possible consequence of this would be a difference in the overall lengths of descriptions generated, with the more 'abstract' types of description being shorter in terms of the number of utterances employed in the description, or in terms of the number of words comprising each utterance. The reason for this difference would simply be that matrix- and line-type descriptions (the more 'abstract' types of description) employ a standard format for all descriptions, whereas path- and figural-type descriptions are lengthened because of subjects having to consider the particular features of the area being described (for example, having to acknowledge gaps in some of the pathways which have to be mentally 'traversed' in the execution of the scanning procedures corresponding to the description). The brevity of the standard description formats of the more 'abstract' types of description reaches its logical extreme in the two-place coordinates of matrix-type descriptions. Thus, a reasonable hypothesis would be that matrix- and line-type descriptions are shorter than are path- and figural-type descriptions.

It was also argued in chapter 6 that, since matrix-type descriptions are more abstract and brief, they require more 'work' on the part of both speaker and listener in encoding and decoding: thus, we hypothesize more pauses (i.e. periods of time during which speech is absent) during both the generation and the acknowledgement of understanding of matrix-type descriptions. For the same reason (that is, because more cognitive 'work' is hypothesized to be involved during the generation and comprehension of matrix-type descriptions) instances of explicit counting aloud should be more prevalent in the case of matrix-type descritions as compared to the other three types of description.

The final section of the present chapter will be devoted to an exploration of the details of, and some of the potential consequences of, the hypothesis that the four different types of description observed in the data from the maze game studies correspond to four ways of conceptualizing the maze shape, and are in fact manifestations of different mental models of the maze. This interpretation of the descriptions raises questions which will be examined in the final three chapters of this thesis.

b) <u>A comparison of the four types of location description in terms of their brevity and the relative prevalence of pausing, hesitation and explicit counting in each.</u>

In chapter 6, a broad outline was given in informal terms of the four basic categories of location description generated by subjects in the course of playing the maze game. These were termed the path-, line-, figural- and matrix-types of description. The actual descriptions themselves were interpreted in procedural terms; it was argued that the descriptions constituted a set of tacit instructions regarding visual scanning operations to be carried out by the listener. To recapitulate briefly:

(i) <u>Path</u> descriptions involved the subjects generating a description by means of firstly identifying a perceptually prominent point in the maze structure and then describing the desired location relative to that prominent point in terms of visual scanning movements along available paths (for example, "from the bottom left hand corner, I'm one along, one up"). The listener is thus tacitly instructed to firstly find the relevant prominent point (in this case, the bottom left hand corner of the maze) and then to 'move' (i.e. visually scan) in specified directions through specified distances.

(ii) Figural type descriptions involve the speaker identifying a particular region of the maze by means of drawing the listener's attention to a figure or shape within the network of pathways comprising the maze, and then specifying the desired location relative to that shape. For example, "At the top left of the maze there's an L shape: I'm at the last part of the L, if you see what I mean". In this description, the listener is tacitly instructed to look for a particular configuration of paths; having found that, his attention is then drawn to a particular part of the figure, and he is then told that this is where the speaker's position marker is located. Again, the description is readily interpretable in procedural terms.

(iii) Line type descriptions involve the speaker firstly identifying a particular line of nodes (a row or a column of nodes) within which the location being described is located, and then specifying a scanning 'movement' along that line until the desired location is reached. For example: "I'm on the third row from the bottom, third from the left". The listener is first directed to find a particular row of nodes (in this case, the third row of nodes from the bottom of the maze) and is then directed to scan along that row a certain distance from a specified edge (in this instance, the left-hand edge of the maze).

(iv) Matrix-type descriptions are implicitly procedural in nature. In this type of description, both subjects agree upon a coding system for denoting individual rows and columns, and the descriptions are given as two-place coordinates (for example, "A5"). When given such a description, the listener's task would be to firstly locate the relevant row and column relative to a previously agreed-upon origin, and then find the intersection of the two.

The matrix class of description, like the other broad classes of description, is not a unitary class; sub-variations on the basic principle exist. In fact two different types of matrix-type description can be distinguished. The first operates in the manner outlined above (the listener firstly finds the relevant row and the relevant column, and then seeks their intersection). The second type of matrix description functions in a similar manner to the Cartesian coordinate system of mathematics in that the two procedures corresponding to the two numbers of a given coordinate description are executed <u>contiguously</u>. For example, the description "(3, 2)" would be interpreted in this 'Cartesian' matrix scheme as follows: first a scanning movement is made in the horizontal plane through a distance of three nodes from an agreed-upon starting position (for example, the bottom left-hand corner of the maze shape). This would result in the scanner being focussed upon the third node from the left on the bottom row. Immediately afterwards, a vertical scanning movement of a distance equivalent to two nodes (or one path movement) is executed. In this case, there would be no need to seek the intersection of the corresponding row and column; contiguous execution of the two procedures would result in the arrival of the scanner at the desired node. This is identical to the manner in which the Cartesian coordinates of mathematics are used, and is rather like a highly elliptical path-type description, except, of course, that, as is the case with all matrix-type descriptions, gaps are 'crossed' without acknowledgement.

This 'Cartesian' type of matrix system contrasts with the first-mentioned type inasmuch as the first-mentioned type involves finding the relevant row relative to the origin, returning the scanner to focus upon the origin, finding the relevant column relative to the origo, and, having established the relevant column and the relevant row, their intersection is then sought. That is, the first-mentioned matrix system involves a return to the origin when establishing the extension of the second coordinate, whereas in the 'Cartesian' scheme, the procedures are executed contiguously, without a return to the origin when seeking the extension of the second coordinate, and without the necessity to seek an intersection of the two axes. As such, the 'Cartesian' system is simpler.

Thus, like the other classes of location description, the matrix-type descriptions are not a unitary class.

Having thus examined briefly the characteristics of each type of location description, several points are worthy of note.

Firstly, path- and figural-type descriptions are inherently less abstract than are line- and matrix-type descriptions. Path- and figural-type descriptions are constrained by the actual configuration of paths in the maze, whereas line-type descriptions involve treating the maze shape as if it were an array of parallel lines of nodes. Gaps within those lines and connections between the lines are ignored for the purposes of generating location descriptions. Matrix-type descriptions are even more abstract and involve treating the maze as a two-dimensional array of nodes, ignoring the actual configuration of paths entirely.

For this reason, we can hypothesize that path- and figural-type descriptions would be lengthier than line- and matrix-type descriptions. This would be because path- and figural-type descriptions are constrained by the configuration of paths in the maze structure involving the speaker in, for example, having to acknowledge gaps, or having to give a complex alternative description rather than a simple direct one which procedurally 'crosses' a gap (examples of this are given in chapter 6). Line- and matrix-type descriptions, on the other hand, have a 'standard format' which can be used regardless of local constraints in the path structure of the maze.

Secondly, since line-type descriptions abstract the horizontal (or vertical) parallel lines of the maze structure and effectively involve ignoring the vertical (or horizontal) connections between these lines, they are less abstract than are matrix-type descriptions, which involve ignoring the configuration of paths in both planes of orientation and instead involve treating the maze as an array of nodes. Thus, it can be argued that path- and figural-type descriptions are the least abstract types of description and are most constrained by the configuration of paths in the maze, line-type descriptions are of an intermediate level of abstraction from the maze structure, and matrix-type descriptions represent a conceptualization of the maze that is abstracted from its actual structure to the greatest degree.

A further difference exists, it could be argued, between the matrix-type descriptions and the other three types of description, and this difference is related to the abstractness of the matrix descriptions. The difference is one of what Marslen-Wilson, Tyler and Levy (1982) refer to as the _on-line_ _resolvability_ _of_ _referential_ _devices_. Put more simply, this means that a given referential device is understandable at the point at which it occurs in the speech stream.

This is most clearly exemplified by line-type descriptions. For example, "I'm third row from the bottom, third from the left". It is possible that a listener would have time to execute the scanning procedures implicit in this description _as the description is spoken_. Thus, as soon as he has heard the first clause of the description, he could set about finding the relevant row. Indeed, if he _expects_ a row-type description, it is possible that he could begin to execute the row-finding procedure as soon as he hears the phrase 'third row'. Thus, he can interpret the description as he is actually hearing it. This is, in principle, also possible for the path- and figural-types of description, provided that the prominent point from which the path description commences procedurally, or the 'shape' of the figural-type description, is stated first in the description.

It is argued here that this 'on-line resolvability' is not possible for the matrix-type descriptions. The principal reason for this is that the matrix descriptions are so brief in duration: the entire description could be spoken before even the first of the procedures corresponding to the description can be executed. Furthermore, in the case of the '<u>non</u>-Cartesian' matrix-type description, that is, that which does <u>not</u> involve contiguous execution of the two procedures corresponding to the two coordinates (see above), a strain is placed upon the listener by the very nature of this type of description. This is because he has to find the relevant row relative to an agreed-upon origin, 'go' back to the origin, find the relevant column relative to the origin (bearing in mind which row was the relevant row) and then seek the intersection of the relevant row and column. Thus, the listener has to remember the results returned by the execution of the procedures corresponding to the earlier parts of the description whilst he executes the procedures corresponding to the later part of the description. He has to keep the description in memory as he executes the procedures corresponding to it, since the description is too brief for 'on-line' execution of the procedures. This is in fact true for both types of matrix description ('Cartesian' and 'non-Cartesian'): the actual verbal descriptions are too brief to permit the on-line execution of the relevant procedures, even if they were to be executed contiguously.

For this reason, it is hypothesized that associated with matrix-type descriptions will be a greater number of pauses (i.e. speechless gaps of more than 1 second's duration) and more frequently-occurring instances of explicit counting out (this latter being a verbal manifestation, it is hypothesized, of the visual scanning procedures being executed). The pauses, it is hypothesized, would correspond to a subvocal version of the explicit counting: the pauses would be periods during which the procedures corresponding to the description are being executed by the listener. In the other three types of description (line-, path-, and figural-types) it is in principle possible for the listener to execute the procedures 'on-line', and fewer instances of explicit counting would be necessary.

Thus, the procedural interpretation of the meanings of the location descriptions has testable hypotheses associated with it. The above hypotheses (that line- and matrix-type descriptions should be shorter in terms of the number of utterances in each description and/or the number of words in each utterance than path- and figural-types, and that the matrix type of description should be the least verbally fluent of the four types of description in terms of their being associated with the highest frequencies of pausing, hesitating and explicit counting) were tested on the data base of those dialogues which had been clearly identified by the analyses reported in chapter 7 as having employed path-, line-, figural-, or matrix-type descriptions predominantly.

The analyses presented in tables 7-5 and 7-6 showed that six clear groupings of individual subjects consistently emerged across different cluster analyses of individual subjects, and of these six groups, three used line-type descriptions, one used path-type descriptions, one used figural-type descriptions, and the sixth group used matrix-type descriptions (see tables 7-5 and 7-6). Those dialogues in which both interlocutors were repeatedly grouped together by different cluster analyses as belonging within a group of subjects which was identifiable as having used a particular type of description were selected for analysis. Thus, for example, cluster (1) of subjects in table 7-5 was comprised of five individuals - 10BA, 10BB, 10GA, 10GB, and 41AA. These five individuals were grouped together by both sets of cluster analyses ((a) and (b) in table 7-5). Furthermore, this cluster of subjects was identified (see table 7-6) as having used line- (specifically, row-) type descriptions. The descriptions from dialogues 10B and 10G were selected as exemplars of row-type descriptions, since both subjects in each game clustered together on the basis of the descriptions they generated within a cluster of subjects identifiable as having used row-type descriptions (see table 7-6 and section d) of chapter 7). Dialogue 41A, on the other hand, was not included in the analysis, since only one of the participants in dialogue 41A (subject A) had clustered within a group of subjects using row-type descriptions and dialogue 41A may therefore have been more heterogeneous with respect to the types of description used by both subjects than would justify its inclusion within the group of dialogues identified as comprising subjects who predominantly used row-type descriptions.

Thus, the earlier analyses were used to aid the selection of the most prototypical examples of different description types in different dialogues. On this basis, a total of 26 dialogues were selected and all location descriptions spoken by both subjects in these dialogues were analyzed: both subjects in five of the dialogues used path-type descriptions predominantly, both subjects in three of the dialogues used figural-type descriptions predominantly, both subjects in nine of the dialogues used line-type descriptions predominantly, and both subjects in nine of the dialogues used matrix-type descriptions predominantly (see table 8-1, below). In addition, a further four dialogues were identifiable on this basis as having involved the predominant use by both subjects of matrix-type descriptions (dialogues 42A, 42B, 44A and 44B). However, these four games were 'reassignment' games, which have already been shown (see chapter 5) to have been significantly more difficult for subjects than other, otherwise equivalent, types of game. To have included the above four dialogues among the dialogues identified as having used matrix-type descriptions could have resulted in an unfair comparison between dialogues using matrix-type descriptions and the other dialogues using the other types of description, since it is possible that one manifestation of the greater degree of difficulty experienced by subjects in the reassignment games could have been a lesser degree of verbal fluency on the part of these subjects. This would, if it were the case, have artificially inflated the mean values of the 'pausing' and 'hesitation' measures (see below) for matrix-type descriptions. For this reason, the four reassignment games in which the participants were identified as having both predominantly used matrix-type descriptions were excluded form the present set of analyses.

(i) Analyses of the lengths of the different types of description

The first hypothesis to be tested was the question of whether line- (or, more specifically, row-) type and matrix-type descriptions were indeed shorter than were path- and figural-types. Two measures of description length were taken of all descriptions in each of the dialogues identified as cases in which both interlocutors predominantly used one of the four types of description (see table 8-1). These measures were:

(a) The number of utterances in each description, starting from and including the wh-question initiating a location description (for example, "Where are you now?"), if any, to the final acknowledgement of comprehension, if any. The number of utterances comprising each description was averaged for each dialogue in a class (e.g. for all dialogues using path-type descriptions, etc.) and for the class as a whole;

and _(b) The mean length of each utterance within a description_, in words or word units, including filled pauses ("uhm" . ., "er . .") and instances of counting (each increment to the count being recorded as a separate word). The mean length of the utterances within a dialogue's descriptions was obtained for each dialogue within a class,and these means for each dialogue were themselves averaged to give a value for the class as a whole.

The summary statistics of those two measures are shown in table 8-2:

As can be observed from table 8-2, there are differences in the predicted direction between path- and figural-, as opposed to row- and matrix-type descriptions, such that path- and figural-type descriptions appear to involve more utterances per description on average (with mean values of 3.84 and 4.98, respectively) than do row- and matrix-types of description (with mean values of 2.647 and 2.507, respectively). These differences were tested using Mann-Whitney U tests, and were found to be significant: row-descriptions involve significantly fewer utterances, on average, than do path-type descriptions (Mann-Whitney $U = 9$, $n_1 = 5$, $n_2 = 9$, $p < .05$ (1 tail)). Likewise, matrix-type descriptions involve significantly fewer utterances, on average, than do path-type descriptions (Mann-Whitney $U = 8$, $n_1 = 5$, $n_2 = 9$, $p < .05$ (1 tail)), and significantly fewer utterances than figural-type descriptions (Mann-Whitney $U = 1$, $n_1 = 3$, $n_2 = 9$, $p < .01$ (1 tail)).

The differences between the two sets of description types (path and figural- as opposed to row- and matrix-types) were much smaller in the case of the 'mean length of utterance' measure, and in fact, none of the predicted differences were significant on this measure. Thus, matrix-type descriptions were <u>not</u> significantly briefer in the mean length of the utterances comprising the descriptions than were path- (Mann-Whitney $U = 12$, $n_1 = 5$, $n_2 = 9$, NS) or figural-type descriptions (Mann-Whitney $U = 12.5$, $n_1 = 3$, $n_2 = 9$, NS). This latter result may seem surprising initially, especially when considered in relation to the apparent brevity of the two-place coordinates which comprise matrix-type descriptions. However, it should be borne in mind that matrix-type descriptions are often considerably prolonged by the explicit counting aloud corresponding to the execution of scanning procedures (e.g."I'm: <u>A,B,C,D</u>: <u>1, 2, 3</u>: D5") and, as a result, are not any shorter than the other types of description in the final analysis.

Thus, our initial hypotheses that row- and matrix-type descriptions are more brief than are path- and figural-types, was confirmed for the 'mean number of utterances per description' measure, but was not confirmed for the 'mean length of the utterances within a description' measure. Thus, row- and matrix-type descriptions <u>are</u> shorter than are path- and figural-type descriptions insofar as they involve briefer verbal interactions overall, but are not shorter in terms of the length of the utterances comprising the descriptions. The mean number of utterances per description in row- and matrix-type descriptions, of approximately 2.5 utterances per description, suggests, perhaps, that a typical description might take the form 'question-description-acknowledgement' with the 'question' or the 'acknowledgement' sometimes being omitted, giving a mean length of descriptions of around 2.5 utterances. The lengthier path- and figural-type descriptions (of approximately 4.5 utterances per description on average) could be lengthier because the 'description' component of the above hypothesized sequence is split into two utterances (rather than being achieved in only one utterance as is usually the case with row- and matrix-type descriptions) with acknowledgement of understanding <u>in between</u> the two utterances of the 'description' component.

Whilst this interpretation is speculative in nature, there are large numbers of descriptions in the data which do conform to these suggested patterns, which indicates that this interpretation is not an unreasonable one.

(ii) Manifestations of difficulty on the part of the speaker in the generation of descriptions: an analysis of pauses, hesitations, explicit counting, and self corrections in the different types of location description

As noted earlier, the procedural interpretation of the maze location descriptions has associated with it certain hypotheses, notably the hypothesis that certain manifestations of difficulty on the part of the speaker should be in evidence most frequently in the case of matrix-type descriptions as compared to the other types of description. Specifically, it is hypothesized that pausing (gaps in the speech stream of longer temporal duration than one second) and instances of explicit counting should occur most frequently in the case of matrix-type descriptions as opposed to the other three types of description, and these manifestations of difficulty should be significantly more prevalent in the case of matrix-type descriptions than is the case with the other three types of description. Other manifestations of difficulty on the part of the speaker are also found in the transcripts, for example, self-corrections, and the hypothesis is postulated that these latter indicators of 'trouble' should appear equally frequently in the different types of location description: there would seem to be no compelling *a priori* reason to expect self-corrections to be significantly more prevalent in any of the four types of location description as opposed to the others.

To test the above hypotheses, five measures of evidence of difficulty on the part of speakers were taken for all of the descriptions in the dialogues representing usages of the four different description types (see table 8-1). The measures were:

(a) The number of hesitations per utterance. A hesitation is defined as a brief pause in the speech stream of less than 1 second's duration, and which is therefore too brief to be timed using a digital watch during transcription of the tapes of dialogue. The number of such hesitations were summed for all the descriptions within a given dialogue, and then averaged across the total number of utterances in all of the descriptions of that given dialogue.

(b) The number of pauses per utterance. A pause is defined as a gap in the speech stream of more than 1 second's duration, and which is therefore able to be timed during transcription. Pauses were timed to the nearest half second, and the number of pauses was summed for all descriptions within a given dialogue and averaged across all utterances of the descriptions in that dialogue. (The introduction section of the present chapter summarizes the reasons why this relatively crude treatment of time was

adopted in the present analysis).

(c) <u>The number of filled pauses per utterance</u>. A filled pause is an expression such as 'er' or 'ehm' which is used to verbally 'fill' a pause. Such non-lexical expressions did occur sometimes during descriptions. Their incidences were totalled and averaged across all utterances comprising all of the descriptions within each dialogue.

(d) <u>The number of instances of counting per utterance</u>. This measure is simply an average, across all utterances comprising all of the descriptions in a given dialogue, of the number of instances of explicit counting out (which, of course, we have hypothesized to correspond to the execution of scanning procedures).

(c) <u>The number of self corrections per utterance</u>. This measure is an average of the number of times a speaker corrected himself or herself during each utterance involved in generating or acknowledging one's understanding of a location description. The total number of such occurrences of self corrections in all of the utterances in all of the descriptions of a given dialogue was averaged to yield a value corresponding to the average number of self corrections per utterance in each dialogue. An example of such a self-correction would be:

(27AA) "If I move again I'll be r-: in the bottom right hand corner".

The underlined part of the above example is the self correction: the speaker begins to articulate a word, stops doing so, and corrects herself by articulating a rephrased version of what she wishes to say.

The above five measures are all hypothesized to be evidence of processing difficulty of some sort on the part of the speaker: all involve breaks in the fluency of the speech being produced, and this could reflect some processing difficulty which the speaker is experiencing.

The five measures were computed for all of the descriptions in all of the dialogues representing uses of the different types of description (see table 8-1). The result is tabulated below (see table 8-3).

Significant differences between matrix-type descriptions and the other three types of description were hypothesized on the 'counting' and 'number of pauses' measures, but not on the 'hesitations', 'self corrections' or 'number of filled pauses' measures. (The reason that significant differences are hypothesized to exist between matrix- and the other three types of description on the 'pauses' measure but not on the 'hesitations' measure is that only pauses (that is, gaps in the speech stream of longer duration than one second) are sufficiently long to allow the hypothesized subvocal counting out corresponding to the scanning procedures implicit in matrix-type descriptions to take place).

Examining the respective scores of the dialogues representing the different types of description on the three measures on which no significant differences are hypothesized ('hesitations', 'self corrections' and 'filled pauses') first, the uniformity of the mean scores of all four description types on the 'self corrections' and 'filled pauses' measures is noticeable. Indeed, no significant differences were found between matrix- and the other three types of description on these two measures. On the 'mean number of hesitations per utterance' measure, there is again a uniformity in the mean scores of three of the description types, the exception to this uniformity being the score for the row-type descriptions, with a value on this measure (0.37 hesitations per utterance) which is considerably lower than those of the other three types of description (which are 0.871 hesitations per utterance for the path-type descriptions, 0.784 hesitations per utterance for the figural-type descriptions, and 0.7057 for the matrix-type descriptions). Matrix-type descriptions proved to involve significantly more hesitations per utterance than did the row-type descriptions (Mann-Whitney $U = 14$, $n_1 = 9$, $n_2 = 9$, $p < .02$ (2 tail)). Matrix-type descriptions did not, however, involve a significantly different number of hesitations per utterance than did path- or figural-type descriptions.

This is a slightly surprising set of results, which might suggest that row-type descriptions are spoken with a greater degree of fluency (as far as the number of hesitations in the flow of speech goes) than is the case with the other three types of location description. This apparently greater fluency of speech in the case of row-type descriptions was also evident in the case of the scores on the two measures of interest (i.e. those on which differences between different types of description had been hypothesized, namely, the number of pauses per utterance and the number of instances of explicit counting per utterance).

On those two measures of greatest interest in the present context ('pausing' and 'counting'), the mean value for the matrix-type descriptions as compared to those of the other three types of description were as predicted: matrix-type descriptions showed the highest rate of pausing per utterance (with a mean rate of pausing per utterance of 0.253 as opposed to 0.019, 0.15 and 0.19 in the case of the other three types of description) and the highest mean number of instances of counting per utterance (0.099 as opposed to 0.012, 0.026, and 0.032 in the case of the other three types of description). Conversely, as was the case with the 'number of hesitations per utterance' measure, row-type descriptions involved the lowest 'difficulty' scores, 0.019 pauses per utterance on average, and 0.012 instances of counting per utterance on average. The low scores associated with row-type descriptions on these two measures again points to row-type descriptions being associated with greater verbal fluency than is the case with the other three types of description.

Turning to statistical comparisons on these two measures, the matrix-type descriptions involve significantly more frequent rates of pausing than do row-descriptions (Mann-Whitney $U = 1$, $n_1 = 9$, $n_2 = 9$, $p < .001$ (1 tail)). However, although the mean values on the 'number of pauses per utterance on average' measure lie in the correct directions, matrix-type descriptions are not significantly different from path-type (Mann-Whitney $U = 14$, $n_1 = 5$, $n_2 = 9$, NS) or figural-type (Mann-Whitney $U = 8$, $n_1 = 3$, $n_2 = 9$, NS) descriptions on this measure. Thus, the overall result is that whilst the mean values on the 'number of pauses per utterance' measure were as predicted, the matrix-type descriptions did not involve significantly more pausing per utterance on average than did path- or figural-type descriptions. They were found, however, to involve a significantly higher rate of pausing than the 'row' type of descriptions.

On the measure 'the number of instances of explicit counting per utterance', again, the mean values lay in the direction predicted, with the matrix-type descriptions involving a rate of counting out three times as high as was the case with any of the other three types.

When statistical comparisons are drawn between matrix-type descriptions and the other three types of description on the 'number of instances of explicit counting' measure, the results substantially confirm the hypothesis. Thus, matrix-type descriptions involve a significantly greater prevalence of explicit counting than do path-type descriptions (Mann-Whitney $U = 7.5$, $n_1 = 5$, $n_2 = 9$, $p < .05$ (1 tail)) and row-type descriptions (Mann-Whitney $U = 5$, $n_1 = 9$, $n_2 = 9$, $p < .001$ (1 tail)). In the comparison between matrix- and figural-type descriptions on this measure, the result approaches but does not quite reach significance at the .05 level (Mann-Whitney $U = 4$, $n_1 = 3$, $n_2 = 9$, NS, required $U = 3$). The small number of figural transcripts involved in this comparison may have been problematic here.

Overall, our hypotheses regarding the greater prevalence of pausing in matrix-type descriptions as compared to the other three types of description was not confirmed; the only significant difference obtained was in the comparison between matrix- and row-type descriptions. The hypothesis that matrix-type descriptions should involve more frequent instances of explicit counting than the other three types of description was confirmed for the comparison between matrix- and row-type descriptions, and it was also confirmed for the comparison between matrix- and path-type descriptions. It was not confirmed, however, for the comparison between matrix- and figural-type descriptions. The partial confirmation of these hypotheses despite the small data base involved suggests that the hypotheses were not unreasonable and would merit further investigation on a larger scale.

The unexpected difference which these studies suggest is the apparently greater fluency associated with row-type descriptions as compared to the other three types of description. This fluency was manifested in the form of lower rates of hesitating, pausing, and the explicit counting out of procedures. The explanation for the apparently superior fluency of row-type descriptions would most probably be quite complicated and involve a number of factors; however, one possibility which might have an influence and which has been discussed before is the level of abstraction from the actual details of the stimulus which the row-type descriptions represent. Earlier in the present chapter, it was argued that matrix-type descriptions were most 'abstract' in the sense that they involve the ignoring of all of the structural features of the maze shape. Indeed, it was their abstractness and brevity which led to the proposal of the hypotheses that greater amounts of pausing and counting should accompany their use.

On the other hand, path- and figural-type descriptions are highly constrained by the actual shape of the maze stimulus, and this lack of abstraction could be a source of difficulty for subjects - the problems posed by gaps in the path structure of the maze have already been noted. The choice of figures with which to metaphorically describe parts of the maze structure could also present problems for speakers, since a large variety of shapes are interlocked in the network of the maze paths (T shapes, L shapes, upside-down F shapes, rectangles, squares and so on: see, for example, the diagram of the asymmetrical maze in appendix I), and this decision could conceivably interfere with the fluency of a speaker's discourse as he generates it.

The row-type descriptions, on the other hand, constitute a medium level of abstraction. They present a readily-applicable description format which is neither too brief nor too abstract and does not involve the subject in awkward decisions about the best choice of description to make. It is a ready-made descriptive framework which is, in principle, resolvable 'on-line', as noted earier.

Whilst this explanation is speculative, it could well be partly true, and would fit the observed data: whilst matrix-type descriptions involve the operation of abstract procedures which are encoded in a very brief verbal format, path- and figural-type descriptions (particularly, figural-type descriptions) could result in the opposite problem arising i.e. that of of not abstracting enough from the particularities of the maze being described.

Row-type descriptions, on the other hand, suffer fom neither of these problems and it could be partly for this reason that row-type descriptions are associated with a greater degree of verbal fluency than are the other three types of description.

In the next section, we go on to a more general discussion of the semantics of the location descriptions.

c) Mental models and the description of maze locations

When characterizing the four different types of location description (path-, line- figural- and matrix-types: see chapter 6), it was briefly mentioned that these four different ways of describing the maze could reflect four distinguishably different ways of conceptualizing the complex stimulus of the maze shape. In particular, it is suggested in the present interpretation of the data that the four different types of location description which have been discussed in this thesis are associated with different underlying types of <u>mental model</u> of the maze shape (cf. Johnson-Laird, 1980, 1981a, 1981b; see chapter 2 of the present volume). Thus, for example, a set of descriptions generated by a subject which are all of the 'row' type could reflect his underlying mental model of the maze shape which takes the form of a series of parallel horizontal rows; he has paid attention to the horizontal lines formed by those pathways in the maze which run horizontally from left to right, treated them as prototypical features of the maze shape, and built a model accordingly.

The present section of this chapter will explore the possible consequences, both theoretical and empirical, of interpreting the maze location descriptions in this fashion.

Firstly, it is worth briefly recapitulating the most important features of mental models as Johnson-Laird (1980, 1981a, 1981b) conceives of them. The most important point about the structure of a mental model is that it mirrors that of whatever is being described. Thus, one may build a mental model of the room one is reading about in a novel. Any description of such a spatial layout is liable to be radically indeterminate, i.e. an infinity of mental models could be built which would satisfy the described layout. Mental models embody arbitrary assumptions built in by the thinker, and this does not matter provided that the important structural properties of the room are mirrored in the model. Thus, to some extent, a particular model constructed on the basis of a given piece of discourse is arbitrary in nature. Different listeners construct different models on the basis of the same piece of discourse; this is possible because of the radical indeterminateness of any given description of (e.g.) a spatial layout. These different models would all share the property of mirroring the structure of what is being described (yet they could all be different from one another). It is important that the model that the listener builds on the basis of the discourse he hears does mirror the structure of what is being described.

Secondly, mental models are intermediate between language and the world. Linguistic input results in change in a mental model, and need not necessarily result in changed behaviour in the listener. The mental model account of semantics thus dissociates the understanding of an utterance from the taking of action by the listener in response to the linguistic input.

Johnson-Laird concentrates his attention on the modelling by a listener or reader of a visually absent stimulus which is being described. He does, however, point out that models can be built of stimuli which are present in the model-builder's visual field at the time he is building the model. In such a case, attention is paid to particular features of the stimulus, these are treated as prototypical, and an idealized model is built, which is in fact an assemblage comprising a number of these prototypical features. For example, the hypothetical interlocutor discussed above who builds a row-type mental model of the maze does so by paying attention to the horizontal lines formed by the maze pathways which run in the horizontal plane, treating this 'horizontal line' feature as being prototypical, and the resulting mental model has as its basis six parallel horizontal lines.

This model would, like those discussed by Johnson-Laird, provide semantic interpretations of linguistic input. Thus, if given the following description

"I'm third row from the bottom, second from the left",

our hypothetical interlocutor would conceptually 'mark' (in some way) a particular location _within his mental model_. The 'marked' model representation would be used as the basis for conducting visual scanning operations on the actual stimulus, for example, if the listener were asked the following further question by the speaker who had just generated the above description:

"Are there any Ss" (switch nodes) "on your screen near me?"

Similarly, if he were asked to describe the location of his own position marker on the maze, our hypothetical subject would consult both the maze stimulus on the screen and his mental model of it, since his mental model embodies the semantics for generating, as well as interpreting, descriptions.

As can readily be appreciated from the above discussion, mental models are intermediate between language and the world in a very real sense; the understanding of an input utterance results in a change _to the model_, and a request to generate a description results in consultation _of the model_ as well as the actual physically-present visual stimulus. The structure of the model would assist the speaker in the generation of location descriptions, since it would constitute a framework with respect to which descriptions could be couched (e.g. the rows of the row-type model providing a framework in relation to which descriptions could be couched). This point will be explored more fully in chapter 11.

An interesting case study which relates, in some respects, to the present interpretation of the maze location descriptions as reflecting different types of underlying mental model of the maze was conducted by Pailhous (1984). Pailhous studied the mental models of the spatial topography of the city which taxi drivers build and use to assist them in planning journeys. Pailhous examined a) the <u>use</u> by Parisian taxi drivers of their mental models of the city in order to plan journeys, b) the relation between freehand maps drawn by the taxi drivers and professionally constructed maps of the city which faithfully represented the spatial interrelations between streets, etc.; and c) how the taxi drivers <u>build up</u> a mental model of a town in which they had not travelled before.

Both the analysis of journeys (in which taxi drivers were set spatial problems in the form of getting to various destinations without resorting to the aid of a map) and the comparison of the maps drawn by the taxi drivers with the professionally-constructed maps pointed to the existence of two components in the drivers' mental models, these being termed the <u>primary network</u> and the <u>secondary network</u>. The primary network involves the faithful representation of localities with respect to their true geographic position and spatial interrelationships with one another, and when using the primary network to plan a journey, the driver would use an algorithmic procedure to assist him in attaining his goal, the algorithm involved being that he take a route which makes the minimal angle with the destination. The secondary network is represented with considerable geographic distortion, is indexed to the primary network by the relation of proximity, and is associated with the use of a heuristic rule (namely, the regaining of the primary network) when being used to plan journeys. The heuristic procedure associated with the secondary network is not always effective; it increases the distance to be covered and the mistakes made increase with the distance of the point of departure from the primary network.

Similarly, when building his mental model of the spatial layout of a town in which he has not driven before, the driver firstly builds up a primary network and then indexes points to the crossroads and pathways of the primary network, and these localities form the secondary network.

Clearly, although Pailhous writes of 'images', he is describing the building and use by subjects of a species of mental model of the topography of a city. Associated with these models are procedures (heuristics and algorithms) which assist the planning of a journey. This is comparable in some ways to the subjects in the maze game studies modelling the maze and using the resulting model and its associated procedures to assist in the generation and comprehension of location descriptions.

Pailhous's work serves to emphasize again the variety of types of mental model and the variety of uses to which mental models can be put.

The present interpretation of the location descriptions generated by subjects in the course of playing the maze game (that is, as reflecting four different underlying types of mental model of the maze) suggests that the four different underlying models would involve different semantic interpretations being assigned to a given location description. Indeed, one could envisage situations in which, if two subjects engaged in playing the maze game were simultaneously entertaining radically different mental models of the maze shape, then a location description generated by one subject (for example, a location description of the matrix type) could entirely fail to be interpreted by the other player (who might, for example, be currently entertaining a figural-type model of the maze). It would thus be of importance that both interlocutors should entertain similar types of model for even partial understanding (or indeed, misunderstanding) to occur. As such, one could hypothesize that there would be some tacit pressure exerted by both subjects on one another to model the stimulus in similar ways in order for successful location description generation and comprehension to occur.

Thus, although the choice of one type of mental model of the maze rather than another is to some extent arbitrary, as are the precise structural details of a given individual's mental model, it is important that both participants in a dialogue entertain models which are similar to the extent that they return similar semantic interpretations for a given expression.

The advantages of the use of mental models in the maze game situation would be considerable. Provided that both participants in a dialogue entertained mental models of a similar type, which returned identical semantic interpretations for terms like 'row' and count values, the task of description generation and interpretation would be greatly simplified. One need only attend to those features of the maze which are modelled (for example, the arrays of nodes if one is currently entertaining a matrix-type model) and could readily ignore the other features of the maze structure (such as gaps in the network of pathways, figures formed by the particular spatial configurations of some of the pathways, and so on) for all practical purposes.

Interestingly, the present conception of a mental model resembles the 'models' of model-theoretic semantics in its functional aspects: that is, the 'model' as conceived of here functions as a semantic system with respect to which particular expressions are interpreted. The common model entertained by both participants in a dialogue is, from this perspective, a tacit agreement as to the extensional semantics of particular expressions (for example, that 'three along' refers to the third node position in the horizonal plane, counting from the left: cf. 'model structure lexical rules' in the extensional metalanguage of model-theoretic semantics).

The present experimental maze game task differs from the experimental 'communication conflict' task used by Blakar and his colleagues (see chapter 4) in one very important respect: in the present maze game task, both interlocutors are trying to build a common mental model of the <u>same</u> stimulus, whereas in the 'communication conflict' task the two subjects are developing models of <u>different</u> stimuli (the stimuli are different in that one of the maps has an extra street). In the maze game task, there is a possibility that the two individuals playing the game might (initially at least) build different mental models of the maze; however, such a difference would be a <u>systematic</u> difference (the subjects employing different designations for the same referents). This is the type of difference to which a solution could readily be negotiated; the subjects are modelling the same referent and their task would be one of arriving at a 'common language' with which to describe that same referent. The subjects would have to change, as it were, the 'rules of designation', (for example, agreeing that count values had as their extension node positions ordered from left to right rather than path movements ordered from left to right). In this case, a general rule is being changed. In Blakar's 'communication conflict' task, on the other hand, the two subjects suffer from the problem that their respective stimuli (and hence, their respective models) differ in a <u>non</u>-systematic way, and it would be extremely difficult, if not impossible, to develop a common model associated with which there is a 'common language' when the two stimuli being modelled by the two individuals differ in a non-systematic way. There is no general rule that could be changed which would help these subjects develop a common model. Hence the difficulty of the 'communication conflict' task for subjects.

The intepretation of the maze location description data in terms of mental models of the maze shape, which is an essentially mentalist position, can be contrasted with what could be termed a 'behaviourist' approach to the location description data. The radical behaviourist would not entertain the theoretical concept of a 'mental model'; for him, such a concept would be redundant. He would prefer instead to think of the procedures corresponding to the location descriptions as being executed directly on the actual visual stimulus on the computer terminal screen, without any intermediate 'modelling' stage taking place.

A corollary of the 'behaviourist' approach would be, however, the notion that the most convenient description to generate at any point in time would be that which is most consonant with the location being described on the stimulus. Thus, figural-type descriptions would be used when one was describing a location which was part of an obvious shape in the maze structure, row-type descriptions would be generated when one described a location on the bottom or top horizontal lines of one of the symmetrical mazes, and so on. That is, if the 'behaviourist' account were accurate, there would be no compelling reason for subjects to use one type of description with great consistency, and there would certainly be no reason for subjects to use matrix-type descriptions when describing locations on an asymmetrical maze, in which many of the locations implicit in a matrix-type coordinate system of notation are entirely absent.

Thus, two corollaries which follow from the 'mental models' account but which do not necessarily follow from the 'behaviourist' account are a) that both interlocutors should entertain the same, or very similar, models and hence generate similar descriptions, and b) that the interlocutors should generate descriptions which are consonant with their mental model rather than those which are consonant with local features of the area being described.

If we accept, for the moment, that the four types of location descriptions _do_ reflect different underlying mental models of the maze shape, the question naturally arises as to how the two subjects within a dyad playing the maze game come to share very similar mental models. By means of what mechanisms is such commonality achieved? This question will be explored further in chapter 10.

In chapter 11, the whole question of mental models of the maze will be examined again in more detail. In the next chapter, an experiment will be described in which an attempt was made to ascertain whether the four classes of location description could be replicated in an experimental situation in which subjects described locations on mazes, but independent of playing the maze game.

	Type of Description			
	Path	Figural	Line	Matrix
Cluster of subjects using description (see also Table 7-5)	(4)	(5)	(1),(2),(3)	(6)
	37B	35B	10B	19F
	37A	33A	10G	20B
	27B	40B	13A	20D
Dialogues	39B		13C	23B
	15E		17F	26B
			17G	27B
			18F	27D
			18D	28B
			31C	28C

Table 8-1 Dialogues identifiable as involving the use of path-, figural-, matrix-, and line (particularly, row-type descriptions. In these games, both interlocuturs were identified as having consistently used the same type of description. N.B.: Reassignment games were excluded from the above classification of games (see text).

Type of Description	No. of utterances in the descriptions		Mean length of each utterance in the descriptions (in words)	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Path (N=5)	3.837	0.841	6.514	0.714
Figural (N=3)	4.985	0.906	5.859	0.447
Row* (N=9)	2.647	0.58	6.404	0.834
Matrix (N=9)	2.507	0.654	5.858	0.918

Table 8-2. Summary statistics on lengths of descriptions (in utterances) and lengths of utterances (in words) in each of the four types of description

*Footnote: All line-type descriptions within this category were in fact row-type descriptions, and hence they will be described as 'row-type' in the subsequent discussion.

Type of Description	No. of Hesitations per Utterance		No. of Pauses per Utterance		No. of filled pauses per Utterance	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Path. (N=5)	0.871	0.191	0.1919	0.125	0.109	0.041
Figural (N=3)	0.7839	0.2415	0.1524	0.073	0.068	0.034
Row (N=9)	0.372	0.2195	0.019	0.033	0.09	0.026
Matrix (N=9)	0.7057	0.226	0.253	0.123	0.099	0.061

Type of Description	No. of Instances of Explicit Counting per Utterance		No. of Self Corrections per Utterance	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Path (N=5)	0.0032	0.0197	0.166	0.071
Figural (N=3)	0.026	0.019	0.154	0.1049
Row (N=9)	0.012	0.019	0.15	0.0408
Matrix (N=9)	0.099	0.066	0.1183	0.0504

Table 8-3. Summary statistics on five indicators of difficulty on the part of speakers in the four different types of location description

CHAPTER 9

THE MAZE-GAME INDEPENDENT LOCATION DESCRIPTION TASK

a) Introduction

In the last three chapters, the nature of subjects' descriptions of maze locations generated during the course of playing the maze game task were discussed. The broad conclusion of both the informal and the formal analyses presented in chapters 6, 7 and 8 was that subjects' location description strategies could be grouped into four classes, which have been named the path-, line-, figural- and matrix classes of location description. In chapter 8, some analyses were presented which illustrated the differing properties exhibited by the four classes of description, and some speculations were advanced as to the wider significance of the different classes of description (specifically, that they may reflect four distinguishably different ways of conceptualizing, or mentally modelling, the maze shape).

It was decided to attempt to replicate the above result (i.e. that there exist four classes of location description) in an experiment in which subjects had to describe locations within mazes _but independent of playing the maze game_. If the above four-fold typology of descriptions were to emerge again in the results from a study which did not involve the subjects playing the maze game, this would support the view that the observed structuring in the data is better explained in terms of the semantics of describing stimuli of the sort used in both studies (i.e. maze shapes) than in terms of the structuring in the data being a specific outcome of the game task.

An experiment was therefore designed in which subjects described locations within mazes which were presented to them in the form of photocopied diagrams* of the maze shapes (similar in nature to the illustrations used in the present thesis). The experiment involved three conditions, of which results from the first two conditions will be discussed in the present chapter. The results from the third condition will be discussed in chapter 10. The three conditions were as follows:

<u>i) Interactive condition</u>. In this condition, a pair of subjects had to describe locations within mazes alternately to one another. Five mazes were involved and eight locations required to be described in each maze. The procedure involved was as follows: first, subject A successively described the eight locations in maze 1 to subject B, then subject B successively described the eight locations in maze 2 to subject A, and so on in alternating

---

* Hence the subsequent references in this chapter to this experiment as 'the paper-and-pencil experiment'.

fashion.

<u>ii) Non-interactive condition</u>. In this condition, the same five mazes with eight locations in each had to be described, but in this case one subject exclusively described all of the locations on all of the mazes, whilst the other subject acted as listener throughout.

<u>iii) Monologue condition</u>. In this condition, one subject at a time described all locations on all five mazes but without a listener being present.

The three conditions allow differing degrees of interaction between the subjects. In condition i), both subjects were involved in describing locations to one another. In condition ii), only one of the subjects generated descriptions, but (as was the case in condition i)) the listener was allowed to ask questions of the speaker if he failed to understand the description generated. In condition iii), of course, no listener was present at all.

The stimuli involved were drawings of two of the maze shapes used in the maze game studies (one symmetrical and one asymmetrical), and three other maze shapes which had been specially-constructed for the experiment. These latter employed similar principles in their construction as did the mazes used in the game task (consisting of arrays of nodes connected by paths), but exhibiting certain deliberately manipulated features (notably, the distance between adjacent nodes was not uniform in the case of two of these mazes, and diagonally-oriented paths were included in the third maze: see Appendix V). These features were manipulated in an attempt to ascertain whether they had any effect upon the nature of the location descriptions generated.

The listener's task was to place a mark for each location (1 - 8 or X_1 - X_8) described by the speaker on a blank diagram of the maze in which the locations were being described. Listener and speaker were located in the same room but with a screen placed between them to prevent the possibility of eye contact and non-verbal communication between subjects taking place. This helped to ensure that the maze game and the 'paper-and-pencil' experiment were comparable with respect to this possibility, since eyecontact and non-verbal communication were not possible in the maze game (the subjects being located in separate soundproofed booths).

The present experiment aimed to replicate the results noted earlier (i.e. to replicate the four description types found in the maze game studies). Subjects' descriptions were analyzed using the same thirty-eight lexical categories as had been used to describe the raw data of the maze game experiments (the thirty-eight lexical category classification was designed to be able to describe data from <u>both</u> sets of studies and was in fact a modified version of an earlier twenty-seven lexical category classification which had been used on the maze game data but which would not have adequately described the data from the 'paper-and-pencil' study because of the unusual features of the mazes employed in this study). The use of the same descriptive classification for both maze game and 'paper-and-pencil' experimental data ensured comparability of the two analyses.

As was the case with the maze game, the described data from condition i) and ii) of the present study were subjected to hierarchical cluster analyses in order to ascertain which groupings of lexical categories would emerge. It was hypothesized that similar clusters of lexical categories would emerge as had emerged from the maze game data - in short, it was expected that the familiar path-, line-, figural- and matrix-type location descriptions would be replicated.

b) Method

i) Subjects

The subjects were recruited from the same pool of first year undergraduate psychology students as had been involved in the maze game studies. Subjects were selected and assigned to same-sex pairs, and were paid £1.00 each for participating in the study. Sixteen pairs of subjects were involved, four pairs of males and four pairs of females participating in each of conditions i) and ii), and eight single subjects (four males and four females) were involved in the monologue study.

ii) Apparatus and Stimuli

The apparatus used in the present experiment was the same TEAC A-3340 four-channel tape-recorder which was used in the maze game studies. Two recording channels were used, one channel for recording the speech of each subject. Subjects were seated side by side in the same room but with a wooden screen placed between them to prevent eye contact. The stimuli took the form of forty photocopied diagrams of maze shapes kept in a ring-binder folder. Each subject described maze locations in blocks of eight descriptions of different locations on the same maze. One such maze location was depicted on each page of the folder containing the stimuli. The listener had a similar ring-binder folder containing diagrams of the mazes in which locations were being described by the speaker but on which none of the locations was marked in any way. The listener's task was to place appropriate marks (1 - 8 or X1 - X8) in the locations in his blank maze which corresponded to the speaker's eight descriptions.

iii) Design and Procedure

As noted earlier, the design was a complex one, involving 5 stimuli in each of three conditions. The interactive condition (in which both participants were involved in describing locations) and the non-interactive condition (in which only one participant was involved in describing locations) were both analyzed together in the present analyses. The sequence of mazes being described was altered for each pair of subjects in order to counterbalance the possible effects of order on the way in which locations were described.

In the interactive condition, the two subjects alternately described blocks of eight locations on a given maze, as follows:

Subject	Maze 1	Maze 2	Maze 3	Maze 4	Maze 5
A	describes	listens	describes	listens	describes
B	listens	describes	listens	describes	listens

In the non-interactive condition, one subject (subject A) described blocks of eight locations on each of the five mazes.

In both of these conditions, listeners were allowed to ask questions of the speaker in the event of failure to understand a description.

All speakers (including those in the monologue study) were instructed to generate descriptions which were precise enough to permit a listener to uniquely identify a given location solely on the basis of the speaker's description of it. All speakers were aware that their listener had an identical, but blank, maze on which to mark the relevant locations.

The procedure was as follows. Subject A (for example) had a folder containing eight diagrams of a maze shape, on each of which was marked a different location. Subject B, correspondingly, had one diagram of the same maze shape on which none of the locations was marked in any way. Subject A's task was to describe the locations in succession in enough detail to enable subject B to find the relevant locations unambiguously. In the interactive condition, the speaker/listener roles would switch from maze to maze, whereas in the non-interactive condition the same speaker described all locations on all five mazes. Each dyad thus generated a total of forty descriptions, which were themselves described in terms of the same thirty-eight lexical categories as had been employed to describe the location descriptions in the maze game studies (see chapter 6). The classificatory technique of using hierarchical cluster analyses of differing types and ascertaining which groupings of lexical categories consistently occur across different analyses was again employed. The results are discussed below.

<u>c) Results</u>

<u>i) The cluster analysis of lexical categories</u>

A casual inspection of the raw data reveals two striking features, namely:

a) Path-type descriptions predominate in frequency of occurrence as opposed to other types of description. This is in contrast to the data obtained in the maze game, in which matrix- and line-type descriptions predominated (compare the numbers of subjects in clusters 6, 1, 2 and 3 as opposed to the number in cluster 4 in the cluster analysis of subjects in the maze game data (see table 7-5). Clusters 6, 1, 2, and 3 of subjects are comprised of those subjects using matrix- and line-type descriptions, whereas cluster 4 is the group of subjects who use path-type descriptions. Matrix- and line-type descriptions clearly predominate in the maze game data).

b) Matrix-type descriptions <u>are</u> <u>entirely</u> <u>absent</u> from the data of the present experiment. No subjects used matrix descriptions at all. A small number, however, used the related strategy of generating descriptions in terms of specifying <u>both</u> row <u>and</u> column, for example:

"The sixth: X: is on the sixth: row: third column:" (4.5s.)

This method of describing locations is similar in principle to the coordinate systems employed by some of the subjects in the maze game studies, but is rather less elliptical in nature.

The apparent prevalence of path-type descriptions and the absence of matrix-type descriptions in the present experiment contrasts strongly with the results from the maze game studies (see chapter 7).

The actual path- and row-type descriptions generated in the present experiment are similar in nature to those generated by subjects during the course of playing the maze game. For example, a typical path-type description generated in the experiment is:

"Ehm:start at the bottom left-hand corner of the page: (2 s) er: move up two boxes: (1.5 s) turn right: and its the: (2 s) third box: (4 s)."

The principle behind this path-type description is identical to that behind the path descriptions generated during the course of playing the maze game (see chapter 6) - a prominent point in the maze structure is identified, and a location description is generated by describing the location in question relative to the prominent point in terms of path movements from the prominent point to the location in question.

Interestingly, the presence of unusual features in some of the mazes (namely, irregular node spacings and diagonally-oriented pathways) did not prevent locations within these mazes being described using the classic path description format. Indeed, there were cases where, for example, path-type descriptions were generated which involved scanning along such unusual diagonal pathways. For example (see figure 9.1 for an illustration of the

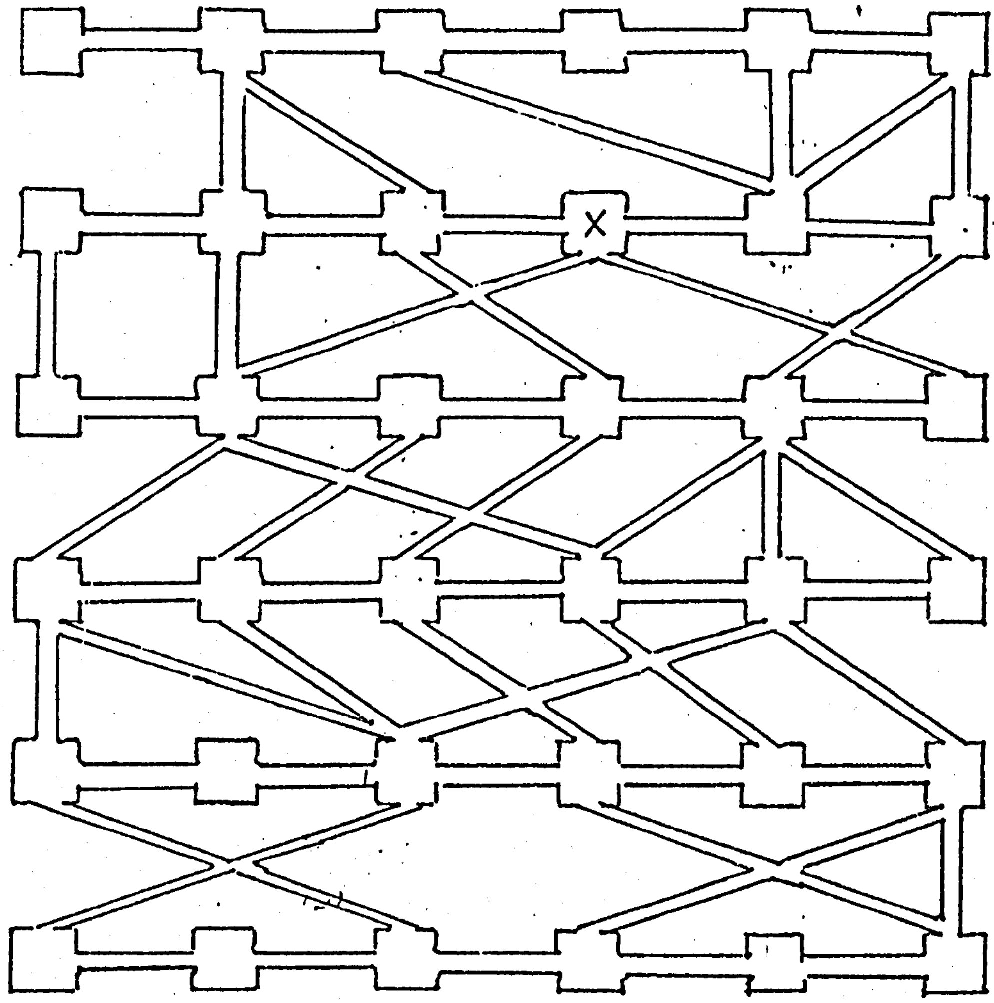


Figure 9-1. The location described by a subject in the 'paper-and-pencil' experiment as "The top:right . . . across diagonally downwards... towards the left . . . straight along towards the left the: cross is in there".

location being described):

A ehm: the top:right.

B O.K.

A Across diagonally: downwards

B Right

A Towards the left: along: ehm: straight along towards the left the: cross is in there:

Thus, a path-description was employed in this instance which involved mentally 'traversing' a diagonal path. These unusual features did not, therefore, preclude the use of the same principles for generating path-type maze location descriptions as were used by subjects in the maze games.

Likewise, the row-type descriptions generated by subjects in the present experiment are identical in principle with those used by some subjects in the maze game. For example:

"Third row down: (1 s) second from the left (3 sec)".

The same principle of firstly identifying a row and subsequently identifying a location within that row is also employed by the subjects who use row-type descriptions in the maze game.

A small number of figural-type descriptions were also generated by subjects in the present experiment. As was the case with the maze game data, however, the occurrence of figural-type descriptions in the present experiment was infrequent in nature.

Thus, a casual inspection of the transcripts of dialogues generated in the experiment suggests that three of the four description types noted in chapter 6 have been replicated: path-, row-, and a few figural-type descriptions were generated by subjects in this 'paper-and-pencil' experiment. Matrix-type descriptions were absent from the data of the experiment, although the related strategy of generating descriptions couched in terms of stating both the row and the column of nodes whose intersection yields the relevant location was noted to occur (albeit infrequently).

The unusual features of some of the mazes in the 'paper-and-pencil' experiment (irregular node spacings and diagonally-oriented pathways) thus did not appear to greatly affect the semantics of the descriptions generated.

The raw data generated in the 'paper-and-pencil' experiment were subjected to the same analyses as were the data from the maze game studies: the descriptions were themselves described in terms of the same thirty-eight lexical categories, and the resulting data matrix was cluster analyzed (several cluster analyses being conducted, manipulating both the measure of similarity on which the clustering was based and the clustering criteria, comparisons of the results from which yielded groupings of lexical categories which repeatedly clustered together across different analyses). The results obtained after cluster comparisons were conducted on the results from the two sets of cluster analyses (across variation in the measure of similarity on which clustering is based and across variation in clustering criteria) are presented in table 9-1 (overleaf).

As was the case with the results from the maze game data analyses, the two sets of cluster analyses of lexical categories in the present experimental data show a considerable degree of similarity to one another, suggesting that there is indeed a definite structure in the data.

The largest groupings of lexical categories are the clusters numbered (1) in both analysis (a) and analysis (b) of table 9-1, and they agree in their content very closely. They are made up of a set of lexical categories which would be used when generating path-type descriptions - corner specification terms, movement terms ('move/go'), terms directing directions of movement ('up', 'left', 'down', 'right' and 'along') and cardinal counting for counting out node positions or path movements during the execution of the scanning procedures. Both analysis (a) and analysis (b) (see table 9-1) contain the same lexical categories in their respective clusters numbered (1), except that cluster (1) of analysis (a) includes the movement-denoting category 'along' where cluster (1) of analysis (b) does not.

The clusters numbered (2) in both analyses agree on the inclusion of three lexical categories - 'left', 'right' (edge names) and 'node'. The clusters numbered (3) in both analyses are in perfect agreement on the inclusion of all three constituent lexical categories ('order', 'to edge', and 'row'). Taken together, clusters (2) and (3) comprise a selection of lexical categories which correspond to lexical items which are usable in row-type descriptions. This is in agreement with the analysis of the maze game data (see table 7-3) in which the lexical categories 'order', 'to edge', 'row', 'left', 'right' (and also 'top', 'bottom' and 'line') clustered together. (Taking clusters (2) and (3) together in the present analysis is justified by the fact that subjects actually sample the clusters of lexical categories in this fashion: subjects' sampling of the lexical categories is dealt with in the next section).

The clusters numbered (4) and (6) in the present set of results group together a set of lexical categories corresponding to lexical items which are usable in figural-type descriptions ('shape' and 'column' in cluster (4), and 'middle', 'limb', 'symmetry', 'extreme' and 'diagonal' in cluster (6)).

The clusters numbered (5) in both analyses consistently group together two semantically-related lexical categories, 'horizontal' and 'vertical', whilst the clusters numbered (7) in both analyses consistently group together two lexical categories, 'ownother' (i.e. previously specified position of self or other) and 'junction'. Both of the lexical categories in cluster (7) correspond to lexical items which are usable in path-type descriptions, for example, when describing a location in terms of a path movement from a previously specified position or towards the junction of two paths.

The classification of the set of lexical categories by means of cluster analysis thus yields consistently-occurring clusters which, since they are uncovered repeatedly by different cluster analyses, could be held to reflect the structuring of the data. The next question of interest is the question of how the subjects actually use the lexical categories within each cluster. The analysis of the maze game data (see chapter 7) demonstrated that subjects' sampling of the lexical categories was highly systematic such that, in the case of five out of the six groups of subjects identifiable from the cluster analysis of subjects, most of the lexical categories in one or two of the clusters of lexical categories were used, but few or none of the other lexical categories from other clusters were used. Furthermore, subjects' sampling of the lexical categories was readily explicable in terms of the four-fold classification of description types. It was, therefore, of interest to examine whether a similar type of sampling from among the lexical categories by subjects could be detected in the present experimental data. The results of this analysis are set out in the next section.

ii) The analysis of subjects' usages of the lexical categories

The first part of the attempt to ascertain the manner in which subjects selected among the lexical categories involved a cluster analysis of individual subjects. This was achieved in the same manner as was the case in the cluster analysis of individual subjects in the maze game studies: the data matrix of 38 lexical categories X 80 subjects* was transformed through 90^{o} to yield a data matrix of 80 subjects X 38 lexical categories, each subject was treated as a variable, and the 80 variables (subjects) were cluster analyzed using BMDP program P1M. Once again, the results from different cluster analyses were compared, and a number of consistent groupings of individual subjects' data was found to obtain. In fact, 3 major groupings of subjects were identified (containing 12, 8, and 16 subjects, these being clusters a), b) and c) respectively) and a total of seven lesser groupings, each of four or five subjects, were identified (clusters d) to j)).

---

* Each dyad (16 in all) described five mazes in total. Each maze's data was treated separately, resulting in 16 x 5 = 80 'subjects'.

The frequencies with which each subject within each of these groupings used lexical items corresponding to each of the lexical categories were summed across all subjects within a given cluster of subjects and averaged for the group. This allowed an assessment of which lexical categories were used by each group of subjects as a whole. Thus, for example, each subject in cluster a) of subjects used the category 'top' an average of 2.5 times in describing eight locations on one maze, the category 'row' an average of 7.166 times, the category 'order' an average of 7.92 times, and so on. When this was completed for all subjects in all groups, those categories which were used at least once on average by each subject were noted, and the results from this analysis are summarized in table 9-2 (overleaf).

Table 9-2 shows which lexical categories within the various clusters of lexical categories yielded in the analysis described in c)i) above were used by each grouping of subjects ((a) - (j)). A casual inspection of the distribution of ticks in table 9-2 reveals that subjects' sampling of the lexical categories within the different clusters of lexical categories is systematic in nature. Consider the three major groupings of subjects, (a), (b) and (c).

Cluster (a) of subjects uses the following lexical categories - 'top', 'handside', 'move' 'along' and 'down' (from cluster (1) of the cluster analysis of lexical categories, 'bottom', 'left', 'right', and 'node' from cluster 2), and 'order', 'to edge' and 'row' from cluster (3). These lexical categories are of the type used by subjects in generating row-type descriptions, for example:

"It's third row down from the top and third

(ORDER) (ROW (DOWN) (TO EDGE) (TOP) (ORDER)

node along".

(NODE) (ALONG)

Notice that this cluster of subjects do not use any figural terms (corresponding to the lexical category named 'shape'), nor do they use any path movements away from specified corners ('count' 'left', 'right', 'corner'). Thus, this group of subjects would appear to be consistently using those lexical categories from the first three clusters of the cluster analysis of lexical categories which are readily usable in row-type descriptions.

Cluster (b) of subjects use a similar but smaller group of lexical categories, namely 'top', 'down', 'bottom', 'left', 'right', 'order', 'to edge', 'row' and 'extreme' and would thus appear to be using row descriptions of a less specific nature than those used by the subjects comprising cluster (a) of subjects. This interpretation is borne out by examination of sample descriptions from the data generated by one of the subjects in cluster (b) of subjects:

"Third row down: second from the left".

(ORDER) (ROW) (DOWN)(SECOND)(TO EDGE)(LEFT (EDGE))

"Second row down: third from the right:" (5 sec)

(ORDER)(ROW)(DOWN)(ORDER) (TO EDGE) (RIGHT (EDGE))

Cluster (c) of subjects (see table 9-2) use all of the lexical categories comprising clusters (1) and (2) of the cluster analysis of lexical categories, except for the lexical category 'rowpcol' in cluster (2), plus the lexical category 'order' from cluster (3). This collection of lexical categories includes counting of both ordinal and cardinal types, path movement terms, edge names, and 'corner', but <u>not</u> 'row' or 'shape' (the categories which would seem to be diagnostic of row - and figural-type descriptions, respectively). This would suggest that cluster (c) of individual subjects comprises those individuals who use path-type descriptions, and this interpretation is supported by examination of typical descriptions generated by subjects within this group: for example-

"Er: (2.5 s) start at the top:right hand corner of the page: (1.5s) come down to the third box: (1 s) turn left: (1 s) and its the first box: (6 sec)".

This cluster contains the largest number of subjects (16) and thus supports the earlier-stated impression of the predominance of path-type descriptions.

Of the seven lesser clusters of individual subjects, cluster (f) would again appear to contain subjects who use path-type descriptions, as would cluster (i) since both groups of subjects use most of the lexical categories used by the subjects comprising cluster (c) without the use of either 'row' or 'shape', the critical categories diagnostic of row- and figural-type descriptions, respectively). Cluster (h) of subjects comprises a group of subjects who use edge names, prepositions ('to edge'), figural terms ('shape') and also 'column'. These subjects would thus appear to be using some figural-type descriptions. Cluster (g) of subjects comprise five subjects' data and these five subjects described locations in terms of 'row plus column' ('rowpcol'). For example:

> "My second: X: is in the second: row: and the fourth: column:"

The analysis of how subjects sample from the lexical categories confirms the initial casual impression that three of the four description types (path-, row- and figural-types) were used by subjects in the present experiment. As was the case with the data from the maze game task, clusters of lexical categories which consistently occur across several cluster analyses are identifiable in the data from the present experiment, and groupings of subjects (which groupings also consistently emerge across different cluster analyses) can be shown to sample from the clusters of lexical categories in a highly systematic fashion which is explicable in terms of the description types summarized in chapter 6 as path-, row- (or line-) and figural-type descriptions.

d) Discussion

The analyses presented in this chapter indicate that a partial replication of the four types of location description which had been observed in the maze game data has been obtained in the present 'paper-and-pencil' experiment. Path-, row- and figural-type descriptions have once again been observed in the present 'paper-and-pencil' experiment; however, matrix-type descriptions are conspicuous by their absence, an interesting departure from the maze game results. The relative prevalence of the three types of description which were replicated in the 'paper-and-pencil' experiment was different as compared to that in the maze game data. In the 'paper-and-pencil' experiment, path-type descriptions were the most common with row-type descriptions being the next most popular and with a small number of figural-type descriptions being generated. In the maze game data, on the other hand, matrix- and row-type descriptions were most prevalent, with path-type descriptions being less prevalent and with a small number of figural-type descriptions being generated.

Thus, the principal differences between the results from the maze game studies and the present experiment were that in the maze game data, four different description types were observed, whereas only three of those four description types were replicated in the present 'paper-and-pencil' task. Furthermore, the relative popularity of the three different description types was different in the 'paper-and-pencil' experiment as opposed to their relative popularities in the maze game data; path-type descriptions were most popular in the 'paper-and-pencil' experiment, yet they rated third most popular in terms of the number of subjects consistently shown to be using them in the maze game experiments. The replication in the 'paper-and-pencil' experiment is therefore a partial replication in two senses.

The reason why there should be only a partial replication of the maze game data in the 'paper-and-pencil' experiment is an interesting question. The answer to this question may in fact be rather a complex one.

The original motivation behind the 'paper-and-pencil' experiment had been to investigate the nature of maze location descriptions which had been generated independently of playing the maze game, using similar mazes (with slight modifications, as detailed in the introduction to the present chapter) and similar subjects. Thus, the 'paper-and-pencil' experiment was designed primarily with the semantics of the task in mind; from a purely semantic point of view, one might have expected similar mazes to be described in similar ways, regardless of whether the descriptions were generated within the context of a computer game or not. To some extent, this point of view was borne out: three of the four location description types which were found in the maze game studies were replicated, despite slight changes in the nature of the mazes. However, a fourth type of description was not replicated, and furthermore, the prevalence of the three description types which were replicated was different in the 'paper-and-pencil' as opposed to the maze game task. This result requires an explanation, and it will be argued that the explanation lies in the differing natures of the maze game and the 'paper-and-pencil' tasks both in terms of their cognitive aspects and in terms of the differing types of social situation which they constitute. It is further argued that these two aspects, cognitive and social, are inextricably linked.

In removing the location description subtask from the maze game scenario and setting it as an experimental task in its own right, three significant aspects of the maze game task have been altered; these are (i) its inherently collaborative nature, (ii) the spontaneity of the situation, and (iii) the <u>dynamic</u> nature of the task. These three aspects of the task have, it is argued, both cognitive and social consequences and all three are intimately interrelated, making it difficult to discuss them separately.

The first point of note is that the collaborative element of the maze game task is largely removed in the 'paper-and-pencil' location description task. The two subjects are no longer (more or less) equally involved in a joint effort to solve a single complex problem in the 'paper-and-pencil' task. Instead, subjects have fixed roles: at any given time, one subject is a 'describer' and the other is an 'understander'. This assignment of roles effectively means that there is less incentive for the 'understander' to make an active, collaborative effort to assist the 'describer' in his task (by generating alternative 'checking' descriptions as a check that he has correctly understood the description generated by the speaker, for example). The (initial, at least) equality and mutual interdependence of the subjects in the maze game is absent in the 'paper-and-pencil' experiment; the two subjects are more independent than interdependent. Furthermore, the potentially deleterious consequences of failing to correctly understand a location description in the maze game task (that is, being blocked off from one's chosen route to goal or even being exposed to the risk of being 'eaten' by the monster if one has misunderstood one's partner's location and agreed to his making a move which in fact has harmful rather than beneficial consequences for one) are absent in the 'paper-and-pencil' task. There is thus less <u>incentive</u> to collaborate in the 'paper-and-pencil' task, and indeed less necessity to do so.

The second aspect of the maze game task which is different vis-a-vis the 'paper-and-pencil' task is the greater degree of spontaneity inherent in the maze game as opposed to the 'paper-and-pencil' task. In the maze game task, the question of which of the two players would describe a given location and at which point in the game he should do so was a choice entirely up to the subjects. In the 'paper-and-pencil' task, on the other hand, these variables were prescribed by the experimenter beforehand. <u>Who</u> it was that described a location, and <u>when</u> they did so was not under the control of the subjects in the 'paper-and-pencil' experiment. This lack of spontaneity in the 'paper-and-pencil' experiment allows the subjects less opportunity for negotiation. In the maze game, the choice of a location description type and the extensional semantics of the chosen description format are both open for negotiation (see chapter 10). In the 'paper-and-pencil' experiment, on the other hand, because the subjects' roles are to a greater extent prescribed by the experimenter, and because there is less necessity and incentive to collaborate, it is more likely that the describer's mental model of the maze will prevail unnegotiated (or at least negotiated to a much lesser extent). Thus, priority is given to the describer's mental model of the maze rather than allowing (or encouraging) the subjects to arrive at a joint model.

The third difference noted above was the difference between the two tasks in terms of their dynamic aspects: the maze game is inherently more dynamic in nature than is the 'paper-and-pencil' task. In the maze game task, the subjects are involved in moving their respective position markers within the maze from starting positions to goal positions - indeed, this is the essence of the game as far as the subjects are concerned. Subjects in the maze game task occasionally describe their position as a subgoal of the larger, ongoing task of getting to the goal position. Subjects in the 'paper-and-pencil' task, on the other hand, have to describe a series of different locations one after the other. These different locations are not connected in any way; the task set to the subjects is one of discretely describing separate locations - these successive separate locations do not represent progress towards a longer term goal as they do in the maze game task. Thus, the 'paper-and-pencil' task is effectively a series of discrete tasks, whereas the maze game task involves the generation of descriptions as part of a single larger ongoing process. As such, it is not unreasonable to suppose that the subjects are not cognitively operating on the mazes in the same fashion in the two tasks; the 'paper-and-pencil' task is more like map-reading, whereas the maze game task is more like travelling within the area represented by a map <u>plus</u> occasional 'map-reading'.

It should be emphasized that the differences between the maze game and the 'paper-and-pencil' tasks in these three aspects (collaboration, spontaneity, and dynamics) are differences of <u>degree</u>; for example, although collaboration is not required to the same extent in the 'paper-and-pencil' as opposed to the maze game task, <u>some</u> collaborative effort (e.g. to ensure that correct understanding has taken place) is still necessary in the 'paper-and-pencil' task.

Some of the potential social consequences of removing the location description subtask from the maze game scenario and treating it as an experimental task in its own right have been touched upon above (for example, the fact that the two subjects are more independent than interdependent in the 'paper-and-pencil task). Other potential social consequences are implicit in the above discussion. Specifically, in the maze game task, the subjects are jointly involved in a collaborative problem-solving task. This task, it was noted above, is of a spontaneous nature: how the subjects set about solving it, when they describe locations and how they do so (etc.) is under <u>their</u> control. As such, the possibility that one subject might exert overall control over the choice of moves, location description format, (etc.), can arise in the maze game. This is less of a possibility in the 'paper-and-pencil' task, because of the fixed roles of the subjects; by prescribing their roles and when descriptions are to be generated, the opportunity for spontaneous negotiation of who (if anybody) should be in overall control of the task is removed.

Another aspect of interest in the present context is this: so far, we have explored differences between the maze game and the 'paper-and-pencil' task in terms of spontaneity, collaboration and dynamics. There are also differences between the 'interactive' and 'non-interactive' conditions of the 'paper-and-pencil' task. In the interactive condition of the 'paper-and-pencil' task, both subjects are involved in generating location descriptions. In the non-interactive condition, on the other hand, only one subject is involved in generating descriptions, whereas the other subject is exclusively involved in listening to descriptions. There is thus a greater opportunity for the two subjects to negotiate the choice of a description format in the interactive condition as opposed to the non-interactive condition. There are in fact examples in the data from the interactive condition in which there does appear to be some negotiation of the choice of a description format. In the following extract of dialogue, for example, player

A had previously described the eight locations on the first maze and (it being the interactive condition) it was then player B's turn to describe the eight locations on the second maze (which is the symmetrical maze with irregular node positions: see Appendx V).

B Ehm: the first X: is ehm: (5 sec) ehm: (2 sec) oh: right: we'll stick to the rows and columns.

A {Mmm-hmm

B {Even although some:bits are missing, O.K.?

A Right: (laughs) Fine:

In this extract, player B offers to continue using the description format used by player A to describe the previous set of eight locations (which she did in terms of stating row position plus column position), despite the fact that this format of description is less appropriate for a maze with an irregular spacing of nodes and having gaps in the path structure. Player A agrees to this. In the non-interactive condition of the 'paper-and-pencil' experiment, on the other hand, only one player is involved in describing the maze and therefore the other interlocutor never has the opportunity to choose between continuing with the previous speaker's chosen description format and changing to another format. In <u>both</u> conditions, of course, the extensional meaning of a profferred description is open to negotiation (see the discussion of negotiation in chapter 10).

In addition to the above-noted differences between the maze game and 'paper-and-pencil' tasks (i.e. in terms of spontaneity, dynamics and collaboration) it is not unreasonable to suggest that there exist further, perhaps more subtle, differences of a social psychological nature. The discussion of these differences will necessarily be somewhat speculative in view of the fact that these differences were not directly tested for. Nevertheless, such possible differences are potentially of importance for the data produced.

In essence, it is argued that the subjects are likely to represent the two tasks (maze game and 'paper-and-pencil' tasks) rather differently, that is, what the subjects understand their task to be is different in each case. This relates to the literature on 'demand characteristics' (Orne, 1962) and that on 'social representations' (Abric, 1984; Codol, 1984). Orne (1962) discusses the experiment as a social situation and notes an apparent willingness on the part of experimental subjects 'to be good subjects', that is, to help validate the experimental hypothesis. Orne defines demand characteristics as ' . . . the totality of cues which convey an experimental hypothesis to the subject . .' (Orne, 1962, p. 779). Orne notes that a subject's behaviour in an experiment will be determined by two sets of variables: (i) those which are traditionally defined as experimental variables, and (ii) the <u>perceived</u> demand characteristics of the experimental situation. It is argued that the perceived demand characteristics are different in the maze game studies as opposed to the 'paper-and-pencil' experiment, and that this difference may partly account for the difference in the results of these two sets of studies.

It is suggested here that the subjects who participated in the maze game studies most probably understood their task to be one of playing a game - that is, solving a problem. This problem (getting to the goal position in the maze game in as few moves as possible and incurring as few penalty points as possible in the process) requires the generation of location descriptions as an incidental part of the whole task. It is thus argued that in the maze game task, the subjects are oriented towards solving the problem that the experimenter appears to have set them (i.e. solving the maze game as efficiently as possible) and consequently do not dwell on the location description aspect of the task. Their concern is with strategy in the game and the achievement of a maximally efficient solution. Generating location descriptions is, for them, a purely instrumental matter which helps them achieve a solution to the larger problem of the game task. It is thus highly likely that they generate descriptions in an unreflecting fashion - indeed, it is this facet of the game task which ensures true spontaneity and naturalness of the location description data, as was argued in chapter 4.

In such situations, where the generation of location descriptions is a purely incidental matter to the solution of the task in hand, any 'short cut' which appears to aid the subjects in generating location descriptions quickly and efficiently will be regarded as useful. Matrix-type descriptions could function in this fashion, being an apparently convenient, brief (in terms of the lengths of the utterances in the descriptions in words) coding system which appears to simplify and speed up the location description process, thus allowing subjects to concentrate on the 'real' task. (The phrase '<u>apparently</u> convenient brief . . coding system' is used because the results presented in chapter 8 suggest that the impression that matrix descriptions are more brief and convenient is somewhat illusory).

In the 'paper-and-pencil' experiment, on the other hand, subjects are explicitly aware from the outset that the study is concerned with the location descriptions that they will generate. (Even although they are not explicitly told this, it is the sole task set to them). Subjects are thus likely to be aware that the generation of location descriptions is the object of the experimenter's interest, and their attention is directed to this task. Furthermore, subjects in the 'paper-and-pencil' experiment are not under pressure of time, and there is no 'monster' to harass them, and so there is less necessity for them to employ 'short cuts' of the type exemplified by matrix-type descriptions.

Thus, it is argued that the two tasks differ in their peceived demand characteristics such that the subjects in the 'paper-and pencil' experiment focus their attention on the location description task, whereas the subjects in the maze game studies pay attention to the problem-solving aspects of the game, employing 'short cut' descriptions to enable them to devote more attention to the 'real' task.

The importance of how subjects represent the experimental task in which they are engaged is emphasized by Abric (1984) and Codol (1984). Abric (1984) has demonstrated that how the subject represents the experimental task, the other person engaged in the task, and the context of the task is critical to how the subject behaves in the task. For example, in the Prisoner's Dilemma game (a mixed-motive, two-person, experimental game derived from games theory in economics, in which the subject can behave either cooperatively or competitively) different behaviours are evoked from the subject depending upon whether he has represented the task as a _game_ (in which case his behaviour is rather more competitive) or as a _problem-solving activity_ (in which case his behaviour is rather more cooperative). Similarly, varying the context of the task has effects on the subject's behaviour. For example, in one case, the Prisoner's Dilemma game is given a 'work scenario' context in which the subject has to put himself in the place of one of the contestants in a work scenario (a trade unionist or a manager) and in another case, he has to place himself in the position of one of the contestants in a casino-gaming scenario. The results obtained confirmed the effect of the context. Abric concludes that it is not the logical structure of the situation which determines subjects' choices in the experimental game task but the meaning of the situation for the subjects.

Thus the work in French social psychology on 'social representations' underlines the importance of how subjects represent the various aspects of the experimental situation (the experimental task, themselves, the other subject(s), and the context of the experimental task) for the outcome of the experiment. This emphasis is entirely in accord with the above interpretation of the difference between the maze game and the 'paper-and-pencil' task in terms of a difference in perceived demand characteristics and a consequent change in the subjects' behaviour.

Implicit in the earlier comparison of maze game and 'paper-and-pencil' tasks in terms of differences in perceived demand characteristics (see above) is another possible source of differences between the two experimental situations: this difference is one of differing degrees of conscious awareness on the part of the subjects of the language they produce in the two experimental situations. It was argued earlier that in the maze game task, the subjects' attention is focussed upon the problem-solving aspects of the task. Subjects in the maze game studies are thus likely to be in that state of awareness known as 'awareness-of-self-as-subject' (or agent) and are less likely to be consciously aware of the language they produce when generating location descriptions than is the case with subjects in the 'paper-and-pencil' experiment. These latter subjects, it was argued above, are more likely to be aware of the fact that the location descriptions that they generate are the object of the experimenter's interest. They are thus more likely to pay conscious attention to the location descriptions which they produce and be 'aware-of-self-as-object' than is the case with the subjects in the maze game studies. It is therefore argued that the two experimental tasks differ in the extent to which subjects are consciously aware of the language that they produce such that subjects in the 'paper and pencil' task pay more attention to the descriptions they generate (and thus generate those descriptions in a less spontaneous and natural fashion) than is the case with the subjects in the maze game studies.

In sum, several differences between the maze game and 'paper-and-pencil' tasks were noted which could account for the observed differences in the results yielded by the two tasks - specifically, the maze game involves more collaboration between the subjects, involves greater spontaneity on the part of the subjects, and is more dynamic than is the 'paper-and-pencil' task. Other differences between the tasks were hypothesized - notably that the perceived demand characteristics, and consequently the subjects' behaviour, differ in the two tasks, that the two tasks might be represented differently by the subjects, and that the subjects are more likely to be 'aware-of-self-as-object' and consequently conscious of the language they produce in the case of the 'paper-and-pencil' task. How subjects did in fact represent the tasks, and what they perceived the demand characteristics of the tasks to be, were not measured in these experiments and hence these latter hypotheses are somewhat speculative. Nevertheless, they do indicate the complexities of the change from game task to 'paper-and-pencil' task from a social point of view (in addition to the, perhaps more obvious, changes from a cognitive point of view).

Thus, the differing natures of the two tasks as social situations was highlighted. Also touched upon were the potential differences from a social point of view between the different conditions within the 'paper and pencil' task. If the points raised in this discussion were correct, then a very interesting series of experiments could be set up using the 'paper-and-pencil' task and varying the nature of the social situation. For example, one could have the two subjects generating individual location descriptions alternately, giving more scope for negotiation; one could introduce penalties of some sort for failure to correctly understand one another's descriptions, thus increasing collaboration; and one could change the context of the task by describing the same location description task as involving the identification of locations within a maze, locations within a street map, and so on. One could also alter the monologue condition by instructing the subject to generate location descriptions which were precise enough for <u>he himself</u> to later identify the described position.

In conclusion, the failure to replicate all of the location description types found in the maze game studies in a task in which maze locations were described independently of playing the maze game was explained in terms of many factors, including changes in the nature of the social situation which the two tasks constitute.

Table 9-1 Clusterings of lexical categories which consistently emerge across different cluster analyses - of the data from the game-independent location description task

(a) Across variation in measure of similarity/ distance on which clustering is based	(b) Across variation in cluster linkage criteria
1. TOP HANDSIDE CORNER MOVE ALONG UP LEFT (Direction) RIGHT (Direction) COUNT DOWN	1. TOP HANDSIDE CORNER MOVE UP LEFT (Direction) RIGHT (Direction) DOWN COUNT
2. BOTTOM LEFT (Edge) RIGHT (Edge) NODE ROWPCOL	2. LEFT (Edge) RIGHT (Edge) NODE
3. ORDER TO EDGE	3. ORDER TO EDGE

ROW

4. SHAPE
COLUMN

5. HORIZONTAL
VERTICAL

6. MIDDLE
LIMB
SYMMETRY
EXTREME

7. DIAGONAL
OWNOTHER
JUNCTION

8. BOTTOM
ROWPCOL

ROW

4. SHAPE
COLUMN

5. HORIZONTAL
VERTICAL

6. LIMB

SYMMETRY
DIAGONAL

7. OWNOTHER
JUNCTION

Table 9-2. Subjects' usages of individual lexical categories in the maze-game independent location description task

Clusters of lexical categories	a N=12	b N=8	c N=16	d N=5	e N=4	f N=5	g N=5	h N=5	i N=4	j N=4
1. TOP	✓	✓	✓	✓	✓	✓		✓	✓	
HANDSIDE	✓		✓		✓	✓			✓	
CORNER			✓			✓			✓	
MOVE	✓		✓	✓		✓			✓	
(ALONG)	✓		✓	✓	✓			✓	✓	
UP			✓	✓		✓			✓	
LEFT (D)			✓	✓		✓				
RIGHT (D)			✓	✓		✓			✓	
COUNT			✓				✓	✓	✓	
DOWN	✓	✓	✓	✓	✓	✓		✓	✓	✓
2. (BOTTOM)	✓	✓	✓	✓	✓	✓		✓	✓	✓
LEFT (E)	✓	✓	✓	✓		✓		✓	✓	✓
RIGHT (E)	✓	✓	✓	✓		✓		✓	✓	✓
NODE	✓		✓	✓	✓	✓		✓	✓	✓
(ROWPCOL)							✓			
3. ORDER	✓	✓	✓	✓	✓		✓	✓	✓	✓
TO EDGE	✓	✓			✓			✓		✓
ROW	✓	✓								
4. SHAPE								✓		
COLUMN								✓	✓	

Table 9-2 Cont

Clusters of lexical categories	a N=12	b N=8	c N=16	d N=5	e N=4	f N=5	g N=5	h N=5	i N=4	j N=4
5. HORIZONTAL				✓				✓		
VERTICAL										
6. (MIDDLE)										
LIMB										
SYMMETRY										
(EXTREME)		✓		✓						
7. (DIAGONAL)									✓	
OWNOTHER						✓			✓	
JUNCTION						✓				

N.B. Those categories marked with a tick (✓) are used by subjects within a cluster, on average, once or more than once. Those in parentheses are included in the cluster they are in in only one analysis, either analysis (a) or analysis (b) (see table 9-1) Categories with (D) written alongside in parentheses are direction - specifiers, those with (E) alongside are edge names.

CHAPTER 10

PRAGMATICS AND MENTAL MODELS

a) Introduction

In the previous three chapters, formal analyses were presented which confirmed the informal four-fold classification of location descriptions which was first presented in chapter 6, and some further properties of the four description types were also explored. In chapter 8, it was suggested that a potentially profitable approach to the data, theoretically speaking, would be in terms of interpreting the four location description types as reflecting different underlying types of <u>mental model</u> of the maze shape.

Some of the semantic aspects of this interpretation of the data were explored in chapter 8, in which it was pointed out that mental models of the maze shape, as conceived of in the present theoretical treatment, would be truly intermediate between language and the 'world' under discussion (i.e. the maze shape on the screen). Thus, a description of a location would be translated into a procedure (to pursue the procedural semantics metaphor) which would, in the first instance, be executed in relation to the model and, having thus conceptually 'marked' the relevant location on the model, only then would scanning of the 'world' take place.

This theoretical approach has one important advantage: if a model is constructed by the subject, then the greater the simplicity of the model as compared to the original stimulus, the easier it is to generate and interpret descriptions. A 'row'-type model of the maze, for example, would simply consist in a series of parallel lines representing the horizontal pathways of the maze. Vertical pathways, and gaps in the horizontal pathways on the actual maze, would be ignored on the model, thus simplifying description generation and interpretation.

The present chapter will focus on pragmatic, rather than semantic, considerations in relation to the 'mental models' interpretation of the data. It is perhaps best to state clearly at the outset of the present chapter what the present interpretation of the data is. It will be argued that the data are best interpreted in <u>pragmatic</u> terms: we shall emphasize heavily the importance of the meanings that a particular speaker and listener attach to a given set of expressions, and we shall argue that these meanings will, in all probability, be different from those attached to the <u>same</u> expressions by a <u>different</u> speaker and listener.

The largest section of this chapter is devoted to a discussion (with reference to examples taken from the data) of one possible mechanism by means of which an individual speaker and listener can come to select one interpretation (from the many possible interpretations) of their chosen type of location description: it is argued that speakers and listeners tacitly (or sometimes explicitly) <u>negotiate the meanings of the expressions that they use</u>.

Thus, for example, a particular speaker and listener may both conceptualize the maze shape in terms of a 'matrix' type of mental model, and consequently generate expressions like '(3, 2)' when describing maze locations. It is argued here that the extensional meanings of the procedures corresponding to that expression are negotiated by the speaker and listener. The expression '(3, 2)' could refer to a variety of locations depending upon the location of the origin of the coordinate system and depending upon whether the count procedures corresponding to the two numbers in the description involve counting <u>path movements</u> or <u>node positions</u> (see the discussion in chapter 6 and figure 6-12 for an explanation of this distinction). Thus, within each class of mental model, there is great scope for variation in the extensional meanings of the procedures implicit in the location description of that model type, and it is argued that <u>the extensions of the procedures implicit in the location descriptions are tacitly negotiated and agreed-upon by particular speakers and listeners</u>.

A number of examples from the data will be discussed and, it is argued, will illustrate these points: <u>both</u> the choice of a description type, <u>and</u> the meanings of the expressions comprising location descriptions are negotiated outcomes.

This view of mental models, as social representations in which both the choice of model and the meanings of its procedures are negotiated outcomes, contrasts with the views expounded by Johnson-Laird (1980, 1981a, 1981b). Johnson-Laird emphasizes that function of language which enables us to gain access to another person's experience of the world: models of perceptually absent objects, persons and situations are built by a listener (or reader) on the basis of the discourse which he hears (or reads). As such, Johnson-Laird's emphasis is less explicitly social than is the present notion of a 'mental model'. The implications of the present notion of a mental model and the contrast between the present account and that of Johnson-Laird will be further explored in the following, final chapter.

This consideration of the data from the maze game (that is, in terms of negotiated models) will allow an examination of the reassignment games from the maze game studies and the monologue condition of the 'paper-and-pencil' experiment in the light of the hypothesis that the extensional meanings of the location descriptions are negotiated.

In the following chapter, the potential theoretical implications of the results from the present set of studies will be explored.

b) The negotiation of the meanings of the location descriptions

Given that there appears to be a tendency for both interlocutors participating in a given dialogue to use the same type of referring expression as one another (see chapter 7, section c)) the question naturally arises as to how the choice of a particular description type by a pair of subjects is achieved.

Another, related, question also demands an answer: given the variety of interpretations of a given expression (e.g. 'row') across dialogues, how do a given pair of subjects using that expression manage to use it in such an apparently unambiguous, trouble-free fashion? Each expression does appear to have a variety of possible interpretations. Why are subjects not in difficulty in the face of such ambiguity of meaning? For example, the term 'line' has been observed to have (at least) four possible intepretations in the context of the maze game. These are: (i) 'Line' could be used to refer exclusively to horizontal lines of nodes (rows) or, alternatively, to refer exclusively to vertical lines of nodes (columns). Both of these 'narrow' usages are observed in the data of different dyads. (ii) 'Line' could have a more generic interpretation: it could be taken to refer to any line of nodes within the maze structure, regardless of orientation or length (it need not refer to lines of nodes which span the entire distance across the maze structure as 'rows' usually do). (iii) 'Line' could refer to an imaginary line of symmetry dissecting the maze. (iv) More unusually, the term 'line' was once used by a pair of subjects playing a 'monster' game on a symmetrical maze to refer to an imaginary pair of diagonal lines in the maze structure from bottom left hand corner to top right hand corner, and from top left hand corner to bottom right hand corner. These imaginary diagonal lines were used to generate descriptions of the following type (see figure 10-1 for an illustration of the point being described):

A Yeah: {I'm: s-: two up along the diagonal: from
{ bottom left.
B {Yeah:
A To top right: . . .

This last interpretation, of an imaginary <u>diagonal</u> line, whilst more idiosyncratic in nature than interpretations (i) - (iii) of 'line', is no less intuitively reasonable.

Thus, there is a great deal of variation in the extensions of a given term, and it is a matter of great interest to attempt to ascertain how it is that subjects use such highly ambiguous expressions without any apparent difficulty of interpretation. For example (see figure 10-2 for an illustration of the points being described):

1) A I'm on: the second row, second box along.

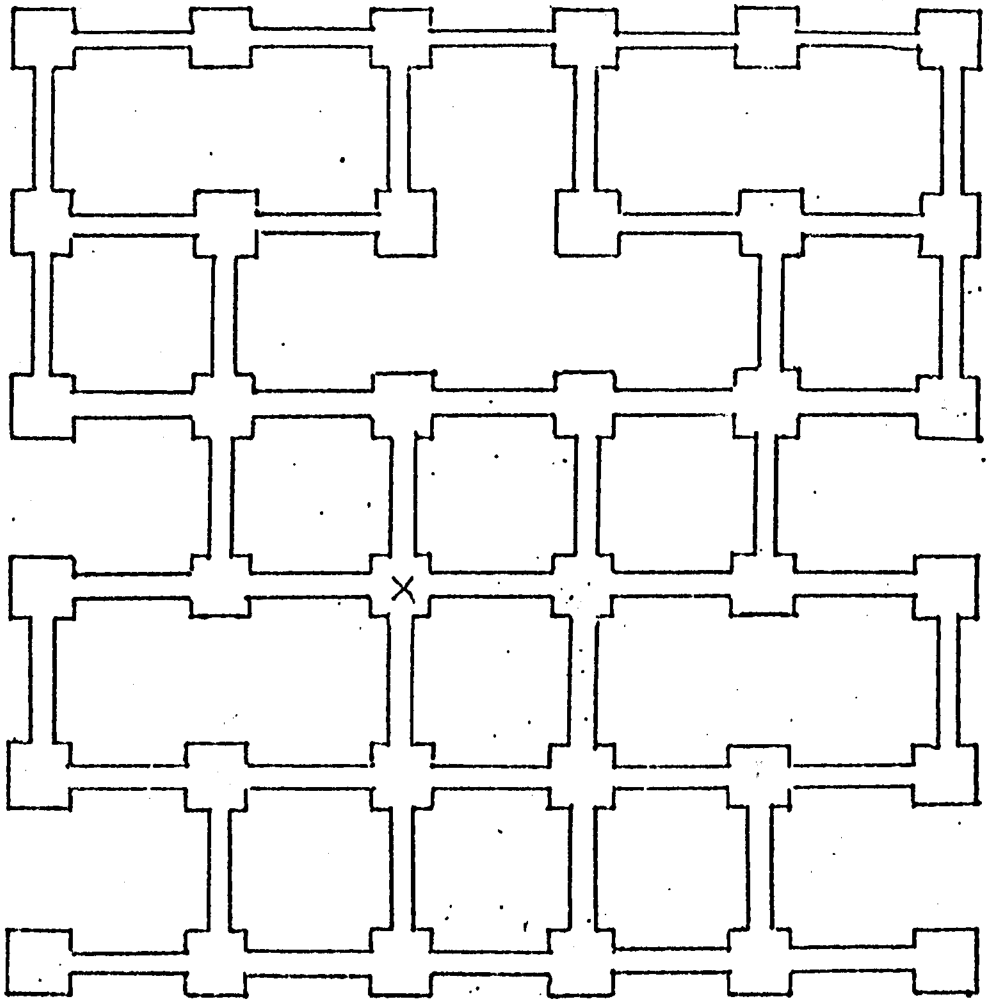

X

Figure 10-1; the location (X) described as "two up along the diagonal frombottom left to top right".

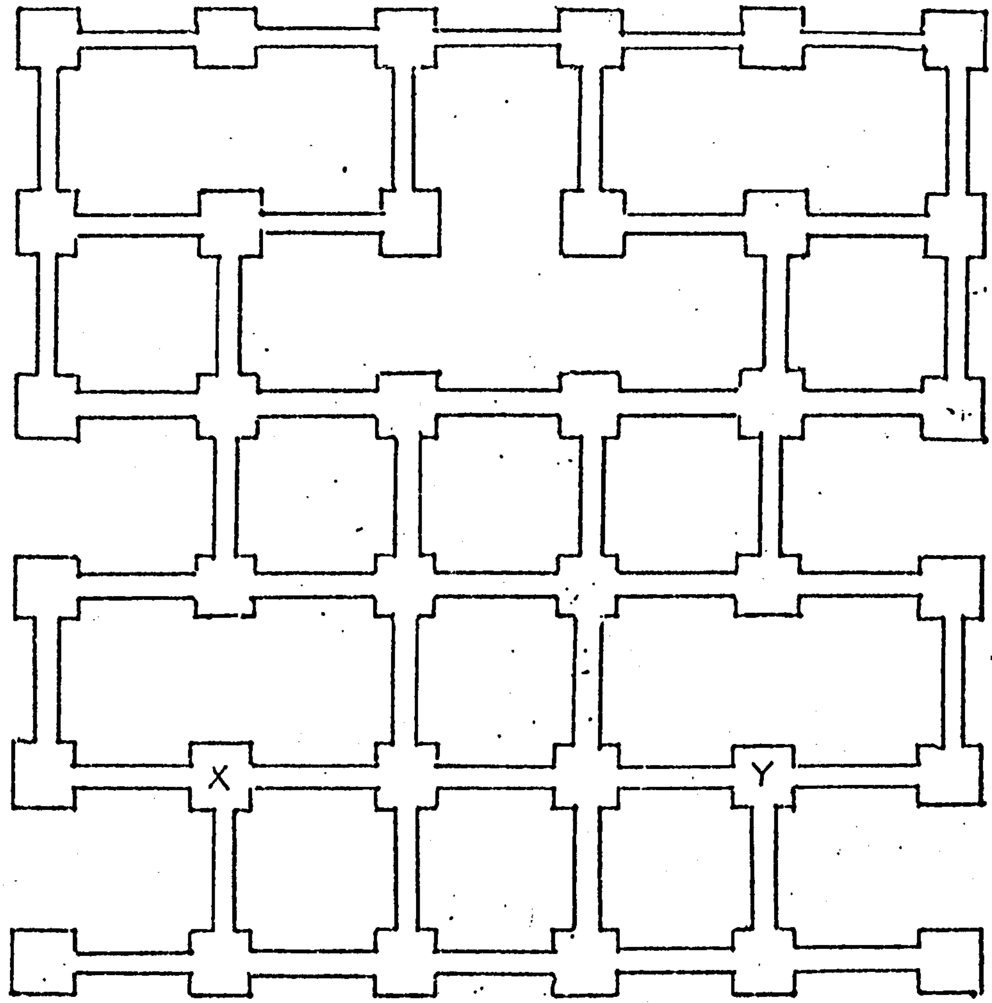


Figure 10-2: The point(X) described as on "the second row, second box along" and the point (Y) described as "the second row, second last box along . . . from the left".

2) B And I'm on the second row on the second last box along.

3) A O.K.

4) B From the left.

In this extract, player A generates a highly ambiguous description (second row from where? Second box along from where? Is 'second' the second <u>node position</u> or the second <u>path movement?</u> etc.) Player B does not appear to ponder on the potential ambiguity of the description profferred by player A, but instead immediately (without any delay whatsoever) generates a similar description of her own.

In the present section of this chapter, a discursive exposition of some of the description sequences in some transcripts of dialogues will be given and the interpretation placed upon the data will be that each pair of subjects tacitly (or explicitly) <u>negotiates</u> the extensional meaning of their chosen type of location description. A discursive, rather than a statistical, hypothesis-testing approach will be adopted, because the data base from the maze game studies is insufficiently large for a proper statistical test of the hypothesis that negotiation of meaning takes place. Nevertheless, the qualitative data presented do provide interesting support for a hypothesis couched in terms of the negotiation of meaning.

Essentially, it is argued that both subjects come to entertain similar mental models of the maze. This is achieved by their negotiating the choice of a description type and subsequently negotiating the extensional semantics of the chosen description type, until a tacit agreement is reached as to how the expressions comprising the descriptions are to be interpreted. This effectively eliminates the potential ambiguity of the referring expressions, and thereafter, rapid interpretation is achieved of what are, at face value, rather ambiguous expressions.

The sort of negotiation of meaning adumbrated above is readily apparent in some of the transcripts of practice games (and those games played on full-size mazes which were first-played games). An example of the negotiation of the choice of a type of description comes from a game in which two female subjects were playing a 'baseline' game (see figure 10-3 for a depiction of the locations being described). The players describe their locations as follows:

1) B So where are you aiming for?

2) A I'm aiming for: right. If you take each wee box as a coordinate position, (3, 4).

3) B Oh.

4) A (3, 5) rather.

5) B What's three? Three is - (A interrupts)

6) A Right, O.K., say the top left and you go along two and down one.

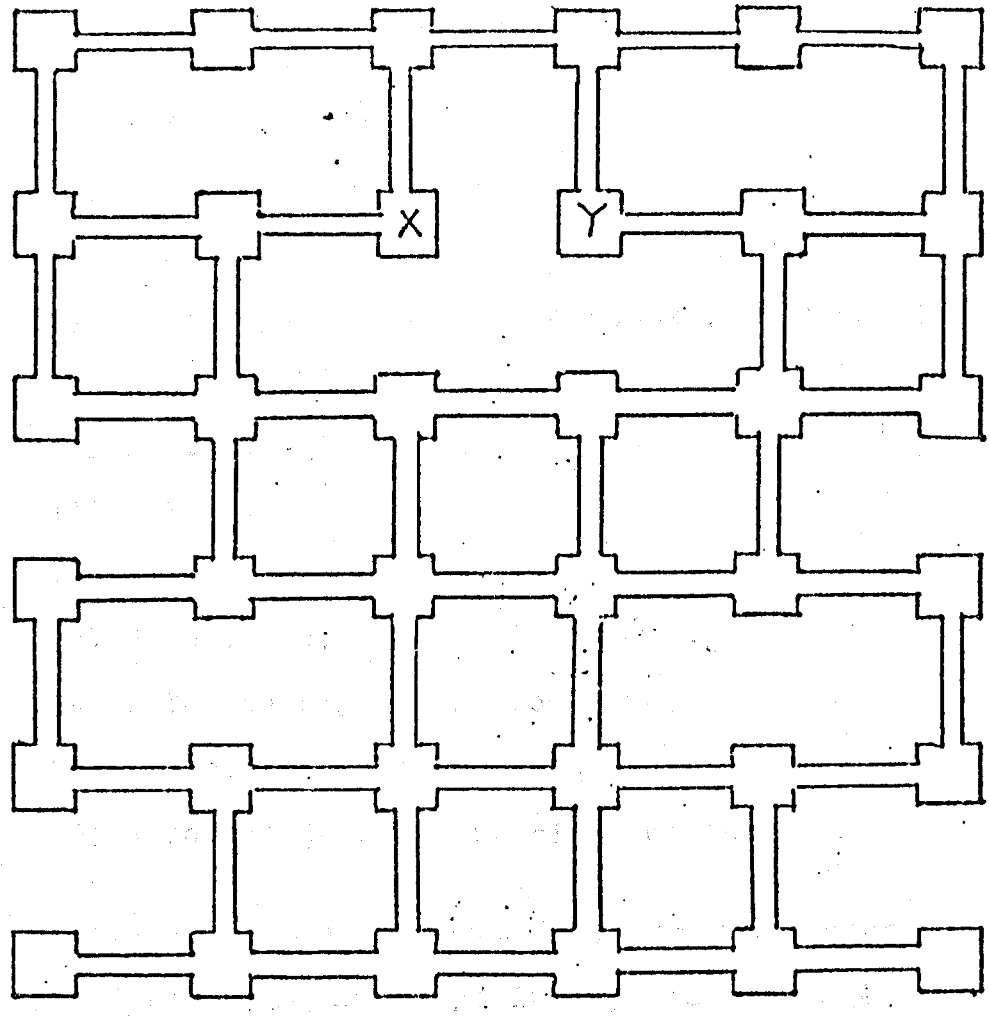


Figure 10-3: The point (X) described both as "(3, 5)" and "the top left and you go along two and down one" . . "you're (Y) just in the one to the right of that."

7) B Yeah - oh yes.

8) A Right? And you're just in the one to the right of that.

9) B Have you got - wait a minute, you're going along two -

10) A Right - Top left.

11) B Right.

12) A Go along two.

13) B One, two.

14) A To the right, and down one.

15) B One. Yeah, I'm just to - (A interrupts)

16) A You're just to the right of that.

This extract is interesting. In utterances 1) - 4), player B asks player A's position, and player A generates a matrix- or coordinate-type description which has as its origin the bottom left-hand corner node of the maze, and which involves the counting of node positions (as opposed to path movements) in the execution of the procedures implicit in the description (see figure 10-3). Player B's initial response (utterance number 3)) is non-committal but in utterance 5), she attempts to clarify for herself the nature of the procedures implicit in player A's description (specifically, she needs to know the location of the origin of the matrix system and whether the scanning procedures involve counting node positions or path movements). Player B seems ready to accept this form of description provided its extensional semantics are made clear; she seems willing enough to negotiate. In the next utterance, however (utterance 6), player A does not take up player B's implicit offer to negotiate, and instead tries an entirely different type of description, a path-type description which counts path movements in its scanning procedures, in contrast to the previous matrix-type description generated by the same player (see figure 10-3). In utterance 7), player B seems to understand this path description ("Yeah-oh yes") and player A follows the path description up in utterance 8) with a further confirming and clarifying description in terms of player B's (previously specified) position. However, player A's path-type description of utterance 6), although specifying the prominent point from which scanning movements are to be made, does not specify whether, in the execution of the scanning procedures, node positions or path movements are to be counted. Consequently, player B is heard to commence the generation of a further clarifying description in utterance 9). In utterances 10) - 15), a decomposed path type description is given by player A, and player B is heard to execute the relevant scanning procedures as the parts of the description are given by player A. A final checking of the description is made in utterances 15) and 16), where player B begins to describe her own location relative to that which has just been described, but is interrupted. Thereafter, and also in a succeeding game played by the same dyad, path-type descriptions were used consistently.

An interesting point is worth noting as an aside: on reading the preceding extract of dialogue and considering it carefully, one gets the impression that player A is in some sense in overall control of what is going on at this point (cf. the discussion in chapter 3 of this thesis of Lewis's notion of control). Player A initiates more than one type of description and interrupts player B three times (in utterances 5)/6), 9)/10), and 15)/16)). Little more can be said about the notion of control, except that the number of 'initiatives', in the sense of introducing new terminology, and the number of interruptions made by each interlocutor might also be indicators of which player is 'master' and which is 'slave'.

This extract of dialogue illustrates how the choice of a type of description to be used throughout the dialogue is negotiated; more than one type of description is used initially, one of these types is selected for use and the details of the extensional semantics of its implicit procedures are clarified, and thereafter the negotiated description type is used predominantly throughout the game.*

* It is used only predominantly, rather than exclusively, because players often describe their locations, particularly towards the end of a game, in terms of path movements away from their, previously specified, goal positions. This type of description, relative to a previously specified position, is used in the case of all types of game and all types of predominant description type.

It was argued above that not only the choice of a description type to be used predominantly is negotiated, but the details of the chosen type of description's extensional semantics are also negotiated. Perhaps the most obviously negotiated type of description is the matrix-type: being of an abstract nature, the clarification of the extensional semantics of the scanning procedures of matrix descriptions by means of negotiation is often a necessary practice. The following extract of dialogue comes from a practice game (which was played on the small '4 x 4' practice maze) by two male subjects. Again, the choice of a description type to be used overall had been under negotiation: both row and path-type descriptions had been used by the subjects.

In the following exchange, player A begins by clarifying a previous statement to the effect that he would move in an upward direction (see figure 10-4 for a diagram of the location being described).

1) A And that'll put me: (1s) on: the third row up: first square from the left:

2) B O.K.

3) A Second square: you know:

4) B Aye

5) A On: on the right if you see:

6) B (laughs)

7) A If you see what I mean (laughs)

8) B You're getting me confused here.

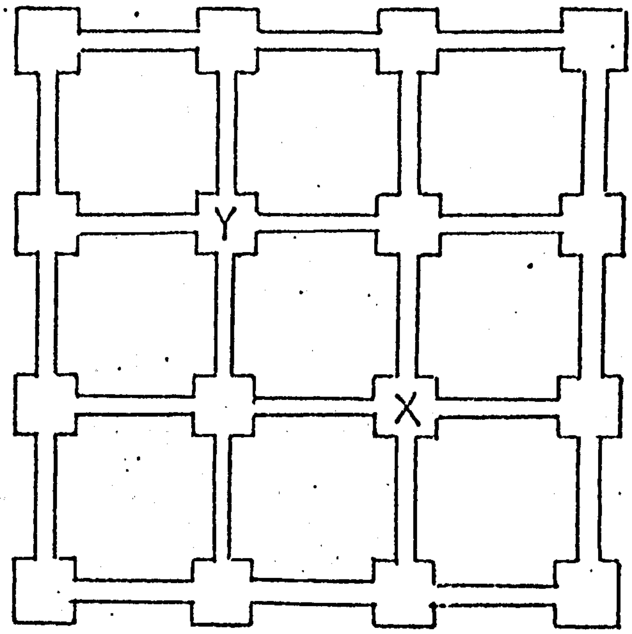


Figure 10-4: the location (X) described as ". . . . . .first square from the left", and as (3, 1): and the location (Y) described as "(2, 2)". Note the difference in the count procedure corresponding to the two coordinates; one, the horizontal coordinate, counts node positions, the other coordinate counts path movements.

9) A We should have a little - we should: O.K. right: make an imaginary grid: starting from the [bottom left: that's 1, 2, 3, 4, y'know?

10) B [O.K.

bottom left:

11) A Bottom left and along.

12) B 1, 2, 3, 4.

13) A So always the: the first number you say is the bottom, right?

14) B 1, 2, 3, 4: 5, 6, 7, 8: 9, 10, 11, 12: 13, 14, 15, 16:

15) A No, no, no, no, just make it a grid:

16) B O.K. Ah, right. [O.K.

17) A You know, like Cartesian [thingies: and- and the same-starting from the bottom: er: right hand corner, 1, 2, 3, 4 up:

18) B O.K. Right: [Got you now]

19) A [Y-: y-: see:] so I am on: let's think: (2.5 sec) (laughs) (3, 1).

20) B (3:1 O.K.: right, I'm on: (1s) oh thr-: (1s) A-: aye (3,1) means you're on the bottom grid then?

21) A Oh, it depends which way you're reading it I guess: O.K.: Y coordinates first: that - Y is going down along the bottom, right, imagine along the bottom, 3 along and 1 up, O.K.?

22) B Count 3 along: and 1 up:

23) A Starting from the left:

24) B Right.

25) A (3, 1)

26) B You're in one of my asses-: my-s-asses.

27) A Yeah, that's right, that's how I changed it.

28) B And: I'm on: (3 sec) 2: (1s) (2, 2): (3 sec)

29) A Oh you're: damn high

30) B Yeah.

(A discussion of S box positions follows, after which the following exchange takes place).

31) B 1, 2, 3 and then: oh, do you count the bottom square as one?

32) A Yeah: you must.

33) B Right, one, two, three, yeah: (3, 3):

The above lengthy extract illustrates clearly the problems posed by the negotiation of the extensional semantics of a matrix system of descriptions. In utterances 1) - 8), player A attempts to clarify the extension of the 'count along row' scanning procedure which he employed as part of the row description which he has just generated. It is a node- rather than a path-movement-counting scanning procedure (see figure 10-4 and utterances 1) - 8) above). In utterances 9) - 13) player A suggests 'making an imaginary grid' and proceeds to clarify the extensional semantics of the scanning procedure corresponding to one of the coordinates of the matrix-type description (that coordinate which denotes locations on the horizontal axis of the coordinate system, that is, that axis corresponding to the X-coordinate of the Cartesian system - although note that player A describes it as the Y coordinate in utterance 21)). In utterance 14), player B effectively suggests numbering each node of the maze individually, a suggestion which is rejected by player A in utterances 15) and 17), in which he explicitly mentions the Cartesian-type coordinate system which he has in mind. He wishes to set up a two-place coordinate system which numbers nodes from 1 to 4 from the left in the horizontal plane and from 1 to 4 from the bottom in the vertical plane. At this point, a satisfactory coordinate system appears to have been set up.

However, further confusions follow: player A describes his location as (3, 1) in utterance 19) when he should in fact have decribed it as (3, 2) according to the coding system which he has just set up with player B (see figure 10-4). Player B correctly deduces that this description '(3, 1)' would place player A on the bottom horizontal line of nodes if A's description were true and the coordinate system worked as they had just set it up to work. In utterances 21) - 27) the actual position of the location is clarified. Notice that this involves a path-type description (in utterance 21): 'three along and one up'). Notice also that, in clarifying the working of the coordinate system, the system itself is modified: the nodes are still numbered 1-4 in the horizontal plane but are now numbered 0-3 in the vertical plane; that is, path movements are being counted in the vertical plane. Player B also uses this modified system to describe his location in utterance 28).

Following a short discussion of S box locations, the above inconsistency in counting is resolved in utterances 31) - 33), where player B questions the counting procedure in the vertical plane, which has been a source of inconsistency in previous utterances. Having thus finally clarified the situation, the matrix system is used again in the subsequent dialogue until a solution to the game is achieved. However, this dyad subsequently discarded the coordinate system in the 'monster' game that followed the practice game from which this extract of dialogue came: path-type descriptions were used instead.

The above extract does illustrate the negotiation of the semantic system which the matrix-type mental models of the maze constitute: both the selection of a particular type of description and the clarification of the extensional semantics of the procedures corresponding to the descriptions are <u>negotiated</u> outcomes.

The final extracts from transcripts illustrate the negotiation of the extensional semantics of procedures corresponding to particular description types.

The first such extract involves a game played on the 'column biased' maze in which the two female subjects were generating line-type descriptions which were of the 'column' type. The third description of the game involves negotiation of the 'move along line' procedure of the column-type description (see figure 10-5).

1) A Right: well I'm in the: left hand side: (1s)

2) B Uh-huh:

3) A And I've got: an S: uhm: (B interrupts)

4) B No wait a- which- which- which column of
[the six? (A interrupts)

5) A [On the very: on the very: first: [line:right?]

6) B [uh-huh:yeah]

7) A I'm not in the- (1s) not the ground line, the
first floor line: say:

8) B O.K.: right.

9) A No: I'm in the second floor line, sorry:

10) B So you're: three: one, two, three from the: (1s)

11) A From the: [thirty-nine moves, where it says moves]

12) B [Bottom Yes: uh-huh]

13) A One, two, three, right?

In the game from which this extract of dialogue is taken, the choice of 'column'-type descriptions has already been settled upon. As can be seen from figure 10-5, it is quite a natural choice. In utterance 1), player A initially fails to completely specify which column her asterisk is located in; it is in fact located in the

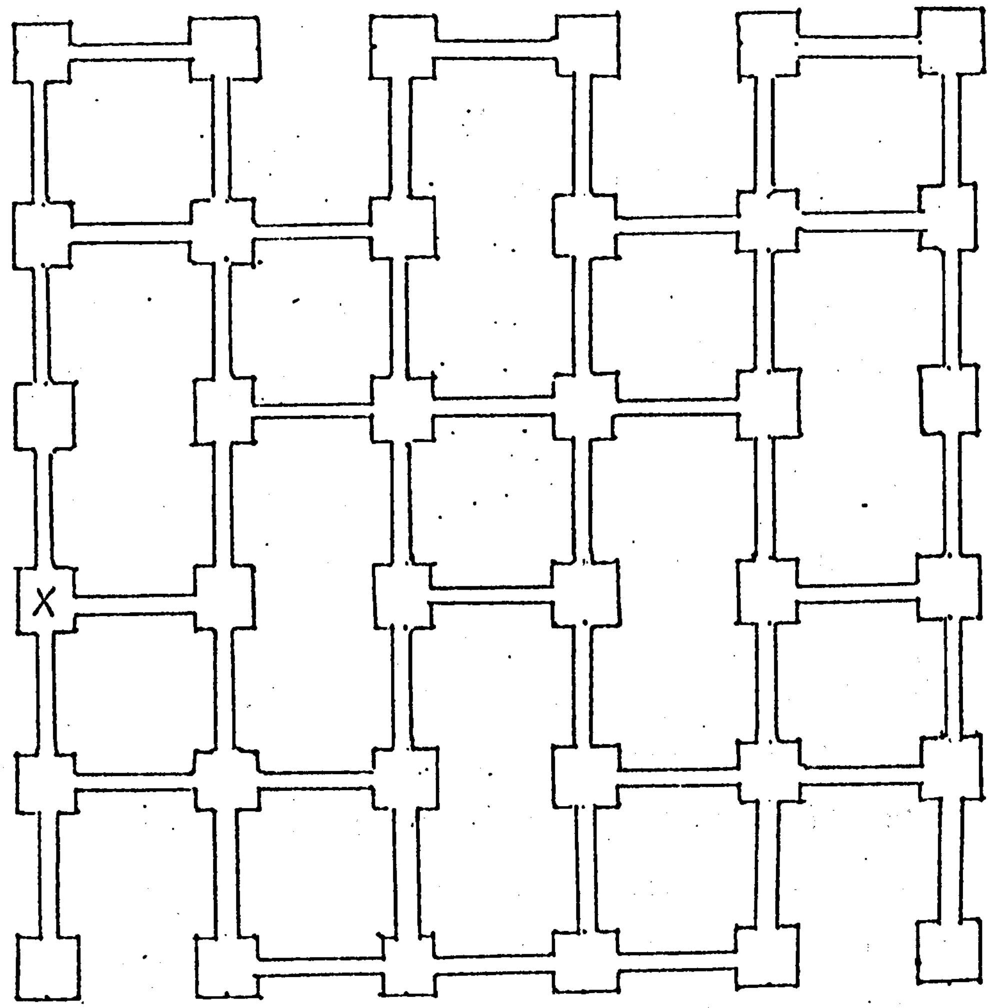


Figure 10-5: The location described as "the very first line: . . . not the ground line, the first floor line . . . No: I'm in the second floor line, sorry."

leftmost column. Utterances 2) - 5) involve clarification of which is the appropriate column. Notice that in utterance 5), player A refers to this as 'the very first line' ; the notation which denotes columns has already been tacitly agreed by both subjects. The columns are termed first, second, third, etc. in order from left to right.

In utterances 7) and 9), player A employs a fairly ingenious solution to the problem of whether she is specifying path movements or node positions when counting out the 'move along line' scanning procedure, a problem which bedevilled the male subjects whose dialogue was discussed above, in the previous dialogue extract (see also figure 10-4). Her solution is to describe the vertical node positions by means of metaphorical reference to the floors of a building, with the 'ground floor' horizontal line of nodes being the bottom-most horizontal line of nodes, the 'first floor' line being the line above it, and so on. In utterance 10), player B checks the description profferred by player A by counting nodes, and, in order to establish the node-counting procedure firmly, player A makes reference in utterance 11) to the tally of moves and penalty points on the screen display, which is located immediately below the bottom of the maze on the screen. Thus, the two subjects establish that the 'move along line' scanning procedure counts nodes rather than path movements, each subject doing so in her own terms in a negotiation process. The negotiation involves one subject generating a description, and the other subject restating the description in different terms as a check that the first description was properly understood and that both descriptions did refer to the same maze location.

This same phenomenon of negotiation by means of checking the extension of a profferred description by generating a slightly different 'checking' description is again exemplified in the final extracts of dialogue, which are from a practice game with female subjects. The following extracts of dialogue also illustrate strikingly the confusions which can arise because the extension of a count-type scanning procedure has not been adequately clarified (i.e. whether node positions or path movements are being counted). The extracts again point to negotiation as the solution to the problem (see figure 10-6 for a diagram of the location which is being described):

1) B Right: I'm in the: I'm: what: (2s) second from the right:
2) A Wai-: oh: (laughs)
3) B Second from the bottom (laughing)
4) A Second from the right: sec-: so you're in: still in the top: left hand corner?
5) B No, I'm in the bottom right hand corner.
6) A (laughs) You're joking. Wait a minute (1s) Oh second from the left?: second- (B interrupts)

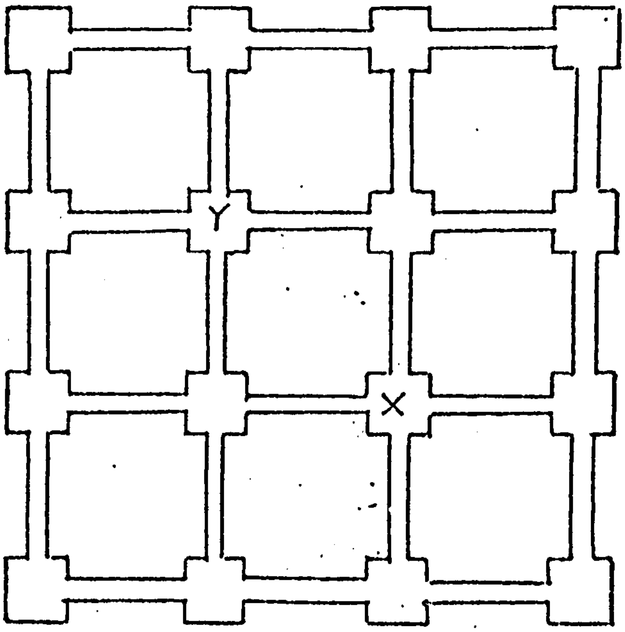


Figure 10-6: the location (X) which player B describes as "second from the right . . . second from the bottom" (Y) is where her partner believes her to be.

7) B No second from the: second from the right.

8) A (laughs) Wai-: take it from the right hand corner.

9) B Right

10) A And: right, go along and up. Tell me what you've done.

11) B One along, one up.

12) A One al-: yeah that's: (2 sec) right:

13) B O.K.?

Later in the same game, the following lengthy extract again illustrates the fact that player B counts nodes when scanning the maze and player A counts path movements (both extracts are discussed below): (see figure 10-7):

14) A . . . and where are you then?

15) B I'm in the: second from the l: left and I'm the: second from the top.

16) A Second from the left, second from the top: right: that's one of my Ss.

17) B Is it? (1 s) It's one of mine.

18) A Wait a minute. Second: second from the top?

19) B Uh-huh.

20) A No you're not: second . . .

21) B So you're still in the bottom: left-right hand corner, the corner of it?

22) B No: {yeah

23) A {Oh: yeah

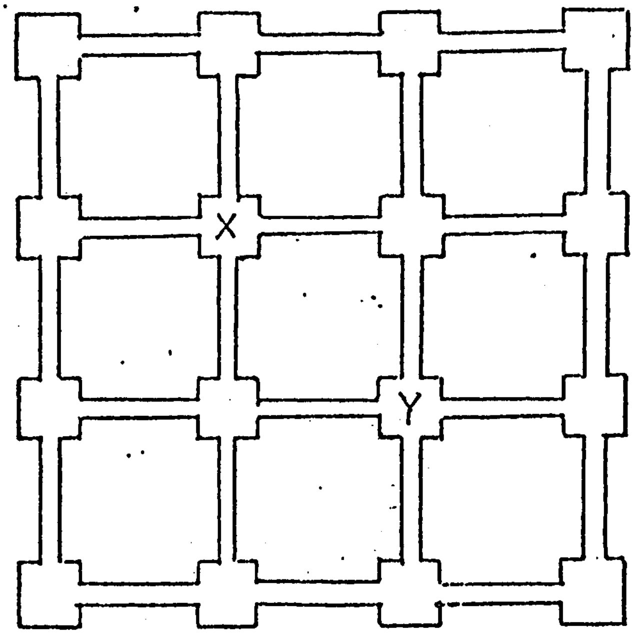


Figure 10-7: The location (X) described by player B as "second from the left . . . second from the top", and the location (Y), which is where player A understands player B's position marker to be located.

24) B I'm in the bottom right hand corner of the top left hand square: (laughing) is that what you mean?

25) A No (laughs)

26) B (laughs)

27) A Right: take it: (1 sec) take it from the left,

27) A right? {like - (B interrupts)

28) B {Right I'm: one along: {and two up}

29) A {One along: } and two up. One a-: right. So you're really: diagonally opposite the asterisk?

30) B I'm: well: your asterisk, aye.

In the first extract, player B describes her location in utterances 1) and 3) in terms of distances from two edges, an infrequently-used type of description, the procedures of which operate in a manner similar to those of matrix descriptions (yielding an intersection of two lines of nodes). Player B is describing location X in figure 10-6; she is counting <u>node</u> <u>positions</u>. Player A, however, evidently counts <u>path</u> <u>movements</u> in her execution of the scanning procedures corresponding to this description. This results in her believing player B to be occupying the position Y in figure 10-6: that this is the case is evidenced by her description of B's location as 'in the top: left hand corner' in utterance 4) and B's reply that her position marker is in fact located in the bottom right hand corner (it is indeed located in the lower right area of the maze) takes player A by surprise ('You're joking'). In utterances 8) and 10), player A attempts to clarify the position for herself by asking her partner to generate a path-type description, which she does in utterance 11). This description appears to succeed, although notice that player B in utterance 13) ('O.K.?') continues to check that player A is sure she has understood the description fully.

In the second extract from this dialogue, the same problem arises again: player B again generates a description couched in terms of distances from two edges, the relevant distances being specified in terms of <u>node positions</u> (see figure 10-7 for an illustration of the location being described). Again, player A understands player B's location to be diagonally opposite where it actually is (in figure 10-7, X is player B's position marker's actual location, and Y is where player A thinks player B's position marker is located). In utterance 22), player B appears to be confused by player A's clarifying question in utterance 21), and consequently generates a further 'checking' description which is intended to clarify the situation. This description (utterance 24) - 'the bottom right hand corner of the top left hand square') is a figural-type description: the 'square' described is a square in the path structure of the maze shape. Player A rejects this clarifying description and in utterance 27) again requests a path-type description, which player B generates in utterance 28). This is checked by player A by means of a further checking description phrased as a question in utterance 29). This is again a figural-type description.

Perhaps not surprisingly, this pair of subjects used path-type descriptions in the subsequent discourse, since other types posed problems for them. In fact, some of the path-type descriptions subsequently used by them were highly elliptical, resembling matrix-type descriptions:

A So you're like: (1, 3) from the left hand side

(2 sec)

B Yeah

(See figure 10-8).

This is a path-movement-counting path-type description of such a highly elliptical form that the horizontal scanning movement (of one path length) and the vertical scanning movement (of three path lengths) in the path description are stated as two numbers. However, the use of path-type descriptions, regardless of how elliptical they are, is seen to be a negotiated outcome in this instance: path-type descriptions are the least problematic type of description for this pair of subjects and come to be used by this pair of subjects for that reason.

The foregoing extracts of dialogue and the discussion

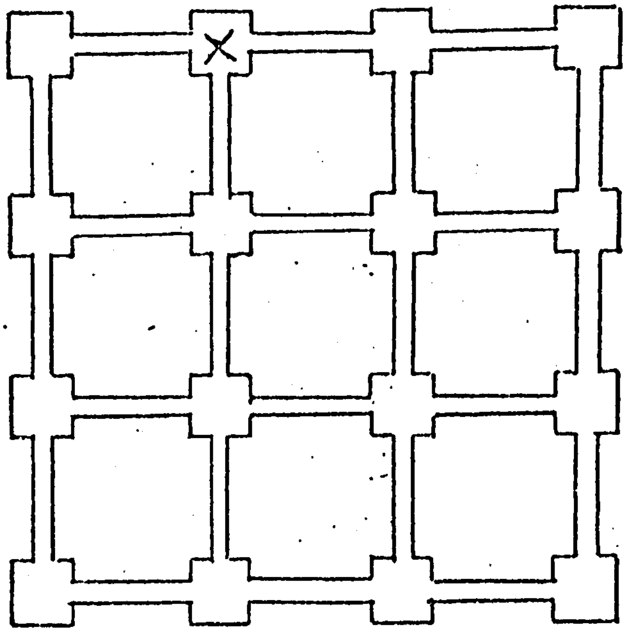


Figure 10-8: The location described as "(1, 3) from the left hand side".

of them serve to underline the hypothesis that the meanings of the location descriptions, and indeed the choice of location description type, are negotiated by each pair of subjects.

The extracts also underline the point that both subjects within a dyad need to entertain very similar interpretations of the expressions they use: the last extract in particular illustrated the confusions that result when the two subjects execute their scanning procedures slightly differently.

The examples chosen involve fairly explicit negotiation; in other transcripts, the negotiation is much more subtle, and in some cases it appears to be absent entirely. This latter eventuality is possible where subjects are describing locations on a simple maze (for example, the practice maze) and both players hold similar interpretations of what the count procedures actually count (node positions or path movements) from the outset. In such cases, the commonality of interpretation occurs more by happy accident than by negotiation.

As can be readily appreciated from the discussion of the extracts of dialogue presented in this chapter, a hallmark of the hypothesized negotiation process is the clarification of what the extensional semantics of the procedures corresponding to the descriptions are. Such '<u>procedure clarification statements</u>' (or questions) would appear to some extent to be diagnostic of negotiation.

The classic example of such a procedure clarification question is 'Do you count the bottom row as one?' (which, in effect, is an effort to ascertain whether the count procedures implicit in a speaker's description count <u>node positions</u> or <u>path movements</u>). If the hypothesis that the extensional semantics of the maze location descriptions are negotiated is correct, one would expect such statements and questions to be more prevalent in earlier-played rather than later-played games, and one would also expect such statements and questions to appear earlier within a given game rather than later. This latter distribution of these statements could reflect a process of negotiation taking place during the early part of a given game, and the comparative dearth of such statements later in the game would suggest that the interpretation of the procedures had been tacitly agreed upon earlier and were not, therefore, called into question later in the game.

In order to shed some light on the negotiation hypothesis, twenty-four first- and second-played games were examined for the presence of such statements and questions. It was observed that eleven games contained at least one such statement. Thus, roughly half of the games considered contained such statements: that is, the sort of (fairly explicit) negotiation exemplified in the extracts of dialogue considered in this chapter is in evidence in roughly half of the games. Furthermore, the hypothesis that such statements should appear earlier within a given game rather than later was tested by comparing the number of such statements occurring during the first 50% of the descriptions within a given game as compared to the number occurring during the second 50% of the descriptions within the game. Ten of the eleven games containing such statements conformed to the expected pattern (i.e. having more 'procedure clarification statements' earlier in the game rather than later in the game). The eleventh case had as many procedure clarification statements during the first 50% of descriptions as during the second 50% of descriptions. Thus, there were no cases which actually went <u>against</u> the hypothesis. A sign test provided confirmation that this was a significant result ($N = 10$, $S = 0$, $p < .01$ (1 tail)). Thus, although this is a crude test of the negotiation hypothesis, the result does suggest that the procedures implicit in the descriptions are clarified by subjects earlier rather than later within a given game, a pattern consistent with a hypothesis of early negotiation and subsequent tacit agreement as to the extensional meanings of the procedures implicit in a description.

The hypothesis that each pair of subjects entrain one another linguistically, and that the uniformity of their description types (and the agreed interpretation of those descriptions) is achieved by means of negotiation, has several important consequences. The general consequences of the possibility that meaning is, to some extent, negotiated in natural dialogue is explored more fully in the next chapter. In the next section of the present chapter, however, the consequences of the hypothesis for the difference (see, for example, the review of Lomov's work in chapter 4 of the present thesis) between monologue and dialogue, and the possibility that the apparently increased difficulty of the reassigment games of study IV of the maze games studies could have been in part due to difficulties of negotiation, will be considered.

c) The reassignment and monologue tasks re-viewed

The view that the extensional meanings of the location descriptions is negotiated by each pair of subjects is an interesting perspective with which to approach the reassignment games of the maze game studies and the monologue condition of the maze-game-independent location description experiment. One would expect that in the case of the reassignment games (in which the pairs of subjects playing the games had been made up on the basis of each individual within a pair having used a different type of description in a previous game than did his current partner), the negotiation of a common mental model of the maze would be more difficult than would be the case if the same two subjects were starting 'from scratch', as it were. Likewise, one would expect the impossibility of negotiation in a monologue to affect the type of description generated. These propositions will be examined in turn.

The reassignment games were found to present significantly more difficulty for subjects than did the baseline games (see chapter 5), this despite the fact that the subjects playing the reassignment games had had more experience of playing the game than had subjects in the baseline study. This result suggests that the location description aspect of the task is an important one.

Of the six reassignment games, in five games the subjects predominantly used matrix-type descriptions whilst in the sixth game, row-type descriptions predominated. On inspecting the transcripts of dialogues of reassignment games, one or two unusual features are apparent. For example, consider the following extract from a reassignment game involving two female subjects; player A had previously used matrix-type descriptions, whereas player B had previously used row-type descriptions (see also figure 10-9 for an illustration of both locations being described):

1) A Where are you? (laughs)

2) B Well- uhm: I am on the fourth:block from the: (2 sec) what, the right or left, I can't see: hold on: 1, 2, 3, yeah the fourth block.

3) A Uh-huh

4) B Ehm: (1 sec) where are you?

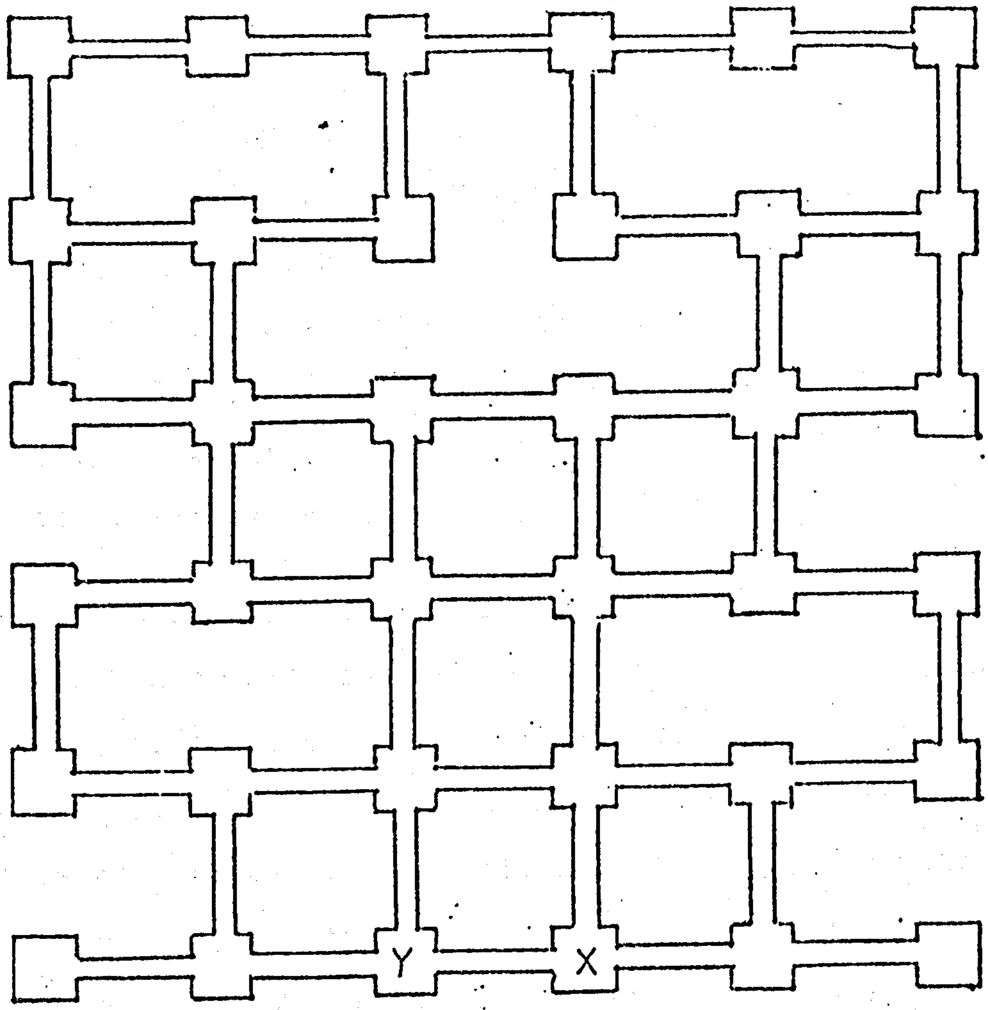


Figure 10-9: The point (X) described as "the fourth block from the left on the bottom row" and the point (Y) described as "(3,1)".

5) A The fourth block?

6) B You know: fourth from: (A interrupts)

7) A O-on the bottom row?

8) B Yeah, on-on the bottom row.

9) A Yeah, well I'm in (three:one): (1 sec)

10) B Uh-huh:O.K.

11) A That's the third one:

In this extract, what is rather odd about the exchange as a whole is the fact that it violates the usual tendency to adopt a similar description style to that of one's partner. Player A's matrix description in utterance 9) has been preceded by no discussion of a coordinate system at all, and it follows a row-type description (part of which player A herself supplied). Player A's clarifying comment in utterance 11) ('that's the third one') is of little help in clarifying the matrix system she is using, since she does not specify on which row her position marker is located at 'the third one'.

In the subsequent discourse, it becomes clear that player B misunderstood player A's matrix-type description to mean that player A's position marker was located on the second row from the bottom when in fact it was located on the bottom row. The two players then decide to refer to the bottom row as 'zero' rather than 'one' in the coordinate system and use the coordinate system thereafter. Notice that player B assumed she correctly understood player A's description and did not question it. This is rather odd, given that the coordinate description was generated 'out of the blue', so to speak.

In another reassignment game, in which two male subjects were playing on an asymmetrical maze, the two subjects spoke a total of no less than fifty-eight utterances in describing their initial positions and goal positions alone. In this instance, they set up a coordinate system for generating descriptions, but described their locations in terms of a mixture of coordinate and figural descriptions, with one or two path-type descriptions in addition. The following extract (see also figure 10-10) illustrates the mixtures of descriptions produced. In this extract, the two players are discussing their goal positions. Player A had used path-type descriptions in a previous game, whereas player B had previously used matrix-type descriptions.

1) A . . Now: we'll find out: (1 sec). My: goal
   box: would be: (1 s) B2
2) B B2
3) A B2, it's right under nothing (laughs)
4) B Aye: (laughs) Right, so:so another:
   another way of describing it would be to:
   from the top:
5) A {Uh-huh
6) B {Left hand corner, go down one and one: one in.
7) A That's it exactly, {right:}

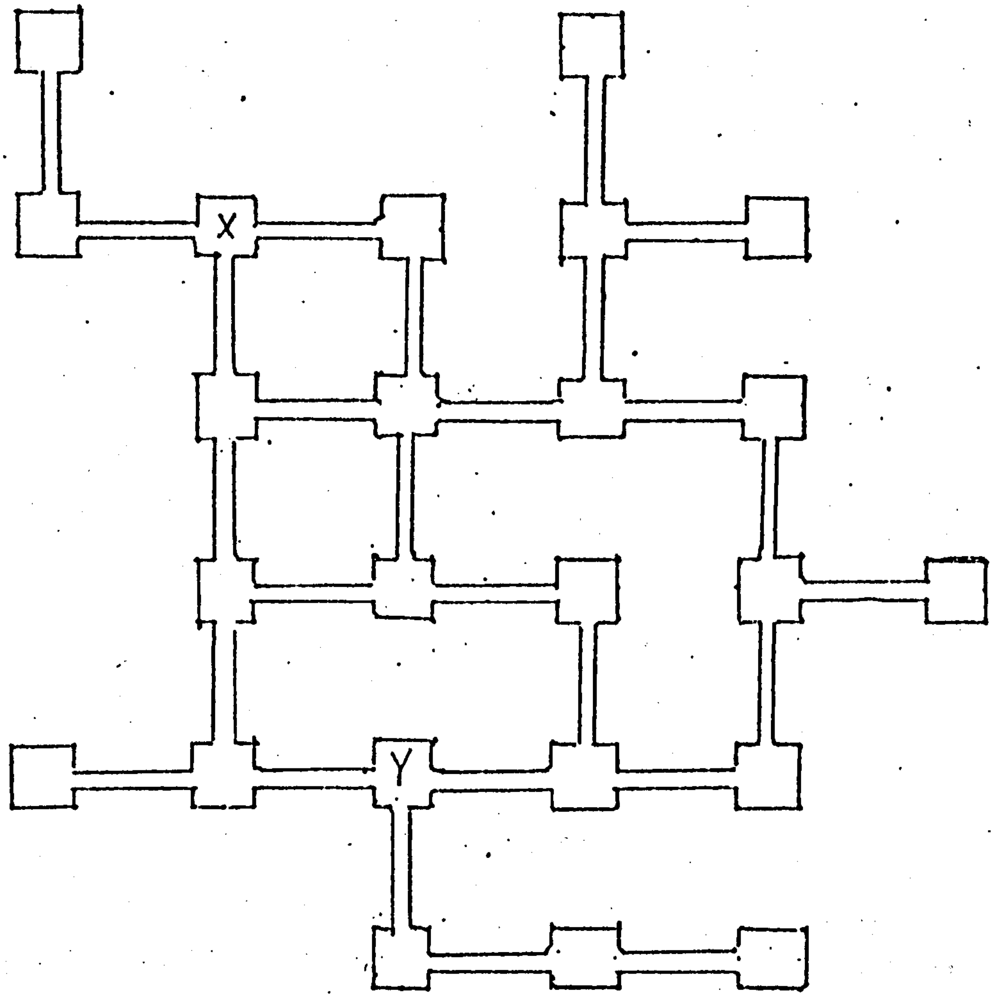


Figure 10-10: the point (X) described as "B2" and "from the top left hand corner, go down one and one in" and the point (Y) described as "C and two up from the bottom", "or three in and two up if you like".

8) B {Right:} now: my goalbox
is:er: right: taking the A sort of stuff
it's: (1 sec) C: (2 sec) C and
two up from the bottom:

9) A C in:

10) B Two up from the bottom:

11) A Two: one-: (1 sec) So: or: or (B interrupts)

12) B Or three in and two up if you like (1 sec)

13) A Three in: two up, so i- from the right
hand co- the right hand corner,
it's three in: two up? (1 sec)

14) B The bottom left hand co-: well, bottom-
bottom left or bottom right, its
in the middle.

15) A Aye oh ah-O.K. right I: I know where you are,
exactly in the middle?

16) B Uh-huh:

17) A Right and - and you've nothing above you?
(1 sec) Well y- (B interrupts)

18) B Er: no {there's no path,} I can only go
sort of

19) A {There's no path }

20) B left or right from there.

The mixture of description types (cf. utterances 11) and 12)) and the protracted checking descriptions (e.g. utterances 13) - 20)) are interesting. This pair of subjects did settle down to the subsequent predominant usage of matrix-type descriptions, but as can be judged from the above extract, the first four descriptions involved protracted negotiation.

It is difficult to draw any general conclusions from the small data base of six dialogues as to what might have caused the observed significant slowing of the rate of moving in the reassignment games as compared to the baseline games. Certainly, the two extracts considered, the first featuring insufficient negotiation and consequent misunderstanding in the subsequent discourse, and the second featuring somewhat protracted negotiation, do seem to point to unusual features in the negotiation of meaning in these games. Definitive conclusions would probably require a larger data base for adequate testing. The 'negotiation of meaning' hypothesis and its potential consequences for a 'reassignment'-type experimental manipulation would be a very interesting matter for further detailed study.

Turning to the monologue condition of the maze-game-independent location description experiment from the perspective of the 'negotiation of meaning' hypothesis, it is interesting to note the general characteristics of monologue-type descriptions as opposed to dialogue-types (in which a listener is present who can ask appropriate questions if there has been a failure to understand a description). Subjects taking part in the monologue study were instructed to generate descriptions which would be sufficiently precise for a listener to be able to uniquely identify the appropriate node purely on the basis of their descriptions.

Of the eight subjects who generated descriptions in the monologue condition (i.e. with no listener present) six of the subjects generated path-type descriptions, although these varied in their degree of ellipsis. For example:

Er:one: (3 sec) one up: one across (3 sec)

This type of description could be expected to fail to identify a particular individual location within a large maze, since the starting point for the path movement is not specified, nor is the counting procedure clarified. One would reasonably expect that, if the above description were the first in a maze game, the listener would ask questions of the speaker to clarify the situation for himself and more detailed descriptions would subsequently be profferred. In the monologue study, the speaker who generated the above description (it was the very first description he gave) did not have the benefit of such feedback and went on to give a further thirty-nine such descriptions.

Some of the other path type descriptions generated in the monologue study would have a better chance of success. For example: ". . er: (1 sec) top: top right hand corner: (1 sec) three down: and one along to the right".

This description, being more explicit, would stand a better chance of comprehension (provided, of course, that the listener were entertaining a mental model of the maze of the path type).

Another type of description which was found in the monologue data was the row type of description. This had the classic format found in row-type descriptions generated in the dialogue studies. For example:

> Er: the cross is on the: (1 sec) second row from the bottom, second in from the left.

Again, this would stand a reasonable chance of success provided that the listener was not in doubt as to the extensions of the count procedures.

Two rather unusual description strategies were encountered, both of which could reasonably be expected to fail utterly in the task of uniquely identifying a particular node. One was of a figural type and took the format exemplified by the following description, the referent of which is depicted in figure 10-11:

> Er: point number one's: surrounded by three paths:er (1sec) one coming from the west, one from the north, one from the east: there's nonet- coming from the south: (2 sec) er:

This description would be expected to fail since it

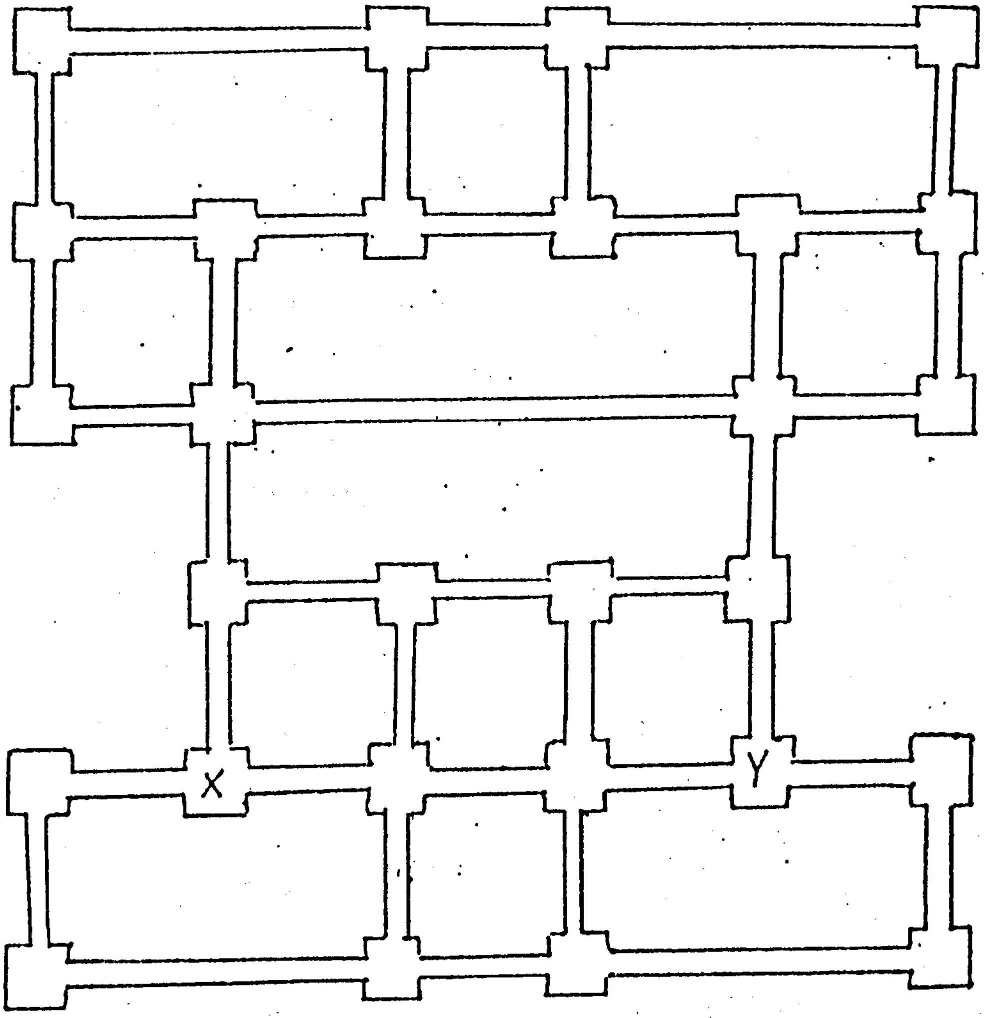


Figure 10-11: The point described as "surrounded by three paths: er: one coming from the west, one from the north, one from the east: theres none: coming from the south". This description does not uniquely identify the referent; point Y also fits the description.

does not uniquely describe its intended referent: other nodes in the maze structure also fit this description. One would expect a listener to question this description immediately if he were present.

The other unusual description strategy which, it is hypothesized here, would fail to uniquely identify a particular node, was generated by a subject who was having difficulty in describing locations on the asymmetrical maze. Her solution to this problem took the following format (see fig. 10-12):

> On the second: image: the X: is: (1.5 sec) ehm: (2 sec) twenty-:five: (1 sec) percent in from the right hand side: (1 sec) and: (2.5 sec) two-thirds up from the bottom: (3 sec).

The distances being specified in this case are distances from specified edges _of the page on which the maze diagram was printed_. The ambiguity of what twenty-five per cent of the width of the page actually refers to would again be expected to evoke questioning from a listener, and probably also an alternative description framework.

The examples considered illustrate the general point about monologue-type location descriptions as compared to those generated under conditions where the speaker is a participant in a dialogue; the speaker's personal, idiosyncratic way of conceptualizing and describing the

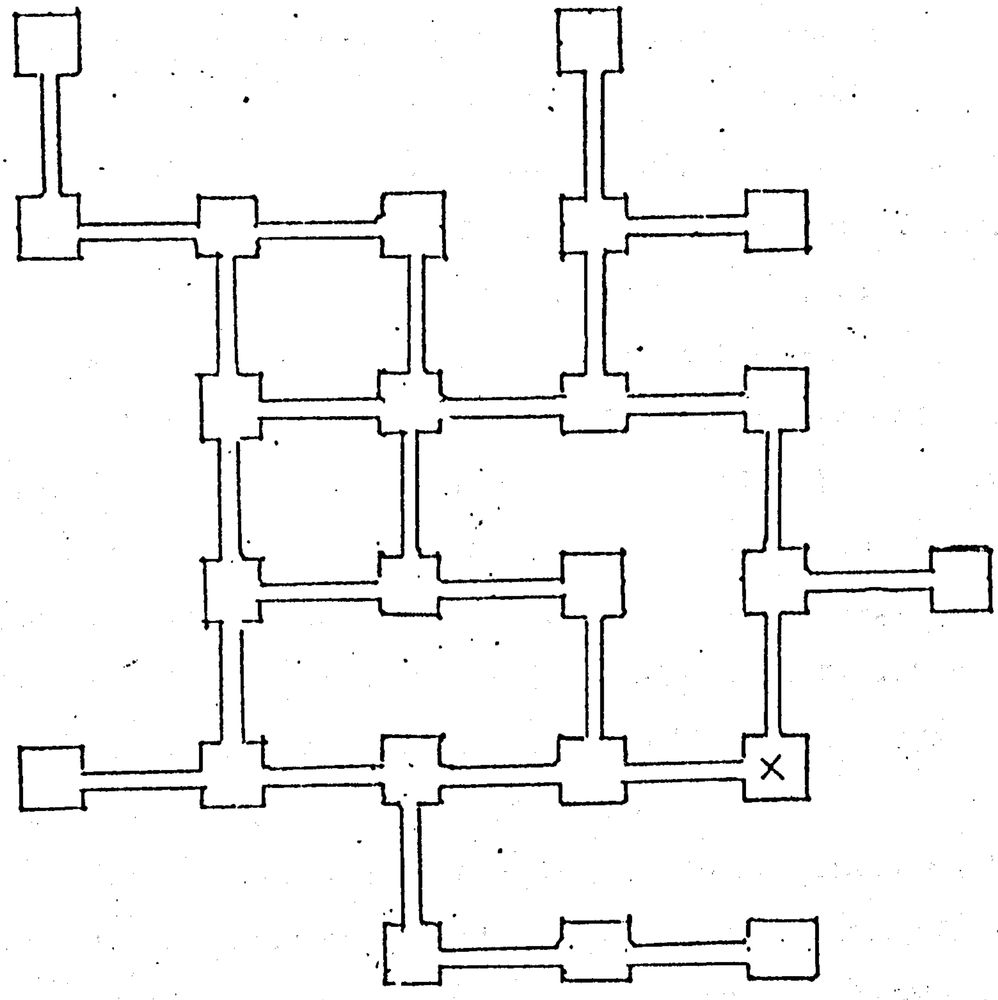


Figure 10-12: The point (X) described as "twenty-five percent in from the right hand side and two-thirds up from the bottom" (of the page).

maze is allowed to prevail unchecked and unaltered by negotiation. From this perspective, the feedback of the listener's responses to profferred descriptions is vital to the subsequent modification of the speaker's descriptions. The speaker in a monologue does not have the benefit of feedback from his listener; indeed, given the pragmatic interpretation of the present data, it could be argued that a monologue location description task is a slightly unfair task for a subject in that it is one which he would be expected to fail (in the sense of failing to uniquely specify the referent so that any listener could find the described location purely on the basis of the speaker's description of it).

The present approach to the data is in agreement with Segalowitz's comments on Lomov's work (see chapter 4 of this volume) that among the important differences between dialogue and monologue is that the participants in a dialogue have the benefit of feedback; and feedback, according to the present interpretation, is the basis for the negotiation of the meaning of the descriptions.

d) Conclusions

The examples considered in the present chapter do lend support to the view that the meanings of the location descriptions generated by each pair of subjects are negotiated products. This perspective throws interesting light on the difference between monologue and dialogue (which is seen as partly involving a difference in the ability to negotiate meanings and description formats) and also on the reassignment task (the special difficulties of which could in part be due to difficulties of negotiating a common mental model of the maze, given that both subjects start with a preference for <u>different</u> types of model).

What has been implicit in the foregoing discussion has been the proposition that it is necessary for both participants in a dialogue to entertain highly similar mental models of the maze (similar enough to return identical extensions of a given description to both interlocutors) in order for the semantics of the location descriptions to be effective. The 'mental model', as conceived of here, is thus a form of social representation in that it is entertained by both participants in the conversation: it is a negotiated (and therefore social) product.

The very notion that meaning is to some extent negotiated is a very interesting one and is one which could potentially throw some light upon certain theoretical dilemmas. Notably, Rommetveit (1974, see chapter 3 of the present volume) writes of 'contextually appropriate optional elaborations' of the meaning of a word or phrase (the example he chooses is the sentence 'The old man is poor'). Rommetveit does not, however, suggest how the options of elaboration are selected among by the participants in a dialogue. Likewise, in his notion of drafts of contracts concerning shared strategies of categorization and attribution, negotiation could be the mechanism by means of which the drafts of such contracts are, to continue the metaphor, 'drawn up'.

Once negotiated, however, the meanings of utterances have very precise interpretations: speakers and listeners do not appear to ponder over potential ambiguities in what is said. Thus, language is seen to be <u>both</u> inherently flexible <u>and</u> inherently precise in its functioning; negotiation would allow interlocutors to select among the many potential meanings of a given utterance, and, following this, the selected, agreed-upon meaning functions in a precise way: there is no ambiguity.

Thus, the hypothesis that meanings are, to some extent, negotiated products not only goes some way towards explaining many of the features of the present data but also provides insights into some intriguing theoretical issues, some of which will be explored in the following, final chapter.

CHAPTER 11

CONCLUSIONS

The research reported in the present thesis constitutes an attempt to examine certain aspects of the meanings of natural language expressions. Specifically, the work reported here focusses upon the semantics and pragmatics of descriptions of spatial location in spontaneous task-oriented dialogue. Certain tentative conclusions can be drawn on the basis of the work reported in this thesis, and these conclusions are discussed below. Perhaps more interesting, however, are the questions raised by the data and results reported here, and these questions are discussed later in the present chapter.

Considering, firstly, the semantics of the descriptions of maze location generated by subjects in the course of playing the maze game task, the data and the formal analyses from the present set of studies (see chapters 6 and 7) suggest that there exists a comparatively small number of strategies for generating descriptions of locations within such a spatial network. Four basic ways of generating descriptions of spatial location were observed in the maze game experiments, and two other, rather less frequently occurring, types of location description were also noted. All of these different ways of describing locations within a spatial network were interpreted to be implicitly procedural in nature, taking the form of tacit instructions to the listener as to how and where to look to find the desired location.

The four major classes of location description each comprise several subvariations, but each generic class can be characterized in a manner which accounts for all subvariations. The four generic description types which were observed in the data are:

(a) Line-type descriptions, which involve the speaker and listener treating the maze shape as if it were an array of parallel lines of nodes and generating descriptions by first identifying the particular line of nodes within which the location being described is located, and then identifying the position on that line which corresponds to the relevant location;

(b) Path-type descriptions, which involve the speaker generating descriptions by firstly identifying a perceptually prominent point in the maze structure, such as a corner, and describing the desired location in terms of path movements from the prominent point to the location in question;

(c) Figural-type descriptions, which involve the speaker identifying a region or area of the maze by means of referring metaphorically to a shape formed by part of the network of paths in the maze structure, such as an L- or T-shape, and then identifying the desired location relative to that shape; and

(d) Matrix-type descriptions, which involve the speaker and listener in treating the maze as a regular two-dimensional array of nodes and developing a coding system for denoting positions on each dimension, thus generating descriptions taking the form of two-place coordinates. In fact, two distinct subvariations on the matrix theme can be discerned on the basis of how the speaker and listener execute the relevant scanning procedures. One is rather like the Cartesian coordinates of mathematics in that the procedures corresponding to the two coordinates are executed <u>contiguously</u>. For example, a horizontal scanning movement through a specified distance is made, followed immediately by a vertical scanning movement through a specified distance, after completion of which the scanner is located at the relevant node. This is rather like the 'move along X then move up

Y' (where X and Y are specified distances) strategy employed in path-type descriptions, except, of course, that gaps in the pathways are mentally 'traversed' without acknowledgement in the case of the matrix-type descriptions. The other, 'non-Cartesian', type of matrix description also involves the use of two-place coordinates as a standard description format, but the corresponding procedures to be executed by the listener are rather more complicated. On hearing a description like '(3, 4)', for example, the listener's task is to scan horizontally through a distance of three nodes (for example), thereby establishing the relevant column of nodes in which the location being described is situated, then go back to the (previously agreed-upon) origin, scan vertically through a distance of four nodes, thereby establishing the relevant row of nodes in which the location being described is situated, and, having established the relevant row and the relevant column of nodes, he then has to find their intersection. The location being described is, of course, situated at that intersection. The scanning procedures involved in this latter case are more complicated than is the case with the 'Cartesian' type of matrix description: whereas the 'Cartesian' type of description involves two scanning procedures operated contiguously, the 'non-Cartesian' type of matrix description involves seeking the relevant row and column and obtaining their intersection.

The two less frequently occurring classes of location description which were observed in the data could in fact be argued to be akin to the matrix-type of description in their procedural aspects. These two types of description are:

(e) The description of locations within the maze in terms of specified distances (either in terms of node positions or path movements) from two of the edges of the maze structure. For example: "I'm two from the right, two from the top". This is rather akin to the 'non-Cartesian' matrix type of location description, discussed above, in that two lines of nodes (a column and a row, respectively), are specified, and the listener's task is to locate these lines of nodes and to seek their intersection.

(f) The description of locations within the maze in terms of a specified row and a specified column; for example "Third row, fifth column". Once again, an individual row and an individual column are specified relative to a previously established origin, and the listener has to locate these two lines of nodes and seek their intersection.

Thus, four major classes and two minor classes of description can be distinguished, all of which are (it is argued) implicitly procedural in nature. The procedural interpretation of the meaning of the location descriptions was supported by informal observation of subjects in the 'paper-and-pencil' experiment reported in chapter 9, in which listeners were observed to execute the procedures hypothesized above with the aid of ostensive gestures using their forefinger or a pen during the course of interpreting the descriptions.

It was argued in chapter 8 of the present thesis that each of the four broad classes of location description could in fact correspond to a class of underlying conceptualizations, or <u>mental models</u> of the maze shape (cf. Johnson-Laird, 1980, 1981a, 1981b). According to this interpretation of the data, the mental modelling of the maze shape would involve each subject in extracting some feature of the maze structure, treating it as prototypical, and setting up a representation which represents the maze as an assemblage of a number of those features. Thus, for example, a given subject might pay particular attention to the horizontally-oriented lines of pathways in the maze structure, treat this feature as prototypical, and represent the maze as a set of parallel horizontal lines (rows). This mental model would thus represent the maze shape in a simplified fashion and serve as an intermediate representation in terms of which descriptions could be produced and understood.

This model would be 'intermediate' between language and the 'world' under discussion in a very real sense. According to the 'mental models' interpretation of the location descriptions, the procedures implicit in a description would be executed initially on the internal mental model of the maze shape, and the appropriate location would be 'marked' (in some sense) on the model. This 'marked' model, in turn, would be the basis for the subject's visual scanning of the maze display on the computer terminal screen.

This would obviate the subject having to consult the screen every time he heard a location description; he would have the option of 'marking' the relevant location on his model and taking action later (for example, consulting the actual maze display on the screen in order to ascertain whether he possesses any switch nodes near to the recently-described location).

There is another sense in which such a mental model is intermediate between language and the world: the mental model constitutes a kind of 'framework' which is intermediate between the complex perceptual stimulus of the maze display on the computer terminal screen and the sorts of semantic distinction encoded in the lexicon, and location descriptions can be more readily formulated with respect to the mental model than is the case with the actual stimulus. Thus, for example, the asymmetrical maze presents a complex shape to the subject, and the description of any given location within it is rather a difficult task for a speaker. If the speaker builds a row-type mental model of the maze, however, his model consists in a series of parallel horizontal lines which are simplified vis-a-vis the original stimulus but which are also more language-encodable: terms like 'row', 'line', 'horizontal', and so on, are more consonant with the model than with the original stimulus.

Johnson-Laird (1980, 1981a, 1981b) emphasizes the benefits conferred upon the language understander by the building of mental models; however, as can be judged from the above, mental models also confer considerable benefits on the speech producer: models represent perceptually and conceptually complex entities in a manner which more readily permits the formulation of descriptions.

Thus, by representing the original stimulus in a somewhat idealized way, the model simplifies both language production and language understanding.

As noted already in chapter 8, Johnson-Laird emphasizes the processes whereby a reader or listener constructs a mental model of an object (or person, or situation) which is <u>not</u> perceptually present on the basis of discourse about that object. This 'listener's discourse model' is a considerable aid to comprehension and inference (see chapter 2). The above formulation, in which a mental model of a perceptually-present object (in this case, the maze display on the computer terminal screen) is built as an aid to language production and understanding, is not at variance with Johnson-Laird's account. Indeed, in his 1980 paper, Johnson-Laird hypothesizes that a mental model of a perceptually present object could be constructed in precisely the manner indicated above, and for the reasons outlined above. Where the present account does differ from Johnson-Laird's formulation is in the detail of its application of the notion of a 'mental model' to the dynamic social situation that is the dialogue.

It was argued in chapter 8 that, in order for location descriptions to be satisfactorily produced and interpreted, it is necessary that both of the participants in a dialogue should model the maze in a similar way. This is necessary because, as was argued in chapter 8, different models of the maze return different semantic interpretations of the same referring expression, and it is necessary in the context of the present game studies that both subjects should interpret a given referring expression in an identical fashion. Thus, it is hypothesized that there should be pressure on both interlocutors to develop similar models of the maze, and this should manifest itself in the form of subjects constraining one another's choice of description. Having chosen a particular type of description to be used predominantly in the subsequent discourse, there should then be some negotiation of the extensional semantics of the descriptions. A corollary of this 'common mental model' hypothesis, therefore, is that a certain amount of tacit, or even explicit, negotiation of the meanings of the location descriptions should take place in order that the two subjects can modify their (probably) initially different representations of the maze in such a way that they come to share a common mental model.

In chapter 10, some extracts of dialogue which appeared to lend support to the hypothesis that subjects negotiate a common mental model (and consequently a common set of interpretations of referring expres1sions) were discussed. The hypothesis that the meanings of referring expressions are, to some extent, negotiated products was not tested formally in the present set of experiments and remains an intriguing speculation for future research to examine empirically.

Thus, the interpretation that was placed upon the maze location description data is as follows. It was argued that each of the four major types of location description reflect different types of underlying mental model of the maze shape. It was also argued that, in order for referring expressions to function satisfactorily (i.e. in order that the same interpretation should be placed upon a given referring expression by both speaker and listener), both interlocutors should entertain highly similar mental models of the maze. Thus, the mental model, in the present interpretation of the concept (i.e. as applied to the dynamic, ongoing social situation that is the dialogue), is a kind of social representation which is shared by both interlocutors. Furthermore, it was suggested (and several examples from the transcripts of dialogue are readily interpretable to be supportive of this suggestion) that the mechanism by means of which a particular type of mental model is selected for use and the precise details of its extensional semantics are agreed is the mechanism of negotiation.

The semantics of the maze location description data were therefore explained in the present thesis in terms of mental models of the maze shape, and the pragmatics of the experimental situation were discussed in terms of the mechanisms by means of which particular speakers and listeners jointly select a particular mental model with respect to which referring expressions can be produced and interpreted. The social-pragmatic emphasis underlying this interpretation of the data is clear; the question of greatest interest is that of how particular speakers and listeners negotiate such necessarily highly similar representations of the same object, and it also becomes necessary to discuss what particular individual speakers and listeners understand by a given referring expression, as opposed to what its 'literal' or 'public' meaning, so to speak, might be.

The possibility that the meanings of expressions are, sometimes at least, negotiated has several interesting implications. Among these are the possible insight that this notion offers into the potentially problematic theoretical dilemma between 'literal meaning' and 'meaning potential'. On the one hand, some theorists, notably those following in the tradition of Chomsky's work, advocate the analysis of meaning in terms of semantic components, selection restrictions and 'literal meanings' (see chapter 2 for a brief review of this work). These theorists therefore emphasize the precise functioning of language, and, as was noted in chapter 1, their analyses _do_ seem to capture some aspects of meaning. On the other hand, some theorists, who emphasize the importance of pragmatics and the inherent flexibility of language, particularly Rommetveit (1974), have pointed out that a given expression or word has a large variety of potential meanings, some aspects of which are _not_ describable in terms of 'propositional content' or 'literal meaning', and that a proper assessment of the meaning of a given expression should include details of the particular speaker and listener involved, the context, and the particular use to which the expression was put, in the act of communication.

The possibility that each pair of interlocutors engaged in dialogue negotiate a particular set of interpretations for the expressions that they use could resolve this dilemma. Having negotiated a tacit agreement to interpret the expressions that they use in a particular way, the language would thereafter function in the precise manner which the advocates of 'literal meaning' would hypothesize. The possibility of meanings being negotiated thus allows a theoretical compromise to be made between the two positions adumbrated above, according to which both of the theoretical positions are correct. Words and larger linguistic expressions do have an inherently great flexibility of meaning but the negotiation of the meaning of a particular expression by two interlocutors would allow the selection of one of the potential meanings of the expression to be made, and the expression could thereafter be used unambiguously in their discourse.

Thus, if we examine the different meanings understood by, for example, a particular row-type description across different pairs of subjects who use row-type descriptions, the expression would seem highly ambiguous. Yet each pair of subjects could be using the expression in a highly precise way, having previously negotiated the extensional semantics of the component parts of the complex expression. Hence the pragmatic emphasis of the present interpretation of the location description data, and the advocacy of studying what particular dyads of subjects take a given expression to mean.

Turning from the semantics and pragmatics of the experimental maze game task to the importance of the social situation which any given experimental task constitutes, the experiment reported in chapter 9 indicated that the nature of the experimental task, both in cognitive terms and in terms of the social situation which the experimental task constitutes, is vitally important to the nature of the data produced by subjects engaged in that task. Thus, in an experiment in which subjects had to describe locations on mazes similar to those which had been used in the maze game studies but independent of playing the computer-controlled maze game, the resulting distribution of types of description was different as compared to that obtained in the maze game studies, with one of the most popular types of location description in the maze game studies - the matrix-type- being entirely absent in the game-independent location description task. In removing the maze location description task from its setting within the context of an ongoing experimental game and setting it to subjects as an experimental task in its own right, a number of complex and subtle changes had been made which could have produced the observed change in the distribution of the different types of description in the data. In addition to the, perhaps more obvious, changes of a purely cognitive nature (see chapter 9) there may also have been a number of more subtle changes of a social psychological nature when setting the location description task to subjects as a 'paper-and-pencil' type of experiment.

For example, the subjects' representation of the task could have changed in the 'paper-and-pencil' task vis-a-vis the maze game task, resulting in a difference in the perceived demand characteristics of the two tasks. Furthermore, there are good reasons for hypothesizing that subjects in the 'paper-and-pencil' experiment would be more consciously aware of the language they produce than would be the case with subjects in the maze game. Setting the location description task within the context of an ongoing computer game would result in subjects becoming more task-involved and the location descriptions would be produced in a genuinely spontaneous, natural way. Indeed, this was listed as one of the advantages of studying task-oriented dialogues in chapter 4. On the other hand, subjects in the 'paper-and-pencil' task have no such task-involvement to draw their attention away from the location-description aspect of their experimental task.

It should be emphasized that these analyses in terms of variation in the two experimental tasks as social situations (in terms of differences in the degree of conscious awareness on the part of subjects of the language produced, differences in the perceived demand characteristics of the two experimental tasks, and so on) are <u>post hoc</u> interpretations: such social variables were not measured in either of the two tasks. However, they are discussed here because they could indeed have borne an influence in the experimental situations, and because the discussion of them highlights the sheer complexity both of dialogue itself and of the study of dialogue.

Thus, the importance of the experimental task used to obtain linguistic data as a social situation comes into focus and would appear to merit further investigation.

The present set of experimental studies have constituted an attempt to do on a limited scale what Marslen-Wilson et al. (1982) strongly advocate, namely, to study language <u>both</u> as a cognitive process <u>and</u> in terms of its natural communicative function in normal interaction. Both of these facets of language are fascinating in their own right, but as Marslen-Wilson et al. argue, they are interdependent facets. The present set of studies have involved an examination of both of these facets in relation to but one limited aspect of the data produced by the dialogue maze game, yet the complexity of the issues raised, both from a cognitive and a social-interactional point of view, is considerable. Indeed, the present set of studies have no more than scratched the surface of the rich data produced by the experimental game task, but in so doing a number of important questions are raised which would merit the attention of future research.

For example, the maze location descriptions were interpreted to reflect mental models of the maze. There are a large number of testable hypotheses which follow from this interpretation. For example, do subjects' eye movements during the course of intepreting descriptions match what a mental model-based set of hypotheses would predict? A second aspect of the 'mental models' hypothesis which is well worth empirical study is the question of whether the subjects memory of the maze shape (either recall or recognition) would exhibit deviations from a 'literal' recall or recognition in directions predictable from the 'mental model' theory. An example of this would be a case in which a subject who (we hypothesize) built a 'row' type model misremembers complete rows and provides only vague recollections of the vertical connections between the rows in his recall of the maze shape after having played the game. This would appear to be an interesting question for future research to tackle.

Similarly, a whole host of questions pertaining to the domain of pragmatics spring to mind in connection with the maze game. The first question is Lewis's notion of control, a notion of great interest to social psychologists. It would be of great interest to compare the participants of a given dialogue in terms of how many tactical plans each member of the dyad initiates, how many location descriptions each generates, how many questions each asks, how many interruptions each initiates, and so on. This would be exceedingly interesting in relation to the hypothesized negotiation of meaning, and one could also devise a number of intriguing social psychological experiments (for example, taking two dyads of subjects, establishing which subject in each dyad in 'master' and which is 'slave', and then pairing the two former 'slaves' together). Similarly, a much more extensive set of 'reassignment' studies would prove of great interest in relation to the notions of control and negotiation.

The hypothesized negotiation itself is, of course, a highly worthwhile topic for further study. A set of rather more 'longitudinal' studies, with each pair of subjects playing perhaps five games in succession on difficult-to-describe asymmetrical mazes of different types would perhaps produce an appropriate data base with which to appropriately test the negotiation hypothesis. The 'negotiation of meaning' hypothesis is perhaps the most contentious hypothesis postulated in the present thesis, and would certainly prove an interesting issue to tackle. A further question which naturally arises with regard to this hypothesis is whether such negotiation (if conclusively demonstrated in the data from the maze game task) does in fact take place in other types of natural conversation concerning different domains of discourse as compared to the present maze game task (for example, whether it would be in evidence in the data from other different types of experimental speech-production tasks).

The present research also has implications for the computer simulation of natural dialogue. The use of model-based theories in such work is, of course, not new. However, the <u>building</u> of a mental model during the course of a dialogue would be an intriguing (if exceedingly complex) project for simulation. Such a project could provide theoretical insight into some of the phenomena of pragmatics (particularly that of audience design) from the artificial intelligence perspective. Similarly, it would seem possible in principle to simulate some of the hypothesized mechanisms of negotiation postulated in chapter 10. One would then be in a position to account for 'audience design' in terms of different models being constructed by the processor when in 'conversation' with different others.

Finally, the work reported in this thesis underlines the importance of studying language which is in use for a purpose by those who generate and interpret it, a point made most forcefully by theorists such as Rommetveit and Wittgenstein. It is argued that experimental tasks which involve the subjects in producing natural dialogue in order to achieve some purpose is an excellent way of framing and investigating a large variety of hypotheses, many of which simply could not be studied otherwise. This is not to say that other experimental tasks are inferior by comparison; the point to be emphasized is that the data produced by such tasks as the experimental maze game certainly merit detailed study even if such tasks are (perhaps) less well controlled empirically than is the case with other experimental tasks which investigate language production and understanding.

As is the case with all research into the phenomena of language use, the present set of studies raise many questions. These questions, in particular that concerning mental models as social representations, and the possibility that meaning is, to some extent, a negotiated product, would appear to be two particularly interesting questions which provide theoretically intriguing and potentially fruitful hypotheses for future studies.

REFERENCES

Abric, J. C. (1984). A theoretical and experimental approach to the study of social representations in a situation of interaction. In R. M. Farr and S. Moscovici (Eds.), Social representations, Cambridge: Cambridge University Press (in press).

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., and Trollip, S. R. (1976). Instantiation of general terms. Journal of Verbal Learning and Verbal Behavior, 15, 667-79.

Austin, J. L. (1962). How to do things with words. Oxford: Oxford University Press.

Bar-Hillel, Y. (1967). Dictionaries and meaning rules. Foundations of language, 3, 409-14.

Bartlett, F. C. (1932). Remembering. Cambridge: Cambridge University Press.

Blakar, R. M. (1973). An experimental method for inquiring into communication. European Journal of Social Psychology, 3, 415-425.

Blakar, R. M. and Pedersen, T. B. (1980). Control and self-confidence as reflected in sex-bound patterns in communication: an experimental approach. Acta Sociologica, 23, 33-53.

Blakar, R. M. and Valdimarsdottir (1981). Schizophrenia and communication efficiency: a series of studies taking ecological variation into consideration. Psychiatry and Social Science, 1, 43-52.

Bransford, J. Barclay, J., and Franks, J. (1972). Sentence memory: a constructive versus interpretatave approach. _Cognitive Psychology_, 3, 193-209.

Broadbent, D. E. (1973). _In defence of empirical psychology_. London: Methuen.

Buhler, K. (1982). The deictic field of language and deictic words. In R. J. Jarvella and W. Klein (Eds.), _Speech, place and action: studies in deixis and related topics_. Chichester: Wiley, 9-30.

Carnap, R. (1956). _Meaning and necessity: a study in semantics and modal logic_. Chicago: University of Chicago Press.

Cherry, C. (1957). _On human communication: a review, a survey and a criticism_. New York: Science Editions.

Chomsky, N. (1957). _Syntactic structures_. The Hague: Mouton.

Chomsky, N. (1968). _Language and mind_. New York: Harcourt Brace Jovanovich.

Clark, H. H. and Clark, E. V. (1977). _Psychology and language: an introduction to psycholinguistics_. New York: Harcourt Brace Jovanovich.

Clark, H. H., and Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, I. Sag, and B. Webber, (Eds.), _Elements of discourse understanding_. Cambridge: Cambridge University Press, 10-63.

Clark, H. H., and Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. Le Ny and W. Kintsch (Eds.,) _Language and comprehension_. Amsterdam: North-Holland, 287-299.

Codol, J. - P. (1984). On the system of representations in an artificial social situation. In R. M. Farr and S. Moscovici (Eds.), _Social representations_. Cambridge: Cambridge University Press (in press).

Collins, A. M., and Quillian, M. R. (1972). How to make a language user. In E. Tulving and W. Donaldson, (Eds.), _Organization and memory_. New York: Academic Press, 310-349.

Craik, K. (1943). _The nature of explanation_. Cambridge: Cambridge University Press.

Davies, D. J. M. and Isard, S. D. (1972). Utterances as programs. In D. Michie (Ed.), _Machine intelligence_, _7_, Edinburgh: Edinburgh University Press.

Deutsch, M. (1973). _The resolution of conflict: constructive and destructive processes_. London: Yale University Press.

Deutsch, M. and Krauss, R. M. (1960). The effect of threat on interpersonal bargaining. _Journal of Abnormal and Social Psychology_, _61_, 181-189.

Deutsch, M. and Krauss, R. M. (1962). Studies of interpersonal bargaining. _Journal of Conflict Resolution_, _6_, 52-76.

Dixon, W. J., Brown, M. B., Engelman, L., Frane, J. W., Hill, M. A., Jennrich, R. I., and Toporek, J. D., Eds.) (1981). _BMDP statistical software_. London: University of California Press.

Donnellan, K. (1966). Reference and definite descriptions. _Philosophical Review, 75_, 281-304.

Ehrlich, K., and Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. _Journal of Verbal Learning and Verbal Behavior, 21_, 296-306.

Farr, R. M., and Anderson, A. (1983). Beyond actor-observer differences in perspective: extensions and applications. In M. Hewstone (Ed.) _Attribution theory: social and functional extensions_. Oxford: Blackwell, 45-65.

Fodor, J. (1978). Tom Swift and his procedural grandmother. _Cognition, 6_, 229-247.

Fodor, J. D., Fodor, J. A., and Garrett, M. F. (1975). The psychological unreality of semantic representations. _Linguistic Inquiry, 4_, 515-531.

Frege, G. (1952). On sense and reference. In P. T. Geach and M. Black (Eds.), _Translations from the philosophical writings of Gottlob Frege_. Oxford: Blackwell.

Garfinkel, H. (1967). _Studies in ethnomethodology_. Engelwood Cliffs, New Jersey: Prentice-Hall.

Garnham, A. (1979). Instantiation of verbs. _Quarterly Journal of Experimental Psychology, 31_, 207-14.

Gordon, A. D. (1981). _Classification_. London: Chapman and Hall.

Greene, J. (1972). Psycholinguistics: Chomsky and psychology. Harmondsworth: Penguin.

Grice, H. P. (1975). Logic and conversation. William James Lectures, Harvard University, 1967. In P. Cole and J. L. Morgan (Eds.) Studies in syntax III. New York: Academic Press.

Grosz, B. J. (1977). The representation and use of focus in dialog understanding. Technical Note 151, Artificial Intelligence Center, Stanford Research Institute, Menio Park, California, July 1977.

Hall, E. T. (1966). The hidden dimension: man's use of space in public and private. London: The Bodley Head.

Hartigan, J. A. (1981) P1M: cluster analysis of variables. In W. S. Dixon, M. B., Brown, L. Engelman, J. W. Frane, M. A. Hill, R. I. Jennrich, and J. D. Toporek, (Eds.) BMDP Statistical software 1981. Berkeley: University of Californian Press, 448-455.

Hollan, J. D. (1975). Features and semantic memory: set-theoretic or network model? Psychological Review, 82, 154-5.

Jarvella, R. J. and Klein, W. (Eds.), (1982). Speech, place and action: studies in deixis and related topics. Chichester: Wiley.

Johnson-Laird, P. N. (1977a). Procedural semantics. Cognition, 5, 189-214.

Johnson-Laird, P. N. (1977b). Psycholinguistics without linguistics. In N. S. Sutherland, (Ed.), _Tutorial essays in psychology_, Vol. 1, Hillsdale, N. J.: Erlbaum.

Johnson-Laird, P. N. (1978). What's wrong with Grandma's guide to procedural semantics: a reply to Jerry Fodor. _Cognition, 6_, 249-261.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. _Cognitive Science, 4_, 71-115.

Johnson-Laird, P. N. (1981a). Mental models of meaning. In A. Joshi, B. Webber and I. Sag (Eds.) _Elements of discourse understanding_. Cambridge: Cambridge University Press, 107-126.

Johnson-Laird, P. N. (1981b). Comprehension as the construction of mental models. _Phil. Trans. R. Soc. Lond., B195_, 353-374.

Johnson-Laird, P. N. (1982). Formal semantics and the psychology of meaning. In S. Peters and E. Saarinen (Eds.), _Processes, beliefs, and questions_. Dordrecht: D. Reidel Publishing Company.

Johnson-Laird, P. N. and Garnham, A. (1980). Descriptions and discourse models. _Linguistic Philosophy, 3_, 371-393.

Johnson-Laird, P. N. and Quinn, J. G. (1976). To define true meaning. _Nature, 264_, 635-6.

Johnson-Laird, P. N. and Steedman, M. J. (1978). The psychology of syllogisms. _Cognitive Psychology 10_, 64-99.

Johnson-Laird, P. N. and Wason, P. C. (1977). _Thinking:_

readings in cognitive science. Cambridge: Cambridge University Press.

Jones, E. E., and Nisbett, R. E. (1972). The actor and the observer: divergent perceptions of the causes of behaviour. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, and B. Weiner (Eds.), *Attribution: perceiving the causes of behaviour*. Morristown: General Learning Press, 389-420.

Kaplan, D. (1977). Demonstratives: an essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. Paper presented for Symposium on Demonstratives at a meeting of the Pacific Division of the American Philosophical Association.

Katz, J. J. (1972). *Semantic theory*. New York: Harper and Row.

Katz, J. J. (1973). The realm of meaning. In G. A. Miller (Ed.) *Communication, language and meaning: psychological perspectives*. New York: Basic Books, 36-48.

Katz, J. J. and Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39*, 170-210.

Katz, J. J., and Nagel, R. (1974). Meaning postulates and semantic theory. *Foundations of Language, 2*, 311-340.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, New Jersey: Erlbaum.

Klein, W. (1982). Local deixis in route directions. In

R. J. Jarvella and W. Klein (Eds.), _Speech, place and action: studies in deixis and related topics_. Chichester: Wiley, 161-182.

Levelt, W. J. M. (1982). Linearization in describing spatial networks. In S. Peters and E. Saarinen (Eds.) _Processes, beliefs and questions_. Dordrecht: D. Reidel Publishing Company, 199-220.

Lewis, D. (1979). Scorekeeping in a language game. _Journal of Philosophical Logic, 8_, 339-359.

Lidz, T. (1963). _The family and human adaptation_. New York: International University Press.

Linde, C. & Labov, W. (1975). Spatial networks as a site for the study of language and thought. _Language, 51_, 924-939.

Lomov, B. F. (1979). Mental proceses and communication. In L. H. Strickland (Ed.), _Soviet and western perspectives in social psychology_. Oxford: Pergamon Press, 211-225.

Lyons, J. (1968). _Introduction to theoretical linguistics_. Cambridge: Cambridge University Press.

Lyons, J. (1977). _Semantics_ (vol. 1). Cambridge: Cambridge University Press.

Mandler, J. M. (1978). A code in the node: the use of a story schema in retrieval. _Discourse Processes, 1_, 14-35.

Mani, K. and Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. _Memory and_

Cognition, 10 (2), 181-187.

Markova, I. (1982). Paradigms, thought and language. Chichester: Wiley.

Marslen-Wilson, W., Levy, E., and Tyler, L. K. (1982). Producing interpretable discourse: the establishment and maintenance of reference. In R. J. Jarvella and W. Klein (Eds.). Speech, place and action: studies in deixis and related topics. Chichester: Wiley, 339-378.

Mead, G. H. (1934). Mind, self and society: from the standpoint of a social behaviourist. Edited and with an introduction by Charles W. Morris. Chicago: University of Chicago Press.

Miller, G. A. (1982). Some problems in the theory of demonstrative reference. In R. J. Jarvella and W. Klein (Eds.) Speech, place and action: studies in deixis and related topics. Chichester: Wiley 61-72.

Miller, G. A. and Johnson-Laird, P. N. (1976). Language and perception. Cambridge: Cambridge University Press.

Miller, G. A. and McKean, K. E. (1964). A chronometric study of some relations between sentences. Quarterly Journal of Experimental Psychology, 16, 297-308.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (ed.) The psychology of computer vision. New York: McGraw-Hill.

Montague, R. (1974). Formal philosophy (Ed. R. H. Thomason). New Haven: Yale University Press.

Morris, C. W. (1938). _Foundations of the theory of signs_. In international encyclopaedia of unified science series, Vol. 1, No. 2, Chichago: University of Chicago Press.

Mossige, S., Pettersen, R. B., and Blakar, R. M. (1979). Egocentrism and inefficiency in the communication of families containing schizophrenic members. _Family Process, 18_, 405-425.

Nunberg, G. D. (1977). _The pragmatics of reference_. Doctoral dissertation, City University of New York. Reproduced by the Indiana University of Linguistics Club, Bloomington, Indiana.

Ogden, C. K. and Richards, I. A. (1972). _The meaning of meaning._ London: Routledge and Kegan Paul.

Olson, D. (1970). Language and thought: aspects of a cognitive theory of semantics. _Psychological Review, 77_, 257-273.

Orne, M. T. (1962). On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. _American Psychologist, 17_, 776-783.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). _The measurement of meaning_. Urbana: University of Illinois Press.

Pailhous, J. (1984). The representation of urban space: its development and its role in the organization of journeys. In R. M. Farr and S. Moscovici (Eds.) _Social_ _reprsentations_, Cambridge: Cambridge University Press (in press).

Pearson, G. (1975). _The deviant imagination_. London: MacMillan.

Pedersen, T. B. (1980). Sex and communication: A brief presentation of an experimental approach. _In_ H. Giles, W. P. Robinson, and P. M. Smith (Eds.) _Language: social psychological perspectives: selected papers from the first International Conference on Social Psychology and Language held at the University of Bristol, England, July, 1979_. Oxford: Pergamon Press.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.) _Semantic information processing_. Cambridge, Massachusetts: MIT Press.

Quine, W. V. (1971). The inscrutability of reference. In D. D. Steinberg and L. A. Jakobovits (Eds.) _Semantics: an interdisciplinary reader in philosophy, linguistics and psychology_. Cambridge: Cambridge University Press.

Rommetveit, R. (1974). _On message structure_. London: Wiley.

Rommetveit, R. (1979). On the architecture of intersubjectivity. In R. Rommetveit & R. M. Blakar (Eds.) _Studies of language, thought and verbal communication_. London: Academic Press, 93-107.

Rommetveit, R. (1984). The role of language in the creation and transmission of social representations. In R. M. Farr and S. Moscovici (Eds.) _Social representations_. Cambridge: Cambridge University Press (in press).

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.) _Categorization and cognition_. Hillsdale, N. J.: Erlbaum.

Sacks, H. Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. _Language, 50_, (4), 696-735.

Sanford, A. J., and Garrod, S. C. (1981). _Understanding written language: explorations in comprehension beyond the sentence_. Chichester: Wiley.

Savin, H. B. and Perchonok, E. (1965). Grammatical structure and the immediate recall of English sentences. _Journal of Verbal Learning and Verbal Behavior, 4_, 348-353.

Schank, R. C. (1975). _Conceptual information processing_. Amsterdam: North Holland.

Schank, R. C. and Abelson, R. (1977). _Scripts, plans, goals and understanding: an enquiry into human knowledge structures_. Hillsdale, New Jersey: Erlbaum.

Schegloff, E. A., Jefferson, G., and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. _Language, 53_ (2), 361-382.

Schegloff, E., and Sacks, H. (1973). Opening up closings. Semiotica 8, 289-327. Reprinted in R. Turner (Ed.) Ethnomethodology, 233-264. Harmondsworth: Penguin (1973).

Slobin, D. I. (1971). Psycholinguistics. Glenview, Illinois: Scott Foresman.

Solvberg, H. A. and Blakar, R. M. (1975). Communication efficiency in couples with and without a schizophrenic offspring. Family Process, 14 (4), 515-534.

Tarski, A. (1956). Logic, semantics, metamathematics: papers from 1923 to 1938. Translated by J. H. Woodger. Oxford: Clarendon Press.

Turner, R. (Ed.) (1974). Ethnomethodology: selected readings. Harmondsworth: Penguin.

Ullmer-Ehrich, V. (1978). Linguistic aspects of static arrangement descriptions in German livingroom descriptions. Paper presented at Max-Planck Conference on deixis and language production, Nijmegen, 1978.

Vygotsky, L. S. (1962). Thought and language. Cambridge, Massachusetts: M.I.T. Press, 185-210.

Winograd, T. (1975). Frame representations and the declarative-procedural controversy. In D. G. Bobrow and A. Collins (Eds.), Representation and understanding. New York and London: Academic Press.

Woods, W. A. (1975). Syntax, semantics and speech. In D. Raj Reddy (Ed.), Speech recognition: invited papers presented at the 1974 IEEE symposium. New York:

Academic Press.

Woods, W. A. (1981). Procedural semantics as a theory of meaning. In A. Joshi, B. Webber and I. Sag (Eds.), _Elements of discourse understanding_. Cambridge: Cambridge University Press, 300-334.

Wunderlich, D. and Reinelt, R. (1982). 'How to get here from here'. In R. J. Jarvella and W. Klein (Eds.) _Speech, place and action: studies in deixis and related topics_. Chichester: Wiley, 183-201.

Appendices

Appendix I. The basic maze shapes used in the maze game studies.

a) Maze 1A, the practice maze used in studies 1 and 2.

b) Maze 5B} The symmetrical

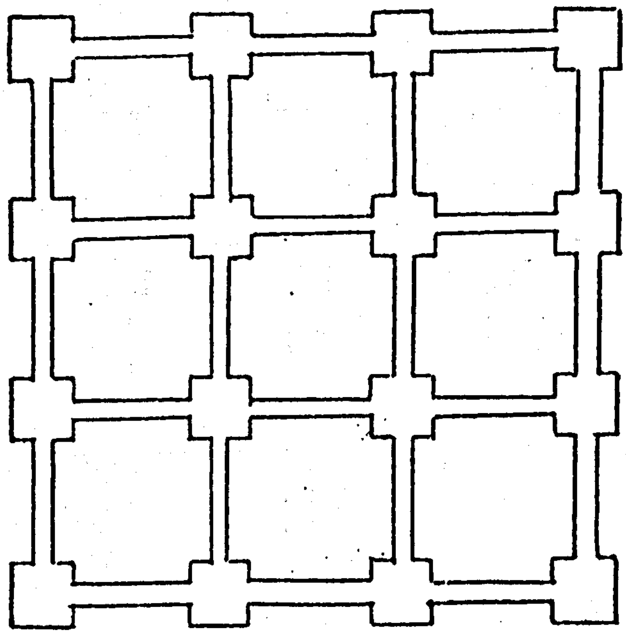
c) Maze 6B} mazes which were

d) Maze 7A} used in studies

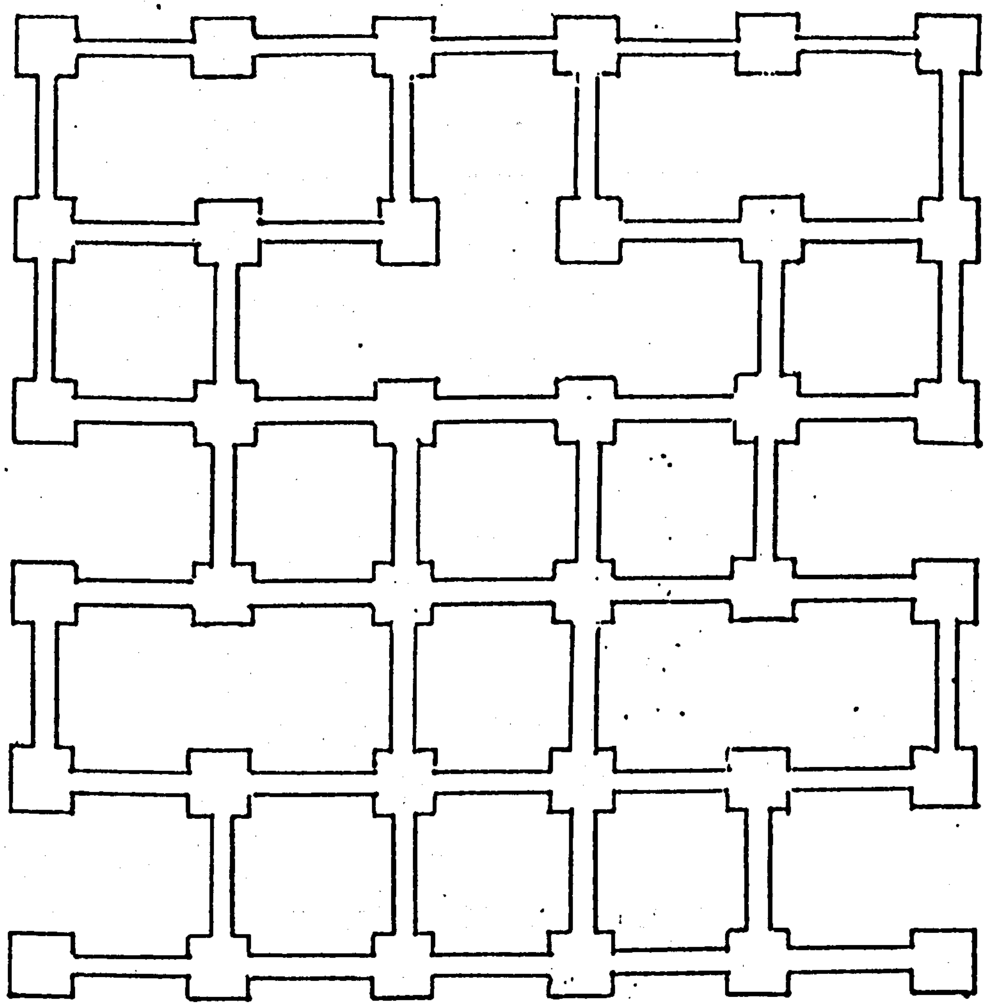
e) Maze 8A} 1, 2 and 4

f) Maze 9B the symmetrical 'column biased' maze which was used in study 3.

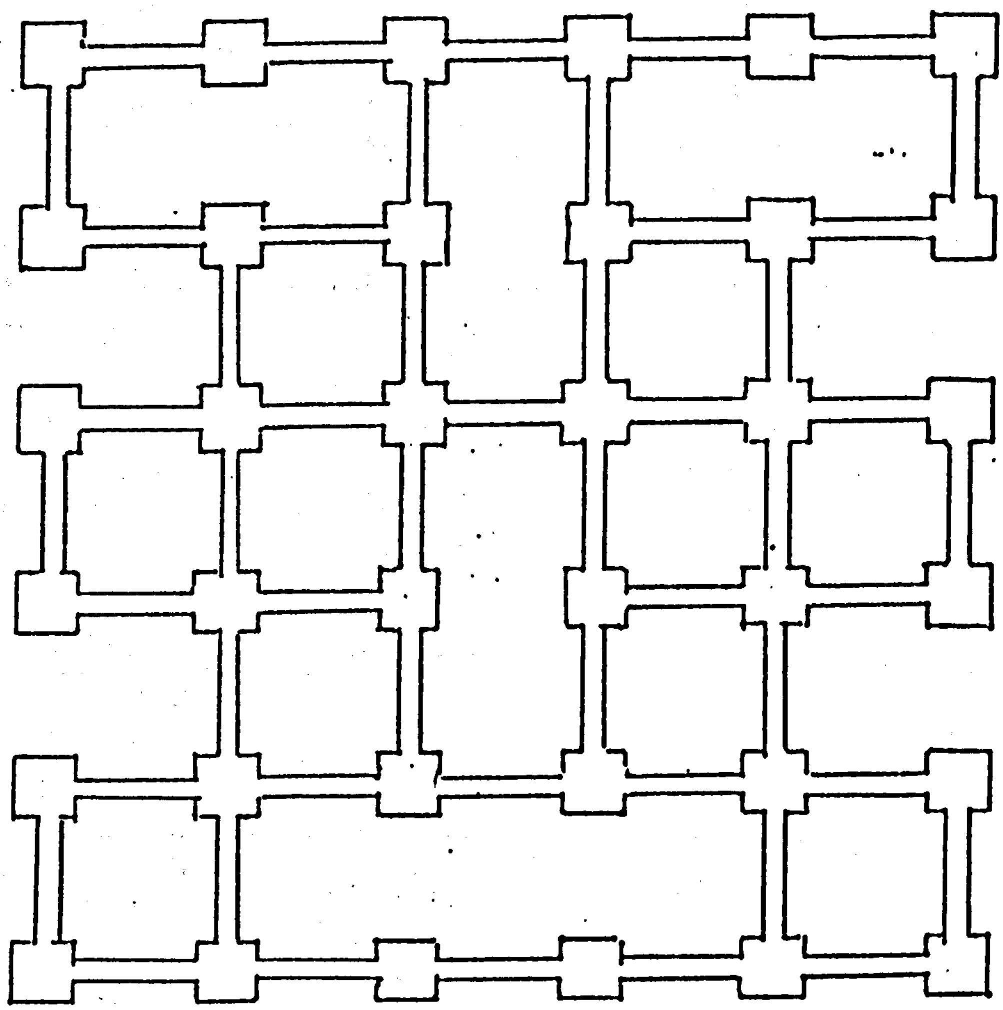
g) Maze 10A, the asymmetrical 'specially-constructed' maze usedin studies 3 and 4.



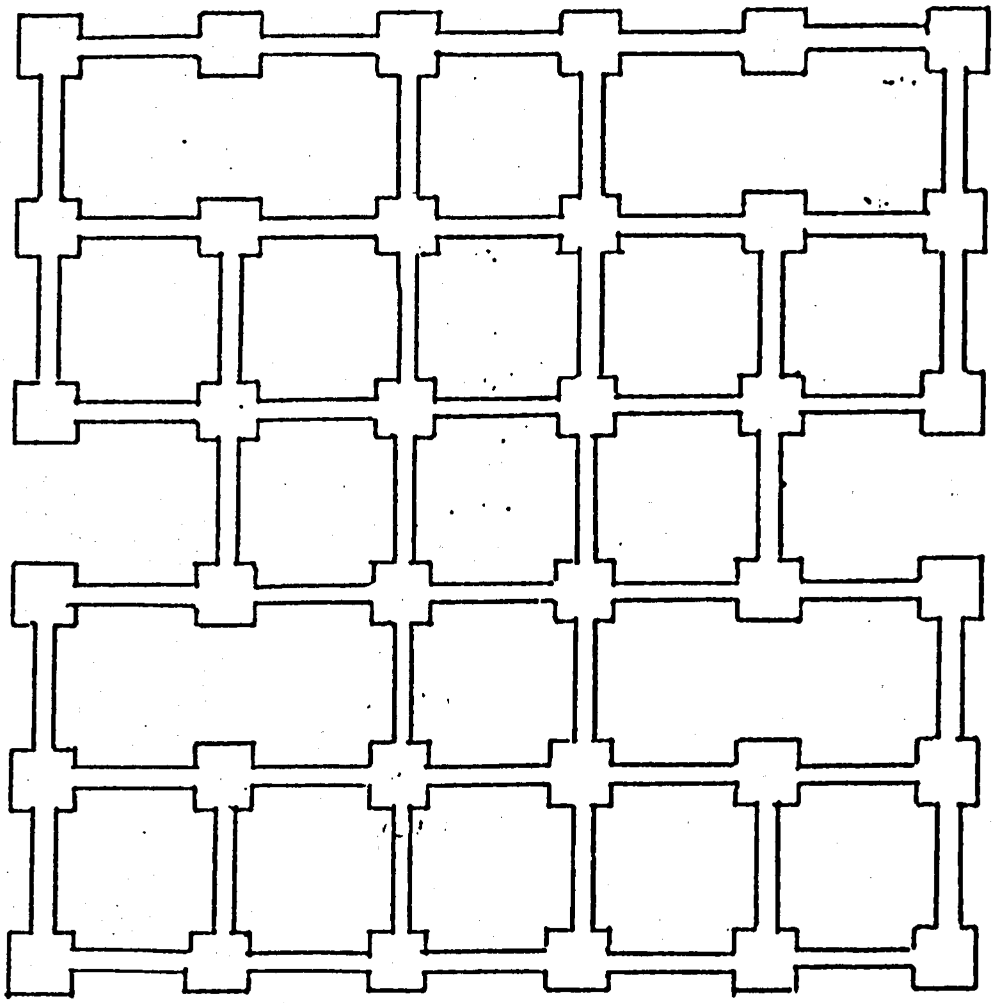
Appendix Ia) Maze 1A, the symmetrical practice maze (based on the smaller 4 x 4 matrix of nodes) used in studies 1 and 2.



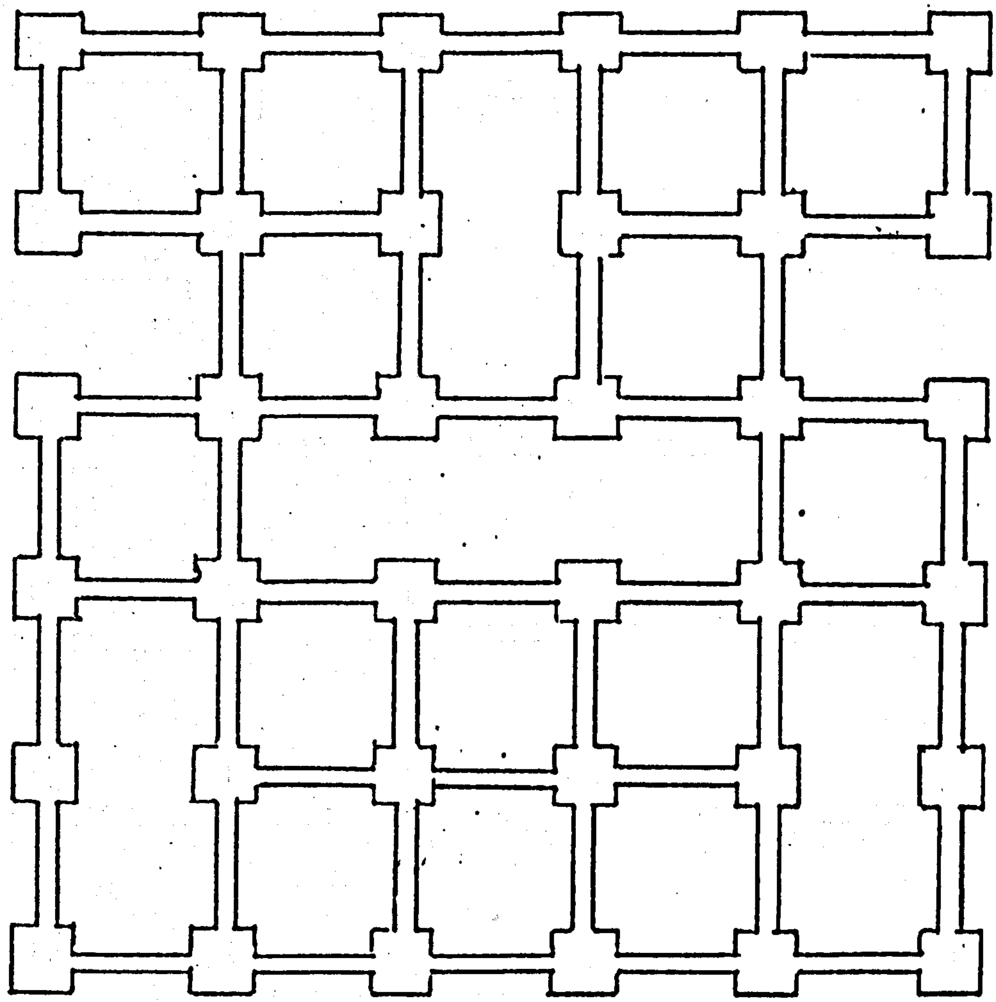
Appendix Ib) Maze 5B, the symmetrical maze (based on the larger 6 x 6 matrix of nodes) used in studies 1 and 2.



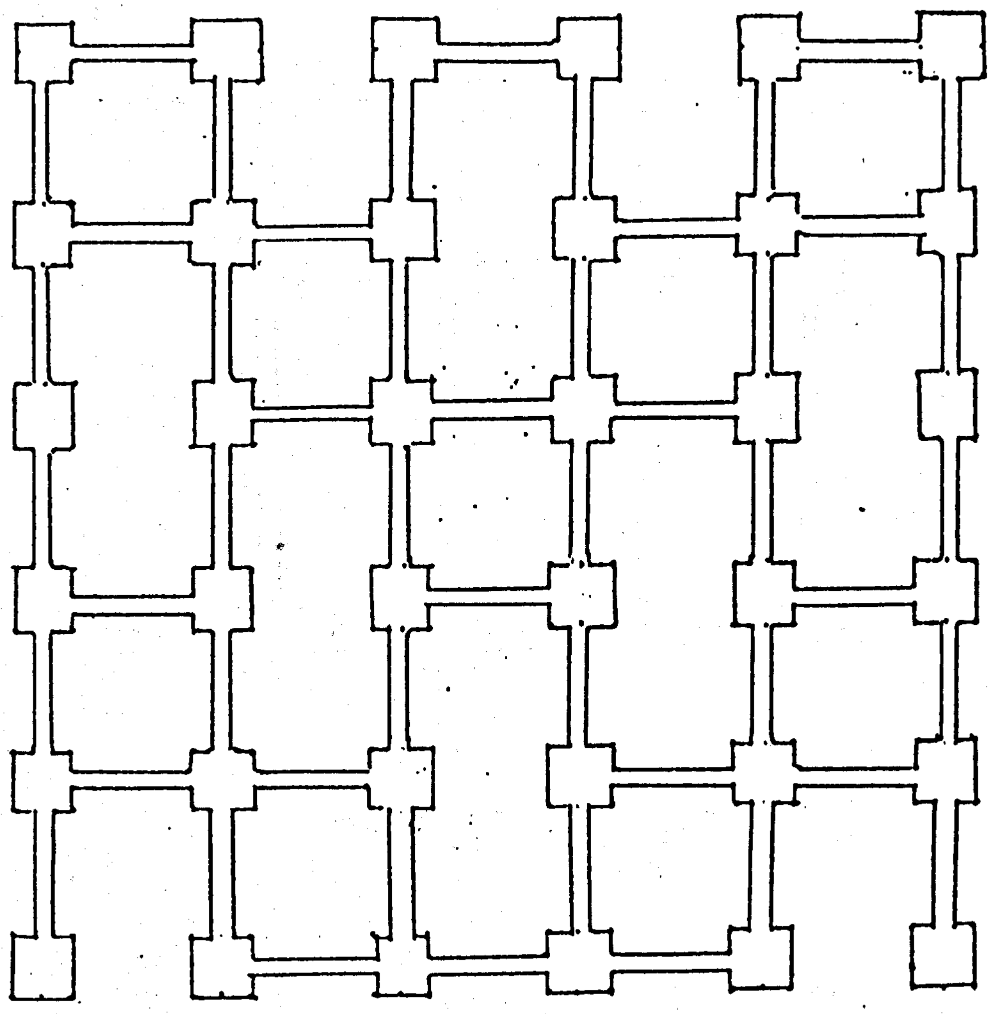
Appendix Ic) Maze 6B, the symmetrical maze (based on the larger 6 x 6 matrix of nodes) used in studies 1 and 2.



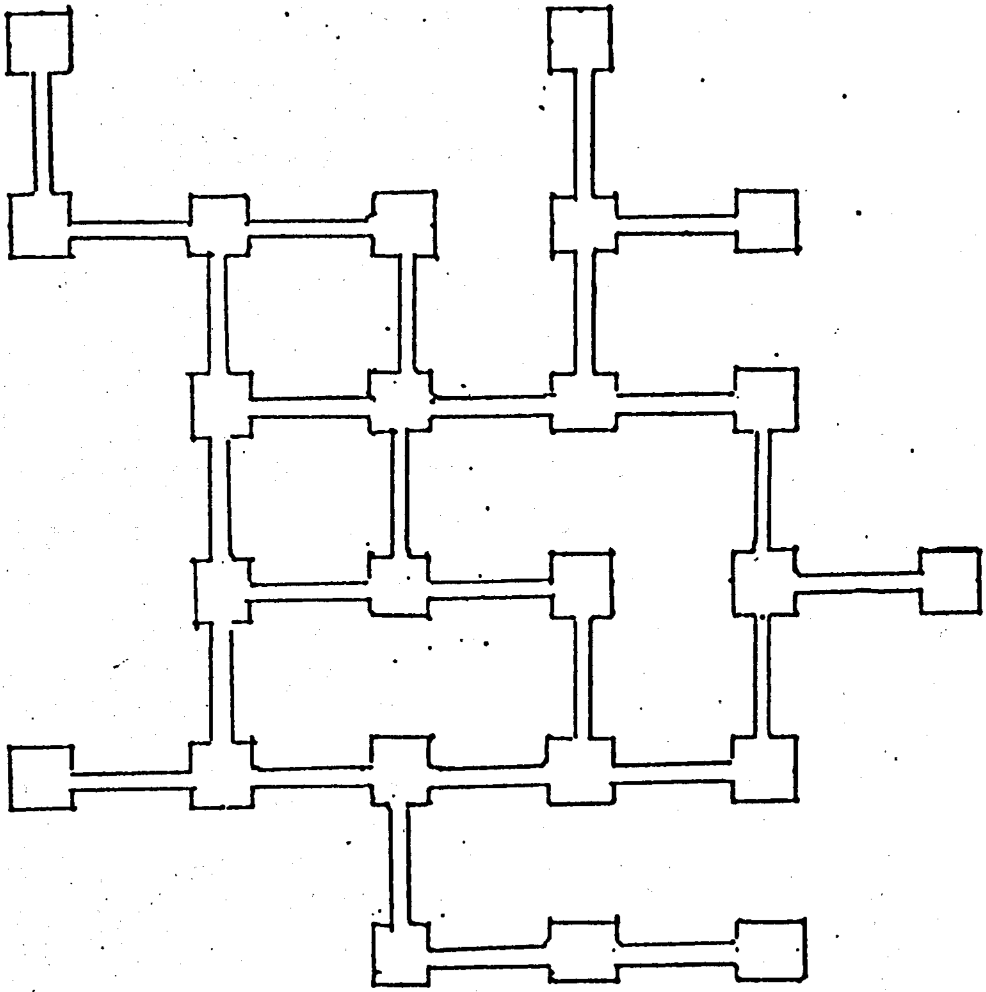
Appendix Id) Maze 7A, the symmetrical maze (based on the larger 6 x 6 matrix of nodes) used in studies 1, 2 and 4.



Appendix Ie) Maze 8A, the symmetrical maze (based on the larger 6 x 6 matrix of nodes) used in studies 1, 2 and 4.



Appendix If) Maze 9B, the symmetrical 'column biased' ('specially constructed') maze used in study 3.



Appendix Ig) Maze 10A, the asymmetrical ('specially-constructed') maze used in studies 3 and 4.

Appendix II. State of game for both players at the start of the various types of game, showing initial position marker locations, and barrier and switch node positions in the following types of game.

a) Practice game (studies 1 and 2)

i) initial game state for player A.
ii) initial game state for player B.

b) Baseline game involving similar directions of movement for both players:

i) initial game state for player A.
ii) initial game state for player B.

c) Baseline game involving opposite directions of movement for the two players.

i) initial game state for player A.
ii) initial game state for player B.

d) Monster game:

i) initial game state and monster position for player A.
ii) initial game state and monster position for player B.

e) Game played on asymmetrical maze (studies 3 and 4).

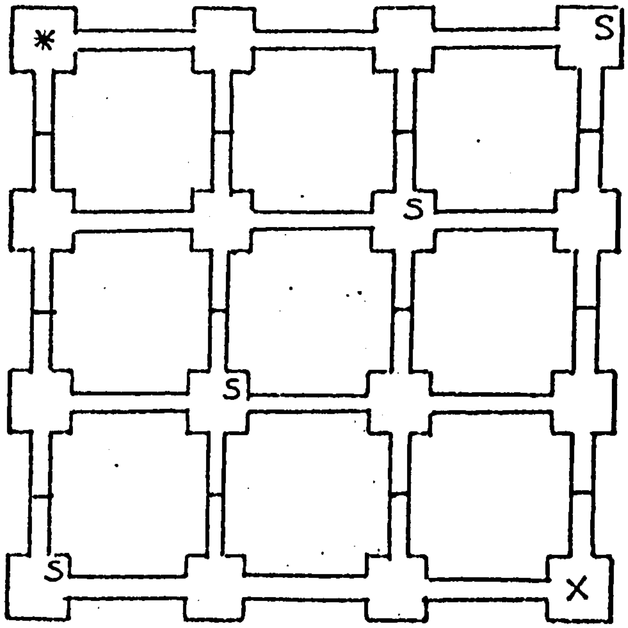
i) initial game state for player A.

ii) initial game state for player B.

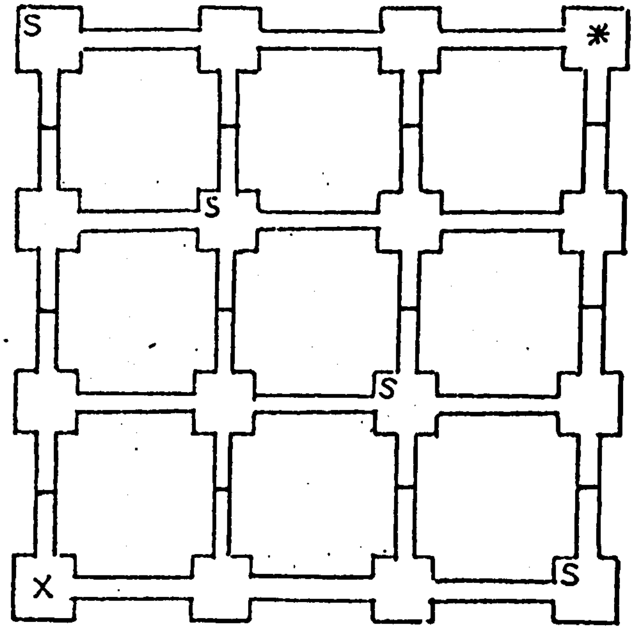
f) Game played on 'column biased' maze (study 3):

i) initial game state for player A.

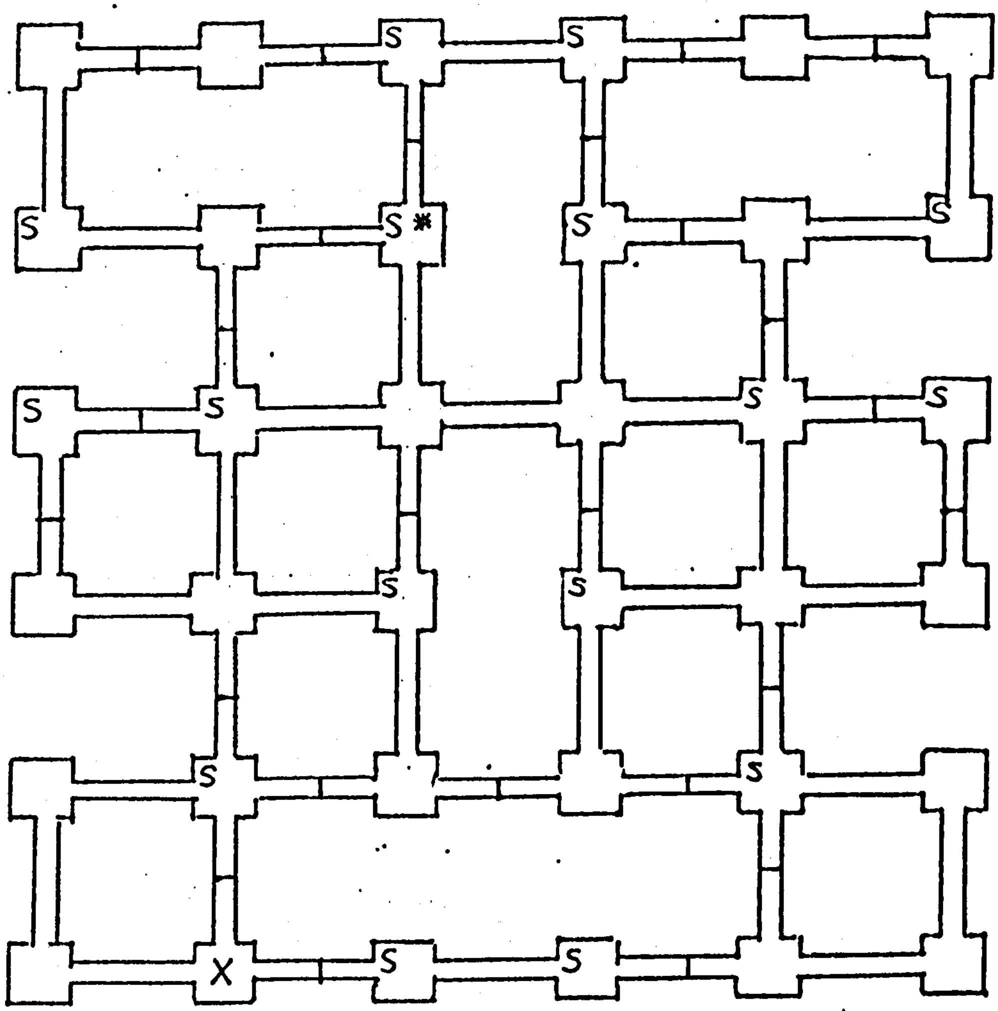
ii) initial game state for player B.



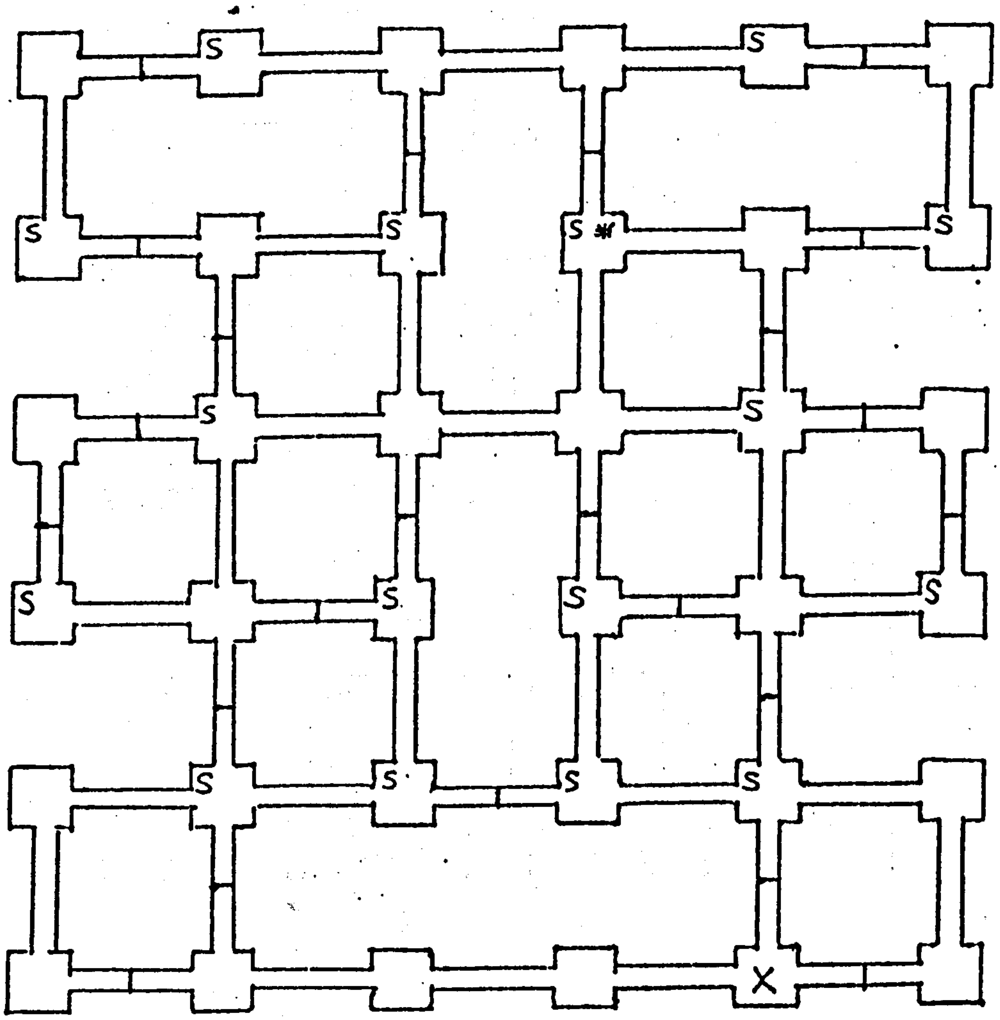
Appendix IIa) i) Practice game (studies 1 and 2): initial game state for player A.



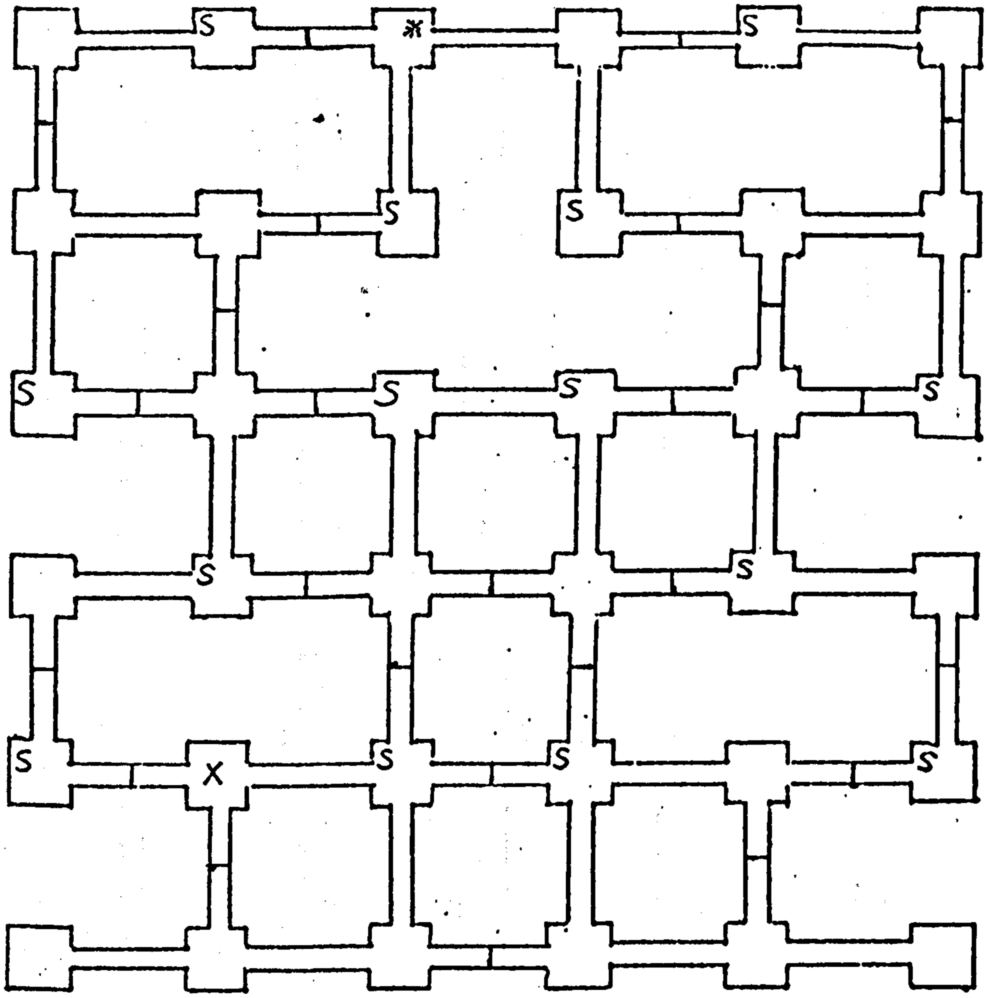
Appendix IIa) ii) Practice game (studies 1 and 2): initial game state for player B (cf. Appendix IIai)).



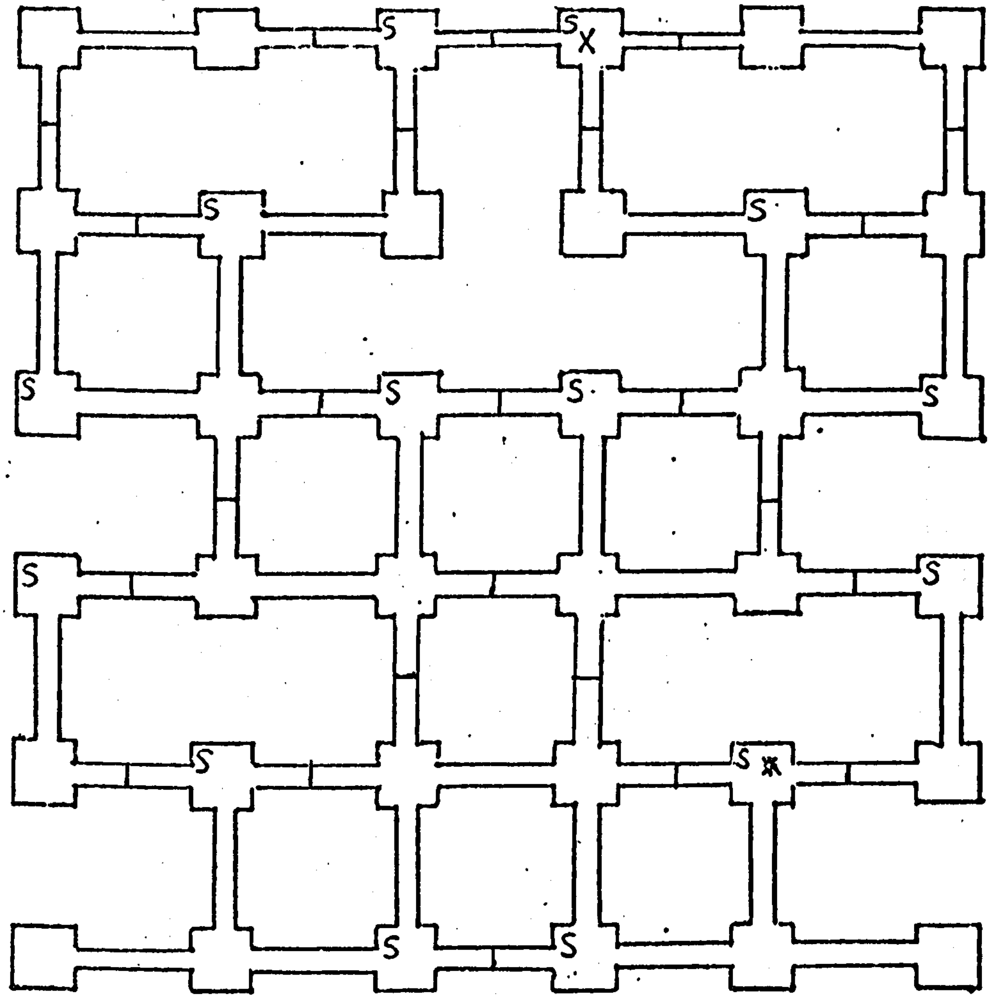
Appendix IIb) i) Baseline game involving similar directions of movement for both players: initial game state for player A



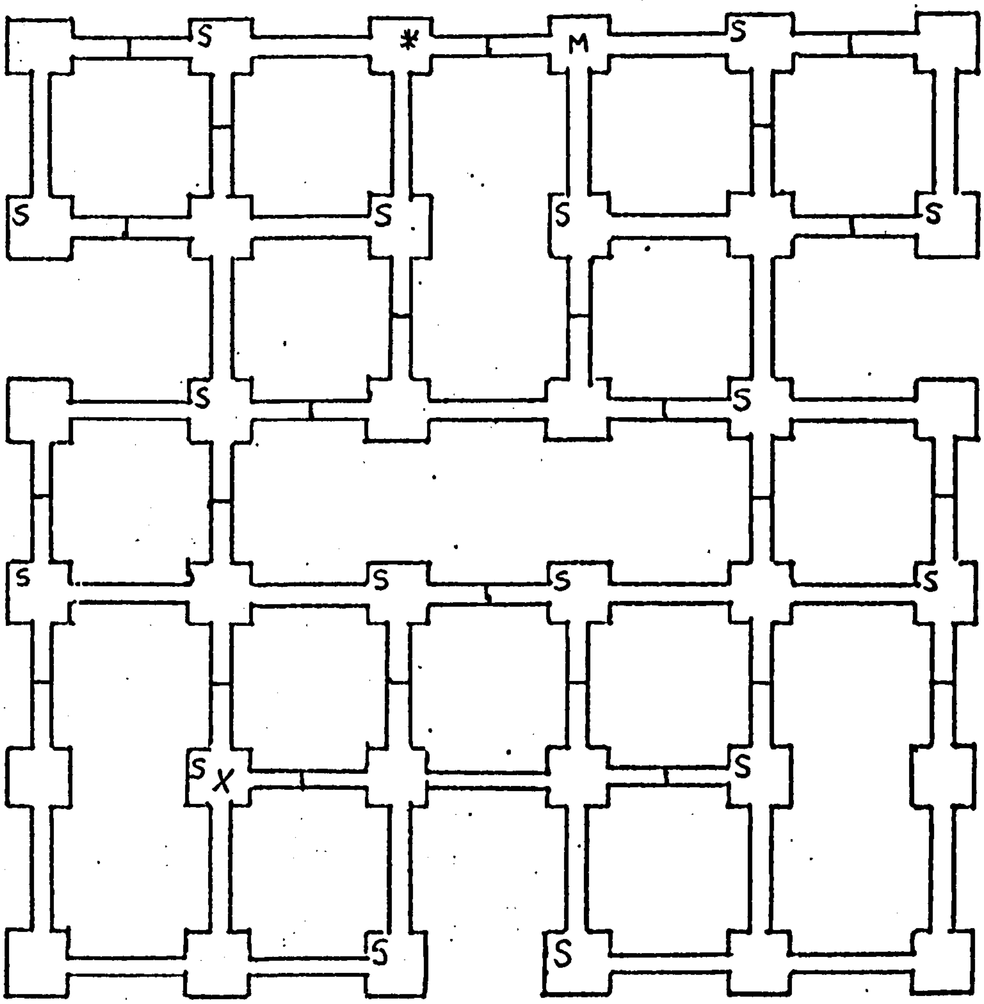
Appendix IIb) ii) Baseline game involving similar directions of movement for both players: initial game state for player B (cf. Appendix IIb) i)).



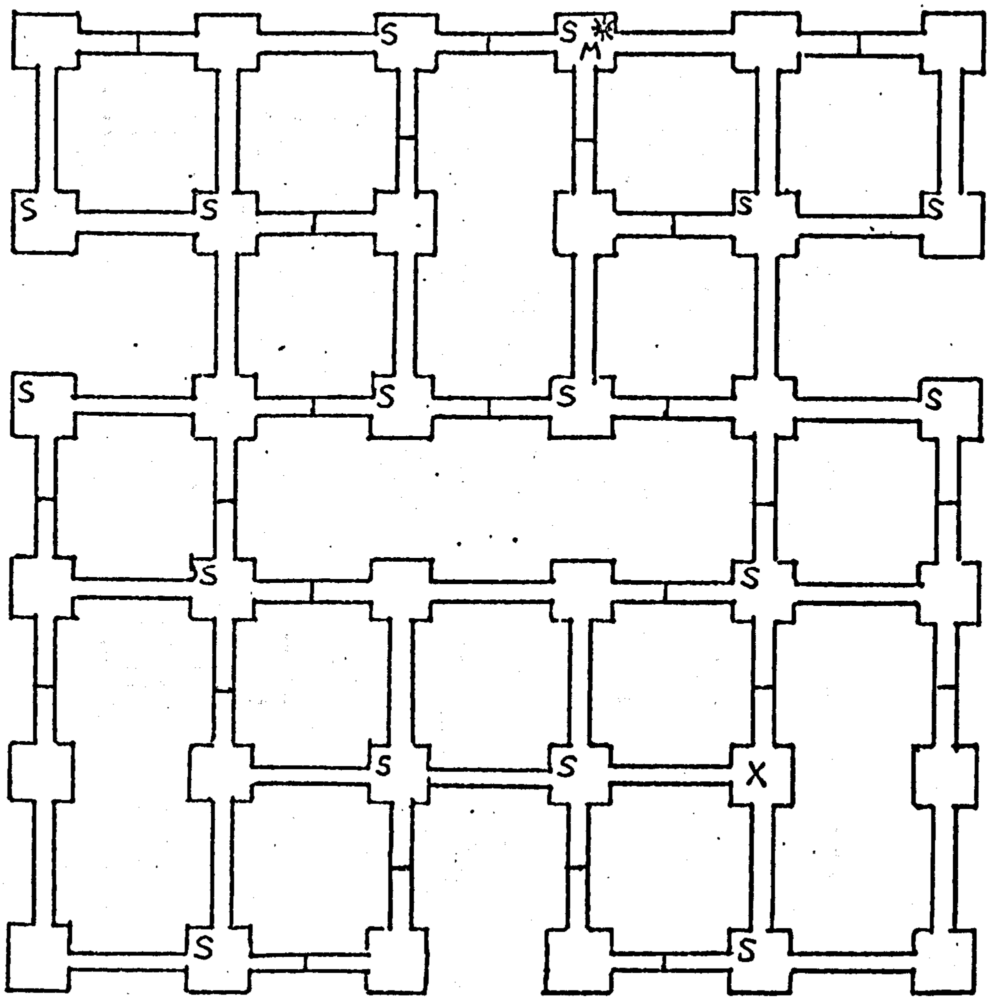
Appendix IIc) i) Baseline game involving opposite directions of movement for the two players; initial game state for player A.



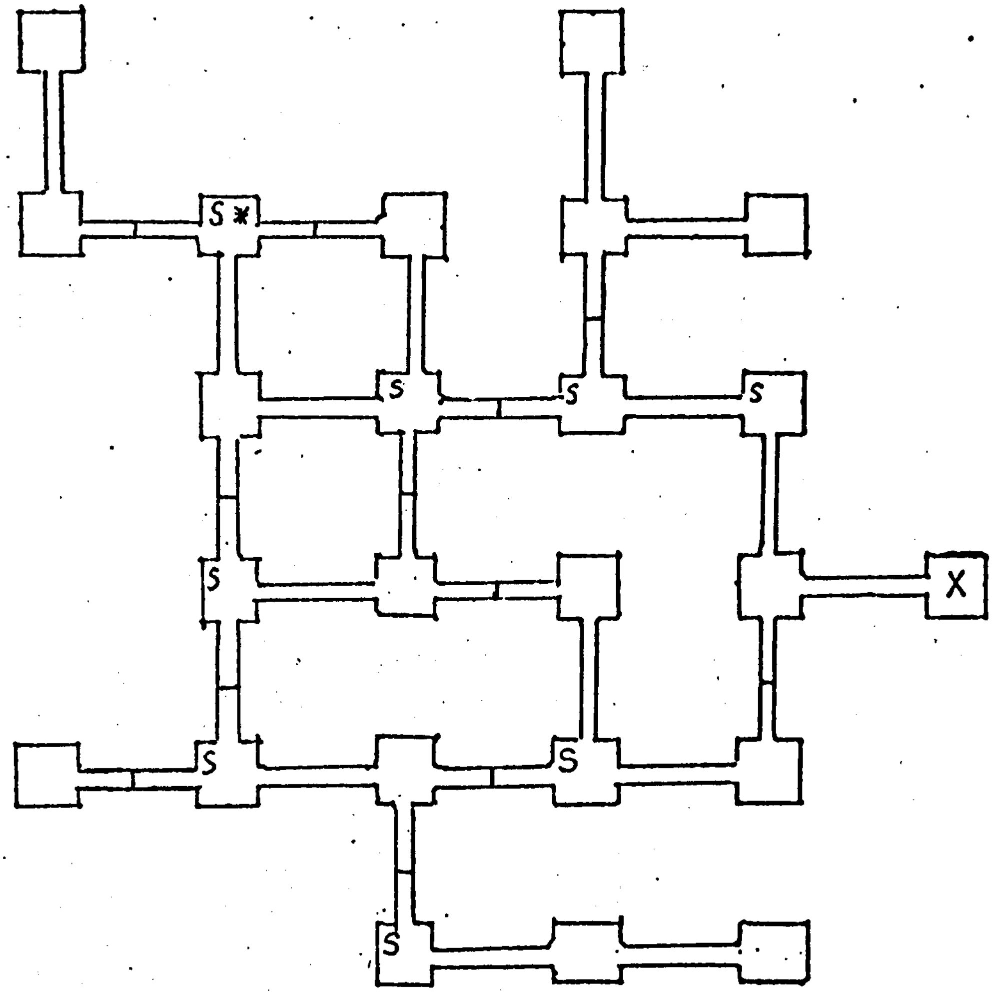
Appendix IIc) ii) Baseline game involving opposite directions of movement for the two players, initial game state for player B (cf. Appendix IIc) i)).



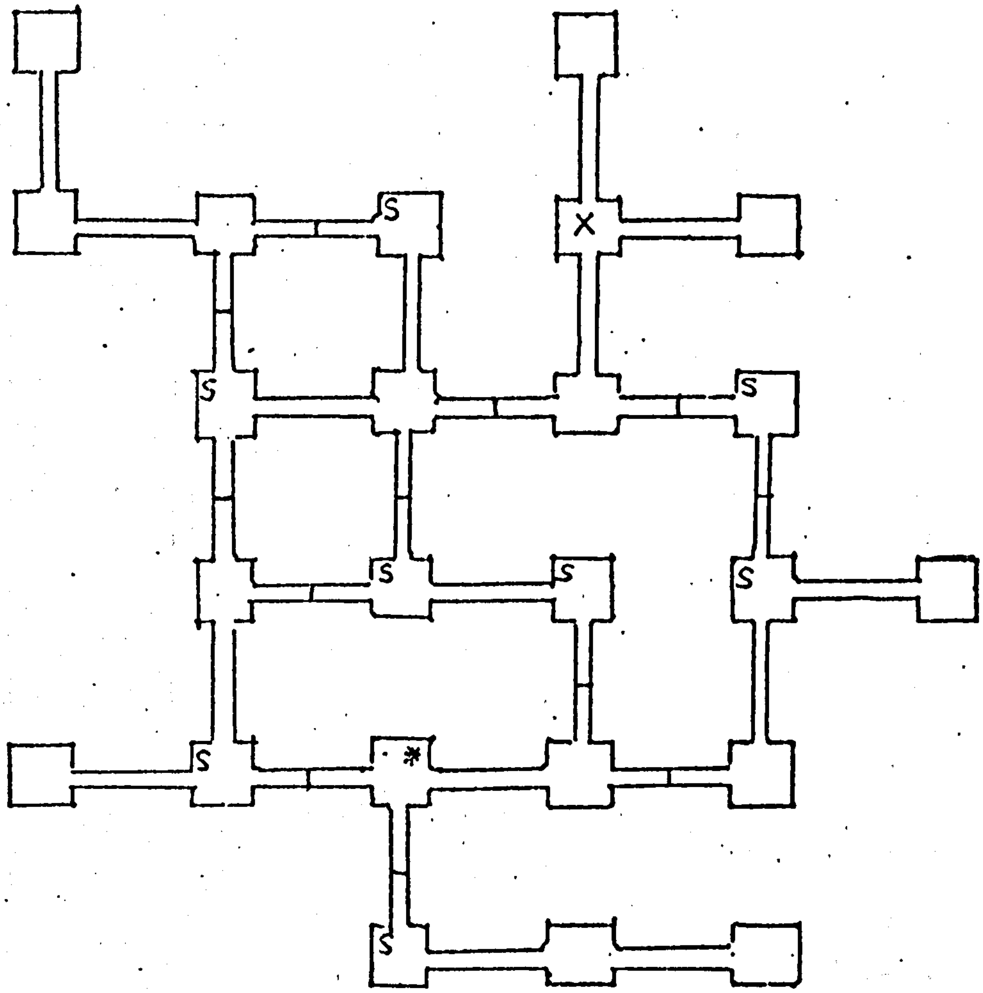
Appendix IId) i) Monster game: initial game state and monster position for player A.



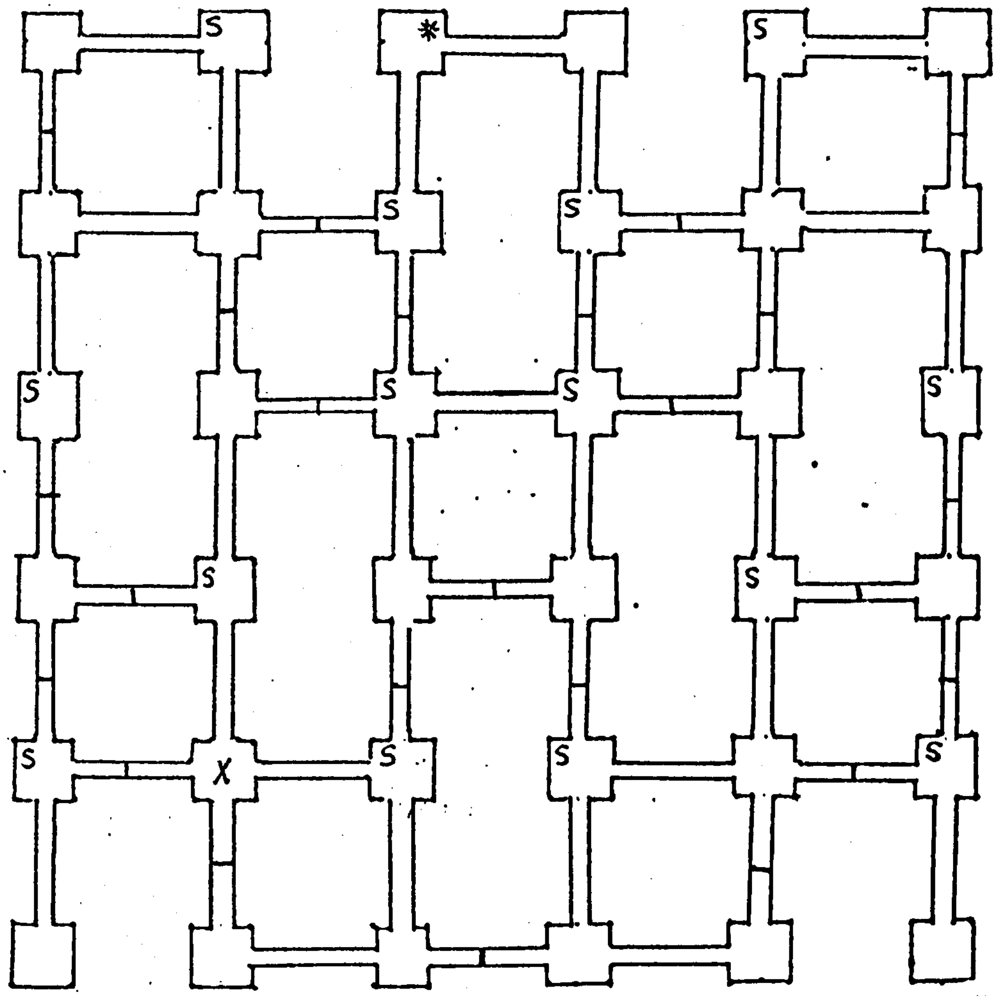
Appendix IId) ii) Monster game: initial game state and monster position for player B (cf. Appendix IId) i)).



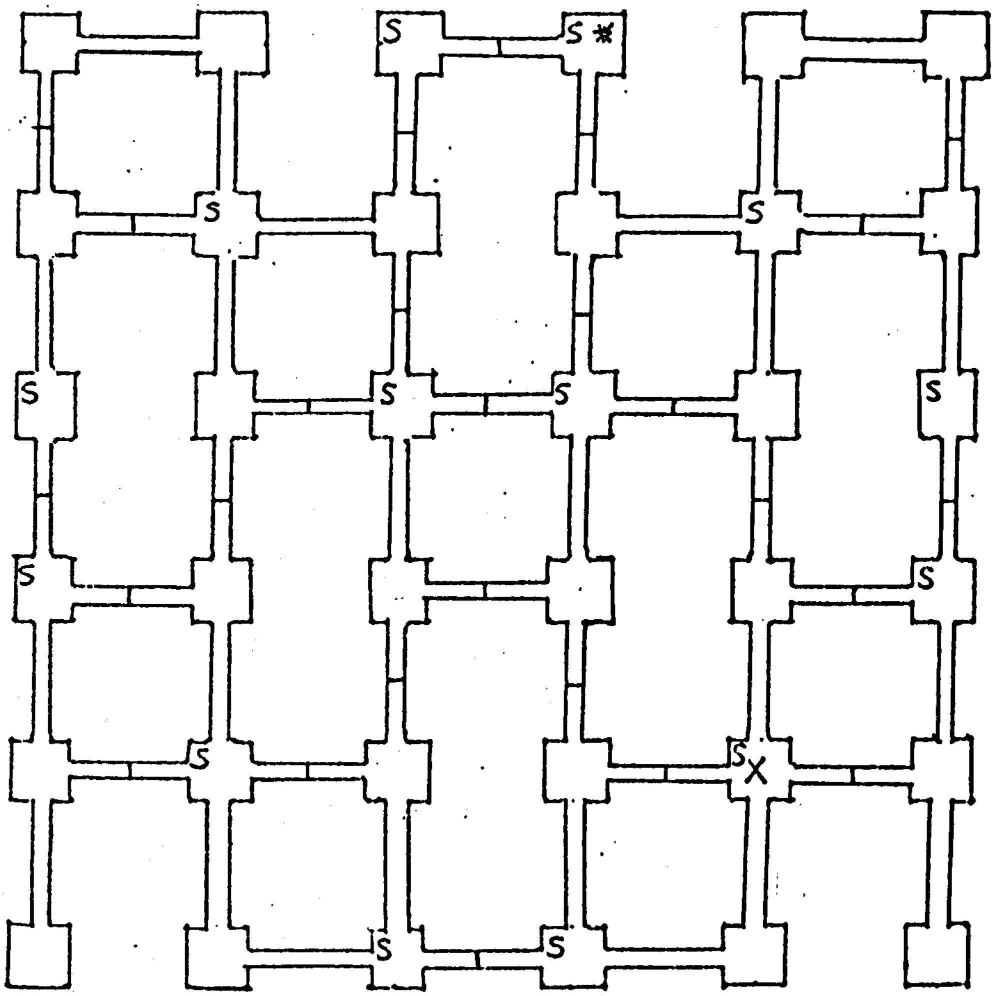
Appendix IIe) i) Game played on asymmetrical maze (studies 3 and 4): initial game state for player A.



Appendix IIe) ii) Game played on asymmetrical maze (studies 3 and 4): initial game state for player B (cf. Appendix IIe) i)).



Appendix IIf) i) Game played on 'column biased' maze (study 3): initial game state for player A.



Appendix IIf) ii) Game played on 'column biased' maze (study 3): initial game state for player B (cf. Appendix IIf) i)).

Appendix III

Symbols employed in transcripts of subjects' speech

When transcribing subjects' speech, care was taken to provide as accurate a representation as possible. The resulting transcripts therefore contained information regarding filled and unfilled pauses, the length of such pauses, the amount of simultaneous speech, and any special intonation contours (such as the occasional shout, or stressing of particular syllables). The symbols employed in representing these various features are listed below. These are the symbols employed throughout the text of the present thesis.

Pauses are represented in two forms. The colon, : , is used to symbolize pauses of less than 1 second duration. Longer pauses were measured as accurately as possible using the second counter of a digital wristwatch, are represented in parentheses, and are accurate to the nearest half second - for example, (3.5 sec.) indicates a pause of approximately 3.5 seconds in duration.

Filled pauses took the form of expressions such as "ehm", "er", "uhm", and the like, and were transcribed and represented in a fashion as close to the original sound as possible.

Simultaneous speech is represented in the transcripts and in the present volume by means of placing the simultaneous speech of the two individuals adjacently vertically, the start and end of the simultaneous speech being represented by means of square brackets. For example:

B . . . if you go to the bottom right corner

[one along, two up]

B [one along] two up.

In this example the phrase "one along" is spoken simultaneously by both players (although player A speaks the phrase a little more slowly and overlaps the following phrase 'two up' of player B).

Cases where subjects employed <u>special intonation</u>, as judged on a purely intuitive basis, were represented by a line drawn <u>above</u> the specially-intoned speech. For example:

B Oh no, one along and two up:

A Two up?

B Well: youknow, counting both squares.

Player A's repetition "two up" takes the form of a question, and his special emphasis in this example is best described as one which reflects incredulity. The representation of such intonational emphasis in the form of a line drawn <u>above</u> the relevant phrase or lexical item avoids confusion with the convention adopted in the present thesis of <u>underlining</u> exemplars of whatever is being discussed (and also the names of lexical categories).

Appendix IV. Technical Appendix: the technique of hierarchical cluster analysis

Hierarchical cluster analysis is a descriptive statistical technique used for the purposes of exploratory data analysis of data sets involving a number of variables. Gordon (1981) presents a lucid exposition of the rationale behind the technique, and in the present appendix some of the important points raised by Gordon from the point of view of the present set of studies will be discussed.

The aim of a classification couched in terms of a cluster analysis is the uncovering of the structure that is present in the data rather than the imposition of some inappropriate structure upon them. It is a widely used technique, being used in a variety of disciplines.

From the point of view of the present set of studies, two of the possible uses of cluster analysis are of relevance, namely:

i) Data simplification. Cluster analysis can be used to detect the important relationships between variables and the patterns of relationships in the data. This is particularly useful where an investigation involves large numbers of variables, which makes it difficult for an experimenter to easily understand the data.

and ii) Hypothesis generation. As a corollary to i) above, if the data have been summarized in a manner which allows the important relationships between variables to emerge more clearly, this can aid the investigator in generating predictions or hypotheses to account for the form of the data.

In the present context, the uncovering of the structure of the data is a desideratum. A summary of the important relationships between the lexical categories in our classification would be useful, and it would also be of great interest to us to assess how well the formal description arrived at by purely mechanical means would accord with our intuitive description which was presented earlier (see chapter 6).

Gordon (1981) describes a number of statistical techniques which are used to uncover the structure of the data in large data sets. He refers to these techniques as classification techniques. They include hierarchical cluster analysis, 'clumping' methods, partitioning methods, and geometrical methods of classification. The present set of studies used the hierarchical cluster analysis technique available in the BMDP (Bio-medical computer programs) statistical software package (Dixon et al., 1981).

Hierarchical cluster analysis works essentially as follows. Firstly, a measure of similarity (or distance) between each pair of variables is computed (for example, the correlation between the variables could be used as a measure of similarity between the variables), resulting in a half-matrix of pairwise similarities (or distances). Next, clusters are formed according to a chosen amalgamation rule, and the amalgamation procedure operates on the computed similarities or distances between variables. Initially, each cluster contains only one variable, and at each step the two most similar clusters are joined to form a new cluster, until all the variables are in the same cluster. The result is summarized in a tree diagram known as a dendrogram (see example dendrogram below).

The resulting hierarchical structure is a best-fit hierarchical structure imposed on the data using the linkage rule, and the tree displays both the groupings of variables within clusters and the relationships between clusters. The hierarchical structure is imposed on the data to the extent that each group in a partition at one level is wholly contained within a single group at a higher

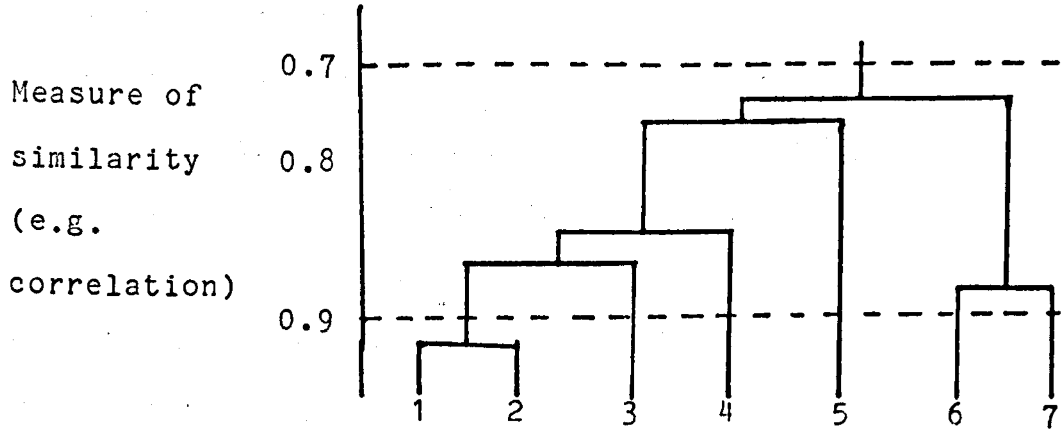


Figure IV-1

Hypothetical example of a dendrogram showing the clusterings between seven variables. The vertical axis represents the scale of similarity (in this case the measure of similarity between variables is the correlation). Thus, variables 1 and 2 are more similar than are variables 6 and 7. The structure is hierarchical in that clusters which are separate at one level (e.g. that corresponding to a correlation of 0.9) form part of a single cluster at a higher level (e.g. that corresponding to a correlation of 0.7; see diagram above and text).

level: as Gordon (1981) points out, it is conceivable that partitions at different levels should not be hierarchically nested. Since hierarchical cluster analysis imposes a best-fit hierarchical structuring on the data, it is possible that the structure in the data is distorted by this procedure, and one should not accept uncritically the results suggested by a single clustering procedure (Gordon, 1981, p. 33-34).

It is possible to compare results yielded using different clustering procedures (that is, different amalgamation or linkage rules for forming clusters). The rationale behind this is as follows. If there really is a distinct structure present in the data, any clustering procedure worthy of use should uncover that structure. Conversely, if the results of several different clustering procedures agree closely, ". . . then one has more confidence in the reality of any group structure which is indicated; it is less likely to be purely an artifact of the classification criteria used" (Gordon, 1981, p. 132). Thus, if the same structure (the same groupings of variables) is revealed by different cluster analysis procedures, we can be more confident that this is indeed part of the structure in the data rather than it having been imposed by the classification (clustering) technique itself.

To summarize: cluster analysis operates by generating a half matrix of pairwise similarities or distances between variables, and the procedure then imposes a best-fit hierarchical structuring upon the variables using a selected amalgamation (linkage) rule for forming clusters which operates upon the half-matrix of pairwise similarities (or distances). Since the hierarchical structuring is to some extent imposed on the data, it is best to examine sets of results yielded using more than one clustering (amalgamation) rule, since greater confidence can be had in results which appear consistently according to different amalgamation criteria: these consistent results are more likely to reflect the structure in the data, rather than being an artifact of the clustering techniques employed.

Gordon (1981) presents a simple method for comparing results from different cluster analyses. In the present context, we are interested in quite 'low-level' relationships, that is, we are concerned with which groupings of individual lexical categories exist, rather than being concerned with the relationships between groupings of categories. In other words, we are primarily interested in the lowest levels of the dendrogram. Gordon's comparison technique is as follows: consider results from two analyses which have cluster analyzed seven variables, numbers 1 to 7. The analysis based on amalgamation rule Q (for example, the sum of squares criterion) yields three groupings: Q_1 (comprising variables 1 and 2), Q_2 (variables 3 and 4) and Q_3 (variables 5, 6 and 7). The other analysis, based on the single link amalgamation rule S_1, yields four groups as follows: group S_1 (variables 1, 2 and 3) group S_2 (variable 4), group S_3 (variable 5), and group S_4 (variables 6 and 7). One analysis thus partitions the variables into three groups, whereas the other analysis partitions the data into four groups. A two-way table is drawn up which summarizes the intersections of the groups, as follows:

Table IV-1:

Comparison of two partitions of a data set (from Gordon, 1981, p. 134).

Single link groups	Sum of square groups			
	Q_1	Q_2	Q_3	Q_4
	(1, 2)	(3, 4)	(5, 6, 7)	
S_1 (1, 2, 3)	(1, 2)*	(3)		
S_2 (4)		(4)*		
S_3 (5)			5	
S_4 (6, 7)			(6, 7)*	

The starred entries indicate the optimal matching of the two data sets.

There are two points of note. Firstly, Q_4 is defined to be an empty set in order to equalize the number of groups in the two partitions. Secondly, the groups are matched in pairs in such a way as to maximize the number of variables in common in the set of matched pairs of groups. Thus, although both groups are shown in the table set out in order $Q_1 \ldots Q_4$ and $S_1 \ldots S_4$, one of these sequences could be rearranged if this resulted in a greater degree of matching between the groups.

The overall result, as can be seen from table IV-1, is that the two partitions agree on five out of the seven variables, these falling into the three groups (1, 2), (4), and (6, 7), with the other variables (3) and (5) being outliers or intermediate between some of these groups. One might have more confidence in the reality of the three groups (1, 2), (4) and (6, 7), since they are yielded in the intersection of two analyses satisfying different clustering criteria.

This technique of comparing cluster analyses has a double appeal in that (a) it is simple and quick to use, and (b) it is readily extended to allow comparison of more than two data partitions. The technique is, of course, also readily extendable to partitions involving larger numbers of groupings.

In the present set of studies, the above cluster comparison technique was employed to, as it were, 'filter out' the consistently occurring groupings of lexical categories. The statistical package BMDP (Dixon et al., 1981) was used to carry out the cluster analyses using program P1M (cluster analysis of variables). With the BMDP P1M program, it is possible to vary (a) the measure of similarity or dissimilarity on which the clustering is based and (b) the criteria on the basis of which clusterings are built up. Both the measure of similarity and the criterion for clustering were varied in the present set of studies (see chapter 7), and the above cluster comparison technique was employed to ascertain which consistent groupings of variables were occurring across different analyses.

Appendix V. The maze shapes used in the maze-game-independent ('paper-and-pencil') location description task.

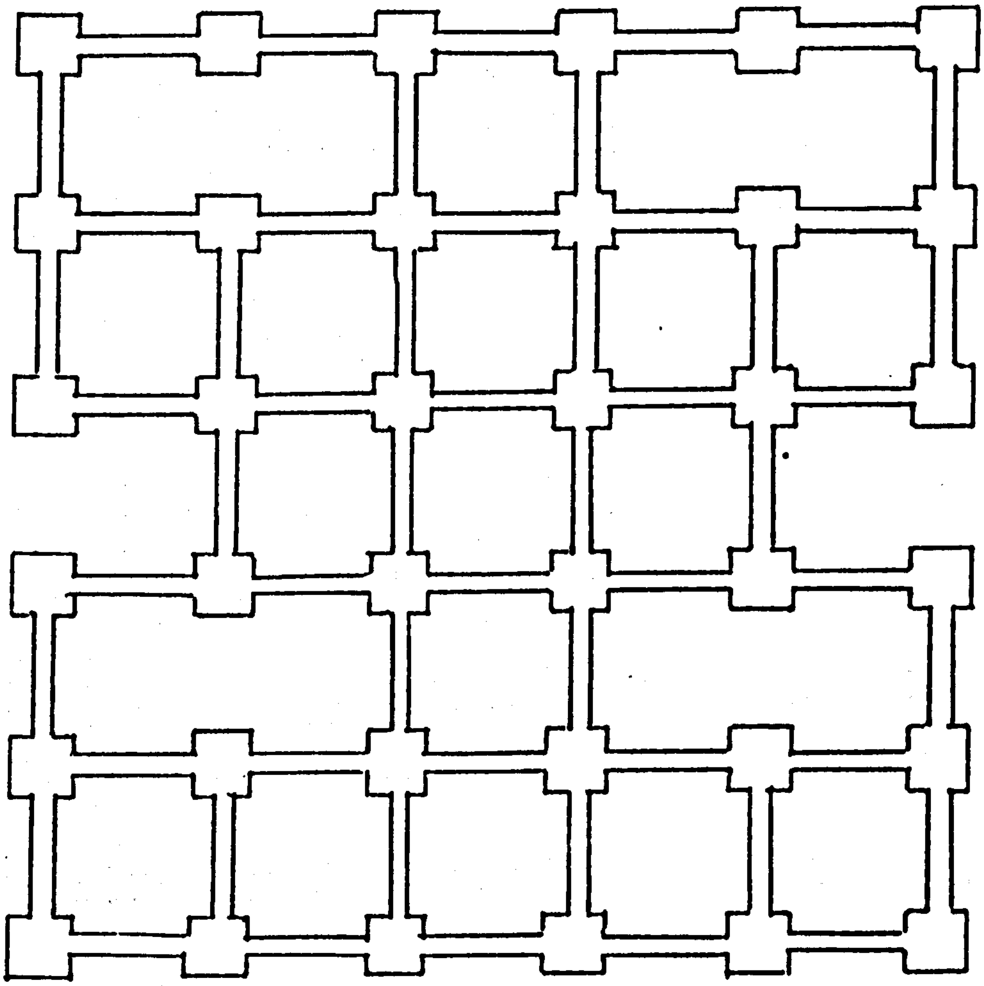
a) The symmetrical maze with regularly-spaced node positions.

b) The symmetrical maze with irregularly-spaced node positions.

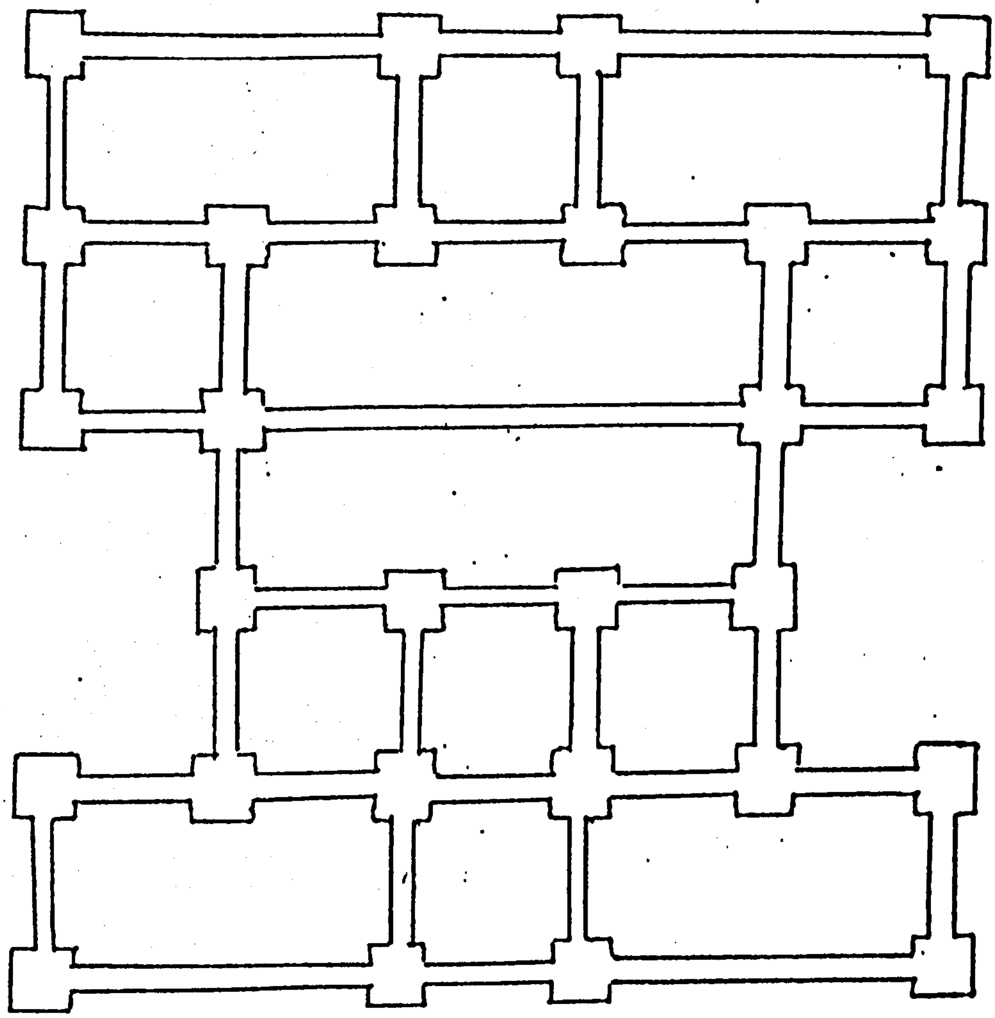
c) The asymmetrical maze with regularly-spaced node positions.

d) The asymmetrical maze with irregularly-spaced node positions.

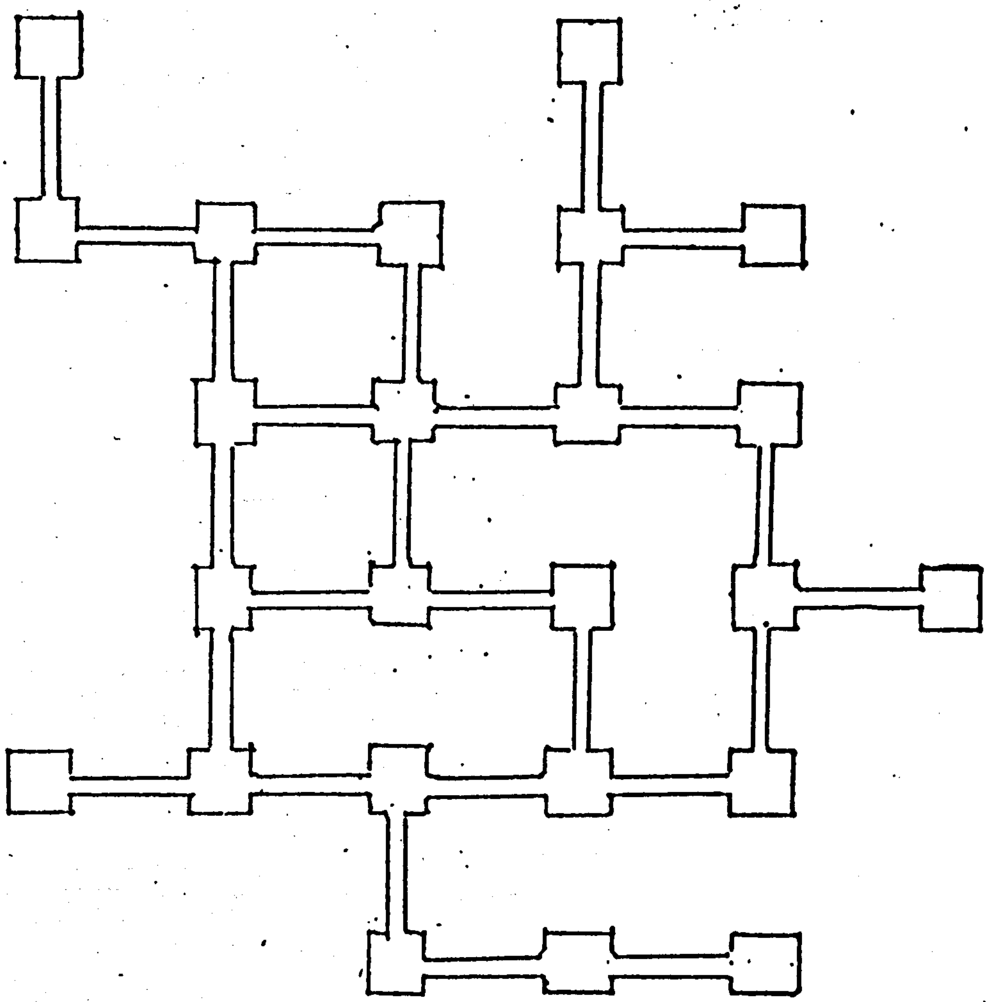
e) The symmetrical maze with regularly-spaced node positions and diagonally-oriented pathways.



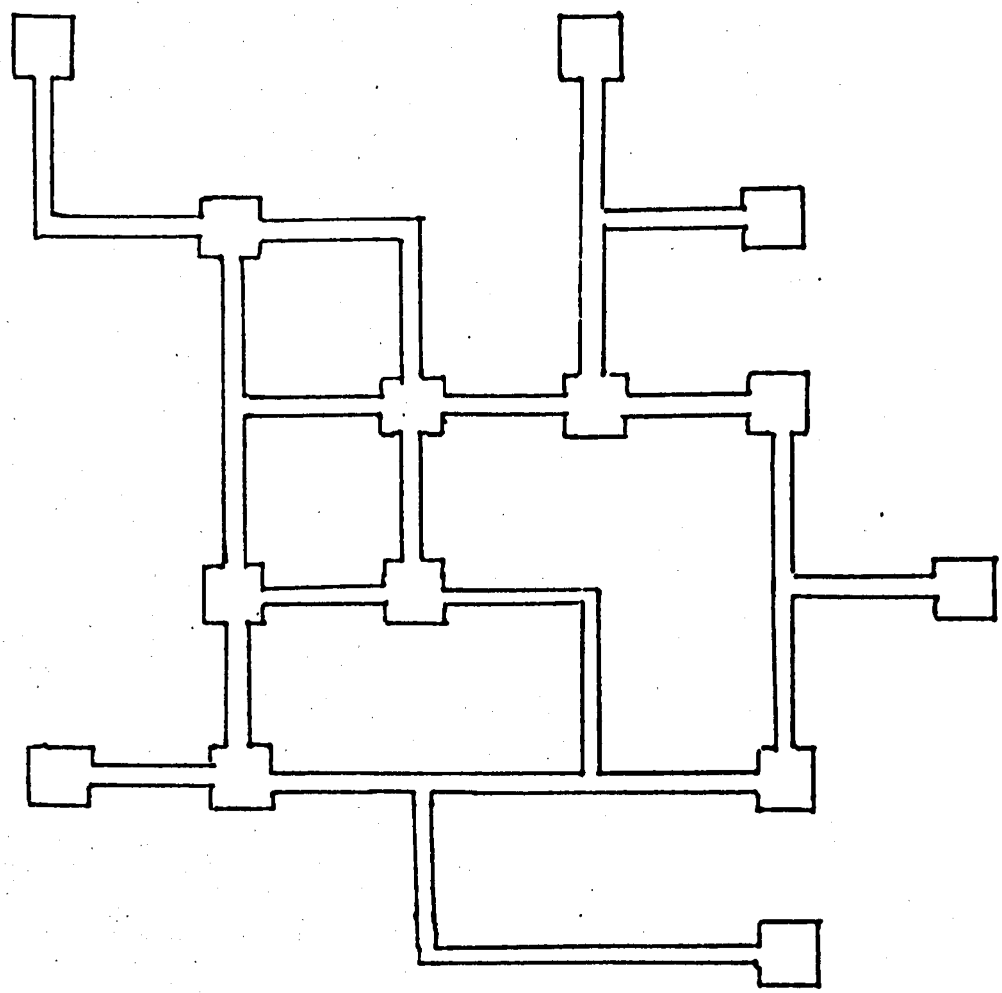
Appendix Va) The symmetrical maze with regularly-spaced node positions (identical to maze 7A, see Appendix Id)) used in the 'paper-and-pencil' study.



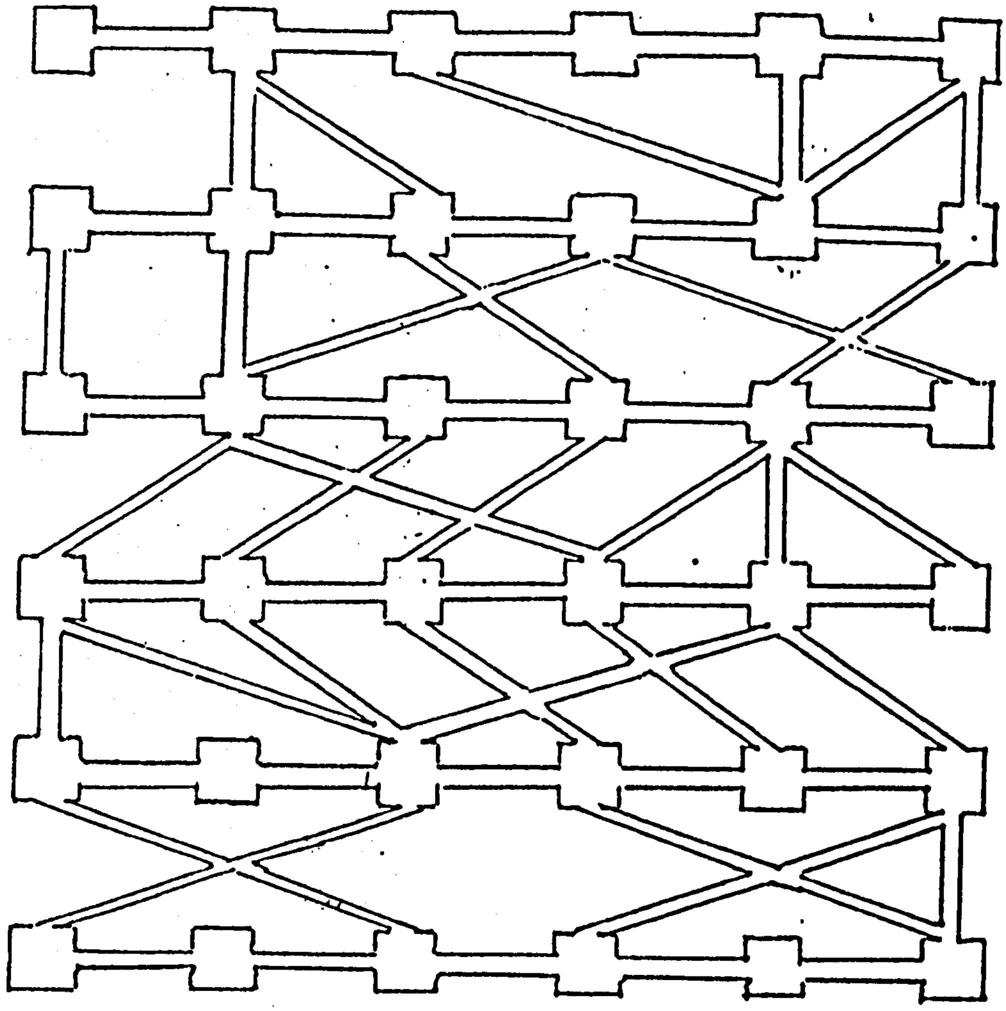
Appendix Vb) The symmetrical maze with irregularly-spaced node positions which was used in the 'paper-and-pencil' study.



Appendix Vc) The asymmetrical maze with regularly-spaced node positions (identical to maze 10A, see Appendix Ig)) which was used in the 'paper-and-pencil' study.



Appendix Vd) The asymmetrical maze with irregularity-spaced node positions used in the 'paper-and-pencil' study.



Appendix Ve) The symmetrical maze with regularly-spaced node positions and diagonally-oriented pathways which was used in the 'paper-and-pencil' study.

APPENDIX VI

THE MAZE GAME PROGRAM: FURTHER DETAILS

As noted already in Chapter 4, the original maze program (i.e. that used in the present set of studies) was written in the appropriate machine codes of the Data General Nova computer and the 'Intel' and 'Imlac' terminals by Mr. J. Mullin, Computer Manager of the Department of Psychology at the University of Glasgow. The maze game itself was devised by Dr. S. C. Garrod of the Department of Psychology at the University of Glasgow.

Since the present set of studies were conducted, the maze game program has been rewritten in the programming language Pascal for use with Apple II microcomputers using a Corvus Constellation disk drive. The updated version of the program contains additional options which had not been available in the original program, namely, the spacing of nodes in the maze can be made irregular (as in some of the mazes used in the 'paper-and-pencil' study) and a series of marks or 'footprints' can be placed in those paths traversed by subjects. The updated version of the maze game program was written by Mr. S. Oliphant, of the Department of Psychology at the University of Glasgow. Enquiries regarding the program can be sent to:

Mr. Stuart Oliphant
Department of Psychology
University of Glasgow
56 Hillhead Street
Glasgow G12