Clarifying spatial descriptions: Local and global effects on semantic co-ordination.

Gregory J. Mills
Interaction, Media and Communication Research Group
Department of Computer Science
Queen Mary, University of London
London E1 4NS
gj@dcs.qmul.ac.uk

Patrick G. T. Healey
Interaction, Media and Communication Research Group
Department of Computer Science
Queen Mary, University of London
London E1 4NS
ph@dcs.qmul.ac.uk

Abstract

A key problem for models of dialogue is to explain the mechanisms involved in generating and responding to clarification requests. We report a 'Maze task' experiment that investigates the effect of 'spoof' clarification requests on the development of semantic co-ordination. The results provide evidence of both local and global semantic co-ordination phenomena that are not captured by existing dialogue co-ordination models.

1 Introduction

Perhaps the shortest possible clarification question is attributed to Oscar Wilde[1]. After sending a telegraph to his Parisian literary agent enquiring about the sales figures of his latest novel he received the response that sales of the book were indeed favourable. Wilde's subsequent telegraph was the single-character "?", to which the agent responded with the equally terse "!". Wilde could, of course, have formulated his clarification question differently, potentially leading to different patterns of response by his agent:

A: The sales are favourable.

1. W: What?
   A: Sales are favourable
   A: Sales are better than expected

2. W: How good ?
   A: 300 pounds
   A: More than your previous book

3. W: Favourable?
   A: Yes. 612 copies
   A: No. Incredible

Clarification requests (henceforth CRs), such as (1)-(3) above, are used to signal potential problems with the interpretation of a previous utterance. They are thus central to maintaining co-ordination in dialogue, as they serve the purpose of bringing the conversation "back on track" (Schegloff, 1992) when inter-subjectivity is threatened.

An account of the mechanisms underlying the use of different CR's, and their effects on the interaction is essential for an adequate understanding of dialogue and important for the practical goal of creating more natural, robust dialogue systems. However, empirical investigations of CRs have generally been limited to post-hoc analysis of corpora. For practical reasons it is difficult to achieve the levels of control necessary to support experimental manipulations of CR's. This has made it difficult to compare the effects of different CR's on conversational trajectories or subsequent semantic co-ordination.

This paper develops an experimental technique described in Purver et al. (2003). It combines an experimental chat-tool with a version of the maze game developed by Garrod & Anderson (1987). This enables the introduction of artificial 'probe' clarification requests into participants' dialogue without causing overt disruption to the conversation. By manipulating the type of probe CR used we can investigate their relative

[1] This has also been attributed to Victor Hugo, seeking the opinion of his publishers on his latest manuscript, and is possibly apocryphal.

brandial 2006: Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue, pages 122–129.

impact on semantic co-ordination in the maze game (cf. Garrod & Anderson, 1987; Garrod & Doherty 1994).

2 Clarification Requests

2.1 Levels of misunderstanding

A common thread running through the CR taxonomies proposed by Schlangen (2004), Gabsdil (2003) and Purver (2003), is that different CRs access different levels of understanding within some form of action hierarchy or 'ladder' (Clark, 1996; Allwood, 1995). An example ladder is:

Level 4. Action recognition.
Level 3. Meaning recognition.
Level 2. Utterance recognition.
Level 1. Securing Attention.

So, for example, the "Favourable?" CR above might typically request further specification of the meaning (level 3) whereas the "What?" CR might typically request clarification of what the initial utterance was (level 1) (although see Drew, 1997). Communication is only fully complete if understanding is secured at all levels (although see Allwood, 1995).

Hearer's choice of clarification type can thus signal the information required for them to reach a higher level of understanding. CRs such as (1) above typically signal low co-ordination as they give fewer clues about the nature of the problem or expected response (Schlangen 2004) than CRs such as (3), which requests further specification of what 'favourable' could mean in this context.

This ordering trades on a pragmatic expectation that people normally design their CR's to give as much information as possible about their current level of understanding. Although "what's" can be used to clarify at higher levels the expectation is that people should produce CR's that signal the highest level of understanding currently available to them. In the collaborative model this is formulated as the "strongest initiator rule" (Clark and Schaefer, 1989) which posits that a "[hearer] ought to index the parts he did hear, or the parts he didn't hear, and request help" in reaching the higher levels of comprehension (see Drew, 1997 for criticism of this formulation).

2.2 Semantic co-ordination

'Ladder' approaches can thus categorize and rank sources of problematic understanding, however they are, in effect, 'semantically neutral'. The different levels of the hierarchy don't address the potential for different forms of co-ordination that depend on semantic differences. This possibility is illustrated by data from the Maze game (Garrod & Anderson, 1987).

In the Maze game (see below) participants are faced with a recurrent problem of describing locations to each other. Over time their spatial descriptions shift from predominantly instance-bound 'Figurative' (Figural/Path descriptions described below) versions that depend on the specific configuration of the current maze to more 'Abstract' (Line/Matrix described below) approaches that invoke a relatively systematic underlying model of the maze that abstracts away from each instance and generalises across instances more easily.

There is evidence that interaction mechanisms, and not simply task experience, play a specific role in this shift toward more 'Abstract' description schemes. Even where task experience is equivalent, pairs' preference for 'Abstract' or 'Figurative' schemes differs depending on the opportunities they have had for interaction (Garrod and Doherty, 1994; Healey, 1997). There is also evidence that participants can only develop a shared 'Abstract' scheme following a prior stage of co-ordination with a 'Figurative' scheme (Healey & Mills, 2006; Healey, et. al. in press). Drawing on data from a different task, Schwartz (1995) argues that the shift towards abstraction emerges as a result of general collaborative processes that are unavailable to solitary speakers.

The experimental evidence thus suggests that differences in choice of description type correspond to differences in the degree of semantic co-ordination developed between dialogue participants. However, these shifts in semantic co-ordination are not readily explained by existing accounts of dialogue co-ordination. For example, they are not due to the kind of 'contraction' of referring expressions observed in many definite reference tasks (see e.g. Clark, 1996). The 'Abstract' descriptions are not reduced versions of Figurative descriptions, they involve a change in the underlying semantic model of the maze that participants are using (Garrod and Anderson, 1987). Also, as Garrod (1999) argues, local entrainment / priming mechanisms of the kind incorporated into the interactive alignment model (Pickering and Garrod, 2004) are conservative and not equipped to address global trends towards abstraction or innovations in description type. When people change schemes, the interactive alignment prediction is that the most frequently used (i.e., primed) prior scheme will predominate. A global shift towards a new scheme thus requires us to consider alternative co-ordination mechanisms.

Similarly, there is a general expectation in both Purver et al's (2003) and Rodriguez & Schlangen's (2004) models that interlocutors will modify their original utterance in response to a CR. However, there are no mechanisms for predicting what kinds of semantic change occur in response.

In summary, there is evidence that interaction contributes directly to the development of semantic co-ordination. Prima facie it seems likely that clarification requests play a key role in this process (cf. Clark, 1996; Pickering and Garrod 2004; Healey and Mills 2006). However, existing models of dialogue do not provide clear ways of interfacing between patterns of clarification request and possible semantic changes that might occur as part of the response. One reason for this is that it has not been possible to systematically investigate the effects of different kinds of CR on dialogue co-ordination.

Two basic empirical questions that arise then are a) whether there is a direct connection between the occurrence of CR’s and semantic co-ordination and b) whether there is a connection between the ‘level’ of CR and the form of semantic co-ordination.

To address these questions, a “Maze Game” experiment was set up using a text-based chat tool. The basic rationale of the experiment was to test the effects of different CR types on the form and content of participants' responses. Before introducing our specific hypotheses we explain the experimental methods in more detail.

3 Methods

The experiment employs a modified version of the "Maze Game", devised by Garrod and Anderson (1987). This task creates a recurrent need for pairs of participants to produce location descriptions. These descriptions can be reliably classified into four broad categories (see below), thus enabling the indexing of semantic co-ordination between participants (Garrod & Anderson, 1987; Garrod & Doherty 1994).

To support turn-level experimental manipulations of the dialogue, a chat-tool technique is used that engages participants in artificial clarification sequences. Before giving details of the procedure we first describe the implementation of the maze game and chat-tool used in the experiment and then explain the generation of these clarification requests.

3.1 The maze game application

The maze application is written in Java and displays a simple maze consisting of a configuration of nodes that are connected by paths to form grid-like mazes (see Fig 1). The mazes are based on a 7x7 grid and are selected to provide both grid-like and asymmetric instances.

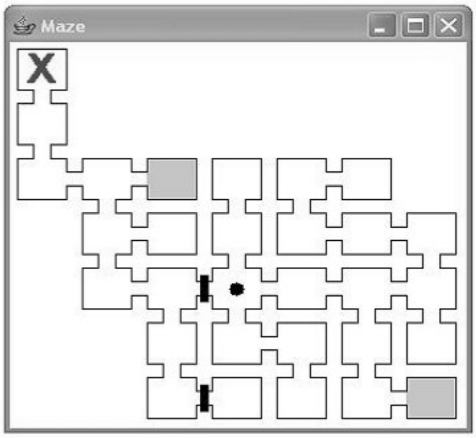

Figure 1: Example maze configuration. The solid black circle shows the player's current position, the cross represents the goal point that the player must reach, solid bars the gates and grey squares the switch points.

Subjects can move their location markers from one node to another via the paths. Each move is recorded and relayed to the server where it is time-stamped and stored. The game requires both subjects to move their location markers from a starting location to a goal that is marked with a cross. Although the maze topology is the same for both subjects, each subject has a different starting location and goal, neither of which are visible to the other subject. They are also not able to see each other's location markers.

Movement through the maze is impeded by gates that block some of the paths between nodes. These gates can be opened by the use of switches (grey coloured nodes). The locations of switches and gates are different on each maze and not visible to the other subject. Whenever a subject moves to a node that is marked as a switch on the other's screen, all of the other subject's gates open. All the gates subsequently close when they move off the switch.

This constraint forces subjects to collaborate: in order for participant (A) to open their gates, A has to guide participant B onto a node that corresponds to a switch that is only visible on A's screen. Successful completion of a maze (when both reach their respective goals) therefore requires subjects to exchange descriptions of their location and the locations of gates, switches and goals. Each new maze has a new configuration, starting points, gates and switches.

3.2 The Chat Tool

All communication takes place via a custom-built java chat tool similar to desktop messaging applications. The display is split into an upper window, a status bar and a lower window. The upper window displays the ongoing conversation, and the lower window is used for typing. All key presses are time-stamped and stored for later analysis. The status bar is a prominent single line of text that is controlled by the server and is similar to the status bar of proprietary messaging tools that display the activity status of the other conversant.

3.3 The Chat Server

In addition to relaying turns between participants, the server monitors the content of the turns in order to generate artificial clarification requests that appear, to participants, to originate from each other.

The server compares each turn with a lookup table of location descriptions obtained from a previous corpus of 10000 maze game turns (Healey & Mills, 2006), combined with rules for detecting misspellings and non-standard "txt" conventions. This ensures that CR's are generated only on turns containing spatial descriptions.

Each clarification request generated by the server is preceded by the other participant's chosen nickname, followed by a colon, and is dynamically modified to mimic spelling and typing speed.

To provide a manipulation of CR type the two classes of CR were selected; Reprise Fragments ('Frags') that echo a word from the target turn and 'Whats' (e.g., "what?" or "sorry?") that query the turn as a whole. These are the two most common forms of CR in ordinary dialogue (Purver et. al. 2003) and they provide two different levels of clarification. Reprise fragments involve direct re-use of a word from the turn and imply that the rest of the turn was understood. By contrast 'Whats' suggest that there were global problems finding a sense for the turn (but see also Drew 1997).

'Frags' (High co-ordination): Repetition of a single fragment of the location description.

'Whats' (Low co-ordination): What? Huh? Sorry? Ehh? Uhh? Where?

Participants' responses to the probe CR's are captured by the server. The probe CR and the response are displayed only in the participant's own chat-window. After receiving a response to the CR, the server sends one of the following acknowledgement turns to the recipient: "ok"; "k"; "ok right" and resumes relaying subsequent turns as normal. During the 'fake' CR exchange the server monitors whether the other participant starts typing. If this occurs, an error message is displayed and further text-entry is prevented until either the CR sequence is finished or a predefined time-out threshold is reached. To ensure error messages do not cue the interventions, a small number of random error messages are also introduced at other points in the dialogue.

3.4 Subjects

21 pairs of native English speaking subjects were recruited, 23 male and 19 female, from undergraduate students. They were recruited in pairs to ensure that they were familiar with each other. Only subjects who had some previous experience of using internet chat software such as ICQ or Microsoft Messenger were selected for the experiment. Each subject was paid £10.00 for participating in the experiment.

3.5 Procedure

Pairs of subjects were seated in separate rooms in front of a desktop PC. On each PC a window containing the maze (same configuration but different features see Fig 1) and a chat-tool window are displayed. Subjects were asked to select a nickname to be used in identifying chat turns and then wait for further instructions.

Subjects were told that the experiment was investigating the effects of a novel chat-tool and computer game on how people interact with each other. They were informed that their interaction would be recorded anonymously for subsequent analysis. Subjects were advised that they could request the log to be deleted and were free to leave at any time but would still receive payment in full.

They were given a written description of the maze game and told that the experiment involved solving twelve mazes. No information was given about the CRs generated by the server. At the end of the experiment the full nature of the experimental interventions was explained.

Twelve mazes were presented in random order to each pair. Artificial clarification requests on turns that involved spatial descriptions were introduced throughout the experiment with a 5% probability of any turn being clarified – consistent with findings from Purver et al. (2003) and Schlangen (2004) that approximately 5% of dialogue turns are CRs.

The experimental group were thus exposed to a within-subjects manipulation of CR type ('Frag' vs. 'What'). A control group of 12 additional pairs, recruited from the same undergraduate population, followed the same procedure but without the manipulation of probe CR's.

Experimental Hypotheses:

1. The introduction of artificial CR's will interfere with semantic co-ordination
2. More severe problems will cause more disruption (i.e., 'Whats' will cause more disruption than 'Frags')
3. People will systematically shift to more Figurative forms of semantic co-ordination (Figural / Path descriptions below) where problems occur.

4 Results

Overall, 246 clarification requests were artificially generated by the server: 109 'Frags', 128 'Whats' and 9 CRs generated for turns that did not contain spatial descriptions. These non-spatial clarifications were excluded from further analysis. On debriefing, nobody in the experimental group reported detecting that the probe CRs did not originate from their partner.

4.1 Description Types

Both the target turns used by the server to generate CRs and subjects' responses were classified according to the criteria developed by Garrod and Anderson (1987). This categorizes location descriptions into four basic classes corresponding to different underlying mental models of the maze:

Figural: a heterogeneous category of relatively concrete descriptions that draw on some specific element of the overall configuration of particular features to identify a target location.

A: "right above the sticking out bit at the top"

Path: involves identifying a route to be traversed through the maze to the target location. Path descriptions are sensitive to the specific layout of boxes and connections in the maze.

A: "From middle go up 1, 2 right, 1 down"

Line: classifies the maze into a set of line elements corresponding to rows, columns or diagonals. The target line is described first, followed by the target box as a position along it.

A:"In the bottom box, 2nd column from right"
A:"The third row, fifth to the left"

Matrix: introduces a Cartesian coordinate system with locations identified via the specification of two vectors either as rows and columns or in terms of numbers or letters for each axis.

A : "My switches are at 4,6 5,4 . I'm on 3,4"
A: "I'm in the 3rd row, 4th column"

Baseline Dialogues Clarified Dialogues

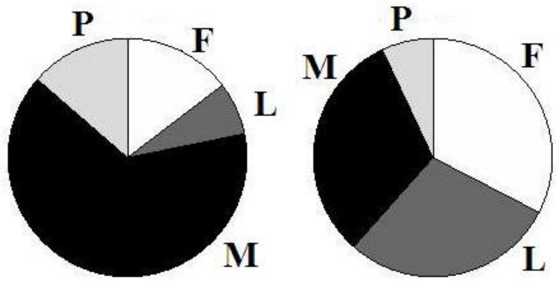

Fig 2: Global distribution of description types in baseline (control) condition and in dialogue queried with clarification requests (F = Figural, P = Path, L = Line, M = Matrix)

4.2 Distribution of description types

Figure 2 above illustrates the contrast in the global distribution of description types in the baseline control condition and in dialogue that is periodically interrupted with artificial clarification requests. The difference in use of description types is reliable (Multinomial Regression: $Chi^2_{(3)}=276$, $p=0.00$). The results show that the probe CRs significantly disrupt co-ordination in the experimental group. The largest category of description type in the experimental group is Figurative whereas in the baseline control group the Matrix descriptions predominate.

To check whether co-ordination was still developing over time (but to a lower level) in the experimental group the distribution of description types used in target turns the first four games was compared with those used in the last four games (see Fig 3).[2] This showed that there was still a significant shift in the use of description types over time (Multinomial Regression: $Chi^2_{(3)}$ =15.1, p=0.00) with participants migrating from 'Figurative' descriptions (Figural/Path) towards 'Abstract' (Line/Matrix) in the later games. This suggests that semantic co-ordination was still developing but at a significantly slower rate than in the control group.

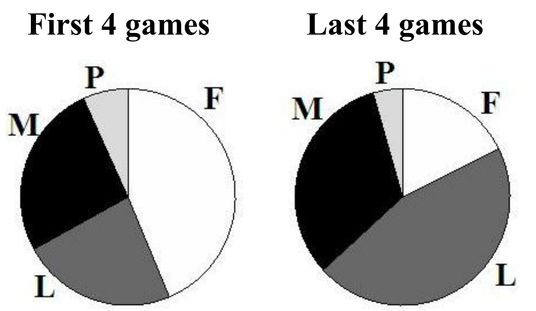

Fig 3: Global distribution of description types in first four and last four games in the experimental (CR) group.

In order to test the effects of the two CR types on the way responses were formulated a focused comparison of the distribution of description types in the responses immediately following the 'What' vs 'Frag' CRs was made. This showed no reliable difference (Multinomial Regression $Chi^2_{(3)}$ =1.68, p=0.64).

To provide an additional test of the third hypothesis –namely that people systematically shift to more 'Figurative' description types as a way of resolving co-ordination problems- we examined the relationship between the description type used in the target turn produced by a speaker and the spatial description type they produced in their response to the probe CR. Out of a total of 142 spatial description responses 101 (71%) responses used the same description type as the target. Of the 29% (41) that changed description type 14 (34%) involved a shift from 'Figurative' (Figural/Path) to 'Abstract' (Line/Matrix) whereas 27 (66%) involved a shift in the opposite direction. Overall, responses to the CR's predominantly used the same description type but where a change occurred it was more likely to involve a change to a more Figurative description type.

[2] Target turns only were selected for this analysis as these would be furthest from the immediate influence of the artificial CRs.

4.3 Other Measures of CR Effects

Times from the log files were used to provide two further comparisons of responses to the 'Frag' and 'What' CR's. Firstly, turn completion time – the time from the onset of typing of a response to its completion. A one-way analysis of variance revealed reliable differences between CR types, ($F_{(1,235)}$=6.5, p= 0.01). Overall, participants took longer to formulate their responses to 'Whats' than to fragments, taking an average of 18 seconds to respond to the former, and 25 seconds to respond to the latter.

The second measure of response time used was typing-onset time: the time between the onset of an intervention and the initial onset of typing the response. A one-way analysis of variance showed no effect of CR type ($F_{(1,235)}$= 0.32, p = 0.57).

In order to provide a measure of the indirect disruption caused by a CR, data from the log files was used to calculate the number of turns between receipt of a CR by a participant and the next turn in which they produced a spatial description (see 5.1 below). A one-way analysis of variance revealed significant differences between CR types ($F_{(1,93)}$=8.46, p=0.02). Overall, 'Frags' caused less disruption (average 3.2 turns before next description) than 'Whats' (average of 5.3 turns before next description).

The log files were also analysed for number of 'deletes' or edits that occurred in the construction of a turn prior to sending it. Although there was no reliable evidence of a relationship between edits and description type in responses ($Chi^2_{(1)}$ =0.881, p= 0.35) there was a reliable relationship between edits of the target turn and description type of the subsequent response to the CR ($Chi^2_{(1)}$ = 9.9, p=0.002). If there were no edits in the target turn there were more 'Abstract' responses (44 Matrix/Line vs. 23 Figural/Path). If the target was edited prior to sending there were fewer 'Abstract' responses (32 Matrix/Line vs. 43 Figural/Path).

5 Discussion

The global distribution of description types reported here (see also Healey and Mills, 2006) replicates the patterns of use observed in spoken Maze game studies (Garrod and Anderson, 1987; Garrod and Doherty, 1994; Healey, 1997). In particular, the pattern of migration from relatively concrete descriptions (Figural/Path) that depend on the specific details of each maze, towards more abstract description types (Line/Matrix) that invoke schemata that generalize across instances (see above) is the same in both modalities.

The advantage of using chat tools with the maze game is that it makes it possible to carry out context sensitive, turn-level experimental manipulations of dialogue; in the present experiment the manipulation of probe CR's. This allows us to address the question, raised in the introduction, of whether CR's have a direct effect on patterns of semantic co-ordination. The experimental results presented above provide strong evidence for such a connection.

While pairs in the control group converge on the 'Matrix' scheme, those exposed to CR's do not. Although their form of co-ordination does change over time it evolves more slowly and they do not converge on the Matrix scheme by the end of the experiment. The marked difference in the distribution of description types between the experimental (CR) and control (baseline) groups thus supports hypothesis 1.

The second question raised in the introduction was whether there is a connection between the particular type or 'level' of CR and form of semantic co-ordination. The results reported here do not provide a clear answer to this question. Hypothesis 2 predicted that 'Whats' would cause more disruption than 'Frags', however no reliable difference was found in the distribution of description types in responses to the two CR types. This is not, however, because the two CR types failed to have any distinct effects. The response time data show participants took longer to formulate their responses to 'Whats' than 'Frags'. In addition, the 'disruption' data indicate that participants took longer to get the dialogue back on track after a 'What' than a 'Frag'.

Overall, participants were sensitive to the difference between the two classes of CR. As expected, the 'Whats' were more disruptive to the dialogue than the 'Frags'. However, while the evidence thus supports hypothesis 2, the results make it more difficult to explain the nature of the connection between CR's and forms of semantic co-ordination.

The third hypothesis considered above was that the local effect of CR's should be to prompt a shift from 'Abstract' (Matrix/Line) to 'Figurative' (Figural/Path) descriptions. The results provide some support for this. In the cases where participants do change description type in response to a CR, there is a greater preference for changing from 'Abstract' to 'Figurative' than vice versa. However, the more striking observation is that in 71% of cases participants do not change type. This local consistency in description type echoes Garrod's original findings (Garrod and Anderson, 1987; Garrod and Doherty, 1994). However it presents a puzzling contrast with the global effects of the CR's. Although the additional clarification questions have a significant impact on overall co-ordination –as indicated by choice of description type– it appears that these effects are not manifest in the immediate context in which the CR's occur.

Perhaps the simplest potential explanation for the apparent contrast between the local and global effects on semantic co-ordination is that the CR's undermine participants' confidence in the interaction as a whole. So, although they are locally consistent in their response to the CR, they subsequently become more generally conservative in their choice of description types. If we treat editing of the target turn as an index of confidence prior to the CR then there is some support for this in the data. Figurative responses are more likely after CR's to an 'edited' target. This is consistent with a view that the CR aggravates the lack of confidence. A 'confidence' explanation, however, still provides no mechanism that can explain the trend towards more abstract forms of semantic co-ordination.

In the introduction we noted some problems of the action 'ladder' approach as a way of analyzing differences in semantic co-ordination. The global character of the effects observed here suggests an additional problem. Rather than pointing to 'vertical' modifications to the ladder they indicate a need for more 'horizontal' co-ordination mechanisms that could operate over larger stretches of interaction.

There are two methodological issues which need to be resolved in future work. First, the experiment was designed to produce CR's with a frequency similar to everyday conversation. In practice this resulted in each participant being exposed to one CR approximately every 40 turns. If it is true that the effects of the CR's are global rather than local it is possible they interfered with each other. In particular it suggests that combining the 'What' and 'Frag' manipulations in a single within-subjects condition is problematic. Second, the main advantage of using the Maze task is that the taxonomy of description types provides an attested way of indexing semantic co-ordination. However, this is still a relatively crude measure. Within each category there is considerable variation in how the descriptions are constructed and used (see e.g., Garrod and Anderson 1987; Garrod and Doherty 1994 for discussion). As a result local 'sub-description type' changes in response to the CR's would not be detected. The global vs. local contrast in the data could thus be an artifact of the measures of semantic co-ordination used.

Nonetheless, the results clearly show that participants reliably distinguish between CR types, and also show that the introduction of CR's into participants' dialogue has a strong effect on the kinds of description used in the maze game. Thus, future studies need to develop a more detailed analysis of the local impact of CR's on semantic co-ordination.

6 Conclusion

The data from the present experiment demonstrate a causal connection between the use of clarification questions and the development of semantic co-ordination in the maze task dialogues. Contemporary models of dialogue co-ordination need to be modified to accommodate these semantic effects. However, further empirical work is required to clarify the mechanisms involved.

References

Allwood, J. (1995). An activity based approach to pragmatics. *Gothenburg Papers in Theoretical Linguistics, 76* , Gothenburg, Sweden.

Clark, H. H. & Schaefer, E.F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259-294

Clark, H. H. & Marshall, C (1981) Definite reference and mutual knowledge. *Elements of discourse understanding,* Cambridge University Press.

Clark, H. H. (1996). *Using Language.* Cambridge: Cambridge University Press.

Drew, P. (1997) 'Open' class repair initiators in response to sequential sources of troubles in conversation. *Journal of Pragmatics* 28, pp.69-101.

Gabsdil, M. (2003). Clarification in spoken dialogue systems. In *Proceedings of the 2003 AAAI Spring Symposium. Workshop on Natural Language Generation in Spoken and Written Dialogue,* Stanford, USA.

Garrod, S. C., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27*, 181-218

Garrod, S. C. & Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition, 53,* 181-215.

Garrod, S. (1999) The Challenge of Dialogue for Theories of Language Processing. In S.Garrod and M.Pickering (Eds) Language Processing, Hove: Psychology Press. 389-416

Healey, P. G.T. (1997). Expertise or expert-ese? The emergence of task-oriented sub-languages. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society.* Stanford, California.

Healey, P.G.T., Purver, M., King, J., Ginzburg, J. and Mills, G. (2003) Experimenting with Clarification in Dialogue. In Alterman, R. and Kirsh, D. (eds) Proceedings of the 25th Annual Conference of the Cognitive Science Society. Mahwah, N.J.: LEA pp.539--544.

Healey, P.G.T. (2004). Dialogue in the degenerate case? (Peer commentary on Pickering & Garrod (2004), *Behavioural and Brain Sciences,* 27(2)

Healey, P.G.T. & Mills, G. (2006) Participation, Precedence and Co-ordination. (forthcoming) In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Vancouver. Canada.

Healey, P.G.T., Swoboda, N., Umata, I. and King, J, (in press) Graphical Language Games: Interactional constraints on representational form. *Cognitive Science*.

Purver, M. & Healey, H. & King, J. & Ginzburg J. & Mills, G. Experimenting with Clarification in Dialogue. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society (CogSci 2003)*, Boston, MA.

Purver, M., Ginzburg, J. and Healey, P.G.T (2003) On the Means for Clarification in Dialogue. In Smith, R. and van Kuppevelt, J. (eds) *Current and New Directions in Discourse and Dialgoue.* ACL, pp. 235-255.

Pickering, M., & Garrod, S. (2004). The interactive alignment model. *Behavioural and Brain Sciences, 27*(2), 169-189.

Rodriguez, K. J. & Schlangen, D. (2004). Form, Intonation and Function of Clarification Requests in German Task-oriented Spoken Dialogues. In *Proceedings of the Eighth Workshop on Formal Semantics and Dialogue.*

Schegloff, E. A. (1992) Repair after next turn: The last structurally provided defense of intersubjectivity in Conversation. In *American Journal of Sociology* (97)5, 1295-1345.

Schlangen, D. (2004). Causes and Strategies for Requesting Clarification in Dialogue. In *Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue*

Schwartz, D. L. (1995). The emergence of Abstract Representations in Dyad Problem Solving. *The Journal of the Learning Sciences,* 4(3), 321-354.