Cognition, 27 (1987) 181–218

Saying what you mean in dialogue: A study in conceptual and semantic co-ordination*

SIMON GARROD
Glasgow University

ANTHONY ANDERSON
Strathclyde University

Abstract

This paper explores how conversants co-ordinate their use and interpretation of language in a restricted context. It revolves around the analysis of the spatial descriptions which emerge during the course of 56 dialogues, elicited in the laboratory using a specially designed computer maze game.

Two types of analysis are reported. The first is a semantic analysis of the various types of description, which indicates how pairs of speakers develop different language schemes associated with different mental models of the maze configuration. The second analysis concerns how the communicants co-ordinate in developing their description schemes.

The results from this study would suggest that language processing in dialogue may be governed by local principles of interaction which have received little attention in the psychological and linguistic literature to date.

Introduction

In both psychology and linguistics there is a long tradition of treating speakers and listeners as isolated individuals from a processing point of view. Yet, on most occasions when we speak or listen we do so within the broader interactional framework of dialogue, where a major goal of both parties is to achieve mutual intelligibility. To communicate effectively, speaker and listener must co-ordinate their respective use and interpretation of the language, within the

*This work was supported by a grant from the Economic and Social Research Council of Great Britain. We would also like to acknowledge the help of Jim Mullin in writing the maze game computer program and the helpful suggestions from Phil Johnson-Laird, Tony Sanford and two anonymous referees who commented on an earlier draft of this paper. Requests for reprints should be addressed to Simon Garrod, Department of Psychology, University of Glasgow, Glasgow, United Kingdom.

context of that particular exchange. They need to establish that they share the same overall conception of what is being discussed and agree upon how each utterance should be interpreted with respect to this conception. In Wittgensteinian terminology, the communicants must ensure that they are both playing the same language game.

This paper examines how people construct such a common conceptual and semantic system during the course of a particular type of dialogue. The dialogues which we consider emerge from a co-operative computer game designed to elicit spontaneous conversation but in a very restricted setting. In the game, two players collaborate in solving a joint maze which is presented in such a way that each player only has partial information about where his partner is located. This means that the dialogues contain a number of location descriptions from both players and it is the analysis of how these descriptions develop which forms the basis of the paper.

The paper divides into four sections. The first consists in a brief discussion of meaning and co-ordination and considers the argument that the meanings for expressions in any particular language are fixed through a system of conventions. However, it will be suggested that global conventions of this sort constitute just one way in which communicators solve the problem of semantic co-ordination and may be insufficient to account for the degree of shared understanding needed to explain successful communication. In the second section, we report the dialogue experiment and show how it enables one to track the various ways in which the players describe locations in their shared perceptual environment. On the basis of this study, two more specific analyses are described in Sections 3 and 4. The first is a semantic analysis of the descriptions, which examines how expressions in the language can be mapped directly into different types of mental model of the spatial domain and suggests that both the models and specific rules for mapping expressions into the models are often established locally through collaboration between the participants. The final section identifies certain mechanisms through which this conceptual and semantic co-ordination may come about which take advantage of the intrinsically interactive nature of dialogue.

1. Meaning, co-ordination and convention

The most convincing account for how meanings come to be established within any linguistic community rests on a view of natural languages as conventional systems, whereby the essentially arbitrary relationship between signs and their meanings is fixed within the community of speakers through conventions of use. However, as Lewis (1968, 1975) has so elegantly demonstrated, conventions are not chipped in stone but depend upon co-ordination of both action and belief among members of the population supported by mutual interest.

Lewis argued that conventions only arise in situations where a community faces a recurrent problem in co-ordinated action, and he suggested that conformity to conventions was one effective means of overcoming such problems. More recently, Clark and his colleagues (Clark, 1985; Clark & Marshall, 1981; Clark & Wilkes-Gibbs, 1986) have argued that many other aspects of human communication can also be analysed in relation to such problems of co-ordinated action and belief.

Co-ordination problems arise in decision-making situations where there is an inter-dependence between participants such that they only benefit when they all make a particular conjoint decision. For instance, if your telephone connection is broken in mid-conversation you will have to decide whether to call back or wait for your conversant to do so. This represents a co-ordination problem, since you only both benefit when one party calls and the other waits. The problem is that you have to make your decision on the basis of what you expect your partner to do, but this expectation will in turn depend upon a higher-order expectation about what your partner might reasonably expect you to do, which in turn will reflect back upon your expectations about their expectations of your choice of action, and so on, ad infinitum.

The effective solution to co-ordination problems therefore depends upon establishing a degree of mutual knowledge[1] of all the participants' expectations, but this cannot reasonably occur by independently verifying each of the infinite chain of higher-order expectations (Clark & Marshall, 1981; Lewis, 1968; Schiffer, 1972). To solve such problems one has to discover good grounds for mutual knowledge; that is establish a premiss or set of premisses which support all the relevant higher-order expectations. Of course, the most straightforward solution is for all to sit down and negotiate a common course of action. But, as in the telephone example, this is often not practicable. So how do we manage to solve such problems? A number of heuristic solutions have been suggested. Schelling (1960) argued that people tend to choose the most salient option, on the grounds that it is the one which everyone would expect the others to choose. Schiffer (1972) suggested that in recurrent situations for which there is a mutually known precedent we should opt for the

---

[1]Mutual Knowledge or Common Knowledge, as Lewis (1968) describes it, can be most readily defined according to the recursive formulation (from Harman, 1977):

A & B mutually know that p
iff
(q) A & B know that p and that q.

In other words, when they both know that p, and both know that they both know that p, and both know that they both know that they both know that p, and so on ad infinitum.

previous course of action. So, in the telephone example you might only call back if that is what you have tended to do before with that conversant. Finally, there is the conventional choice which Lewis (1968) argued takes into account both of these heuristics. The idea is that for regular recurrent situations, any initial precedent will tend to be repeatedly followed by all participants, and as this occurs it will increasingly become the most salient choice. At this point, the regularity of choice will itself come to justify conformity on the grounds that it is in everyone's best interest to conform to the convention and hence establish a general solution to that co-ordination problem within their community. Thus in some communities that suffer from unreliable telephone systems, there is a convention that the one who initiated the call calls back after a breakdown.

Whenever we use language we are faced with such a co-ordination problem. As speakers we have to select expressions to convey what is intended and as listeners we have to select interpretations for those expressions in the hope that they capture the intended meaning. However, it is generally assumed that this semantic co-ordination problem is solved by conventions of language use. Thus Lewis (1968) suggested that any particular language (i.e., function which assigns meanings to linguistic strings) is used by, or is the language of some community by virtue of a convention of truthfulness and trust in the language which holds in that community. By sticking to these conventions, it was assumed that speakers and listeners solve the semantic co-ordination problem, since the meanings of expressions become fixed through conventional use. However, the question remains as to whether simply abiding by general meaning conventions (i.e., those held by the whole population of speakers of that language) is sufficient to account for the degree of semantic co-ordination actually achieved in everyday conversation.

Two outstanding problems exist. The first, identified by Clark and Marshall (1981), relates to reference. As they point out, successful reference seems to require a degree of conceptual co-ordination between communicator and audience which goes beyond that given by conventions of language use. They argue that to produce a felicitous definite reference the speaker must ensure that, at the time of utterance, the relevant facts about the context are mutually known to both speaker and addressee. In other words, communicants cannot simply rely upon their own isolated appraisal of the context of utterance, but must take pains to establish a co-ordinated mutual knowledge of that context.

The remaining problem, recognised by Lewis (1968), concerns the inherent ambiguity and vagueness associated with any natural language. Lewis's general thesis, that usage conventions solve semantic co-ordination problems, depended upon arguments about conventional signalling systems where there is a one-to-one correspondence between expressions and their specific meanings. Yet we manage to communicate effectively with natural languages where ambiguity and vagueness are the rule rather than the exception. In other words, language users achieve a high degree of semantic co-ordination despite the fact that they rely upon languages which effectively underspecify the particular distinctions which they may wish to express.

To the extent that these issues have been considered, it is generally assumed that speaker and listener overcome the problems by inferring the most likely interpretation given their own assessment of the context, which may of course include judgements about the interlocutor's state of mind (e.g., Clark & Marshall, 1981). In effect, such an account assumes that we rely on Schelling's salience heuristic to solve the co-ordination problem. However, it would also seem likely that language users have developed processes which take advantage of the inherently interactive and collaborative nature of dialogue to reduce the complexity of the inferences required for truly co-ordinated understanding in everyday conversation.

In this paper, we will argue that semantic and conceptual co-ordination is often achieved locally within any particular dialogue through a collaborative effort by both parties. In support of this claim, we present evidence to indicate that dialogue participants adopt idiosyncratic 'languages' in Lewis's sense, which depend as much upon local and transient conventions, set up during the course of the dialogue, as they do on the more stable conventions of the larger linguistic community.

In the next section, the dialogue experiment is reported and a preliminary analysis of the location descriptions which emerge is given.

2. The elicitation and preliminary analysis of the dialogues

2.1. The maze game

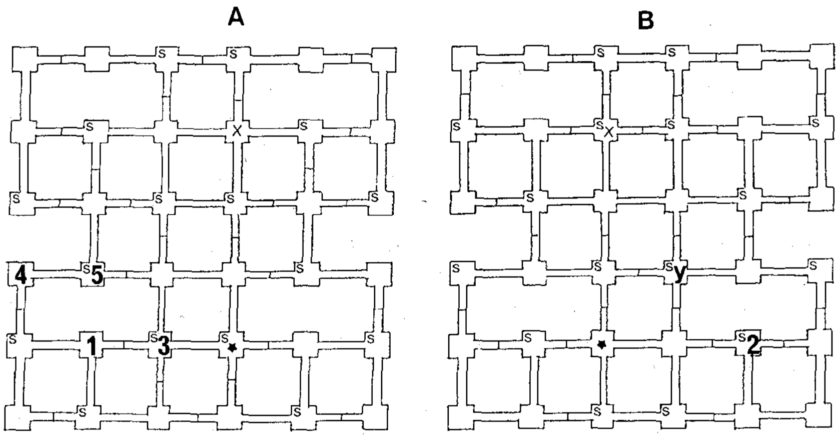
The maze game was designed to elicit natural dialogues containing spontaneously generated descriptions of locations within a predefined spatial network, where the exact positions described could be independently verified by the experimenter.

The essence of the game was as follows. Each player was seated in a different room confronted with a VDU on which a maze was displayed. The mazes consisted of small box-like structures connected by paths along which the players could move position markers (see Figure 1).

The purpose of the game was for the players to move the position markers through the maze (one path link at a time) until they had both reached their respective goal positions. Furthermore, each player could only see his own start position, goal and current position marker.

The co-operative nature of the game arises from two additional features

Figure 1. *Examples of the two mazes for players A and B at the beginning of a typical game. Key: X = player's position, S = switch positions, bar across path = gate. The various letters and numbers refer to positions described by dialogues given in the Appendix.*



of the mazes. First, each maze contained obstacles, in the form of gates, which blocked movement along the paths where they were positioned (see Figure 1). Secondly, there were certain nodes which were marked as switch positions and, like the gates, these were distributed differently for each player. It was in overcoming the obstacles that verbal co-operation was required, since the fundamental principle of the game was as follows. If a given player (say A) moved into one of the switches marked on the other's (B's) screen, then the entire configuration of B's gates would change. All paths that were previously gated would be opened and all those previously open would be gated. Therefore when a player required the gates to be changed, they would have to enlist the co-operation of the other player, find out where he was located and then guide him into a switch node only visible on their own screen.

A further variation in the game came from the optional presence of a maze monster, a computer controlled semi-intelligent third player programmed to pursue one of the subjects (whichever happened to be nearest the monster at the time it moved). If the monster, marked by the letter M on each screen, succeeded in occupying the same node as one of the players, then the player was considered to have been eaten and the game terminated. In practice, the monster proved very difficult for inexperienced players to beat and was useful in eliciting a greater degree of collaboration between the players.

Typically, a game would therefore consist in players attempting to move towards their respective goals with dialogue intervening between moves. The dialogue would contain descriptions of the players' current positions in the maze, switch node locations and goal positions with each speaker's contribution to the dialogue recorded on a separate channel of a four track tape recorder. The remaining two channels were used to record tones triggered by the computer whenever a move occurred or the maze configuration was changed. Thus it was possible to analyse the descriptions against a matching record of the player's positions, gates, switches and so on, which was printed out at the end of each session.

2.2. Design of the study

The corpus of dialogues came from transcriptions of 56 games played by same-sex pairs of undergraduates from Glasgow University. A total of 29 dyads were tested in sessions lasting about an hour. Of these 29 dyads, 21 played two games, 3 completed three games and the remaining 5 pairs only managed to complete one game in the time allotted. It had originally been intended that all pairs should play two games each; however, in the case where three games were played the second one was terminated early because of the monster and so an additional game had to be played in order to elicit a reasonable amount of dialogue.

The corpus came from four types of game:

(1) *Baseline games.* A total of 8 dyads (4 male 4 female) played a set of games in the following sequence. First a practice game was played on a small maze configured as a 4 × 4 matrix of nodes, followed by two games on larger mazes based on a 6 × 6 matrix of possible nodes. None of these games included the monster and all were played on essentially square-shaped mazes symmetrical about the vertical axis (see Figure 1).

(2) *Monster games.* A total of 10 dyads (4 male, 6 female) played two games; one practice game was played on the 4 × 4 maze followed by one game on a larger maze as in (1), but in the second game the monster was present.

(3) *Maze shape manipulation games.* A total of 8 dyads (3 male, 5 female) played 14 games using mazes based on the larger 6 × 6 matrix of nodes, but which were designed to be either asymmetrical or symmetrical about the horizontal axis (see Figure 2). Since this manipulation was included to examine the extent to which maze shape might influence the type of descriptions used no practice game was given.

(4) *Reassignment games.* 2 female and 1 male dyad were formed by taking individual subjects who had played in previous games (one month earlier)

and reassigning them to new partners who had predominantly used a different type of description in the earlier game.

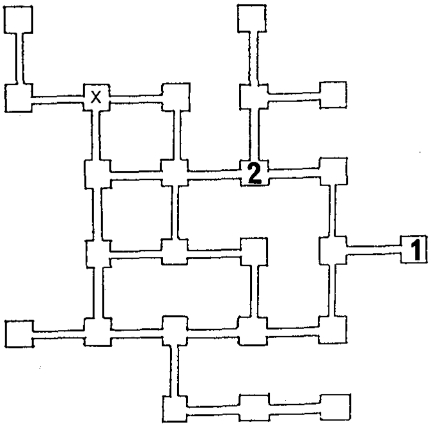
In the present paper, our main interest is in looking at a large corpus of different descriptions rather than considering the specific results of the manipulations in (2), (3) and (4), which have been discussed elsewhere (Anderson, 1983; Anderson & Garrod, 1987). We therefore treat dialogues from all the various types of games as part of the same corpus for the purpose of most of the analyses to be reported.

2.3. Preliminary analysis of the dialogues

When all of the dialogues had been transcribed, each location description was extracted and classified according to its position within the dialogue and the speaker whose location was being described. It was then recorded in the sequence of descriptions from that dialogue against a specification of the actual position in the maze. A complete location description was taken to include everything from the initial request (where that existed) right through to any subsequent exchange which was directly relevant to the description.

Taking the 56 games played, this yielded a corpus of 1396 descriptions, produced at the average rate of 24 per game. Informal analysis of the 56 description sequences revealed a wide range of different types of description and great flexibility in the use of terms to describe components of the maze both between dialogues and, on occasion, across different sections of the same dialogue. For instance, while one pair of speakers might reserve the

Figure 2. *Example of an asymmetrical maze configuration. The numbers refer to positions described in the dialogues in the Appendix.*



term 'row', 'line', or 'column' exclusively for describing vertical arrangements of nodes in the maze, another pair might use an identical expression just for horizontal arrangements of nodes. Nevertheless, against this background of variation it was apparent that pairs of speakers were very consistent in their choice of description within any section of a dialogue. Thus if one speaker described where he or she was by saying:

(1) Third row two along.

the other would typically produce a subsequent description like:

(2) Second row three along.

It was therefore decided to carry out a preliminary classification of the various descriptions in order to see how they were distributed across the different dialogues and try to establish the degree of entrainment between speakers as to choice of description type within any dialogue sequence. This analysis will be presented first by giving an overview of the range of descriptions encountered across all the dialogues and then a more detailed analysis of how such descriptions seem to be clustered within any particular dialogue sequence.

2.4. Classification of description types and their distribution

The classification scheme which we used was a pragmatic one, based on an analysis of systematic differences in the cognitive operations which seemed to be required to produce or interpret descriptions of each type.[2] The rationale behind the analysis will be discussed in detail in the next section, for the present we will just outline the four basic types of description scheme which seemed to emerge and indicate how descriptions of each type were distributed across the dialogues.

In Appendix 1b the description sequences from four dialogues are shown, each illustrating one of the schemes adopted by our subjects. The most common scheme (exemplified by dyad 27B) was what we will call *path* description, where the listener is invited on a tour over the paths in the maze whose destination is the point to be described. Such a description of the node marked 'y' in Figure 1 is illustrated below:

(3) See the bottom right, go two along and two up. That's where I am.

This type of description has often been reported in tasks where people have to describe whole network structures or even layouts of their apartments

[2]This analysis was carried out with the principal aim of constructing a computational model for generating and interpreting descriptions given a representation of the maze. However, the computer model will not be discussed here.

(Levelt, 1982; Linde & Labov, 1972) and it accounted for 36.8% of the corpus.

The next most common type of description (dyad 20B) is quite different from this, in that it depends upon establishing a co-ordinate scheme for the nodes in the maze. An example of a *co-ordinate* description of point 'y' in Figure 1 would be:

(4) I'm on the third row and fourth column.

However, it was more common to see *co-ordinate* descriptions which used an abstract code, as in (5) (see also the example in Appendix 1b):

(5) I'm at C 4.

Where the letters A–F are used to designate vertical lines of nodes and the numbers 1–5 horizontal lines. 23.4% of all descriptions were of this basic *co-ordinate* type.

Co-ordinate descriptions depend upon imposing a structure of vertical and horizontal lines on the maze and then indicating a position as being at the intersection of any two lines. But it is also possible to indicate a position by first describing a particular line of nodes and then describing the position relative to this line. Three examples of this type of description for point 'y' are shown below:

(6) Third bottom line, third box from the right.
(7) Third column from the right, two from the bottom.
(8) I'm one up on the diagonal from the bottom left to top right.

This type of description we will call a *line* description (exemplified by dyad 10G), and it characterises 22.5% of all the descriptions in the corpus.

The final description scheme is illustrated by dyad 40B in the appendix and is what we will call *figural* description since it depends upon first identifying some particular configuration of nodes (e.g., 'a square', 'T shape' or 'a limb sticking out to the side') and then indicating the position with respect to a decomposition of the figure. For instance point 'y' in Figure 1 might be described in the following fashion:

(9) See the rectangle at the bottom right, I'm in the top left-hand corner.

17.3% of all descriptions could be classified as figural.

On the basis of this simple pragmatic analysis, it was possible to classify each location description into one of these four types and calculate for each partner in any dialogue the relative proportion of descriptions in each category.[3] These data are shown in Table 1 for all the dialogue pairs. The data

[3]Notes on the classification criteria and procedure are given in Appendix 1a.

indicate that the various types of description are distributed quite unevenly across the dialogues. For instance, almost half of the 56 dialogues only contain descriptions of two types. At the same time the overall distribution of description types is generally very similar for both speakers in any dialogue.

A simple, if rather crude, way of determining the degree to which speakers in any dialogue adopt the same basic description schemes is to compare the rank orders of the proportions of different types of descriptions used by different speakers. If there was no relationship between two speakers, the expected sum of absolute difference in ranks across all four categories would be 5 whereas with complete agreement it would of course be 0. Using this measure we can obtain an estimate of the variability in the pattern of description choice both within any dialogue and across different dialogues. To establish the across dialogue baseline, one speaker was randomly selected from each of the dyads who played more than one game together and the rank order of description types compared against all those in the rest of the sample. This yielded 190 comparisons for the first game played with a mean difference of 4.1 and 190 for the second game with a mean difference of 4.6, which would suggest that there is little relationship between the patterns of description choice across different dialogues. Against this baseline the mean of the differences within dialogues was 1.87 for the first game and 1.6 for the second with only 1 out of the 40 comparisons yielding differences greater than 4.1 ($p < .001$, sign test). These results indicate that there is a general entrainment of description type between pairs of speakers in the same dialogue.

It is also possible to compare this measure of entrainment between dialogues associated with the first game played versus the second game played by any pair of subjects. There were 20 pairs of subjects for which such a comparison is possible. If we consider only those pairs who have not demonstrated maximal entrainment in the first game, then 10 of the remaining 15 pairs demonstrate greater entrainment in game 2, 3 remain the same and only 2 demonstrate less entrainment ($p < .05$, sign test). Furthermore, of the 20 pairs 25% show maximal entrainment on the first game but by the second game 33% yield no difference in the rank ordering.

This analysis therefore confirms our observation, which is suggested by the examples in the Appendix, that speakers within the same dialogue are likely to be using the same basic types of description as each other. It also indicates that the entrainment is progressive, increasing as the dialogue proceeds. One explanation for this progressive entrainment is that dialogue partners are co-ordinating in some way to achieve a single mutually satisfactory type of description. There is some evidence to support this contention. If one considers the proportion of games where both players use the same description type predominantly (i.e., it represents the most likely description for each player),

Table 1. *Distribution of the various types of description across speakers and dialogues*

	Description type (%)			
Dyad code	Path	Line	Figural	Co-ordinate
10B A	33	66		
B	29	71		
10G A	33	66		
B		100		
11D A	57	43		
B	60	40		
11F A	38	38	24	
B	14	57	29	
13A A	50	50		
B		100		
13C A	46	54		
B		90	10	
15C A	33		50	17
B	33		66	
15E A	83		17	
B	37	31	31	
17D A	23	54	23	
B	44	48	6	
17F A	11	89		
B	25	75		
18D A	19	81		
B	37	58	5	
18F A	30	60	10	
B	11	89		
19F A	30			70
B	47		21	32
20B A	5			95
B	25			75
20D A			12	88
B	50			50
23A A	36	29	21	14
B		29	42	29
23B A	5	5	10	80
B	5		15	80
24B A	62	29	9	
B	71	16	10	3

		Description type (%)			
Dyad Code		Path	Line	Figural	Co-ordinate
25A	A	43		57	
	B	83		17	
25B	A	5			95
	B	31	6		63
26B	A	9		18	73
	B	11		6	83
27A	A	88		12	
	B	71		29	
27B	A	77		23	
	B	75		25	
27D	A	17			83
	B	18			82
28A	A	37		42	21
	B	44	12	32	12
28B	A	33			66
	B				100
28C	A	23			77
	B	4			96
29A	A	38	38	8	16
	B	57	28		15
29B	A	69	5	26	
	B	60		20	20
31A	A	15	23	31	31
	B			100	
31B	A	66			33
	B				100
31C	A	18	64		36
	B	62	12		26
32A	A	43		57	
	B	40	20	40	
32B	A	57	30	13	
	B	52	37	11	
33A	A	30		70	
	B	40		60	
34A	A	12	22	25	41
	B	13	26	13	48
35A	A	100			
	B	25		75	

Table 1. *(continued)*

Dyad code		Description type (%)			
		Path	Line	Figural	Co-ordinate
35B	A	84		16	
	B	77		23	
35C	A	60		40	
	B	85	15		
36A	A	36	64		
	B	38	50	6	6
37A	A	100			
	B	70	30		
37B	A	42	50	8	
	B	36	57		7
38A	A	85		15	
	B	75		19	6
38B	A	64	36		
	B	73	27		
39A	A	31	15	46	8
	B	28	28	28	16
39B	A	62	3	32	3
	B	67	5	28	
40A	A	25		75	
	B	47	18	35	
40B	A	16	16	68	
	B	39	19	42	
41A	A	27	36	36	
	B	46	31	23	
41B	A	11	78	11	
	B	20	60		20
42A	A	18		27	55
	B	14		7	79
42B	A	13		38	49
	B	44		19	37
43A	A	22	78		
	B		85	15	
43B	A	37	12	25	25
	B	20	40		40
44A	A	12	18		70
	B	11	21		68
44B	A	22		28	50
	B	12		12	76

A, B refer to different speakers in the same dialogue. The data from pairs of speakers playi different games are demarcated by a line space. Dyads 42, 43 and 44 are reassignment games.

then, in the first game played, 10 out of the 20 dyads share a predominant description type, while by the second game 19 of the 20 are using the same description type predominantly (Cochran's $Q_1 = 14.72$, $p < .001$). On the basis of the distribution of predominant description types across all these dialogues (see Table 2) it is possible to compare these results against the chance expectation of any two speakers sharing the same predominant description type in any game, which is 28% for the first game and 29% for the second.

Finally, it is interesting to see how the predominant type of description used by any speaker in the first game may not be the same as that adopted in the second. These data are shown in the matrix in Table 2 where the rows indicate the number of speakers who predominantly used a certain description type in the first game and the columns indicate the same data for the second game. Hence any cell of the matrix '$x_{i,j}$' represents the number of speakers for whom description type 'i' was predominantly used in game 1 and description type 'j' in game 2.

Looking first at the distribution of predominant description types one can observe that for the first game (row totals) path descriptions predominate for most speakers (path = 14) while figural and line types of description come next (figural = 9, line = 9), with very few speakers predominantly using co-ordinate descriptions. However, by the second game the situation has changed. Looking at the column totals, it can be seen that only two speakers predominantly use figural descriptions, while there is a substantial increase in the predominant use of line and co-ordinate descriptions. Furthermore, there is quite an interesting pattern of transfer from one type of description to another across the games. This data is represented schematically in Figure 3, which is a directed graph showing the proportion of cases where a speaker

Table 2. *Matrix indicating the predominant choice of description type in the first and second games played*

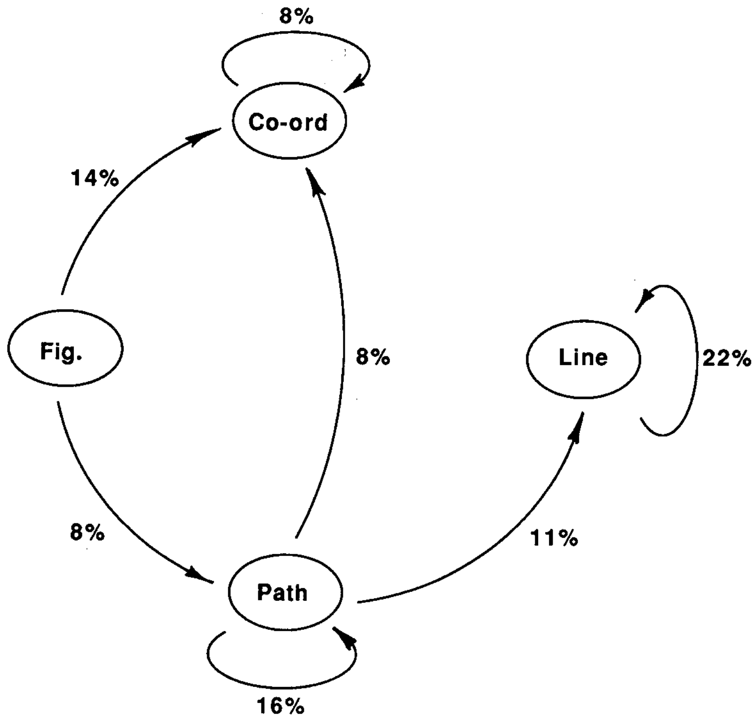
	Second game				
First game	Path	Co-ordinate	Line	Figural	Total
Path	6	3	4	1	*14*
Co-ordinate	1	3	0	0	*4*
Line	1	0	8	0	*9*
Figural	3	5	0	1	*9*
Total	*11*	*11*	*12*	*2*	*36*[a]

[a] Four of the players did not show a single predominant description in one of the games so they were dropped from this analysis.

transfers from one type of description to another across games. It will be seei that while line descriptions seem to be relatively stable there is a genera trend to transfer from figural to path and co-ordinate descriptions as th dialogue proceeds, and a trend to go from path to line type descriptions Speakers seem to initially concentrate on descriptions which rely on percep tual salience (i.e., figural and path) but then move onto descriptions base on more abstract schemes (i.e., line and co-ordinate).

The results from this preliminary analysis of the description sequence therefore suggest that speakers co-ordinate to establish a mutually acceptabl form of description and that this process continues over some time, as th dialogue proceeds from one game to the next. They also indicate that certai types of description are more stable across extended dialogues whereas othe

Figure 3. *Directed graph indicating the proportion of speakers who transferred fro one predominant description type to another across the two games playe (only proportions above 5% are included). The proportions who pr dominantly used each type for game 1 and game 2 are: Path (1) 39% (2) 31% Figural (1) 25% (2) 5%, Line (1) 25% (2) 33% and Co-ordinate (1) 11% (2) 31%.*



tend to be replaced according to a systematic pattern.

However, before examining how such a progressive co-ordination comes about and why one might expect this kind of progression, we need to consider in more detail what it is that differentiates each of these types of description from a processing point of view. In the next section, it will be suggested that the various types of description rely upon quite different conceptualisations of the maze being described, and that the most parsimonious semantic analysis is one which treats each component expression within any description as having an extension within some particular mental model of the maze configuration.

3. Semantic analysis of the various description types

Over the last few years a number of cognitive theorists have promoted the idea that the working mental representations underlying a wide variety of intellectual activities function as models of relevant aspects of the activity domain (Craik, 1943; Gentner & Stevens, 1983; Johnson-Laird, 1983). Such models capture crucial structural relationships between the entities in the domain, where the structure constrains the mental operations that may be performed (Johnson-Laird, 1983). Hence a mental model represents a bounded set of relations between elements in reality. At once, it makes the world intelligible and limits our conceptions of it.

Furthermore, different mental models of spatial domains seem to have different consequences for the interpretation of spatial terms. For instance, with certain relational expressions like 'left of', 'right of' or 'in front', 'behind' the inferences which can be drawn depend upon the nature of the spatial model against which we interpret them. When used according to a deictic frame of reference, these relations are transitive, but when used within an intrinsic frame of reference, which depends upon orientation relative to a model of the object domain, transitivity is not guaranteed, but depends upon the nature of the model (Levelt, 1984). Thus, in a model where the points A. B and C are arrayed on a straight line, if A is to the right of B, and B to the right of C we can correctly infer that A is to the right of C. However, in a model where the three points lie in a circle (say, they represent three people seated around a table) this inference is no longer sanctioned (Johnson-Laird, 1983).

In relation to the maze game dialogues discussed here, it therefore seems reasonable to think of the various different description schemes as depending upon rather different models of the maze, each capturing a rather different overall conception of its spatial and functional organisation. Although each maze was simply made up of a series of nodes connected by paths, the overall configuration can be thought of in a number of different ways. For instance one way of conceiving of the maze is as a series of links between neighbourin nodes, with each link corresponding to a path in a network. Alternatively, i can be thought of as a set of parallel lines arranged in the vertical or horizon tal plane, that is a collection of rows or columns, or it might be seen as collection of discrete, possibly overlapping patterns, which together make u the maze as a whole. Thus, it can be decomposed into T-shapes, squares an limbs that stick out to the side.

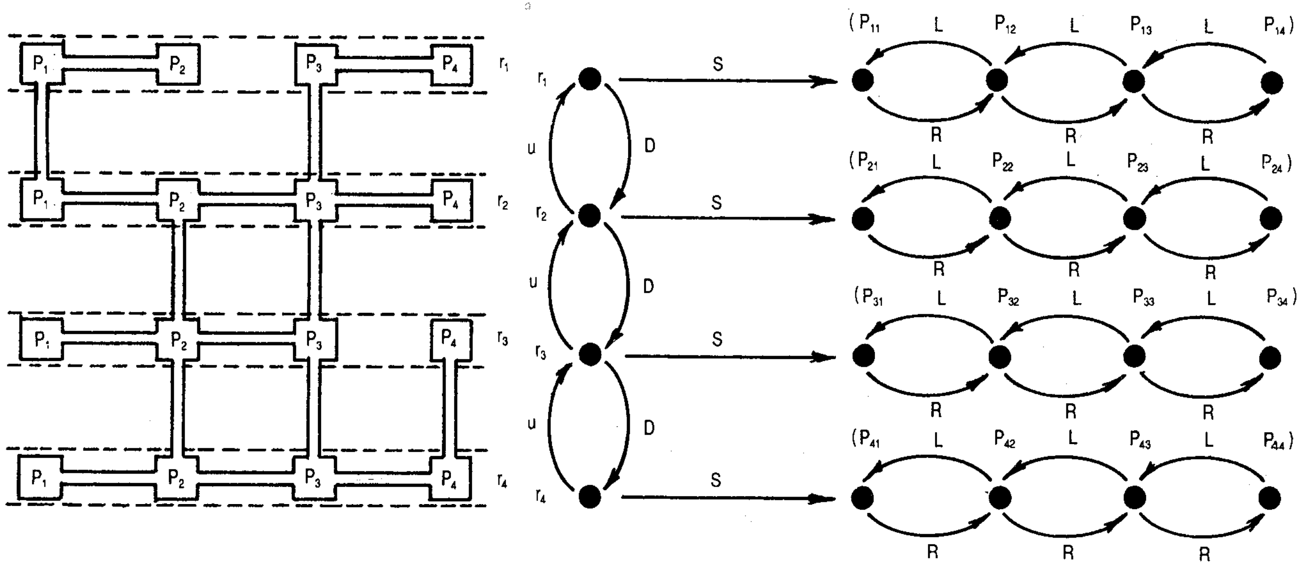
Each such mental configuration corresponds to a different spatial mode of the maze, where each model enables one to carry out simple operation which may then be projected onto the physical configuration of the maz itself. Hence a model can be used as the basis for producing or interpretin descriptions of positions on the maze. To illustrate this, we will consider ho the various description schemes used in the maze game dialogues relate t quite different types of spatial model. It will then be shown how the model impose restrictions on each type of description, both in terms of which point in the maze are easy to describe relative to others and also how the descrip tions may be developed to become more elliptical as the dialogue proceed At the same time, it will become apparent that each dialogue descriptio sequence represents some particular, often idiosyncratic, way of construin the language in that context.

We will begin with the dialogues which use the *line* type of descriptio discussed in the previous section, considering these in some detail. Dialogue illustrating the other description schemes can then be discussed more brief in the light of this analysis.

In giving a line type of description the speaker first indicates a line of nod on which his or her position lies and then describes the position with respe to this line. As was noted in the previous section the line referred to can b oriented in one of three ways, horizontally (e.g., in example (6) above vertically (7) or even diagonally (8), though they are by no means even distributed (75% of such descriptions refer to horizontals, while 22% ref to verticals and only 3% to diagonals). We will concentrate here on describin the model underlying the first type.

In Figure 4, a schematic representation of a horizontal type of line mod is shown alongside the maze being modelled. The model is represented terms of a domain of entities, which can be of two types (1) simple positic elements $p_{i,j}$ which would correspond to possible nodes in the maze, and (complex 'row' elements r_n which correspond to rows or sets of nodes array horizontally. The model also contains a set of spatial relations which captu the configurational properties of the various elements. Thus each model el ment is locally related to its adjacent elements by directional properties. A

Figure 4. *A schematic representation of a simple horizontal line model (right of the figure) shown next to a superimposition of the model over a notional maze (left). Key: r = row, u = up, D = down, L = left, R = right, S = subsumes.*



p elements are related horizontally while all r elements are related vertically. This means that any entity in the model can be defined purely in terms of its relationship to the other entities, thus r_2 can be defined as "above r_3" and as "below r_1" and so on. Alternatively, the same relationships can be captured procedurally in terms of instructions which would enable a "scanning" device to move from one element to another in the model. In this case r_2 could be defined as that element in the model which is arrived at after first "scanning" upwards until no further elements are discovered and then "scanning" down past one element—find the topmost element and then go down one.

In order for such a model to support the generation and interpretation of location descriptions, the speaker and listener will have to establish a restricted language fragment which maps onto the model. For instance, they might introduce some particular term like "row", "line", "level", "layer" or even "floor" to designate the set of r elements in the model. Similarly, complex expressions like "from the bottom" in the description "second row from the bottom" can be taken as denoting a function yielding the ordered set of r elements that emerge when you scan from the bottom to the top. However, it becomes apparent when one considers the exact form of the various types of horizontal line description that there are many subtly different ways of interpreting expressions with respect to such a model.

The prototypical description of this sort is one where the line is described in terms of its position relative to another line at the top or bottom of the maze (e.g., "the third row from the bottom"). However, there are a number of variants on this basic format which we believe throw some light on how a restricted 'language' may be mapped onto a conceptual model of the domain. In particular, we would suggest that descriptions such as "the second row", "the second bottom row" or "row two" which occur throughout the dialogues may reflect either differences in the way speakers are conceptualising the maze configuration or differences in the way they map expressions in the language onto interpretations with respect to this conceptual model.

Table 3 identifies four such horizontal line description sub-schemes. On the left is shown the various ways in which a row in the middle of the maze may be described and on the right is shown the distribution of different descriptions of end rows (those at the top or bottom) which co-occur in the dialogues.

What is most striking about this distribution is how it illustrates that choosing some particular form of description for a row in the middle of the maze seems to constrain the co-occurring descriptions of end rows. For instance, while it is perfectly acceptable to describe the bottom row as just that when producing descriptions for middle rows of the form "the second row from the bottom", (see sub-scheme 1 in the table) this does not seem to be the case

Table 3. *Distribution of the various types of description of rows at the top or bottom of a maze as a function of description schemes defined by descriptions of rows in the middle. Each line of figures shows the distribution within a single dialogue*

Scheme		Frequency per dialogue of each type of description of top or bottom row					
Sub-scheme 1 (7 pairs)		A	B	C	D	E	F[a]
Middle row description:[b]		4					
"the *n*th row from the top/bottom"		2					
		6					
		5					
		2					
		4					
		1					
	Total:	24	–	–	–	–	–
Sub-scheme 2 (2 pairs)							
Middle row description:		1	1				
"the *n*th top/bottom row"			1				
	Total:	1	2	–	–	–	–
Sub-scheme 3 (6 pairs)							
Middle row description:		2		2			
"the *n*th row"		3		1			
		3	1	1			
				1			
					8		
						1	
	Total:	8	1	5	8	1	–
Sub-scheme 4 (1 pairs)							
Middle row description:							
"row *n*"							1
	Total:	–	–	–	–	–	1

[a]Types of top/bottom row description.
A = "the top/bottom row"
B = "the *very* top/bottom row"
C = "the top/bottom *line*"
D = "the *first* row"
E = "the *last* row"
F = "row *one*"
[b]The middle row descriptions indicate the type of description of a middle row used immediately prior to the end row descriptions recorded in the table for that dialogue.

with the more elliptical description "the second row" (sub-scheme 3). Here speakers will typically either choose a description like "the first/last row" which maintains the simple ordinal scheme or introduce a different lexical item such as "line" to yield "the bottom/top line" (see dialogue 10G in the appendix). Similarly, speakers who adopted descriptions like "row two" (sub-scheme 4) always referred to the bottom row as "row one". Perhaps more surprising is sub-scheme 2 which utilises the prenominal forms "top", "bottom" within the description of middle rows (e.g., "The third bottom row"). Here end rows are predominantly described with an intensifier, as in "the *very* bottom row". These distributional regularities would suggest that speakers are adopting particular 'languages' of description in relation to the basic line model, and it is possible to give a reasonable account of these 'languages' if we make a few straightforward assumptions about the form of the underlying model and the semantic structure of the descriptions themselves.

Let us first consider the basic semantic structure of the various descriptions in each sub-scheme and indicate how they can be interpreted against such a model. We will adopt the principles behind Barwise and Cooper's (1981) influential analysis of quantifiers. According to this analysis, all these descriptions can be treated as 'quantifiers' consisting of the quantificational determiners "the" and "second" followed by some set expression "row from the bottom", "bottom row" or just "row". Furthermore, if we want to give a functional compositional analysis of the sort commonly used in model theoretic semantics (Dowty, Wall & Peters, 1981), the ordinal determiner "second" can be characterised as a function from an ordered set or series of elements into the element which is second in that series. In other words, a description like "the second row from the bottom" denotes 'that row which is second in the series of rows as ordered from the bottom to the top'. The same is true of descriptions like "the second bottom row"[4] and under the right circumstances also of the description "the second row". If one accepts this analysis of the descriptions in Level 1–3 of Table 3 it should be clear that the term "row" as used in sub-scheme 1 denotes something different from the same term "row" when used in sub-scheme 3. In the first case, it simply denotes an unordered set of row elements, only given an ordering when qualified by the prepositional phrase "from the bottom". In the second case it must be taken to denote an ordered sequence of row elements, and we would suggest that this sequence is already present in the underlying model of the maze adopted by the speaker. In effect the use of "row" in sub-scheme 3 is similar to the use of "floor" in the description "the second floor"; we do not say "the second floor from the bottom" because there is an agreed con-

[4]Such descriptions were given for rows at the top of the maze, as in "fifth bottom row" used of a row one down from the top of a 6 × 6 maze.

ceptualisation of buildings in which floors are already ordered in this way.[5] A similar problem arises with the interpretation of "bottom" in sub-scheme 2, where it has taken on a restricted meaning in descriptions like "the second *bottom* row" corresponding to something like "ordered from the bottom".

Such an analysis can be used to account for some of the distributional phenomena illustrated in Table 3. Consider, for instance, sub-scheme 3 where speakers commonly introduce an extra term "line" when referring to the bottom or top row in the maze. We would suggest that this is done to clearly differentiate reference to one of the ordered row elements in the model ("rows" in this case) versus reference to unordered row elements ("lines"). If one wants to refer to an end row this can either be done using an ordinal plus a term denoting the ordered set to give the "first/last row", which is an option adopted by two of the dyads, or alternatively use a directional term, "top or bottom", to indicate how to scan across an unordered set of row elements, "lines", which is the option adopted by speakers in four of the dyads. A comparable situation occurs in the case of sub-scheme 2, where we would suggest that the introduction of the intensifier "very" in "the very top row" signals that "top" is not in this case being used according to the restricted interpretation of "the second top row".

Finally we are left with the descriptions in sub-scheme 4, which are probably best treated as complex names made up of a head "row" and an index "one", "two", "three" etc. which serve to designate particular row elements in the line model. Here again the expression "row" is being used in a very restricted manner, quite different from that of "row" in "the bottom row" and for this reason the latter description does not seem to occur in this context.

According to this analysis, *line* type descriptions may depend upon at least three different forms of the line model: one in which the row elements are specified as in Figure 4 and may be ordered either from the top or bottom (underlying descriptions in sub-schemes 1 and 2), a second where the row elements are seriated in only one way (sub-scheme 3) and a kind of nominal model where no spatial relations need be represented (sub-scheme 4).

From this more detailed treatment of the various types of line description two general points emerge. First, there seems to be evidence that dialogue pairs adopt rather specific 'languages' of description, where by 'language' we mean a set of rules for mapping expressions onto interpretations with respect to a common model of the discourse domain. Secondly, the distributional analysis of the descriptions indicates that once speakers have established a

[5]According to this account, if the maze was inverted the "second row" would still be described as that, while it would no longer be "the second row from the bottom".

particular interpretation for an expression in the 'language' they try to avoid any potentially ambiguous use of that expression. In this respect these local 'languages' are similar to Lewis's conventional signalling systems alluded to earlier.

These same principles are also evident in the dialogues reflecting other description schemes, which we will now briefly consider with respect to a mental model type of semantic analysis. Turning first to the *co-ordinate* scheme, the model here is similar to the top level of the line model but containing both a representation of *r* type elements and a representation of *c* or column elements arrayed vertically. This means that the only relations which need to be represented are those between columns and between rows. Hence, any position can be defined as being in the intersection of the two sets of *p* elements in a column and a row. Perhaps more interesting is the way in which the dialogue participants develop simple codes to map onto the model. This is illustrated, for instance, in dyad 20B in the Appendix, where both speakers employ an alphabetic sequence to designate the columns and a numerical one for the rows. In this case, it became apparent from debriefing, that both participants were familiar with the board game "battleships" and had used this common knowledge as the basis for their description scheme.

The remaining two description schemes were *path* and *figural*. Path descriptions seem to depend upon a rather different kind of model from those underlying either line or co-ordinate types, and this model has quite interesting consequences for the relative ease of description of different points on a maze. A path model can be represented as a set of *p* elements related to their neighbours in a network through the actual path connections on the maze. In this respect, path models, like underground maps, do not accurately record the physical distance between locations, instead they represent functional distance, that is distance in terms of the number of neighbouring nodes which need to be traversed to get from one location to another. This means that certain nodes which are physically close to one another may be difficult to describe efficiently within the constraints of such a model. Consider, for instance, the description sequence from dyad 27B in the Appendix, where the points described are shown in Figure 1. The first three descriptions (points 1–3 are marked on Figure 1) all conform to the basic path model scheme, where a speaker starts by identifying a salient point (one of the two bottom corners) and then describes a route through the path network whose destination is the position to be described. Points 4 and 5 however, present problems since they are only connected to a corner by an extremely tortuous route, even though they happen to be physically close to it. One might imagine that the simplest solution would be to ignore the path network and describe point 4 as "two above the bottom left hand corner" and point 5 as "two up and one across", but such descriptions are not interpretable against this model and therefore the speakers have to choose a completely different and somewhat cumbersome solution involving a kind of figural description. These descriptions are shown below:

(point 4) 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 there's a box and then there's another box.
B: Uh-huh.
A: I'm right there.

(point 5) A: I'm one to the right then one up, then there's a gap right.
B: Uh-huh.
A: I'm just in the box above that gap.

Again, this would suggest that the speakers have adopted a particular 'language' interpretable against a particular type of spatial model and this means that they will not suddenly introduce a description in the same language (e.g., in terms of some visual scan description) which violates the locally established rules of interpretation. To this extent, adopting a particular common model seems to constrain the description possibilities available to speakers.

Finally, we come to the *figural* description scheme exemplified by dyad 40B in the Appendix. We would suggest that such descriptions depend upon spatial models where the maze is decomposed into a number of figural subcomponents and that the speakers adopt particular terms to denote these figures. This process is illustrated by dyad 40B who are talking about the maze shown in Figure 2.

Throughout the dialogue the speakers refer to three lines of nodes which stick out on the right-hand side of the maze, and they do so with the term "right indicator". In the first instance, it is introduced by speaker B with the simile "it's like a right indicator", which is then taken as a description of the middle limb. As the dialogue proceeds the description is subsequently extended to cover a "top" "middle" and "bottom" "right indicator". This kind of development is quite characteristic of dialogues containing figural descriptions and is reminiscent of the development of idiosyncratic descriptions reported by Glucksberg, Krauss and Weisberg (1966) in their famous block communication task and more recently by Clark and Wilkes-Gibbs (1986) using a similar task. Again, the adoption of such idiosyncratic languages of description is in line with our general contention that conversants collaborate in some way to establish both a common model of the domain and a local 'language' of interpretation which operates in conjunction with the model.

In describing the more detailed semantics of the various description schemes, we have argued that all the descriptions depend upon an implicitly agreed conceptual model of the maze configuration and a shared 'language'. If this analysis is correct there must be some way in which dialogue pairs can co-ordinate this choice of model and 'language' and it is this which we will discuss in the final section.

4. Co-operative strategies and the output/input co-ordination principle

The problem facing communicants in this situation is clearly one of co-ordination. They must somehow establish mutual knowledge of both conception and 'language'. As was noted in the Introduction, the most sure way of solving such a problem would seem to be for both participants to start out by explicitly negotiating a common conception of the maze and a set of interpretation rules for the descriptions. On occasion, explicit negotiation of this sort does occur in the dialogues, particularly when speakers are trying to adopt a co-ordinate description scheme. However, it does not appear to be either a popular or effective means of achieving co-ordination in practice.

Two observations from the dialogues highlight the problems with explicit negotiation. First, explicit negotiation, when it does occur, only usually happens *after* many descriptions have already been given, and then only after the interlocutors have experienced considerable problems in establishing a mutually satisfactory scheme. For instance, of the 15 games where one can identify explicit negotiation,[6] in 9 cases it occurred only after the dyad had completed one game and in the remainder it occurred after an average of seven descriptions had been produced.

The second observation is that, even in cases where speakers do go to the trouble of negotiating some conceptual/semantic scheme, they very often do not stick to it for long, letting their descriptions wander from the scheme as soon as any problem arises. For instance, in the games where explicit negotiation was observed, the scheme that the speakers negotiated only predicted 59% of the *subsequent* descriptions in that game. Explicitly negotiated semantic plans do not therefore seem to play a major part in determining the

[6]The criterion used for instances of explicit negotiation was the presence of utterances directly concerned with defining the interpretation of terms like 'left', 'row' letters and numerals, or concerned with explicitly characterising the model of the maze like "imagine it as a grid". Of the 15 dialogues containing such utterances, 13 were directed at clarifying co-ordinate schemes and 2 at clarifying line descriptions.

descriptions people actually use in such a task (see Suchman, 1985, for an interesting parallel observation about interactional planning in general). In practice, speakers seem to solve the co-ordination problem in a more flexible and cost effective way.

As Clark and Wilkes-Gibbs (1986) have suggested, a major factor in conversational interaction seems to be the minimization of *collaborative effort*. In other words, the participants should formulate their utterances in such a way that they do not have to spend unnecessary time or effort in ensuring mutual intelligibility. In the context of the dialogues considered here, this can be achieved by following a very simple interactional principle, which we believe may be the basis for much of the co-ordinated activity seen in dialogues in general. We will call this principle *output/input co-ordination*, and it may be simply stated as one of formulating your output (i.e., utterances) according to the same principles of interpretation (i.e., model and semantic rules) as those needed to interpret the most recent relevant input (i.e., utterance from the interlocutor). In effect, such a principle assumes that speakers should be locally consistent with each other, and so long as both speakers abide by this principle, then the chances are that they will quickly establish a mutually satisfactory description scheme with the minimum of collaborative effort.

In dialogue, each participant *must* at some time act as speaker and at other times as listener; each must be capable of both formulating and interpreting utterances on the same dialogue topic. For instance, in the dialogues studied here, a speaker about to formulate a description will almost certainly have encountered a previous description from his partner and in so doing will have had to impose some interpretation upon it. According to the analysis described above, he must select a model of the domain against which the description makes sense and make certain assumptions about interpreting the description with respect to this model. Thus, when formulating his own description, this same scheme should still obtain and form the basis for production. In such a case, the roles of speaker and listener are switched and the original speaker now has to interpret the utterance according to his appraisal of the currently relevant scheme. At this point, any discrepancy will usually become apparent, and if it does the listener has the option of restating the description according to his own assessment of the currently accepted description scheme. Again the roles will be reversed and the comparison process reinstated, and this cycle may be repeated iteratively until both parties have co-ordinated output with input.

Before exploring the communicative utility of such a principle, we will illustrate its application by considering the sequence of descriptions and responses which occur when a speaker violates the principle in one of the maze game dialogues. In this dialogue (dyad 23A), the six initial descriptions from speakers A and B follow the basic horizontal line sub-scheme (version 1) exemplified in descriptions 5 and 6 below:

(5) B: Second row from the top, second box on the right.

(6) A: Third from the left, bottom row.

(7) A: In the third row from the top, *second right.*
(Response) B: *Second box on the right?*

However, in description 7, A introduces a new use of "right" in accordance with description sub-scheme 2, where it is used as a prenominal modifier (it is short for "second right box"). This change violates the output/input co-ordination principle, so B responds with a restatement of that part of the description, but according to the original scheme. A further example of the same sort occurs in A's ninth description, which is shown below:

(8) A: Top row, second from the left.

(9) A: *Third row, at the left.*
(Response) B: You mean *third row from the very top on the very left?*

In 8, A has conformed to sub-scheme 1, but in 9, he tries to introduce a new type of description (sub-scheme 3) and again this instigates a response from B who restates the description according to the original scheme. In fact, from this point on in the dialogue under consideration, all of A's descriptions follow the original scheme and for both speakers output and input are perfectly co-ordinated.

This example serves to illustrate both the advantages and the disadvantages of rigid conformity to the principle. On the positive side, co-ordinating output and input is an efficient means of hunting out the minimal common ground needed to support effective communication. It serves this function in the process of fulfilling the primary purpose of the utterance (e.g., describing locations in this case) without involving a special additional interaction of the sort required when explicitly negotiating a semantic system. Furthermore, it serves the function in such a way that speakers do not have to build up an explicit model of their audience. The model is effectively incorporated in the co-ordinated system. Finally, as a corollary to this, output/input co-ordination has cognitive utility to the extent that it minimises the joint pool of resources which a speaker/listener has to call upon when alternately formulating and interpreting utterances within the same dialogue; when formulating an utterance the speaker only has to refer to the same set of interpretation rules as those needed in understanding one on the same topic.

All of these features would support adopting such a principle as the basis for solving semantic co-ordination problems in dialogues of the sort described here. However, when it is applied rigidly, as in the example shown above, certain limitations become apparent. The most obvious limitation is inflexibility. The speakers in dyad 23A seem to be unable to modify the scheme, which they have established, in order to make it more elliptical. The problem stems from the fact that the principle requires you to be consistent with the *other* speaker. This means that, once co-ordination has been achieved, no modification can be introduced without violating the principle of local consistency.

However, there are ways in which speakers can overcome this problem within the broad constraints of applying the co-ordination principle. One way of doing this is for the conversants to accept a division of labour and control, whereby one takes the role of leader and the other follower (see Lewis, 1978, and his distinction between master and slave). It is then accepted that one participant, the leader, should be the arbiter over the established scheme, while the other participant, the follower, has to abide rigidly by the output/input principle. In this way, one of the participants will adapt to the other, who is free to introduce any modifications in the scheme used by the pair.

Adopting such a modified strategy will enable the orderly development of new language schemes. For instance, consider the description sequence from the beginning of a game played by another pair of subjects:

(1) A: In the second box: em: two from the bottom.
B: Which side?
A: On the—sorry.
B: Two from the bottom?
A: Uh-huh.
B: Second box, two from the bottom ... What do you mean second box?
A: Second ... start at the left.
B: Uh-huh.
A: Second from the left and: two up from the bottom *row*. Where's yours?

In this opening exchange A proffers a description which depends upon a horizontal line model with the positions seriated from left to right. His partner, B, queries this description and eventually forces A to give a fully explicit version in the last utterance, in which the term 'row' is used to designate the horizontal lines. The dialogue continues:

(2) B: Two ... second column from the right and two up.
A: Say that's column four.
B: No it's not that's column five.

(3) A: I'm in the bottom *row (h)*.
B: What column?
A: Three.

(4) B: *Row (v)* five down to the bottom one.

(5) B: Box two, *row (v)* five.

In description (2) it is clear that B is not conforming to the output/input co-ordination principle, but rather introduces a description based upon vertical lines or columns. Furthermore in B's next two descriptions (4 & 5) he uses the same term 'row' to designate the verticals, as indicated by the index 'v' for vertical versus 'h' for horizontal marked next to each occurrence. At this point in the dialogue, A implicitly agrees to follow B's lead and gives his next two descriptions in terms of a vertical line model, using the term 'row' to designate the lines.

(6) A: *Row (v)* one, box ...
B: Three.
A: Three, four I mean. Box four from the top.
B: You're in three ... one, three.
A: One, four.

(7) A: I'm in *row (v)* four, box four.
B: Four, four.

(8) B: Three, three.

(9) A: Five, four.

Finally, in description (8) B introduces a co-ordinate type of description and A follows in (9). In the remaining 54 descriptions from this pair, both speakers conform to this co-ordinate scheme.

What is striking about this dialogue is how co-ordination is only achieved because one speaker, A, is prepared to adapt to the other and not query any of his descriptions, while the other speaker, B, is free to modify the descriptions as he chooses and query any of A's descriptions which do not match his current scheme. So, B questions all of A's descriptions which do not share the same scheme as those assumed in the previous description by B, while none of B's descriptions is questioned. Hence, this pair seem to have adopted the modified output/input co-ordination strategy with one speaker accepted as leader and the other follower.

In fact, close examination of the previous dialogue from dyad 23A also indicates that one speaker is in the leader role and the other follower, but in this case the leader is, so to speak, conservative and uses his role to reject any modification in the description scheme. Thus in 23A, speaker B, as leader, corrects A's descriptions and A, as follower, accepts the corrections.

The second limitation associated with the output/input co-ordination principle, as compared with explicit semantic negotiation, concerns the manner in which it establishes mutual knowledge. The grounds for assuming mutual knowledge in the former case are what might be described as 'falsification definite'; so long as you can interpret the other's utterance using your own scheme, you *presume* that their scheme is in fact identical to yours. However, this may not be the case, particularly when using abstract schemes. For instance, a point in the centre of a five by five matrix of positions can be given the same description 'C 3' according to at least four different versions of a co-ordinate description scheme (i.e., one starting at the top left, another at the bottom right and so on). Hence, even if you know the point being described, the description alone may not be sufficient to identify the particular scheme used, it simply rules out a number of alternatives. As a consequence of this, it might take many descriptions before true co-ordination is achieved when trying to develop an abstract scheme from scratch.

If conversants are relying on the output/input co-ordination principle, one would therefore expect them to be more successful initially with description schemes which depend upon perceptual salience for their interpretation even though these may be less efficient than the more abstract schemes. In fact, just such a tendency was observed in the maze game dialogues (see Section 2) where there was a trend to transfer from the predominant use of figural or path types of description in the initial games to more abstract line or co-ordinate descriptions in later games. Furthermore, explicit negotiation was only observed in relation to the more abstract schemes of description (see footnote 6).

Analysis of the development of particular description schemes therefore gives broad support for the idea that conversants may co-ordinate their 'language' without having to negotiate the underlying semantics or conceptual scheme. The interactive nature of dialogue affords an opportunity to hunt out common ground so long as the participants abide by the principle of co-ordinating their output with the most recent relevant input.

Summary and general conclusions

In this paper we set out to explore the way people co-ordinate their use and interpretation of language within the particular restricted context, produced by the co-operative maze game. For the most part, the paper revolves around analyses of the sequences of location descriptions produced during the course of the game.

The first analysis demonstrated that pairs of speakers adopt very similar forms of description, suggesting a degree of linguistic entrainment between interlocutors. The second analysis indicated that each of the various description schemes relies upon some particular mental model of the maze configuration in conjunction with a set of interpretation rules for expressions in the language. Hence, the schemes represent examples of restricted and unambiguous description 'languages', similar to what Wittgenstein would call language games. It seems that by adopting such local 'languages', speakers are able to achieve a high degree of semantic co-ordination, since the 'languages' are not subject to the inherent ambiguity and vagueness associated with any language which depends upon the general conventions of a larger linguistic community.

In the final section of the paper, we argued that establishing any particular description scheme poses a co-ordination problem for the conversants. However, it was suggested that they did not usually solve the problem through explicit negotiation of the scheme, but followed a general interactive principle of co-ordinating output with input, which represents a good heuristic solution.

The study reported here suggests that general conventions of meaning may serve only as starting points for interpretation, perhaps giving a default meaning which may be overwritten by more local and transient conventions set up during the course of a dialogue. Such a conclusion is consistent with a view that language processing depends as much upon very local semantic considerations as it does upon access to our store of general semantic knowledge.

Appendix

1a: Notes on the classification of the descriptions

Each location description was segmented into sets of descriptions according to the following criteria:

(1) Count as a description unit:

New descriptions i.e., those following several utterances on a different topic, even if incomplete.
Restated rejoinders i.e., restatement by the addressee but in a different format from the original.
Complete restatements by original speaker where different from the original (e.g., includes path descriptions with different origin).

(2) Do *not* count as a description unit:
Repetitions of all or part of the original description proffered.
Procedural clarifications.
Move descriptions.
Descriptions of S boxes.
The second half of a joint effort.

Criteria for classification were both lexical and contextual. For instance, contextual information was used to decide between elliptical path descriptions or line descriptions in cases like the following "I'm third along". The following are some examples of the lexical criteria used:

Path descriptions containing 'along', 'up', 'left', or 'right' as adverbs, specifications of corner start points etc.
Co-ordinate. Numerals, letters, row + column descriptions, distance from two edges.
Line. 'Row', 'column' (alone) 'line', 'level', 'layer', 'floor' etc.
Figural. 'Shape', 'middle', 'limb', 'opposite', 'corner' (without accompanying path description, etc.).

All description units were checked separately by both authors and any ambiguous cases were not included in the analysis. Out of the 1396 description units identified 12 proved unclassifiable.

1b: Examples of place description sequences from the dialogues

Path type. Dyad 27B

(1) *A: Right from the bottom left-hand corner: above the wee purple thing: go along one and up one and that's where I am.

(2) B: Right I'm the same only on the other side.

(3) A: I'm two along, up one now: from the bottom ... from the: left.

B: I'm down one.
A: Oh down one.
B: Uh-huh.
A: What, from where?
B: One along from the right ... I'm one along from the bottom right.

*The numbers refer to positions in Figure 1 (see text).

A: So I'm two along and one up again.
A: I'm one along and one up.
B: From the bottom: what?
A: Left, bottom left.

(4) 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 there's a box and there's another box.
B: Uh-huh.
A: I'm right there.

(5) A: I'm one to the right then one up, then there's a gap right
B: Uh-huh.
A: I'm just in the box above the gap.

Co-ordinate description type (dyad 20B)

A: O.K.? right—er Andy we've got a six by six matrix.
B: Yup.
A: A, B, C, D, E, F.
B: 1, 2, 3, 4, 5, 6.
A: Correct, I'm presently at C5 O.K.

B: E1.

A: I have to get to A, B, B, 1.
B: B1.

A: I take it you have to get to –
B: No.
A: D5, is that correct?
B: Er—A, B, C, D, E: A, B, C, D, E. yeah.

A: So you're now at D1 are you?
B: Uh-huh.

A: And I'm in B5.

Line type description (dyad 10G)

B: I'm on the top line, fourth box along.

A: I'm on the second row, second box along.

B: So I'm fourth box on the top row now.

B: You're on the bottom line second box along, Yeah.
A: Uh-huh.

B: The fourth box on the second row.

A: Second row, first box.

B: Fifth, fifth box fifth row.

B: Fifth box fourth row.

B: Fifth box on the second row.

B: Sixth box on the fourth row.

A: I'm on the second box on the fourth row.

A: That's me on the first box on the fifth row.

Figural description (dyad 40B)
(1)* B: O.K. Stan, let's—let's talk about this. Whereabouts—whereabouts are you?
A: Right: er: I'm: I'm extreme right.
B: Extreme right.
A: Yes, the O- th- there's one er: what d'you call it, there's just one box: to the extreme right.
B: Away out on the right?
A: On my right. I don't know which way you're facing the screen, but I suppose it should be the same.
B: Yeah.
A: You know the extreme right, there's one box.
B: Yeah right, the extreme right it's sticking out like a sore thumb.
A: That's where I am.
B: It's like a *right indicator*?
A: Yes, and where are you?

(2) B: Well I'm er: that *right indicator* you've got.
A: Yes.
B: The *right indicator* above that.
A: Yes.
B: Now if you go along there. You know where the *right indicator* above your's is?

A: Yes.
B: If you go along to the left: I'm in that box which is like: one, two boxes down O.K.
A: One, two boxes down from?
B: From the *top right indicator.*
A: I see.
B: O.K. let's split the screen up, you've got a *right indicator, a right indicator and a right indicator.*
A: Uh-huh, uh-huh, uh-huh.
B: O.K. now th- th- the *bottom right indicator's* got three boxes in it.
A: Yeah.
B: Right O.K.
A: Yeah, yeah.
B: Right O.K. the *bottom right indicator's* got three boxes in it.
A: O.K.
B: O.K. the *top right indicator's* got two boxes in it.
A: Uh-huh.
B: And the box that you're in:
A: Yes.
B: Has got two: in it and you're on the extreme right.
A: Yes.
B: O.K. of the *right indicator.*
A: I see.
B: Of the *middle one* right?
A: O.K. fine right.
B: O.K. so you know where I am.
A: Yes I do.

(3) B: I'm er: well you know how: well it's difficult ehm: you know where the *middle right indicator* is?
A: Yes.
B: Well count that *middle right indicator* as a box.
A: Mm ...
B: Then move to the left, that's one box, two boxes, three boxes, four boxes, five boxes right.
A: Yes.
B: I'm in the fourth box.
A: I'm lost, I'm lost, I'm lost.
B: The *middle right indicator* O.K.?
A: Yes, yes.

B: Move to the left. Count the *middle right indicator*, it's the one, two, three, fourth box along.
A: I see, O.K.

*Indicates the positions described marked on Figure 2.

Key to notation. : = noticeable pause < 1 s.
... = pause > 1 s.

References

Anderson, A. (1983). *Semantic and social-pragmatic aspects of meaning in task oriented dialogue.* Ph.D. dissertation, Glasgow University.

Anderson, A., & Garrod, S. (1987). The dynamics of referential meaning in spontaneous dialogue: some preliminary studies. In R.G. Reilly (ed.), *Communication failure in dialogue and discourse*, 161–183. Amsterdam: Elsevier Science Publishers.

Barwise, J., & Cooper, R. (1981). Generalised quantifiers and natural language. *Linguistics and Philosophy*, *4*, 159–219.

Clark, H.H. (1985). Language use and language users. In G. Lindzey & E. Aronson (ed.), *The handbook of social psychology* (3rd edn.) 179–231. New York: Harper and Row.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, I.A. Sag & B.L. Webber (eds.), *Elements of discourse understanding*, 10–64. Cambridge: Cambridge University Press.

Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22* (1), 1–39.

Craik, H. (1943). *The nature of explanation.* Cambridge: Cambridge University Press.

Dowty, D.R., Wall, R.E., & Peters, S. (1981). *Introduction to Montague semantics.* Dordrecht: Reidel.

Gentner, D., & Stevens, A.L. (1983). *Mental models.* Hillsdale, NJ.: Erlbaum.

Glucksberg, S., Krauss, R.M., & Weissberg, R. (1966). Referential communication in nursery school children: Method and some preliminary findings. *Journal of Experimental Child Psychology*, *3*, 309–316.

Harman, G. (1977). Review of *Linguistic Behaviour* by Jonathan Bennet. *Language*, *53*, 417–424.

Johnson-Laird, P.N. (1983). *Mental models.* Cambridge: Cambridge University Press.

Levelt, W.J.M. (1982). Linearisation in describing spatial networks. In S. Peters & E. Saarinen (eds.), *Processes, Beliefs and Questions*, 199–220. Dordrecht: Reidel.

Levelt, W.J.M. (1984). Some perceptual limitations on talking about space. In A.J. van Doorn, W.A. van de Grind, & J.J. Koenderink (eds.), *Limits in perception*, 323–358. Utrecht: V.N.U. Science Press.

Lewis, D.K. (1968). *Convention: A philosophical study.* Cambridge, MA: Harvard University Press.

Lewis, D.K. (1975). Language and languages. In K. Gunderson (ed.), *Language mind and knowledge. Minnesota Studies in the Philosophy of Science Vol. 7.* Minneapolis: University of Minnesota Press.

Lewis, D.K. (1979). Scorekeeping in a language game. *Journal of Philosphical Logic*, *8*, 339–359.

Linde, C., & Labov, W. (1972). Spatial networks as a site for the study of language and thought. *Language*, *51*, 924–936.

Schelling, T.C. (1960). *The strategy of conflict.* Cambridge, MA: Harvard University Press.

Schiffer, S. (1972). *Meaning.* Oxford: Clarendon Press.

Suchman, L.A. (1985). *Plans and situated actions: The problem of human machine communication.* Palo Alto: Xerox Corporation.

Résumé

Cet article étudie comment les participants à une conversation coordonnent leur utilisation et leur interprétation du langage dans un contexte restreint. Cette étude repose sur l'analyse de descriptions spatiales qui sont apparues au cours de 56 dialogues obtenus en laboratoire en utilisant un jeu de labyrinthe sur ordinateur spécialement conçu à cette fin.

Nous avons effectué deux types d'analyses. D'abord, une analyse sémantique des différents types de description qui indique comment des couples de locuteurs développent différents schémas linguistiques associés à différents modèles mentaux de la configuration du labyrinthe. Ensuite, une analyse de la manière dont les communicants coordonnent la mise sur pied de leurs descriptions.

Les résultats de cette étude nous paraissent suggérer que le traitement du langage au cours d'un dialogue est peut-être régi par des principes locaux d'interaction qui ont reçu peu d'attention de la part des psychologues et des linguistes jusqu'à aujourd'hui.