Co-operativity in Human-Machine and Human-Human Spoken Dialogue

Niels Ole Bernsen, Hans Dybkjær and Laila Dybkjær

Centre for Cognitive Science, Roskilde University

PO Box 260, DK-4000 Roskilde, Denmark

Abstract

The paper presents principles of dialogue co-operativity derived from a corpus of task-oriented spoken human-machine dialogue. The corpus was recorded during the design of a dialogue model for a spoken language dialogue system. Analysis of the corpus produced a set of dialogue design principles intended to prevent users from having to initiate clarification and repair meta-communication which the system would not understand. Developed independently of Grice's work on co-operation in spoken dialogue, these principles provide an empirical test of the correctness and completeness of Grice's maxims of co-operativity in the case of human-machine dialogue. Whereas the maxims pass the test of correctness, they fail to provide a complete account of principles of co-operative human-machine dialogue. A more complete set of aspects of co-operative task-oriented dialogue is proposed together with the principles expressing those aspects. Transferability of results to co-operative spoken human-human dialogue is discussed.

In the last couple of years, we have designed and implemented the dialogue component of a spoken language dialogue system prototype in the domain of Danish domestic flight reservation. We are currently testing the system together with the partners in the project[1]. As the aim of the project is to develop a realistic, application-oriented prototype, the issue of user-system co-operativity has played a central role throughout our work on designing and implementing the dialogue structure. The present paper presents results on co-operativity in spoken human-machine dialogue and comparisons with human-human dialogue.

We argue that dialogue co-operativity is crucial to the design of spoken language dialogue systems (SLDSs). Realistic systems are characterised by limited linguistic skills, a limited vocabulary, limited knowledge of the world and limited ability to leave dialogue initiative with their users. They are not sensitive to prosodic features, such as intonation, vowel elongations, and pauses. They lack the average human's ability to draw inferences. The result is a largely system-directed dialogue rather than mixed initiative dialogue. While functionally adequate for a certain class of well-structured tasks, system-directed dialogue lacks the natural flexibility of the mixed-initiative dialogue that is characteristic of human-human interactions (Bernsen, Dybkjær & Dybkjær, 1994a; Bernsen, Dybkjær & Dybkjær, 1994b).

SLDS designers are currently investigating how to tackle the difficult next step of enabling mixed initiative human-machine dialogue (Bernsen et al., 1994b; Dybkjær, Bernsen & Dybkjær, 1995; Goddeau, Brill, Glass, Pao, Phillips, Polifroni, Seneff & Zue, 1994; Peckham, 1993). It is possible today to approximate system-directed dialogue for fairly complex tasks. Despite being linguistically constrained, such systems behave naturally in terms of task-specific vocabulary, user input understanding that includes natural grammar and appropriate semantics, close-to natural management of discourse and close-to-real-time response. The technology thus enables the construction of usable task-oriented SLDSs which are tolerably inferior to the humans they replace, despite the fact that dialogue with such a system is, in effect, conversation with an idiot savant. In addition to our own system, examples of such systems are presented in Oerder and Aust (1994), Cole, Novick, Fanty, Vermeulen, Sutton, Burnett & Schalkwyk (1994), and Mazor, Braun, Ziegler, Lerner, Feng & Zhou (1994). A crucial point in what follows, however, is that system-directed dialogue breaks down when users ask questions of the system. A key, therefore, to the successful design of system-directed dialogue is to design the dialogue in such a way that users do not need to ask questions of the system. To do this, we claim, requires optimising the dialogue co-operativity of the system.

Given that dialogue initiative lies mainly with the current SLDSs, dialogue designers have to take every possible precaution to minimise the number of situations in which users are inclined to initiate meta-communication for purposes of clarification and repair. Meta-communication is communication on the dialogue itself rather than on the task domain of the dialogue. Human-human dialogue both allows for and is greatly assisted by clarification and repair meta-communication. If we are in doubt as to what our interlocutor said or meant, why a particular topic was raised, why it was raised at that particular point, or why it was raised in a particular way during dialogue, we initiate clarification and repair meta-communication to find out. Similarly, speakers often take advantage of the fact that their partners can demand elaboration at any point. This helps fine-tuning the speaker's contributions and indicates interest from the partner. The standard way of initiating clarification and repair meta-communication is by asking questions of the interlocutor (Schegloff, Jefferson & Sacks, 1977). In largely system-directed dialogue, however, user-initiated clarification and repair meta-communication must be either avoided or restricted to the use of well-defined user commands, such as ‘correct’ or ‘repeat’, because the system is unable to understand unrestricted meta-communication. The achievement of mixed-initiative meta-communication dialogue is an important goal in SLDS design which lies beyond the scope of this paper.

When the dialogue has been designed to optimise co-operativity, users do not need to ask meta-communicative questions of the system in order to understand it (Bilange, 1991; Eckert & McGlashan, 1993). If the system's contributions were already fine-tuned to conform to the human interlocutor's expectations, there would be no need for clarification and repair meta-communication which the user cannot manage by invoking simple mechanisms, such as the keywords 'correct' and 'repeat'. In order to optimise user-system co-operativity, we developed a set of general usability principles to be observed in co-operative human-machine dialogue design. This made it possible to apply the principles in our dialogue design and, just as importantly, to re-use the same principles in other human-machine dialogue design efforts. Having developed and applied the principles, we became aware of the link between our work and Grice's Co-operative Principle and maxims (Grice, 1975).

According to Grice's Co-operative Principle (CP), to act co-operatively in conversation, one should make one's "conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which one is engaged" (Grice, 1975, p.26). Grice proposes that the Co-operative Principle can be explicated in terms of maxims of co-operative human-human dialogue as discussed later. Although Grice's maxims have been conceived with a different purpose in mind, they can be seen as serving the same objective as do our principles, namely that of preventing interlocutor-initiated clarification and repair meta-communication. From this viewpoint, the main difference between Grice's work and ours is that the maxims were developed to account for co-operativity in human-human dialogue rather than in human-machine dialogue, whereas our principles were developed from analysis of a corpus of simulated human-machine dialogues. Another point of potential interest is that, at least superficially, our set of principles is considerably larger than Grice's set of maxims. These differences provide an opportunity to (a) test how Grice's theory of dialogue co-operativity works in the domain of highly restricted human-machine dialogue; (b) compare co-operative human-human dialogue with co-operative human-machine dialogue; and, (c) potentially augment the basis of a theory of spoken dialogue co-operativity.

It would be useful to briefly summarise some main differences between human-human and human-computer task-oriented dialogue. The latter consists of highly specialised mutual-goal exchanges between partners of vastly different skills of language and comprehension (cf. the system limitations noted above). Yet the inferior partner controls the dialogue. This rarely occurs in the human-human exchanges which are closest to the human-computer interactions we are considering. In those exchanges, the interlocutor with superior skills normally takes over the dialogue initiative and simplifies vocabulary, makes own contributions more explicit, and asks more questions. The reason why, in our human-computer interactions, the inferior dialogue partner may control the dialogue is that this partner is the domain expert. This combination of speaker properties seems different from anything found in human-human dialogue, even from an exchange with an esteemed foreign expert. In addition, the human-computer dialogue has no role for prosodic features. Still, the dialogue serves the accomplishment of the task. This is only possible because the human interlocutors understand how to perform the task without the use of prosody and, more generally, display enough flexibility in dialogue to communicate with a partner that is more idiot savant than any human interlocutor could ever be.

The next section describes how our principles were developed. The principles are then presented. The subsequent section presents Grice's theory of co-operativity and situates the comparative analysis to follow. This completes the preparations for providing a detailed analysis of the relationship between the two sets of principles. The result is an account of co-operativity in spoken human-machine dialogue whose applicability to human-human dialogue is discussed in the concluding section. For ease of reference, the term maxim will refer to Grice’s maxims of co-operativity whereas the term principle will refer to individual rules in our analytical scheme.

Developing Principles of Co-Operative Human-Machine Dialogue

The dialogue model for our flight reservation system was developed by applying the method of Wizard of Oz (WOZ) experimental prototyping (Dybkjær & Dybkjær, 1993). WOZ is an iterative simulation technique which is well suited to the testing of dialogue models, including the adjustment of design goals and constraints prior to implementation. During each iteration, a human (the ‘wizard’) simulates the system in dialogue with users who are made to believe that they are speaking to a real system (Fraser & Gilbert, 1991). After the dialogues are recorded, transcribed, and analysed, the results are used to improve the dialogue model. This iterative process continues until an acceptable dialogue model has been achieved. The model is then implemented and tested with subjects who represent the intended user population. In system-directed dialogue design, system co-operativity is a main defining characteristic of an acceptable dialogue model, as argued previously. We performed seven WOZ iterations to achieve an acceptable dialogue model, which yielded a relatively large corpus of transcribed, task-oriented, human-machine dialogues (Dybkjær, Bernsen & Dybkjær, 1993). A total of 125 dialogues were transcribed, amounting to approximately seven hours of spoken language dialogue. Twenty five early dialogues were never transcribed whereas 94 of the transcribed dialogues were recorded during the last two WOZ iterations. These dialogues were performed by external subjects. Each subject received four task scenarios to perform on the phone through dialogue with the system, and was asked to fill in a questionnaire after interaction with the system. A total of 24 different subjects were involved in the seven iterations.

In the last two WOZ iterations, we matched the latest version of the system’s dialogue model against the transcribed WOZ corpus in order to systematically assess improvements in system co-operativity. The dialogue model was represented as a complex state transition network that had system output in the nodes and expected contents of user utterances along the edges (see Figures 1 and 2). Each transcribed dialogue was plotted onto the state transition network. Deviations from the state transition network consisted of unexpected user or system behaviour. Deviations were marked and the reason(s) for the deviations analysed. Also, before the next WOZ iteration, we matched the scenarios to be used against the current dialogue structure in order to discover and remove potential user problems. This matching (plotting) process allowed identification of both actually occurring and potential user problems. These problems were expected to produce user questions in the form of clarification and repair meta-communication. An actual user problem is one which actually occurred during user-system dialogue in the WOZ experiments. Potential user problems are problems discovered by the designers when putting themselves in the place of the actual users.

(Insert Figures 1 and 2 approximately here.)

In addition to the iterative WOZ process just described, the structure of the dialogue model was adjusted in the light of typical structures identified in human-human flight reservation dialogues. We recorded a corpus of 25 Danish domestic flight reservation dialogues in a travel agency, which consisted of approximately one hour of spoken human-human dialogue.

At the end of the WOZ design phase, we began a more theoretical, forward-looking exercise. The actual and potential user problems uncovered during the WOZ experiments were analysed and subsequently represented as violations of principles of co-operative dialogue. Thus, each problem was considered a case in which the system, in addressing the user, had violated a principle of co-operative dialogue. The principles of co-operative dialogue were made explicit, based on the problems analysis. We subsequently analysed how the system’s utterances had been improved to minimise user-initiated clarification and repair meta-communication (Bernsen, 1993; Bernsen et al., 1994a). The WOZ corpus analysis led to the identification of 14 principles of co-operative dialogue. In the original report on the data (Bernsen, 1993), the principles were illustrated by 36 concrete examples of their violation, but the total number of examples in the corpus were 120. Of course, some of the principles were violated more frequently than others were.

To illustrate the WOZ corpus analysis, we present below an example of an identified problem type (a) and the co-operative principle (termed 'design commitment') which has been violated (b). A justification of the principle is provided (c), followed by examples of how the principle was found to be violated (d). Under (d) we note whether a particular example was discovered empirically (i.e. from actual user problems) or analytically (i.e. through design analysis revealing a potential user problem). Finally, a solution to each particular problem is proposed and sometimes discussed. This template (a-d) was applied to each problem that was identified (Bernsen, 1993).

(a) Problem: there is 'semantic noise' in addressing users.

(b) Design commitment: Avoid 'semantic noise' in addressing users.

(c) Justification: Need for unambiguous system response. The design commitment is to reduce the possibilities of evoking wrong associations in users, which in their turn may cause the users to adopt wrong courses of action or to ask questions which the system cannot understand.

(d) Examples:

1. "Are you particularly [stressed by the wizard] interested in making use of special fares?" (WOZ iteration 6). The word 'particularly' was introduced in order to avoid users who did not want to make use of special fares answering 'yes'. Experience with user responses during earlier iterations had shown that this might otherwise happen. The change caused an improvement but the problem did not go away. In the discussion we came across the possibility that users would interpret the system's question (with or

without ‘particularly’) as the question whether they have an interest in travelling as cheaply as possible, which perhaps most people have. Several alternative design options were discussed, including:

Option a. Special fares are offered only after all the information relevant to reservation has been entered into the system’s database. This will not do, however, as users who turn out to be interested in special fares may have to go through most of the reservation dialogue once again.

Option b. At an early stage in the dialogue the system asks if the user’s choice of time of travel depends on the possibility of obtaining special fares.

This example was discovered empirically.

Resolution:

Option b was implemented.

2. To support interactive user-system problem-solving, an ‘Interrupt’ spoken keyword command was introduced which gave users access to the three functions ‘Correct’ (to be used, e.g., when the system had manifestly misunderstood a user request), ‘Change Subject’ and ‘Help’. Unfortunately, many users tended to interpret the ‘Interrupt’ function as the function ‘End the Dialogue Now’ (WOZ iteration 6).

This example was discovered empirically.

We stop the illustration here. The infelicitous term ‘semantic noise’, used to express the principle in (b) above, was chosen ad hoc to denote an open class of terms which were imprecise, ambiguous, misleading or otherwise liable to evoke false associations in users. Many more examples of ‘semantic noise’ were identified. The function described in Example 2 was later removed and replaced by the two user keywords 'correct' and 'repeat' (see Figure 2).

Figures 1 and 2 present two state transition networks from the 6th and 7th WOZ iteration, respectively. The networks represent the same part of the dialogue model and illustrate how the discovery of several user problems led to modifications in the dialogue model. Our technological design constraints were not yet satisfied in WOZ iteration 6. There was a need for more strict and precise instruction to users at the beginning of the dialogue. However, this would lead to more system talk and users who were already familiar with the system might be annoyed by having to listen to the instructions in every dialogue. To resolve this trade-off, the system's introduction was extended and made optional in WOZ iteration 7. The instructions to users on how to interact with the system were made more explicit and mandatory ("The system will only understand you if ..."). The ambiguous meta-communication keyword ‘interrupt’ was replaced by the two more well-defined and self-explanatory keywords ‘correct’ and ‘repeat’ and an explanation of their function was added to the system's introduction. In addition, the system's task domain was extended from a single route (Copenhagen-Aalborg) to all domestic routes. This latter change had nothing to do with user problems, however. As the transition networks show, the WOZ design phase included the three tasks of reservation, change of reservation and information. Of these, only the reservation task was implemented.

The result of the described iterative process of identifying and analysing user problems caused by the dialogue model is a set of co-operative principles for human-machine dialogue. Adherence to each principle would presumably minimise usability problems that frequently occur in SLDS design.

Except for the speech recogniser which was simulated, the implemented system was tested with 12 external users (Bernsen, Dybkjær & Dybkjær, 1995). Each user received four scenarios and a questionnaire. They conducted the dialogues over the telephone as in the WOZ experiments. In addition, subjects received a telephone interview immediately after interaction with the system. The 48 dialogues were recorded and transcribed. The test results show that the co-operative principles have been successfully applied in the design of the dialogue structure. As in the 7th WOZ iteration, very few questions were asked. In the 7th

WOZ iteration, 4 out of 881 user utterances were questions (Dybkjær et al., 1993). In the test, 4 out of 998 user utterances were questions. One of these four questions was asked because the subject had misread the scenario text. The three other user questions all concerned available departure times. This is not surprising since departure times constitute a type of information which users often do not have in advance, but expect to be able to obtain from the system. It may be concluded that the principles developed through the process described above constitute a powerful means of preventing unwanted user-initiated clarification and repair meta-communication, i.e. meta-communication which cannot be handled by using the keyword commands 'correct' and 'repeat'. In the next section, we present and discuss the principles.

A Set of Co-operative Principles for Spoken Language Dialogue Systems

The purpose of the analysis described in the previous section was to develop general SLDS dialogue design principles. These principles would presumably support improved user-system co-operativity in dialogue. Fourteen principles, P1 to P14, were identified based on the WOZ corpus. Each principle presented in Table 1 is accompanied by its justification, which serves the additional purpose of clarifying the meaning and scope of the principle itself. Although not explicitly stated in each justification, we take it to be straightforward that violations of any of the principles may lead users to initiate repair or clarification meta-communication, because this is the strategy naturally adopted in human-human conversation in such cases. To facilitate comparison with human-human dialogue principles, each occurrence of the term ‘system’ in the stated principles may be replaced by ‘speaker’, and each occurrence of the term ‘user’ may be replaced by ‘interlocutor’ or any other favoured term, such as ‘dialogue partner’ or ‘recipient’.

(Insert Table 1 approximately here)

Grice’s Maxims of Co-Operative Human-Human Dialogue

This section presents Grice's maxims and demonstrates commonality of purpose between his analysis and ours. The special role of the maxims of truth and evidence is pointed out. The analysis to follow is contrasted with other strands in the discussion of Grice's work on conversation.

The Co-Operative Principle and the Maxims

Grice's Co-operative Principle (CP) is a general principle which says that, to act co-operatively in conversation, one should make one's "conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which one is engaged" (Grice, 1975, p.26). Grice proposes that the CP can be further explicated in terms of four groups of simple maxims which are not claimed to be jointly exhaustive nor to have been generated on a principled theoretical basis other than the CP itself. The maxims are specified below.

Quantity

M1. Make your contribution as informative as is required (for the current purposes of the exchange).

M2. Do not make your contribution more informative than is required.

Grice observes that M2 is closely related to M5 below. In other words, the maxims are not mutually exclusive but may overlap.

Quality

M3. Do not say what you believe to be false.

M4. Do not say that for which you lack adequate evidence.

Grice notes that M3 and M4 seem presupposed by the other maxims but nevertheless refrains from putting them in a different category from the rest.

Relation

M5. Be relevant, i.e. be appropriate to the immediate needs at each stage of the transaction.

Grice points out that the concept of relevance is in need of further explication (see, e.g. Sperber & Wilson, 1986).

Manner

M6. Avoid obscurity of expression.

M7. Avoid ambiguity.

M8. Be brief (avoid unnecessary prolixity).

M9. Be orderly.

Grice notes that there may well be more maxims in this category.

A Common Objective Between Grice's Analysis and Ours

In preparation of the discussion of the relationship between principles of co-operation in human-machine dialogue and maxims of co-operation in human-human conversation, it must be demonstrated that Grice's maxims serve a purpose which is more or less identical to that served by the principles. It should be noted that the maxims may well have been designed for other purposes. In fact, this may be why Grice did not identify all the aspects of co-operative dialogue to be discussed below. We shall speak interchangeably about 'conversation' and 'dialogue' as these terms denote the same phenomenon for the present purposes.

The principles in our analysis were developed to avoid the need on behalf of the interlocutor (user) to initiate clarification and repair meta-communication through questioning the speaker (system) during task-oriented dialogue. Grice assumes that any particular dialogue serves, to some extent, a common purpose or set of purposes which may be more or less definite and either fixed from the start or that evolve during the dialogue. The maxims are stated, he says, as if the purpose of the dialogue were a maximally effective exchange of information. Task-oriented dialogue would seem to be a prototypical case of purposeful dialogue in this sense. Throughout the dialogue, the interlocutors share one common and specific goal, namely that of completing specific tasks involving flight ticket reservation. Correspondingly, the aim of dialogue design is maximally effective exchange of information. Thus, the CP clearly purports to be relevant to the design of task-oriented dialogue.

Grice claims that adherence to principles, such as the CP and the maxims, is rational in the sense that anyone who cares about achieving the goals that are central to the dialogue must be expected to have an interest in conducting the talk exchanges in accordance with those principles. We accept this claim.

Violation of our principles would normally lead the interlocutor to ask questions of the speaker. Grice does not consider such cases of communication failure and it is clear that he has not developed the CP and the maxims in order to help preventing partner-initiated meta-communication. Rather, his purpose was to investigate ‘conversational implicature’. Conversational implicature is a phenomenon in human-human dialogue in which the speaker manages to imply, or suggest, unstated information under the general assumption of adhering to the CP, and manages to be understood by the interlocutor as doing so. In understanding conversational implicature, the interlocutor is guided by inferences based on the CP and the maxims, according to Grice (Grice, 1975; Grice, 1978).

The large class of inferential phenomena which Grice denotes by the term ‘conversational implicature’ is very important in SLDS dialogue design. Indeed, our work has been very much focused on making the machine put its messages to the user in such a way as to avoid user inferences which may lead to the initiation of clarification or repair meta-communication. However, both the notion and the theoretical issues associated with conversational implicature were absent in our work, at least when it was first developed. Rather, dialogue design focused on making the machine state, as literally as possible, what needs to be stated in context. Grice was not interested in the meta-communication which may arise the moment the speaker fails to adhere to the CP and the maxims, not even in cases where this failure involves (failed) conversational implicature. However, the fact that meta-communication may arise at this point would seem perfectly compatible with Grice's theory. The CP cum maxims, he says, imply that at each stage of dialogue some possible conversational moves would be excluded as conversationally unsuitable. If such moves are nevertheless made, meta-communication is the natural co-operative mechanism which serves to bring the dialogue back on track.

We conclude that the CP and the maxims, as a necessary side-effect of improving understanding and enhancing communication, serve the purpose of preventing the need for clarification and repair meta-communication. Conversely, one way of reducing the need for clarification and repair is to rely on principles that enhance communication. Our principles are of this type.

Truth and Evidence

Grice's maxims of quality have no counterparts among our principles. The reason is that one does not design an SLDS in the domain of air ticket reservation which provides false or unfounded information to customers. In other words, the maxims of truth and evidence are so important to the design of such systems that they are unlikely to emerge during dialogue design problem-solving. This notwithstanding, one of the worst breakdowns during the WOZ experiments actually occurred when the wizard accidentally came up with an inconsistent day of the week/date pair. During system implementation, one constantly worries about truth and evidence. It cannot be allowed, for example, that the system confirms information which has not been checked with the database and which might be false or impossible. As remarked above, Grice did observe the fundamental status of the maxims of quality in general and M3 in particular. Similarly, Searle (1992) has pointed out that the requirement of truthfulness, i.e. M3, is an internal constitutive rule of the notion of a statement. It therefore has a different status from the rest of the maxims. Instead of just adding the maxims of quality to our set of principles, we propose to treat them as basic to our system design effort and ignore them in the discussion to follow.

Situating the Discussion

The CP has been widely discussed in linguistics and cognitive science (for overviews, see Huang, 1991; Sarangi & Slembrouck, 1992). We want to highlight a few important points.

There are still problems over targeting the scope of application of Grice's theory. We assume the apparently uncontroversial position (Sarangi & Slembrouck, 1992) that the CP is at least intended to be valid for dialogues in which the participants share the goal(s) of the dialogue and in which no other goals are involved. This situation does not occur in dialogue in general (Pratt, 1981; Ahluwalia, Agnihotri & Subbarao, 1990; Sarangi & Slembrouck, 1992). However, it does occur in some classes of dialogue (Grandy, 1989). With respect to these, the CP is widely accepted as valid. In spoken human-machine dialogue, full goal-sharing occurs when the user interacts with the machine with the sole purpose of achieving the task for which the machine has been designed.

Several researchers have proposed principles or inference strategies which enable the recipients in conversation to correctly interpret the same class (or more classes) of conversational implicature and other conversational phenomena as do the maxims. These alternative principles are normally fewer in number than Grice's maxims, such as one principle (Sperber & Wilson, 1986) two (Horn, 1984) or three (Levinson, 1987a; Levinson, 1987b). Such approaches have been termed 'reductionist approaches' (Huang, 1991) because they identify fewer principles than Grice did.

The present paper addresses the questions: which dialogue aspects should the partners take into account in co-operative task-oriented spoken dialogue and what are the principles expressing these aspects? These questions differ from (a) the larger, related, issue of making oneself understood, or 'getting one's point across' in dialogue in general; (b) 'reductionist' generalisations of co-operative dialogue principles which fail to make explicit the relevant dialogue aspects and the principles expressing them; and (c) explanatory cognitive accounts of making oneself understood in dialogue. Issue (a) represents a difference in scope, (b) a difference in specificity and (c) a difference between descriptive and explanatory aims.

Several researchers have exposed Grice's theory to data from human-human dialogue. Sarangi and Slembrouck (1992) tested the theory on data from human-human dialogue in various institutional settings. Ahluwalia et al. (1990) examined informal neighbourhood exchanges. In the absence of ways of proving the completeness of a set of principles and aspects of co-operative dialogue, the use of empirical, corpus-based methods would seem to be the only alternative. Moreover, state-of-the-art spoken human-machine dialogue possesses a simplicity and amenability to analysis which is rarely found in human-human dialogue.

Principles and Aspects of Co-operative Task-Oriented Dialogue

In this section we analyse the relationship between the maxims of co-operative human-human dialogue M1, M2 and M5 to M9 and our principles of co-operative human-machine dialogue P1 to P14 (see Table 1). The first aim is to demonstrate that a sub-set of the principles are roughly equivalent to the maxims. We then argue that the remaining principles express non-Gricean aspects of co-operativity. Before discussing in detail the relationship between principles and maxims, let us note some formal commonalities and differences.

First, there are some <u>overlaps</u> among maxims as well as among principles. It is possible that overlaps may be avoided in the final analysis. What is more important at this stage, however, is to arrive at a substantial, if not exhaustive, set of co-operative principles, whereas overlaps among them may be dealt with later. Second, we shall observe a number of relationships of <u>presupposition</u> between the principles, just as Grice observed such relationships between the maxims. Third, it is well-known that the maxims may <u>conflict</u> when applied to actual situations of dialogue utterance production. The same is true of the principles. When maxims conflict during utterance production, the speaker somehow has to decide on the priorities. When principles conflict during dialogue design, the designers have to trade off different design options against one another, with each option having a different weighting of the principles (for examples, see Bernsen et al., 1994a). Fourth, principles, but not maxims, are allowed to be related to one another in a kind-of relationship. In Grice's scheme as well as in the 'reductionist' schemes referenced above, a maxim or principle which is merely a species of another is redundant. This is not the case with principles, whose primary purpose is to guide dialogue design in a way which is sufficiently informative to avoid usability problems. As we shall see, the typical case of principle subsumption is when a generic principle states "Do (make, be, avoid, provide etc.) X" and the subsumed principle states how to do X in a specific type of case. We shall distinguish between generic and specific (subsumed) principles in what follows and mark the specific principles encountered. Finally, a distinction will be made between 'principles' and 'aspects' of co-operativity. Aspects of co-operativity, such as Grice's' quality', 'quantity' or 'manner', are theoretically important because they identify dimensions of co-operativity over and above the level of the co-operative maxims or principles themselves.

Corresponding Maxims and Principles

In view of the fact that the principles P1 to P14 have been developed independently of maxims M1, M2 and M5 to M9, a first question is of course whether the set of principles includes the maxims as a sub-set. We consider the evidence that this is the case by inspecting the relevant principles in numerical order.

P3. Provide same formulation of the same question (or address) to users everywhere in the system's dialogue turns.

P3 represents an extra precaution against the occurrence of ambiguity in machine speech. It can be seen as a special-purpose application of M7 (non-ambiguity), which is not needed in human-human dialogue. P3 is a specific principle.

P5. Avoid 'semantical noise' in addressing users.

P5 is a generalised version of M6 (non-obscurity) and M7 (non-ambiguity). The admittedly non-standard formulation of P5 was due to the fact that we wanted to encompass ambiguity and related phenomena in one principle but failed to find an appropriate technical term for this purpose. P5 may, without any consequence other than improved clarity, be replaced by M6 and M7.

P6. It should be possible for users to fully exploit the system's task domain knowledge when they need it.

P6 has a formulation which lacks any direct correspondence among the maxims. P6 may be considered an application of M1 (informativeness) and M9 (orderliness), as follows. If the system adheres to M1 and M9, there is a maximum likelihood that users obtain the task domain knowledge they need from the system when they need it. The role of orderliness in this context is that the system should address the task-relevant dialogue topics in an order which is as close as possible to the order in which users expect them to be addressed. Eventually saying enough is not sufficient for co-operativity. If the recipient expects some topic to be addressed early on in the dialogue, that topic's non-occurrence at its expected "place" may cause the recipient to question the speaker. In WOZ iteration 3, for instance, the system did not ask users about their interest in discount fares. As a result, a user asked about discount when approaching the end of the reservation dialogue. From WOZ iteration 6 onwards, users are asked early on whether they are interested in discount fares, which blocks impatient questions about discount possibilities. P6 may, without significant loss, be replaced by M1 and M9. This assumes that M9, when applied to the ticket reservation task, would yield the implication that orderliness is defined by the user's expectations.

P10. Avoid superfluous or redundant interactions with users (relative to their contextual needs).

P10 is virtually equivalent to M2 (do not overdo informativeness) and M5 (relevance). Note that Grice observed the overlap between M2 and M5. It appears that P10 can, without any consequence other than improved clarity, be replaced by M2 and M5.

P12. Reduce system talk as much as possible during individual dialogue turns.

P12 is near-equivalent to M8 (brevity).

To sum up, the generic principles P5, P6, P10 and P12 may be replaced by maxims M1, M2 and M5 to M9. These maxims are capable of performing the same task as those principles in guiding the usability engineering of dialogue design. In fact, the maxims appear able to do the better job in view of the facts that (i) M6 and M7 spell out the intended contents of the infelicitously expressed P5, and (ii) M1 and M9 replace P6. This provides corpus-based confirmation of maxims M1, M2 and M5 to M9, i.e. of their stating basic principles of co-operative, task-oriented dialogue between humans and machines. For dialogue design purposes, however, the maxims must be augmented by task-specific or domain-specific principles such as P3. From the point of view of a general account of co-operation in dialogue, such specific principles are of secondary importance. Their significance is primarily tied to specific applications such as SLDS design.

Background Knowledge

The principles discussed in this and the following sections appear irreducible to maxims and thus serve to augment the scope of a theory of co-operativity.

P4. Take users' relevant background knowledge into account.

P4 appears to be a genuine addition to Gricean co-operativity theory, at least as far as human-machine dialogue is concerned. It is expressed at the level of generality of Grice's theory. P4 explicitly introduces two notions central to a speaker's co-operativity in dialogue. The first notion is that of interlocutors' background knowledge, i.e. skill-based knowledge of the natural language in which the dialogue is being conducted, domain knowledge etc. The second is the notion of possible differences in background knowledge between different user populations and different individual users. P4 appears to be presupposed by maxims M1, M2 and M5 to M9 in the sense that it is not possible to adhere to any of these maxims without adhering to P4. Moreover, in order to adhere to P4, it is necessary for the speaker to be able to recognise relevant differences among interlocutors and interlocutor groups in terms of background knowledge. Based on this recognition, a speaker either already has built prior to the dialogue, or adaptively builds during dialogue, a model of the interlocutor which serves to guide speaker co-operativity. Increased user adaptivity in this sense is an important goal in SLDS design (Bernsen et al., 1994b; Dybkjær et al., 1995).

As to possible overlaps with the maxims, we argue that P4 cannot be reduced to M1 (informativeness). First, M1 does not refer to the notions of background knowledge and differences in background knowledge among interlocutors. Second, a speaker may adhere perfectly to ‘exchange purpose’ (cf. M1) while ignoring important elements of the interlocutor’s background knowledge. For instance, in the user test a user wanted to order a one-way ticket at discount price. The system, however, knew that discount is only possible on return tickets. It therefore did not offer the discount option to this user nor did it explicitly remove the user’s misunderstanding. At the end of the reservation dialogue, the frustrated user asked whether or not the price quoted by the system meant that discount had been granted. Design analysis showed that similar cases may arise when, e.g., a user happens to know about a departure which was not offered by the system because the flight was already fully booked, or when a user wants discount on a certain departure but is not offered that departure because it does not allow discount. Third, as argued above, P4 is presupposed by maxims M1, M2 and M5 to M9. Grice, however, does not argue that M1 is presupposed by those maxims whereas he does argue that M3 (truth) and M4 (evidence) are presupposed by the other maxims. For similar reasons, the rather obscure M5, on relevance (see

Behavioral and Brain Sciences, 1987), obviously cannot replace P4. Informativeness and relevance, therefore, are not only functions of the purpose(s) of the exchange of information but also of the knowledge of the interlocutor.

P2. Provide sufficient task domain coverage.

P2 may appear trivial as a principle supporting the design of usable information service systems. However, designers of such systems are continuously confronted with questions, such as: should the system know this? Is this piece of information just within, or barely outside, the system's intended or expected domain of expertise? The system should behave as a perfect expert vis-à-vis its users within its declared domain of expertise. The system is at fault if there is anything it should know which it does not. In WOZ iteration 7, for example, a subject expressed surprise of not having been offered the option of being put on a waiting list in a case in which a flight was already fully booked. We became aware of the problem during the post-experimental interview. However, the subject might just as well have asked a question during the dialogue. Applied to human-human dialogue, some speakers (the experts) have more dialogue obligations in rational dialogue than others: they should not only speak truthfully and based on adequate evidence but are also expected to have complete knowledge in some domain. Since P2 deals with speaker's knowledge, it cannot be subsumed under P4. A solution is to introduce a new generic principle which mirrors P4, namely P15.

P15-NEW. Take into account legitimate partner expectations as to your own background knowledge.

P2, then, is a specific principle subsumed under P15-NEW.

P7. Take into account possible (and possibly erroneous) user inferences by analogy from related task domains.

P7 is a specific principle subsumed under P4 (background knowledge). P7 was developed from specific examples of possible user misunderstandings of the system due to reasoning by analogy. For instance, the fact that it is possible to make reservations of stand-by tickets on international flights might lead users to conclude (erroneously) that this is also possible on domestic flights.

P9. Separate whenever possible between the needs of novice and expert users (user-adaptive dialogue).

P9 is another specific principle subsumed under P4. P9 highlights the fact that interlocutors may belong to different categories with correspondingly different needs for information in co-operative dialogue. For instance, a user who has successfully used the dialogue system on several occasions, will no longer need to be introduced to the system but is capable of launching on the ticket reservation task right away. A novice user, on the other hand, will need to listen to the system's introduction to itself. This distinction between the needs of expert and novice users was introduced in WOZ iteration 7 when several users had complained that the system talked too much.

Dialogue Partner Asymmetry

The responsibility for co-operative human-machine dialogue does not only lie with the speaker (machine). The speaker may impose co-operative conditions on the interlocutor. Dialogue partner asymmetry occurs, roughly, when one or more of the dialogue partners is not in a normal condition or situation. For instance, one of the dialogue partners may have a hearing deficiency or be located in a particularly noisy environment. In such cases, dialogue co-operativity depends on the taking into account of the participant's special characteristics. As the examples show, dialogue partner asymmetry may have nothing to do with differences in background knowledge. For obvious reasons, dialogue partner asymmetry is important in SLDS dialogue design. The machine is not a normal dialogue partner and users have to be aware of this if clarification and repair meta-communication is to be avoided. Two principles lack any counterpart in Grice's theory, as applied to human-machine dialogue, because that theory does not take dialogue partner asymmetry into account.

P1. Provide clear and comprehensible communication of what the system can and cannot do.

Being limited in its task capabilities and intended for walk-up-and-use application, our SLDS needs to protect itself from unmanageable meta-communication by providing users with an up-front mental model of what it can and cannot do. In particular, it should be made clear that the system will only enable users to reserve flight tickets on domestic flights. If the up-front mental model to be acquired is too complex, users will not acquire it; and if the model is too simplistic, its remaining details will have to be provided elsewhere during dialogue. For instance, the system will explain its inability to handle special discounts for groups of more than ten people only if the user states a number of travellers that exceeds ten. Whereas an equally massive asymmetry never obtains in human-human dialogue, related asymmetries do occur when there are marked differences in the task capabilities of human dialogue partners.

P1 adds an important element to the analysis of dialogue co-operativity. P1 does not state what the system should do or avoid doing in order to behave co-operatively. Rather, P1 states that the speaker should inform the interlocutor that, unless the interlocutor takes the non-normality of the speaker into account in dialogue co-operation, dialogue success will be at risk. P1 shows that, at least in human-machine dialogue, dialogue co-operativity is a formally more complex phenomenon than anticipated by Grice. In addition to principles which state how a speaker should behave within the domain of validity of the theory, principles are needed according to which the speaker should consider transferring part of the responsibility for co-operation to the interlocutor. More specifically, the non-normal dialogue partner should inform the dialogue partners of the particular non-normal characteristics which they should take into account in order to act co-operatively.

P8. Provide clear and sufficient instructions to users on how to interact with the system.

P8 has a role quite similar to that of P1 above. What the system actually tells its users is that it will not be able to understand them if they do not answer its questions briefly and one at a time. This phrase which was introduced in WOZ iteration 7, is more emphatic than the one used in earlier iterations. Whereas the design constraint on average user utterance length (3-4 words) had already been satisfied at this point, there were still too many user utterances exceeding ten words. In WOZ 7, the number of user utterances exceeding ten words decreased to 3 from 19 in WOZ 6.

The principles examined in this section introduce two new aspects of dialogue co-operativity, namely partner asymmetry and speaker's obligation to inform the partners of non-normal speaker characteristics. Due to the latter element, P1 and P8 cannot be subsumed under any other principle or maxim. We propose that P1 and P8 are both specific principles to be subsumed under a new generic principle:

P16-NEW. Inform the dialogue partners of important non-normal characteristics which they should take into account in order to behave co-operatively in dialogue.

The term 'non-normal characteristics' refers to communication deficiencies in the speaker and aspects of the environment which impede the speaker's ability to communicate.

Conditional Principles: Feedback and Repair

The principles discussed in this section state what the co-operative speaker should do if certain types of event, specified in the principle, occur during dialogue. The principles may therefore be termed conditional principles.

P11. Be fully explicit in communicating to users the commitments they have made.

P13. Provide feedback on each piece of information provided by the user.

P11 and P13 are closely related and may be discussed together. The novel co-operativity aspect introduced by P11 and P13 is that these principles require the co-operative speaker to produce a specific dialogue contribution provided that the interlocutor has made a dialogue contribution of a certain type, such as a commitment to book a flight. Feedback is a special type of co-operative dialogue contribution in which the speaker explicitly expresses an interpretation of the interlocutor's previous dialogue contribution(s). Corresponding to the use of feedback in our system, one standard use of explicit feedback in human-human dialogue is when the interlocutor makes important commitments vis-à-vis the speaker, such as an important concession during formal negotiation. However, it is far from clear under which conditions it may be maintained that, in human-human dialogue, the co-operative speaker is expected to provide explicit feedback on information provided by the interlocutor. It is also debatable whether conditional principles such as P11 and P13 should be considered specific principles to be subsumed under M1 (informativeness) or whether, due to the new dialogue aspect they introduce, they require a new generic principle. We propose to leave this question open for future investigation.

P14. Provide ability to initiate repair if system understanding has failed.

P14 is a conditional principle which states what the co-operative speaker should do in case of failure to understand utterances made by the interlocutor. Our system adheres to P14 to the extent that it communicates its failure to understand what the user just said. The system currently lacks the ability to express, or otherwise act on, the degree of certainty it has that it correctly understood the user. Whereas the status of P11 and P13 as generic conditional principles may be questionable, P14 clearly is a generic principle of human-machine dialogue. Subsuming P14 under M1 (informativeness) would ignore a basic requirement on the co-operative speaker, i.e. that of initiating clarification and repair meta-communication in case of communication failure. As noted above, Grice ignores the fact that failure to adhere to the maxims may give rise to clarification and repair meta-communication. P14 may be replaced by the slightly revised P14*:

P14*. Initiate repair or clarification meta-communication in case of communication failure.

Concluding Discussion

The Gricean maxims are valid claims that apply to shared-goal, spoken human-human dialogue and in this paper have been empirically validated for task-oriented, spoken human-machine dialogue. Similarly, the generic and specific principles have been empirically validated for task-oriented, system-directed, spoken human-machine dialogue. As to the specific principles, it does not appear warranted to claim their applicability beyond human-machine dialogue. The question to be addressed below concerns the scope of the non-Gricean generic principles. Is their scope similar to that of Grice's maxims?

We have as yet no empirical basis for claiming that the generic principles P4, P14*, P15 and P16 share the scope of the Gricean maxims. Theoretically, however, it may be argued that they do. P16 (asymmetry) appears equally valid for human-human dialogue. If a partner in shared-goal dialogue has important non-normal characteristics of which the interlocutor may be unaware, and the non-observation of which is detrimental to the achievement of the goal, then the interlocutor must be informed about them. Otherwise, dialogue co-operativity will be decreased until the interlocutor discovers those characteristics. P16 is irrelevant to symmetrical human-human dialogue. However, when an asymmetry is present, the principle assumes a fundamental role. Suppose, for instance, that ambient noise prevents me from clearly following my interlocutor's dialogue contributions. If I deliberately omit to inform the interlocutor that this is the case, the conversational implicature is that I do not really care to accomplish the goal of our dialogue. Thus, ignoring P16 is likely to carry rather primitive conversational implicatures such as "I don't care what you are saying" or "I don't care about achieving the goal of our dialogue".

Ignoring a partner's relevant background knowledge (P4) clearly detracts from the speaker's dialogue co-operativity. The same is true when a speaker ignores legitimate partner expectations as to the speaker's background knowledge (P15). In both cases, the conversational implicature is that something else than the shared goal is at stake.

Furthermore, as argued earlier, P4 is presupposed by the Gricean maxims. This implies that P4 shares the scope of the maxims. P14* (repair and clarification) appears to share this scope. The non-observation of P14 carries as negative conversational implicatures as the non-observation of P16 (see above).

In conclusion, when performing shared-goal dialogue, people, just like machines, should communicate their communication deficiencies, take background knowledge into account and initiate repair and clarification meta-communication when needed. Failure to do so detracts, sometimes seriously, from the rationality of the shared-goal dialogue.

On one condition, the maxims and generic principles are in fact symmetrically applicable to SLDSs and their human users. The condition is that users should assign priority to what the machine says according to P16 (asymmetry) and otherwise make their dialogue contributions conform to the maxims and generic principles. As applied to our SLDS, P16 essentially informs users that the system is severely constrained in its language understanding capabilities, background knowledge, inferential abilities and acceptance of user initiative. If these limitations are respected, the system will enable otherwise co-operative users to complete their task.

References

Ahluwalia, N., Agnihotri, R.K. & Subbarao, K.V. (1990). Gricean maxims and conversational goals. International Journal of Dravidian Linguistics 19, 2, 111-25.

Behavioral and Brain Sciences (1987). Précis of relevance, communication and cognition with open peer commentary. Behavioral and Brain Sciences 10, 4, 697-754.

Bernsen, N.O. (1993). Types of user problems in design. A study of knowledge acquisition using the Wizard of Oz. Esprit Basic Research project AMODEUS-2 Working Paper RP2-UM-WP14. In Deliverable D2: Extending the User Modelling Techniques. June.

Bernsen, N.O., Dybkjær, H. & Dybkjær, L. (1995). Exploring the limits of system-directed dialogue. Dialogue evaluation of the Danish dialogue system. Proceedings of Eurospeech '95, Madrid, September.

Bernsen, N.O., Dybkjær, L. & Dybkjær, H. (1994a). Task-oriented spoken human-computer dialogue. Report 6a, Spoken Language Dialogue Systems, Roskilde University, February.

Bernsen, N.O., Dybkjær, L. & Dybkjær, H. (1994b). A dedicated task-oriented dialogue theory in support of spoken language dialogue systems design. Proceedings of ICSLP '94, Yokohama, September, 875-878.

Bilange, E. (1991). A task independent oral dialogue model. Proceedings of the 5th EACL, Berlin, April, 83-88.

Cole, R., Novick, D.G., Fanty, M., Vermeulen, P., Sutton, S., Burnett, D. & Schalkwyk, J. (1994). A prototype voice-response questionnaire for the US Census. Proceedings of the ICSLP '94, Yokohama, September, 683-686.

Dybkjær, H., Bernsen, N.O. & Dybkjær, L. (1993). Wizard-of-Oz and the trade-off between naturalness and recogniser constraints. Proceedings of Eurospeech '93, Berlin, September, 947-950.

Dybkjær, L., Bernsen, N.O. & Dybkjær, H. (1995). Different spoken language dialogues for different tasks. A task-oriented dialogue theory. Human Comfort and Security, Springer Research Report, Springer Verlag (in press).

Dybkjær, L. & Dybkjær, H. (1993). Wizard of Oz experiments in the development of a dialogue model for P1. Report 3, Spoken Language Dialogue Systems, Roskilde University, February.

Eckert, W. & McGlashan, S. (1993). Managing spoken dialogues for information services. Proceedings of Eurospeech '93, Berlin, September, 1653-1656.

Fraser, N.M. & Gilbert, G.N. (1991). Simulating speech systems. Computer Speech and Language 5, 81-99.

Goddeau, D, Brill, E., Glass, J., Pao, C., Phillips, M., Polifroni, J., Seneff, S. & Zue, V. (1994). Galaxy: A human-language interface to on-line travel information. Proceedings of the ICSLP '94, Yokohama, September, 707-710.

Grandy, R.E. (1989). On Grice on language. The Journal of Philosophy 86, 10, 514-25.

Grice, P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), Syntax and semantics Vol. 3: Speech acts (pp. 41-58). New York: Academic Press. Reprinted in Grice, P.: Studies in the way of words. Cambridge, MA: Harvard University Press 1989.

Grice, P. (1978). Further notes on logic and conversation. In P. Cole (Ed.), Syntax and semantics Vol. 9: Pragmatics (pp. 113-128). New York: Academic Press. Reprinted in Grice, P.: Studies in the way of words. Cambridge, MA: Harvard University Press 1989.

Horn, L.R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), Meaning, form and use in context: Linguistic applications (pp. 11-42). Washington, DC: Georgetown University Press.

Huang, Y. (1991). A neo-Gricean pragmatic theory of anaphora. Journal of Linguistics 27, 301-35.

Levinson, S.C. (1987a). Minimization and conversational inference. In Verschueren, J. & Bertuccelli-Papi, M. (Eds.), The Pragmatic Perspective (pp. 61-129). Amsterdam: John Benjamins.

Levinson, S.C. (1987b). Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena. Journal of Linguistics 23, 379-434.

Mazor, B., Braun, J., Ziegler, B., Lerner, S., Feng, M.-W. & Zhou, H. (1994). OASIS - a speech recognition system for telephone service orders. Proceedings of the ICSLP '94, Yokohama, September, 679-682.

Oerder, M. & Aust, H. (1994). A realtime prototype of an automatic inquiry system. Proceedings of the ICSLP '94, Yokohama, September, 703-706.

Peckham, J. (1993). A new generation of spoken dialogue systems: Results and lessons from the SUNDIAL project. Proceedings of Eurospeech '93, Berlin, September, 33-40.

Pratt, M.L. (1981). The ideology of speech act theory. Centrum, New Series, 1, 5-18.

Sarangi, A.K. & Slembrouck, S. (1992). Non-co-operation in communication: A reassessment of Gricean pragmatics. Journal of Pragmatics 17, 117-154.

Schegloff, E.A., Jefferson, G. & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. Language 53, 361-82.

Searle, J.R. (1992). Conversation. In Searle, J.R. et al. (Eds.), (On) Searle on conversation. Amsterdam: John Benjamin's Publishing Company.

Sperber, D. & Wilson, D. (1986). Relevance, communication and cognition. Oxford: Basil Blackwell.

Footnote 1

Center for PersonKommunikation, Aalborg University, and the Centre for Language Technology, Copenhagen.

Acknowledgements

Thanks are due to Judith Ramsay who provided comments on an earlier draft. We also wish to thank the editor, Arthur Graesser, and the three anonymous reviewers who provided many valuable criticisms and suggestions on the draft submissions. The spoken language dialogue system is being developed on a grant from the Danish Technical Sciences Research Council. The analysis of designer problem solving which produced the co-operativity principles was done partly in the Esprit Basic Research project AMODEUS-2. We gratefully acknowledge the support.

Figure captions

Figure 1. The opening part of the state transition network from WOZ iteration 6.

Figure 2. The opening part of the state transition network from WOZ iteration 7. The network changes from WOZ iteration 6 to WOZ iteration 7 were mostly based on identified user problems and violated principles of co-operativity.

Table 1

The SLDS dialogue design principles and their justifications

Principles	**Justification**
P1. Provide clear and comprehensible communication of what the system can and cannot do.	Risk of communication failure in the case of lacking knowledge about what the system can and cannot do. Violation of this principle leads users to have exaggerated expectations about the system's abilities, which again may lead to frustration during use of the system.
P2. Provide sufficient task domain coverage.	Risk of communication failure in case of lacking task domain information. Full task domain coverage within specified limits is necessary in order to satisfy all relevant user needs in context. Otherwise, users will become frustrated when using the system.
P3. Provide same formulation of the same question (or address) to users everywhere in the system's dialogue turns.	Need for unambiguous system response (consistency in system task performance). The principle is meant to reduce the possibility of communication error caused by users' understanding a new formulation of a question as constituting a different question from one encountered earlier.
P4. Take users' relevant background knowledge into account.	Need for adjustment of system responses to users' relevant background knowledge and inferences based thereupon. This is to prevent that the user does not understand the system's utterances or makes unpredicted remarks such as, e.g., questions of clarification, which the system cannot understand or answer.

Principles (continued)	**Justification** (continued)
P5. Avoid 'semantical noise' in addressing users.	Need for unambiguous system response. The design commitment is to reduce the possibilities of evoking wrong associations in users, which in their turn may cause the users to adopt wrong courses of action or ask questions which the system cannot understand.
P6. It should be possible for users to fully exploit the system's task domain knowledge when they need it.	Risk of communication failure in case of inaccessible (or not easily accessible) task domain information. In such cases, users may pose questions which the system is unable to understand or answer.
P7. Take into account possible (and possibly erroneous) user inferences by analogy from related task domains.	Need for adjustment to users' background knowledge and inferences based thereupon. Users may otherwise fail to understand the system.
P8. Provide clear and sufficient instructions to users on how to interact with the system.	Risk of communication failure in case of unclear or insufficient instructions to users on how to interact with the system. Users may become (or remain) confused about the functionality of the system.
P9. Separate whenever possible between the needs of novice and expert users (user-adaptive dialogue).	There are major differences between the needs of novice and expert users, one such difference being that expert users already possess the information needed to understand system functionality.
P10. Avoid superfluous or redundant interactions with users (relative to their contextual	Need for non-superfluous interaction with the system.

needs).	

Principles (continued)	**Justification** (continued)
P11. Be fully explicit in communicating to users the commitments they have made.	Users need feedback from the system on the commitments made.
P12. Reduce system talk as much as possible during individual dialogue turns.	Users get bored and inattentive from too much uninterrupted system talk.
P13. Provide feedback on each piece of information provided by the user.	Immediate feedback on user commitments serves to remove users' uncertainty as to what the system has understood and done in response to their utterances.
P14. Provide ability to initiate repair if system understanding has failed.	In case of system understanding failure, the system should initiate repair meta-communication rather than leave the initiative with the user.