"Can you read this well?"
Error Handling in a Translated Messaging Environment

Sherri Condon and Keith Miller

The MITRE Corporation
(scondon,keith)@mitre.org

Abstract

This paper focuses on the behavior users exhibit when faced with system errors in an instant messaging environment in which the system translates each user's messages into the language of the other participants. Messages were annotated to identify the strategies that participants adopt for managing their interaction, including strategies to repair and adapt to translation problems. Results show that participants employ high proportions of strategies that manage the interaction by explicitly referring to the ongoing communication. These results question the assumption that explicit verification strategies in dialogue systems are dispreferred because they require extra resources.

1. Introduction

This paper focuses on the behavior that users exhibit when faced with system errors and examines how these behaviors can be taken into account in error handling. Many studies of error repair in human-human interaction focus on self-repair strategies such as recycling portions of an utterance, pausing, or using fillers [1,2,3,4]. In contrast, the behaviors of interest to developers of dialogue systems are responses to utterances that have not been self-repaired, typically misunderstandings caused by failures in speech recognition [5,6,7,8]. This study examines human-human, computer-mediated communication in a context that is subject to misunderstandings caused by mistranslation, rather than misrecognition, but which have the same potential for substantial impact on the success of the interaction.

Utterances that function to repair problems in communication can be viewed as part of a larger type of utterances that manage activities required for successful interaction, such as openings and closings. Utterances that perform these discourse management functions contrast with utterances that contribute to the topics or goals that interlocutors share. If the latter are considered to be the talk, then the former are frequently talk about the talk. A reasonable assumption is that participants in interaction will prefer to engage in talk that contributes to their purposes, rather than talk about that talk. Following this reasoning, people will prefer strategies that accomplish management activities with minimal expenditure of cognitive and linguistic resources, such as the formulaic routines that have evolved for openings and closings in ordinary conversation.

Clark and Brennan [9] observe that speakers' preference for self-repair "tends to minimize the cost of a repair" (p.231) because fewer words and turns are required compared to repairs initiated by others. Designers of speech recognition systems have tried to minimize repair costs by using implicit verification strategies rather than explicit ones [10,11], assuming that the preferences motivated in face-to-face interaction are appropriate for machine-mediated interaction. However, Clark and Brennan also observe that strategies may change according to the medium of the interactions due to changes in the costs that are incurred. Consequently, it is useful to examine the strategies that are adopted in machine-mediated media, and this paper focuses on repair strategies in the context of other discourse managements strategies adopted in computer-mediated interaction.

2. Methods

We analyzed logs of interaction conducted using TrIM, MITRE's Translingual Instant Messenger prototype. TrIM is an instant messaging environment in which two or more participants are able to interact by reading and typing in their own preferred languages. The system translates each user's messages into the language of the other participants and displays both the source language and target language versions of each message. TrIM's translation services are provided by the CyberTrans system, which provides a common interface for various commercial and research text translation systems and several types of text documents (e.g. e-mail, web, FrameMaker). It incorporates text normalization tools that can improve the quality of the input text and thus reduce the errors in the resultant translation.

The logs include interactions recorded during an 8-day evaluation conducted in July, 2002, and the logs from 3 of those days were selected for analysis because they contained the highest amounts of sustained, coherent interaction. In all three logs, the languages in use were English, Spanish, and Portuguese. Participants were all native or near-native speakers of the languages they used. Participants on one day included as many as 10 participants, though there were never more than 4 participants communicating simultaneously.

Participants on the second day consisted primarily of 9 users in the morning and 10 in the afternoon. The latter group engaged in a scripted scenario requiring cooperation and information sharing. These participants were in Argentina, Bolivia, Brazil, Columbia, Ecuador, Paraguay, Peru, Uruguay, and the United States. The interaction in the 3rd log was primarily among 4 users and also included a role-playing task. There were about 20 distinct participants in the interactions.

Messages in the interactions were annotated using categories that are primarily based on function. All of the categories are designed to identify messages that function primarily to manage the interaction or messages that work to avoid perceived and potential problems in the communication. Most of the categories involve talk about the talk or *explicit management*, containing expressions that refer to the interaction itself with words like *say* and *mean* or, in these interactions, references to the TrIM system, the network, and

the translations. By tracking all of the messages that contribute to the management of the conversations, rather than only messages concerned with errors, we achieve a perspective that places these behaviors in the context of similar activities.

The unit of analysis is the message, which usually consists of a single utterance. Because messages sometimes consist of more than one utterance and because single utterances are often multi-functional, it was inevitable that some messages functioned in more than one category of interest. This problem was resolved by allowing utterances to be associated with a primary classification and, when appropriate, another secondary classification.

All examples provided here are presented exactly as they appeared to participants, except for the fonts and line breaks. The glosses that are provided are not necessarily those that were produced by the translation system.

3. Results

Table 1 presents the proportions of messages that are annotated with primary categories. *Explicit Management* is used to classify all messages that refer to or foreground the ongoing interaction, but do not satisfy the requirements for any of the other categories in Table 1. These messages comment on the quality of the translation (1a), check for understanding (1b), and describe participants' communicative actions (1c,d).

(1) a. Como esta a traducao para o espanhol?
"How is the translation for Spanish?"
b. Hola, Puedes leer bien este mensaje?
"Hi, can you read this message well?"
c. estoy esperando una respuesta
"I am waiting for an answer"
d. con esto se mejora la comunicacion en la red
"this improves the communication in the chat room"

As Table 1 demonstrates, openings and closings represent a significant portion of the messages in the interactions. The vast majority of these are greetings and formulaic responses to greetings, which are often at the beginnings of interactions when participants log onto the system. As (2) illustrates, greetings can be elaborate, and even simple ones can create problems for the translation engine, as in (2d), where *como estas* is ambiguous without stress.

(2) a. Hola Colombia, tierra del buen café
"hello Columbia, land of good coffee"
b. hola Argentina ... ricas pampas con la mejor carne
"hello Argentina…rich prairies with the best meat"
c. Welcome to all from the Andean Ridge coordinator
d. NESTOR COMO ESTAS?
"Nestor, how are you?"
System translation: NESTOR LIKE THESE?

Messages containing greetings receive only a secondary classification as Openings/Closings if additional language foregrounds the interaction, as in (3).

(3) a. HOLA A TODOS COMO LES VA EN ESTE SEGUNDO DIA DE EJERCICIOS DE TRADUCCION SIMULTANEA???
"Hello to everyone. How are you in this second day of simultaneous translation exercises?"
b. Good morning and welcome to the SURNET chat room
c. Hello and is every one in

Messages like (3) and (4) are categorized as *Orientations to Talk* when they orient participants to the interaction and to each other as participants in the interaction.

(4) a. Buenos creo que estamos todos, no?
"good, I think this is everyone, no?"
b. Quem e van?
"Who is Van? (Van is one of the participants)
c. Is anyone up and running in this chat room?
d. gusto poder conversar contigo
"I like to be able to talk to you"

The primary category *Reference to System* is reserved for messages with specific focus on the TrIM system. These include questions and directions about using the system (5a), comments about using and learning the system (5b), discussion of the capabilities of the system (5c), and scenario instructions that refer to the system (5d).

(5) a. Va a edit, language settings , para fazer traducao no idioma desejado.
"Go to edit, language settings, to set the translation to the desired language"
b. Carlos, we will continue to learn these tools in the morning
c. Pregunto: Qué pasa con los textos incorporados en un archivo JPEG? , tambiéen el traductor los traduce?
"Question: What happens with text incorporated in a JPEG file? Will the translator still translate it?"
d. Over the next 30 -45 minutes you will be engaged in a real time group chat environment to assess the feasibility of using this system to real-time operations between multiple countries

Messages are coded as *Explicit Repair* when participants correct perceived misunderstandings by using forms that explicitly refer to the language such as *say, mean*, and *word*, as illustrated in (6).

(6) a. te dije un cebiche en Lima capital de Peru
"I said a ceviche in Lima the capitol of Peru"
b. CORRECTION DOWNTOWN
c. Que significa PKO
"What does PKO mean?"
d. en la traduccion al Spanish puede haber confusion especialmente en el primer parrafo
"There can be confusion in the Spanish translation especially in the first paragraph"

In many cases, the problems were not caused by translation errors. For example, (6b) corrects a typographical error.

Messages are coded as *Experiments* if they appear to have been formulated primarily to observe how they were translated, and they are classified as *Reformulations* when participants revise a contribution to the interaction in order to achieve a better translation. For example, the revision in (7a) translated much better than the original message and the sender remarks on the strategy in his next message (7b).

(7) a. i would like to be in the meeting to be able to participate (reformulated from *I will like to be in that meeting to monitor progress*)
b. i changed the words
c. el mate es caliente
"the tea is hot"

In contrast, the experiment in (6c) occurs when the system translates *mate*, the word for a tea brewed in Paraguay, as "kill." The participant from Paraguay then produced several different messages containing *mate*. The *Reformulations* also include four messages in which participants reformulated the

translations of messages. It is not clear whether they are experimenting to see how the correct translation is translated

Table 1: Proportions of Messages that Manage the Interaction

Category	Raw Frequency	Proportion of All Messages (n= 506)	Proportion of Managing Messages (n = 288)
Explicit Management	91	.18	.32
Openings/Closings	55	.11	.19
Orientations to Talk	41	.08	.14
Reference to System	35	.07	.12
Explicit Repair	27	.05	.09
Experiments	18	.04	.06
Reformulations	14	.03	.05
Repetitions	7	.01	.02
Total	288	.57	*

*proportions do not sum to 1 due to rounding

or to ensure that participants receive a correct translation. For example, the English request in (8a) was translated awkwardly into Portuguese in (8b), and a participant responded with the Portuguese translation in (8c).

(8) a. Please open your email and receive attachment 1
b. Abra por favor seu email e receba o acessório 1
c. Por favor, abra seu correio eletronico e receba o anexo 1

Finally, there are 7 repetitions of messages, most of which were repetitions of the instructions for a role-playing task. In 2 cases, repetitions were produced by a different participant in a different language.

As Table 1 demonstrates, even if we exclude experiments, reformulations, and repetitions, fully half of the messages are serving management functions that are either openings and closings or explicitly refer to the ongoing interaction. Of course, one goal of the interaction is to learn and evaluate the TrIm system, so that many of the messages are also satisfying the goals of the interaction. Nevertheless, it is clear that participants tolerate high proportions of explicit management strategies in this communication environment.

Table 2 examines the strategies that participants use to manage translation problems. Two of the primary categories from Table 1 are repeated (*Explicit Repair* and *Reformulates*), even though not all of the explicit repairs are for translation errors. Three additional categories appear in Table 2, and most of the messages that were assigned to these secondary categories received a primary classification in the *Explicit Management* category.

The category *Evaluates System's Translation Behavior* is used for messages that were occasioned by a mistranslation, but make a general statement about the system, as in (9a,b) and (1d). This category is also used for evaluations of translation that refer to the system, usually as "it," as in (9c). Finally, messages about whether the system is translating a language were also included in the category, as in (9d).

(9) a. Long sentences cause a problem
b. The computer has problems translating the letter "Capital I"
c. interesting: it translated warm as "caliente"
d. Nao, estou recebendo em espanhol "No, I'm receiving in Spanish"

In contrast, messages coded as *Evaluates Translation*, correct or comment on mistranslations without reference to the system, as in (10).

(10) a. yes, the word welcome is not reception
b. The translation of the last sentence was perfect.
c. boca a boca did not translate
d. caliente es hot "*caliente* is *hot*"

Few messages in this category actually sought to correct specific errors as (10d) seems to.

Finally, the behavior classified as *Adapts* occurs when participants modify their messages in a manner that they know will make the language translate better. For example, as (9b) observes, the system consistently failed to translate "I," but "i" was correctly translated as the English first person singular nominative pronoun. Consequently, some participants began to use "i" instead of "I," as in (7a,b). All 7 messages coded as *Adapts* incorporated this strategy.

Table 2 shows that about one fifth of the messages in the corpus were concerned with correcting, evaluating and managing translations, which is about one third of the managerial messages. However, these proportions are inflated by explicit repair that was not elicited by translation problems.

Table 2: Proportions of Messages that Address Translation Problems

Translation Problem Strategy	Raw Frequency	Proportion of All Messages (n = 506)	Proportion of Managing Messages (n = 288)	Proportion of Translation Problem Messages (n = 92)
Explicit Repair	27	.05	.09	.29
Evaluates System's Translation Behavior	28	06	.10	.30
Evaluates Translation	16	.03	.06	.17
Reformulates	14	.03	.05	.15
Adapts	7	.01	**	.08
Total	92	.18	.30	*

*proportions do not sum to 1 due to rounding **not included in total of managing message

For some of the explicit repairs, it is not clear that translation problems caused the failure to understand. For example, when a participant asks *que significa?* "what does it mean?" in response to instructions for a role-playing task, it isn't clear whether the problem is the translation or the instructions. In other cases, such as the English speaker who asks *what is deiconify?* it is clear that translation has not caused the problem. Consequently, better estimates of the proportion of messages that manage translation issues are probably about 15% of the total messages and about 25% of the managerial messages.

In order to formulate some generalizations about the contexts in which participants adopted the strategies in Table 2, we examined the messages to identify common features. Almost half (13) of the explicit repairs were queries and responses to queries about the meaning of a specific term or acronym. Though the role-playing scenarios comprised only about a fourth of the messages, two thirds of the explicit repairs occurred in the context of a role-playing scenario. Half of the reformulations and half of the translation evaluations also occurred in the context of a role-playing task. In contrast, none of the evaluations of the system's translation behavior occurred during the role-playing scenarios. Half of the latter messages concerned whether the system was translating a particular language.

4. Discussion

The high proportions of explicit management in the interactions coincide with results from other studies of computer-mediated interaction. Higher proportions of explicit management were found when participants performed a decision-making task in a computer-mediated environment compared to a face-to-face environment [12]. There were also higher proportions of explicit management when participants performed a more complex decision-making task in both face-to-face and computer-mediated environments. It appears that participants in interaction tend to adopt explicit management strategies when demands on processing are increased.

Clearly, demands on linguistic and cognitive resources are high for participants in these interactions. They are interacting in an unfamiliar communication environment, and they cannot use strategies that rely on visual or auditory channels, as they can in face-to-face interaction. Much of the interaction involves large numbers of participants, as in chat environments, where the lack of fundamental conversational features such as adjacency makes it difficult to maintain coherence [13]. Moreover, participants must infer intentions from less than optimal translations and at the same time evaluate the translations and the success of the interaction. In some portions of the interaction, they are also asked to engage in a role-playing task.

A major reason why the corpus contains extremely high proportions of messages that explicitly refer to the ongoing interaction is that the goals of the interactions include learning, experimenting with, and evaluating the translation system. This consideration may explain why participants frequently claim that a translation is poor without making any attempt to request or provide a correction. For example, in 16 complaints about poor translations, only 2 messages include the correct forms. This behavior is understandable if participants are discriminating translations that are merely sub-optimal from problems that threaten the success of the communication. The former are evaluated because evaluation is a goal of the interaction, but they are not corrected because they do not compromise understanding.

Given this distinction, we reviewed all of the explicit management identified as *explicit repair, evaluates system's translation behavior*, and *evaluates translation* to determine which messages indicated that the participants were addressing a specific problem of communication. Of the 71 messages classified in those categories, only 23 seemed to be concerned with understanding a specific message, and 21 of these were among the 27 explicit repairs. Therefore, it appears that when participants in these interactions perceive that there has been a failure to achieve common ground, the strategy they select to resolve the problem is to explicitly ask and answer questions about what was meant or said.

The only other strategies that participants adopt to correct translation problems are reformulation and adaptation. Reformulation is never used when it is clear that there is a lack of understanding (while successful adaptation precludes the possibility of misunderstanding). Instead, reformulation only anticipates a potential for miscommunication, and the strategy does not guarantee a successful resolution, unless the user knows that the reformulated expression will translate correctly. (In this case, the strategy would no longer qualify as a reformulation: it would be an adaptation.) In fact, when adopting the reformulation strategy, participants are probably relying on their knowledge that an explicit repair strategy will be used if there is an actual misunderstanding and the reformulation has been unsuccessful. Meanwhile, the extra turn required for a reformulation incurs less cost than the minimum of two turns required to ask and answer a query about the meaning.

Additional evidence that explicit repair and reformulation are the preferred strategies for resolving potentially serious problems of understanding is the fact that these strategies primarily appear when the interaction is most serious and participants have more at stake in the success of the communication. When participants are engaged in role-playing scenarios instead of casual conversation, we find disproportionately high proportions of explicit repair and reformulation. Two thirds of the explicit repairs and half of the reformulations occur in the one fourth of the corpus in which participants are testing the system in scenarios that resemble the strategic operations for which they hope the system will eventually be used. Furthermore, another 29% of the reformulations occur in contexts that have consequences for participants' offline behavior, such as (7a), so that 79% of the reformulations appear when participants are invested in the outcome of the communication. Similarly, 5 of the 7 instances of adaptation occur in these contexts.

Together, the 23 messages that address specific problems of understanding (not all of which were caused by translation problems), the 14 reformulations and 7 adaptations comprise 9% of the messages in the corpus. This proportion is relatively small compared to the other managerial activities in which participants engaged: even if we exclude greetings, all references to the system, references to translation,

reformulations, adaptations, and repetitions, the proportion of messages expended on explicit management is still about 20%, which is the same proportion reported in [12] for complex decision-making tasks performed in a similar computer-mediated environment. The relatively low proportion of messages expended on translation problems is an encouraging result for proponents of machine translation who have expressed the hope that improvements in translation quality and users' experiences will allow machine translation to become fully embedded in applications, running as unobtrusively as any other operation.

Some evidence that participants in these interactions are adjusting to the TrIM environment can be assessed by comparing the proportions of management functions that occur in earlier portions of the interactions compared to later portions in which participants can benefit from the earlier experience. For this comparison, Table 3 presents the interactions in sequence according to the order in which they occurred with each divided in half, so that the sequence can be tracked in half-interaction units. Because greetings typically occur at the beginnings of interactions, they are excluded from the management functions that are summed in Table 3. Otherwise, the frequencies represent all of the functions identified in Table 1.

Table 3: Proportions of Management Functions in Earlier vs. Later Portions of Interactions*

Interaction	First Half of Interaction	Second Half of Interaction
First Day	14 (.63)	9 (.39)
Second Day AM	43 (.63)	28 (.41)
Second Day PM	54 (.42)	52 (.40)
Fifth Day	11 (.32)	22 (.62)
Total	122 (.48)	111 (.44)

*excludes *openings/closings*

Already in the first interaction, managerial activities decrease considerably during the second half of the session. They rise again in the first half of the morning session on the second day, but the 9 participants in that session were more than twice as many simultaneous participants than there were at any time during the first day. The proportion falls again during the second half of the morning session and holds steady when participants begin again during the afternoon. The proportions do not fall during the second half of the afternoon session, but this interaction included a role-playing scenario, which increased the complexity of participants' goals, introduced unfamiliar vocabulary, and motivated greater care in communication.

By the fifth day, management functions have reached a low that is considerably less than the .46 average for these functions in the corpus as a whole. The improvements represented in Table 3 suggest that participants are quickly adapting to the system, except for the large increase in the second half of the fifth day. However, the proportions of management functions in that final unit of the corpus are inflated by two factors. First, participants were engaged in another role-playing scenario and second, there was a lengthy sequence in which one participant asked a series of questions about the capabilities of the system. An example is the query in (5c).

It may be possible for designers of dialogue and translation systems to take advantage of the fact that participants adopt explicit management strategies when processing demands increase and misunderstanding is perceived. The possibility that systems can identify trouble spots in interaction has motivated much of the research on error handling in the hope that systems can use the information to improve recognition and adjust dialogue strategies. Explicit management strategies tend to have characteristic lexical forms and collocations such as *say, mean, type, word,* and terms relevant to the machine-mediated environment like *computer, system*, and *chat room.* Many of the messages that manage translation in the corpus include the term *translation* and evaluative words such as *slow, problem,* and *correct*. Therefore, a direction for further research is to explore methods for identifying trouble spots that make use of these characteristic forms.

5. Conclusions

Designers of dialogue systems that incorporate speech recognition have assumed that implicit verification strategies are preferable to explicit verification strategies, reasoning that the latter "requires extra turns, which users may find annoying" (p. 1423) [10]. However, this corpus supports previous observations that participants in machine-mediated interaction frequently adopt explicit management strategies in spite of the linguistic, cognitive, and interactional resources that such strategies consume. Evidently, the cost of explicit management strategies is considered to be less than the cost of the problems that can result when interaction is pressured by processing demands and potential failures in grounding. In fact, there is evidence that highly explicit confirmation strategies can improve the confirmation process and increase success rates in mixed initiative dialogue systems designed to process spoken queries about train schedules [11].

The study reported in [11] compares typical explicit confirmations (e.g. *Do you want to leave on Monday the 5th of January?)*, two types of implicit confirmations (e.g. *What time do you want to leave on the 5th of January?* and *You are leaving on the 5th of January. At what time?)* and a third strategy that explicitly directs the listener to correct errors *(e.g. You are leaving on Monday the 5th of January. In case of an error, correct me; otherwise, indicate your departure time.)* The two explicit strategies elicited the highest proportions of refutations when errors were present and the lowest proportions of failure to answer questions asked by the system. A system that incorporated these two types of strategies produced the highest rates for successful completion of the transactions compared to systems incorporating the first explicit strategy and one of the implicit ones.

Therefore, we conclude that designers of dialogue systems should not assume that implicit strategies are preferable to explicit ones for confirmation or for any other management functions that the system and participants may need to perform. The data reported here suggest that participants in interaction rely on explicit strategies when demanding goals must be accomplished and when fundamental grounding processes appear to be compromised. Explicit strategies are spontaneously and frequently adopted

by participants in interaction as effective practices for solving the problems inherent in communication.

The data also suggest another possibility that designers may want to consider as users become more familiar with dialogue systems. It appears that alternatives to costly explicit strategies can evolve when cooperative behaviors become familiar routines [14]. For example, the interactional work of opening conversations can be accomplished with a simple *hello*, instead of *I hereby open this conversation with you*. Routines allow interlocutors to accomplish large amounts of interactional work with relatively little expenditure of resources and are highly effective strategies for organizing language behavior.

The strategies for handling translation errors that have been identified here seem to reflect the process by which complex interactional activities can become routinized. When participants choose to reformulate a poorly translated contribution, they anticipate the need for an explicit repair and seek to avoid the cost of that repair using the less expensive reformulation. Of course, the reformulation is only less expensive if it is successful, which cannot be guaranteed, and the strategy is clearly not a routine. However, if the reformulation does guarantee a successful result, such as the use of "i" instead of "I" in the corpus, then the strategy becomes an adaptation which is routinely used to circumvent a translation problem. This kind of routinization may become a more viable option as users become familiar with spoken dialogue systems.

This initial investigation of strategies adopted to manage translation problems is only an exploratory study, and the results must be checked against larger corpora. Nevertheless, the opportunity to examine behavior in the TrIM environment has provided some insight into the strategies that participants adopt in machine-mediated interaction and has suggested some implications for a variety of dialogue and translation systems.

References

[1] Bredart, S. Word interruption in self-repairing. *Journal of Psycholinguistic Research,* 20:123-38, 1991.

[2] Fox, B., Hayashi, M., and Jasperson, R. Resources and repair: A cross-linguistic study of syntax and repair. In E. Ochs, E. Schegloff, and S. Thompson (eds.) *Interaction and Grammar,* pp. 185-237. Cambridge University Press, 1996.

[3] Jefferson, G. Error correction as an interactional resource. *Language in Society,* 2: 188-99, 1974.

[4] Schegloff, E., Sacks, H., and Jefferson, G. The preference for self-correction in the organization of repair in conversation. *Language,* 53: 361-82, 1977

[5] Bear, J., Dowding, J., and Shriberg, E. Integrating multiple knowledge sources for detection and correction of repairs in human-computer dialog. *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, 1992.

[6] Oviatt, S., Bernard, J., and Levow,G. Linguistic adaptations during spoken and multimodal error resolution. *Language and Speech,* 41 (3-4), 419-422, 1998.

[7] Shin, J., Narayanan, S., Gerber, L., and Kazemzadeh, D. Analysis of user behavior under error conditions in spoken dialoogs. *Proceedings of the International Conference on Spoken Language Processing*, 2002.

[8] Swerts, M., Litman, D., and Hirschberg, J. Corrections in spoken dialogue systems. *Proceedings of the International Conference on Spoken Language Processing*, Volume II, pp. 615-618, 2000.

[9] Clark, H. and Brennan, S. Grounding in communication. In L. Resnick, R. Levine, and S. Teasley (eds.) *Perspectives on Socially Shared Cognition,* APA Press, 1991.

[10] Krahmer, E., Swerts, M. Theune, M., and Weegels, M. Problem spotting in human-machine interaction. *Proceedings of Eurospeech99,* Vol. 3: 1423-1426, 1999.

[11] Lavelle, C., Clames, M. and Perennou, G. Confirmation strategies to improve correction rates in a telphonic inquiry dialogue system. *Proceedings of Eurospeech99,* Vol. 3, 1399-1402, 1999.

[12] Condon, S. and Cech, C. Talk about talk in face-to-face and computer-mediated interaction. In E. Nemeth (ed.) *Cognition in Language Use: Selected Papers from the 7th International Pragmatics Conference, Vol. 1.* IPA, 2001.

[13] Herring, S. Interactional Coherence in CMC. *Journal of Computer Mediated Communication,* 4(4) June, 1999.

[14] Condon, S., Cech C., and Edwards, W. Measuring Conformity to Discourse Routines in Decision-making Interactions. *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics,* 238-245, 1999.