9th International Command and Control
Research and Technology Symposium
14 –16 September 2004
Copenhagen, Denmark

Decision Making and Cognitive Analysis

The "Intelligent Listener" in Collaborative Planning.

Authors

Marilyn Cross and Dominique Estival

E-mail:	**marilyn.cross@dsto.defence.gov.au**
Address:	**DSTO,** **Headquarters Joint Operations Command** **14 – 18 Wylde St.,** **Potts Point** **NSW 2011** **Australia**
Fax:	**+61 (0) 2 9359 5832**
Phone	**+61 (0) 2 9359 5748**

ABSTRACT

Collaborative planning in command and control contexts is carried out primarily by talk in the form of `round table planning'. The talk is focused towards a variety of pragmatic goals such as planning some action, assimilating and acting on intelligence and dealing with the logistics of assigning resources. Talk in collaborative planning is fast moving and complex, with planners engaged in putting forward suggestions, accepting, rejecting and modifying proposals, agreeing, disagreeing and carrying out many more types of task in order to progress the planning which is paramount to any military operation. In this paper we report on two studies of planning groups. Following some initial ad hoc observations of a commander's planning group, a linguistic framework was utilized that enabled the analysis of decision making from the actual exchanges of the talk rather than observers' impressions of the meeting. However, the process was extremely time consuming and so a series of experiments was planned to automate the process. The second experiment we report on was conducted with an operational planning group, and is the first trial using speech recognition technology and procedures to automate the capture, analysis and management of the talk of collaborative planning.

1 Introduction

Collaborative planning in command and control contexts is carried out primarily by talk in the form of `round table planning'. The medium of face to face talk in a group enables numerous interactive outcomes, which include putting points of view, giving orders, making requests, soliciting support and making decisions. Talk in collaborative planning is fast moving and complex, with planners engaged in putting forward suggestions, accepting, rejecting and modifying proposals, agreeing, disagreeing and carrying out many more kinds of work in order to progress the planning which is paramount to any military operation. Talk is also ephemeral and that is the problem. While the core planning decisions are captured in a permanent record of some form (e.g. PowerPoint slides), the exploration of alternatives, the justifications, the putting forward and rejection of proposals, the teasing out of information, the judgment of its relevance, the explication of underlying assumptions and the many jobs carried out by the talk of collaborative planning are rarely captured.

Improved management of the talk of collaborative planning will provide access to easily searchable archives of the discussions, which may be aligned with the enduring products from the planning sessions. This will permit more rigorous and accurate auditing and will allow information retrieval from those archives, both for re-use of deliberate and immediate planning products and for the induction and training of incoming staff.

In this paper, we present two studies leading to an analysis of planning talk and to the development of the new technology required to record and use this talk. In the first study of a commander's planning group (CPG), methodologies derived from discourse analysis (Eggins & Slade, 1997) were explored in order to understand the progression of collaborative planning in command and control (Cross and Bopping, 1998a). Utilization of a Speech Function Network (see section 3.3) enabled the fine analysis of planning based on the actual data of the exchanges rather than observers' impressions of what occurred. A critique of that study led to the second study presented here, which is the first of a series of experiments under the "Intelligent Listener" project. The long-term aims of the "Intelligent Listener" are to provide:

1. the ability to <u>recognise and understand communicative intents and speech acts</u> in planning sessions or meetings;

2. the ability to extract and retrieve information from recorded interactions during such sessions or meetings; and
3. the ability to organise this information to present it in a useful fashion.

The particular experiment described in this paper is the first trial of speech recognition technology and procedures to automate the capture of the talk of collaborative planning and to develop a corpus of such talk for analysis. The experiment was conducted in a live operational headquarters during a planning exercise and encompasses a series of operations planning groups (OPGs).

Following the introduction, the paper briefly describes collaborative planning at the joint operational level in section 2. For the first study, the context, theory, methodology and results, are presented in section 3. The critique of the first study which led to the second study is presented in section 3.7. The second study is presented in section 4. The outcomes of the two studies and how they may be taken forward are discussed in section 5 and conclusions are drawn in the final section.

2 Collaborative Planning at Joint Operational Level

Collaborative planning at the joint operational level follows the Joint Military Appreciation Process (JMAP). A single comprehensive process is still being refined to meet the need for a common staff planning process across the three Services (see Figure 1). The process was intended to be simple, robust and adaptable, to include all stakeholders, allow parallel planning, and exploit technology. JMAP was designed to take account of all aspects of operational planning, including risk management. It covers the range from deliberate (contingency) planning that has time to consider all options, to immediate planning that is situation specific and with tighter time constraints.

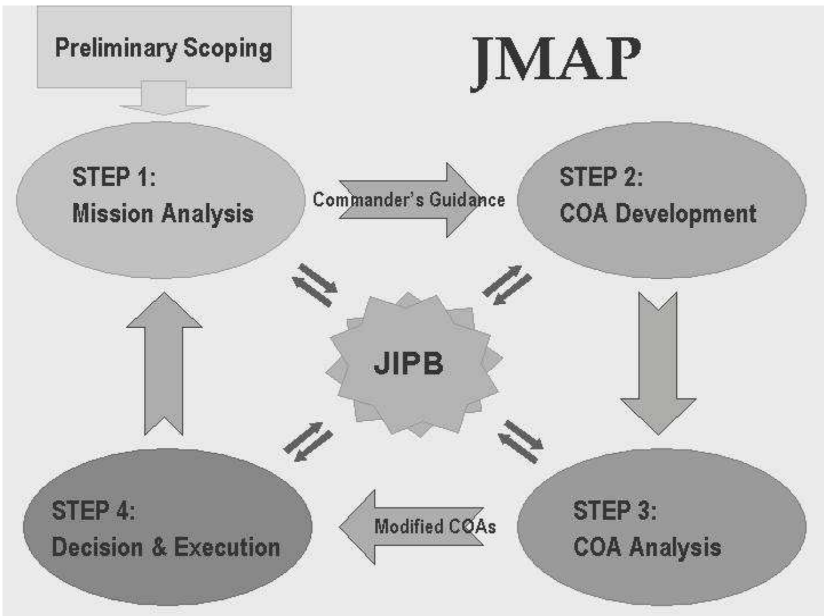

Figure 1: Joint Military Appreciation Process (JMAP), showing the linkage with the Joint Intelligence Preparation of the Battlespace (JIPB)

In studies one and two, the extant structure of the planning groups was that of the JMAP process and the enduring products of the planning reflect that process. However, in the capture and analysis of the talk of collaborative planning an ethnographic approach was taken that studied the actual talk without explicit recognition of the JMAP process. This decision was made in order to look at what was actually occurring as compared with what should have been occurring as prescribed by the JMAP process. At a later stage an overlay of the JMAP process could be applied.

3 Study One: Commander's Planning Group

The information presented in this section was first reported in (Cross and Bopping 1998). The study is discussed in the contemporary context.

3.1 Context

As part of an analysis of Commander's Planning Group (CPG) meetings in an operational headquarters, observers attempted to capture and categorize the elements of talk within the planning process on the fly. More specifically, they sought to gain a sense of how theatre level decision making and problem solving processes could be understood in terms of the talk of which they are composed.

Observers initially assigned their own ad hoc categories to those utterances that they were able to process during the planning meetings. When observers attempted to compare and collate the data, a number of areas of concern were identified which seriously effected confidence in any results that could be obtained. These areas were:

- *Within-rater reliability:*
  An observer could not be certain that they categorized similar speech utterances the same way each time they encountered them.
- *Between-rater reliability:*
  An observer could not be certain that they categorized similar speech utterances the same way as other observers.

In addition to the problem of reliability between and within raters, there were problems of confidence in the observations and the lack of an overall framework which would permit extensibility.

3.2 Aim and Scope

The aim of the study was to develop a more rigorous approach to capturing and categorizing the field observations of collaborative planning than observers attempting to capture and categorize elements of discourse 'on the fly' (Cross and Bopping, 1998). The approach adopted for categorizing the field observations was one which has proved successful in the analysis of casual conversation (Eggins, 1990; Eggins and Slade, 1997). We will refer to this approach as the Speech Function Network (SFN), although it should be understood that the SFN is part of the Systemic-Functional theory (Halliday, 1978) from which are derived a range of techniques for analysing the interactional patterns through which interactants jointly achieve the purposes of discourse. In the current context, the purposes are to make decisions and to progress planning with operational outcomes. The range of techniques include analysing:

1. grammatical patterns at the clause level which indicate power and subordination within interactions;
2. semantic patterns which indicate frequency of contact and familiarity among interactants;
3. patterns of discourse structure which indicate affective involvement and shifting alignments within discourse;
4. the use of text types, for example story telling, which give some indication of shared world views about normality and predictability.

For this study, we focused on the analysis of patterns of discourse structure, as this had the potential for elucidating the cycles of decision making which observers believed typified the planning process. It was anticipated that the analysis of patterns of discourse structure would not only standardize the data and assure higher levels of confidence in any results obtained, but also provide a valuable opportunity to investigate the suitability of this type of method for future use in the analyses of military planning and decision making.

3.3 Theory

Based on the systemic-functional (S-F) approach to discourse structure, Eggins (1990) proposed the SFN as a method suitable for categorizing casual sustained talk in a collaborative setting. Antecedents for this approach are found in the ethnomethodological model of conversational analysis (Sacks, Schegloff and Jefferson, 1974; Goffman, 1961), sociolinguistics (Gumperz and Hymes, 1972), speech act theory (Austin, 1962; Searle, 1969) and systemic-functional linguistics (Halliday, 1978). The S-F approach to discourse structure has also been applied to the linguistic analysis of workplace dialogues for the last 15 years (Bhatia, 1993; Ward, 2003).

3.4 Methodology

Because of security constraints, it was not possible to record the talk during the CPG but it was possible to write verbatim records for short sections of conversation. These records could then be categorized off-line using a rigorous and extensible framework.

In order to perform the analysis, the concepts of *Move*, *Turn* and *Decision Cycle* required definition. *Move* and *Turn* are defined in SF, but *Decision Cycle* was specifically defined for the planning process.

- *Move:*

The systemic functional approach identifies the move as the unit of interactional discourse. The move is tied to the grammatical unit of the clause, however rhythm and intonation patterns are also important in its determination.

- *Turn*:

The systemic functional approach identifies the turn as the point of speaker change. One turn may hence include a speaker making more than one move.

- *Decision Cycle*:

Observers identified a decision cycle as a passage of verbal planning, bounded by topical opening and closure, and consisting of multi-participant discussion. Decision cycles, however, may not always be this clearly identifiable.

Area	*Description*
Opening	moves intending to establish the audience configuration in which an exchange of discourse will unfold.
Sustaining	moves identified when the same speaker as the prior move continues to speak
Supporting	moves reacting to the speaker of the previous move by offering a degree of support or consensus
Confronting	moves reacting to the speaker of the previous move by offering a degree of negativity or non-consensus

Table 1. Speech Function Network: Description of Moves

The moves made by participants can be classified by the SFN as existing within one of four general network areas, as shown in Table 1. A *Move* network can be represented by a diagram indicating the choices available to the speaker at each point in the network, as in the examples given in Figures 2-5 in the next section.

To ensure some degree of between-rater reliability, each extract was categorized in terms of the SFN by two observers.

3.5 Results

The results are represented by extracts from small worked examples, with generalizations derived from more extensive analyses of the data which for brevity are not presented here. The excerpt given in example (1) consists of four moves, within three speaker turns:[1]

(1) Paul: *Have there been any developments in X-land?*
I heard they are moving west.
Joan: *Yes, that's correct.*
Bob: *Why is this important at this stage?*

The first move, "Have there been any developments in X-land?", is an *opening* move that is *floated* to the audience at large and which *initiate*s a request for *information* (not *Goods & Services*) that is *factual*.[2] The network diagram in Figure 2 tracks this move, with alternatives that were not chosen being greyed out.

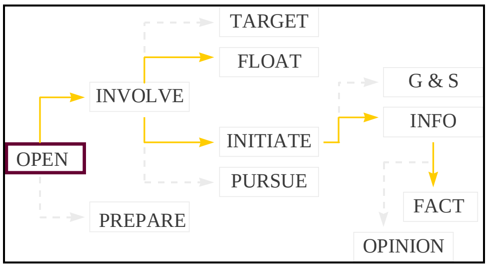

Figure 2: Opening Move

The second move, "I heard they are moving west." *prolong*s the move and *expand*s it by *qualify*ing it. The network diagram in Figure 3 captures the choices made.

---

[1] The names in the excerpt have been changed.
[2] As shown in Figure 2, FLOAT, INITIATE, INFO, FACT and G&S are elements of the network.

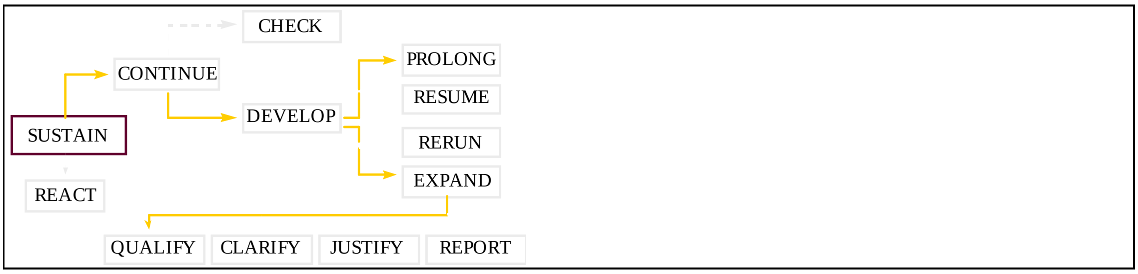

Figure 3: Sustaining Move

The next move, "Yes, that's correct.", is a change of turn and is a *support*ing move that *affirm*s the prior move.

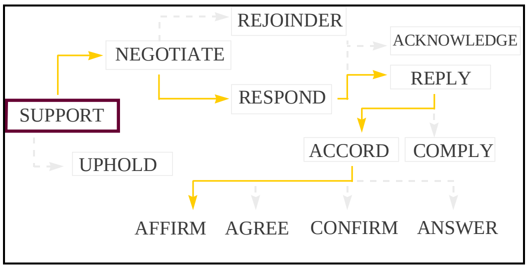

Figure 4: Supporting Move

The final move, "Why is this important at this stage?", is a *confront*ing move that *protest*s (i.e. challenges) the move of the first speaker.

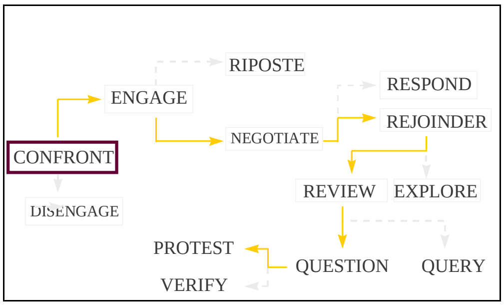

Figure 5: Confronting Move

Using this analysis, our conclusions are summarised as follows:

- Opening moves always used a 'floating' technique that did not target any one member of the group, as in example (2).

  (2) Paul: *Have there been any developments in X-land?*

- Opening moves always 'explored', i.e. moves were 'challenged' or 'countered', as in example (3)

  (3) Paul: *Have there been any developments in X-land?*

Joan: *X-land is not our first priority*

- Decision cycles tended to be balanced, consisting of both 'challenging' and 'pacifying' moves, as in example (4).

  (4) Jeff: *We need to try another option*
  Kay: *Yes, let's not get bogged down*

- Closure was typically marked by 'accordance' or 'complying' moves, as in example (5):

  (5) Paul: *Yes, okay, that's great...excellent choice*

- Decision cycles not closed by accordance or complying moves were invariably re-visited later.

While the first four findings are more descriptive of the linguistic material, the last one is particularly important for an analysis of the planning process. The implication for any decision support system for collaborative planning is that decisions that do not receive closure by accordance or complying moves need to be flagged so that they can either be pursued until closure at that time, or be revisited at a later stage in the planning (for example, when more information is available).

Decision cycles tended to contain smaller 'micro-cycles' which, upon their successful resolution, contributed to construction of the overall decision cycle. The majority of these micro-cycles were organized into clear, but short, accounts of topical opening, discussion and closure.

The final step of these analyses involved mapping decision cycles in terms of the sequential interaction of categories. The sociograms (Cross and Bopping, 1998) which were produced highlighted the importance of exploring (*challenging* and *countering*), qualifying (*adding*) and clarifying (*restating*, *exemplifying*, *amplifying*) moves throughout category interaction. More specifically, it appears as though exploring, qualifying and clarifying moves act as the major nodes which bind much of the decision cycle interaction together.

3.6 Discussion

Utilizing the SFN provided analysts with many useful insights regarding the possibility of its future use in the military planning context. However, there were a number of limitations to the study with regard to both the analysis and the data capture. The original SFN was derived from the analysis of casual sustained 'dinner party' conversation (Eggins, 1990). In this respect, it is not surprising that moves generated from a military planning context did not always fit into this network in a comfortable manner. For instance, analysts agreed that the forced choice between *supporting* and *confronting* moves posed the greatest limitation. As such, it was often this choice which became problematic for moves which analysts regarded as 'neutral' in reaction. Therefore, for future analysis, the SFN would need revision of the meanings or semantics of the interaction based on the characteristics of the data from collaborative planning.

To analyse a one-half hour CPG extract in accordance with the SFN was calculated to take approximately ten person days. This amount of analysis time could arguably be reduced with experience and modification of the network. The gain is that the analysis is based on actual data and not what observers deduce is occurring in the interaction. The question remains however, whether such effort is necessary to explain any additional variance not obtainable by less time-consuming methods. It may be possible to semi automate the analysis process using techniques employed in corpus analysis (Young and Bloothooft, 1997).

The data which could be applied to the Speech Function Network was limited to one CPG extract obtained by one observer and agreed by two others, and consisted of 238 moves, taking 189 turns across 16 speakers. It is important to acknowledge that this extract was deficient in a number of ways. These deficiencies included:

1. the extract being the sole source of data. Other observations did not satisfactorily yield close to verbatim accounts necessary for analysis with the Speech Function Network.
2. the extract being a 'best shot' at manually recording the verbal progression of a CPG. Some utterances were inevitably missed by the observer.
3. the extract lacking intonation or prosodic features that were not captured in its manual transcription.

4 Study Two: Operations Planning Group

4.1 Context

During EX TENDI WALK at the Headquarters Australian Theatre (HQAST), now Headquarters of the Joint Operations Command (HQJOC), from 02 to 06 February 2004, a feasibility experiment was conducted to examine how far, at its current level of development, speech recognition technology could overcome the problem of capturing verbatim talk during collaborative planning (Estival, Cross & Hashemi-Sakhstari, 2004). The feasibility experiment conducted in the live headquarters was the first in a series designed to assess the feasibility of providing an "Intelligent Listener" capability in a new Australian collocated joint operational headquarters planned for 2007/ 2008. The experiments were designed to provide:

1) analytical assessment;
2) assessment and management of security;
3) assessment and management of ethical concerns;
4) assessment of technology for capture; and
5) assessment of processing requirements for captured information.

4.2 Aim and Scope

The aim of the study was to assess the potential use of current speech recognition technology in the context of operation planning and to identify the limitations that will need to be addressed before such technology can be installed and used operationally. Another aim of the project was to collect real conversational data from planning sessions to perform further research on spoken language understanding, information presentation and pervasive computing. In this experiment the aim was primarily to concentrate on assessing the security aspects, the technology for capture and to a lesser degree, the processing requirements for the captured data.

4.3 Science and Technology

Since 1996, when the earlier study was conducted, automatic speech recognition (ASR) and automatic transcription technology have progressed to the point where it is possible to envisage their use in group settings such as planning over the next few years. One such system is the

Automatic speech-to-text Transcriber for Meetings and Interviews (AuTM) which was developed at DSTO (Zschorn et al, 2003). AuTM is a client-server application operating over a TCP/IP network to record and transcribe meetings and interviews. At the time of the experiment, AuTM had been extended to handle up to sixteen meeting participants.

Using a commercial ASR, AuTM automatically collects both textual and audio records. The records are organised according to the structure of the meeting and the audio and textual segments are time-stamped, aligned and linked. The output is a Word or HTML document containing the transcription of each speaker turn, which is linked to the audio file for that segment. Since the text and audio files are already matched, it is very easy to listen to the original audio input and correct the transcript as necessary. The AuTM output can be organised according to an agenda set by the meeting moderator; this facility can be customised to give any transcript a structure more relevant to the particular kind of interaction being recorded. In the case of a collaborative planning session, this could, for instance, follow the stages of the JMAP process.

4.4 Methodology

In the six months preceding the experiment, the goodwill and support of the user population were canvassed and engaged during a series of discussions and demonstrations. The actual feasibility experiment involved setting up the hardware and software within the environment in which it would be used, i.e. the planning room at HQJOC, and organizing how the recordings would be conducted. The week preceding EX TENDI WALK was allocated to set up the equipment and train the users on the ASR.

Dragon NaturallySpeaking is a speaker-dependent ASR, which means it needs to be trained for each user's voice and manner of speaking. This ensures better recognition rates and, although this is not yet quite satisfactory, allows the possibility of recording and transcribing conversational speech. In the current state of the technology, the alternative of a speaker-independent ASR does not allow the transcribing of conversational speech in meetings. Thus, to ensure the best possible results, training of the ASR with the intended users was conducted in the environment and on the machines to be used during the exercise.

From the point of view of conducting the experiment, it was decided that the transcripts produced by AuTM would be saved at the end of each planning session and the data transferred to CD-ROM at the end of each day. During the capture of the planning sessions, at least two DSTO observers were present during each planning session to observe the interactions. This was intended to facilitate later correction of the transcripts and to enable the conversation analysis of the planning sessions.

The hardware equipment consisted of seventeen networked Pentium IV laptops, one of which was used as the server. The server was a Compaq P4 1.7Mhz 1GB Ram 20GB HDD Win XP Pro Laptop with CD writer and the user machines were sixteen Compaq P4 1.7Mhz 512 Mb Ram 20GB HDD Win XP Pro Laptops. The operating and standard software consisted of Microsoft Windows XP Professional and Microsoft Windows 2000. The speech recognition and meeting processing software consisted of DSTO's AuTM and the commercial ASR Dragon NaturallySpeaking v.7. The AuTM server was installed on the laptop selected as the server machine and the AuTM clients on the laptops provided to each member of the planning team around the table. The client machines were to record, process and transcribe all the utterances of the planners during the planning sessions and then send both audio and text files to the server for insertion in the AuTM output transcript. The sixteen user laptops, each with its headset and USB sound card, were placed on the planning room table, one at each user position, and all the machines were networked together.

The environment of the planning room presented a number of challenges for the experiment: acoustically, because of the high level of acoustic noise (air conditioner, slide projector, number of participants); and logistically, because of the lack of space both on the table for all the experimental machines alongside the laptops used by the planners and under the table for all the cables connecting all these machines on two different networks.

It was necessary to use two different networks because the AuTM system had not been accredited to run on the higher security network used by the planners. Otherwise, the software could have been installed on the laptops used by the planners and run on the network already set up in the room. However, while installation on the higher security network would have alleviated the problem of space on and under the planning table, it would have introduced the more serious problem of making the users deal with "on/off" switches on the headsets dynamically. For the current accreditation of that network, only headsets with "push-to-talk" switches are permitted. In effect with a default setting of "off", the users would have had to press the switch every time they spoke in order to capture their voices.

Another challenge was the use of headsets to record speech. Although it could be expected that some users would be accustomed to wearing headsets, during the preparatory discussions users had expressed a certain amount of reluctance to wearing them for the collaborative planning. However, it is widely recognised that speech recognition requires a very high acoustic quality for the input and that in a noisy environment, such as a meeting room, table or ceiling microphones would have been useless for ASR, therefore individual close-talking and noise-cancelling microphones were absolutely mandatory with the current technology. Background or environment noise and reverberation drastically affect the ASR performance. Experiments with lapel microphones in that kind of environment have also yielded disappointing results, so in the reported environment, the only alternative was for users to wear headsets with built-in microphones. The microphones used in the experiment were all *VXI TalkPro* headsets.

4.5 Results

The results we present for this experiment are focused on the methodological findings. The first difficulty was that, although the users had been willing to put their headsets on at the beginning of the first session, when the leader of that session announced that there would first be a presentation by himself and some of the instructors, they all took off their headsets and put them back on the table. The room was crowded, with not only the sixteen participants around the table, but about thirty more people sitting behind them. With about fifty people in the room, it was impossible for the two DSTO experimenters sitting at the back next to the AuTM server to intervene.

The burning question was how would the system handle the overload. It was not long before the answer became obvious: the microphones were all picking up speech signals from the presenters and all the AuTM clients were sending files to the AuTM server. Soon, the network was unable to cope: some clients failed and eventually the server failed. No transcript was produced for that session. A similar problem recurred throughout the week when sessions were broken up in smaller groups, or when people got up and left their headsets on the table. In a number of those cases, no transcript was produced for those sessions.

An additional problem that was unforeseen was that users had access to the transcript during the experiment. Testing the accuracy of the speech recogniser was not the aim of the experiment, the goal was to assess the usability of the system and to identify the issues that have to be addressed before such a system can be used in operations. Because of the low accuracy of the speech recognition, users were distracted by the output. It would have been a simple matter to blank the screens or to install a password-protected screen saver. However, this would have required modifications to the AuTM interface and the risk of making such modifications once the system was installed was too high. This was a good example of a usability problem that only actual experimentation revealed. The solution is to have all the input microphones leading to one single computer where speech recognition is performed; this computer is then neither visible nor accessible to the users, and as a consequence, only the meeting moderator, or the experimenters, are able to access the transcript being produced. This is the solution now implemented in the next version of the system.

In summary, a number of things did not go according to plan once the exercise was under way and recording commenced during the planning sessions. Some difficulties were technical and will be remedied in the next version of the AuTM software (e.g. network architecture, interface, user profile management). Some difficulties were due to the quality of speech recognition at this point in the development of the technology and will require different solutions in the future (e.g. different types of microphones for better quality audio input, better noise cancellation and reverberation cancellation, improved acoustic design for meeting rooms). Other difficulties were due to human factors and should be resolved when the technology becomes an accepted part of the environment (e.g. remembering to put on the headset and/or turning the microphone on).

5 Discussion

The first study concluded that the method of analysis for identifying decisions made in the talk of collaborative planning was viable but would need revision of the semantics of the interaction based on the characteristics of the data. That conclusion was not tested in the second study. Nor has it yet been possible to test and develop methods for automating the analysis process. What was attempted in the second study was to overcome the problems of capturing valid verbatim talk over extended periods. To that purpose, speech recognition technology and procedures to automate the capture of the talk of collaborative planning were trialed. Despite the technological difficulties, recordings were obtained for most of the planning sessions, even for the sessions for which there were no transcripts. When fully transcribed, the recordings will form the beginnings of a corpus of the talk of collaborative planning.

One of the premises for the second study was that speech recognition technology works adequately for individuals and that it is getting to the point of being useful for groups. The technical problems identified in the second experiment, regarding networking machines for the AuTM server and the AuTM clients and the related problems of network overload, are now solved in the new AuTM version which has been implemented with a lightweight messaging system. However, even more important than the technical information allowing us to improve the technology was the information gathered about users and users' acceptance of the technology in their environment. Performing the experiment helped ascertain that users are willing to use the technology, even with the relative inconvenience of having to wear a headset, and that they are already convinced of the benefits of automatically recording planning sessions: keeping an audit trail, improving training and induction and enabling the re-use of planning products.

6 Conclusions

In this paper the results of two studies on the capture and analysis of the ephemeral but critical talk of collaborative planning have been reported. For the first study, following some initial ad hoc observations of a commander's planning group, a linguistic framework was utilized that enabled the analysis of decision making from the actual exchanges of the talk rather than observers' impressions of the meeting. From the analysis it was concluded that the semantics of the interaction required a richer model which could only be based on more extensive data. Without some automation, the data capture was grossly inadequate and the analysis extremely time consuming. Therefore, a series of experiments was planned to automate the process as much as possible. The second experiment we report on was conducted with an operational planning group, and is the first trial of using speech technology and automating procedures for the capture, analysis and management of the talk of collaborative planning. The sustained effort to set up and conduct the experiment was described and results pertaining to the technology and its usage in the headquarters reported. By the time the symposium takes place, some further results on the processing and analysis of that data will be available.

For the future, it is not too fanciful to predict that speech recognition technology and natural language processing tools to capture the decision making inherent in collaborative planning will be pervasive – a part of every well equipped planning room – perhaps even with the discretion of an intelligent human listener. Eventually such tools and the accompanying human processes will be embedded organically into all C2 environments.

References

Austin J. (1962). *How to do Things with Words.* Oxford: Clarendon Press.

Bhatia V. K. (1993). *Analysing Genre in Professional Settings.* London: Longman.

Cross M. and Bopping D. (1998). "Collaborative Planning Processes in Command and Control". *Proceedings of ICCRTS, 1998, Sweden.*

Eggins S. E. R. (1990). Keeping The Conversation Going: A systemic functional analysis of conversational structure in casual sustained talk. PhD Thesis, University of Sydney.

Eggins S. E. R. and Slade D. (1997). *Analysing Casual Conversation*. London: Cassell.

Estival, D., Cross, M. and Hashemi-Sakhtsari, A. (2004). "'Intelligent Listener': Experimentation in a Live Headquarters", *Defence Experimentation Symposium*, DSTO Edinburgh.

Goffman, E. (1967). *Interaction Ritual*. Chicago: Aldine.

Gumperz J., and Hymes D. 1972 *Dimensions in Sociolinguistics*.

Halliday M. A. K. H. (1978). *Language as Social Semiotic*. London: Edward Arnold.

Sacks, H., Schegloff E.A. and Jefferson G. (1974). A Simplest Systematics for the Organisation of Turn-taking for Conversation. *Language, 50, 4*:696-735

Searle J. (1976). A Classification of Illocutionary Acts. *Language in Society. 5: 1-23.*

Young S. and Bloothooft G. (1997). Corpus Methods in Language and Speech Processing. Dordrecht: Kluwer.

Sladek, J., Zschorn, A. and Hashemi-Sakhtsari, A. (2002). "Speech to Text Transcription in Support of Pervasive Computing", *Ubiquitous Computing: Trends, Tricks and Traps, and the Inaugural Asia Pacific Forum on Pervasive Computing,* October/November 2002, Adelaide Convention Centre, Australia. This paper was published in *Volume 25 - Advances in Computer Science Selected Papers from the ACS State Conferences,* edited by Brenda Ainsley.

Ward M. (2003). Systemic Dis-empowerment in Labour Union Contract Negotiations: a Preliminary Investigation. Paper presented at *ASFLA03*, 17 – 19 July, 2003.

Zschorn A., Littlefield J. S., Broughton M., Dwyer B. and Hashemi-Sakhtsari A. (2003). *Transcription of Multiple Speakers Using Speaker Dependent Speech Recognition*. DSTO Technical Report, DSTO_TR_1498.