

Conversation with and through Computers

SUSAN E. BRENNAN

*Dept. of Psychology, State University of New York,
Stony Brook, NY 11794-2500*

(Received 30 May 1990; in final form 21 August 1990)

Abstract. People design what they say specifically for their conversational partners, and they adapt to their partners over the course of a conversation. A comparison of keyboard conversations involving a simulated computer partner (as in a natural language interface) with those involving a human partner (as in teleconferencing) yielded striking differences and some equally striking similarities. For instance, there were significantly fewer *acknowledgments* in human/computer dialogue than in human/human. However, regardless of the conversational partner, people expected *connectedness* across conversational turns. In addition, the style of a partner's response shaped what people subsequently typed. These results suggest some issues that need to be addressed before a natural language computer interface will be able to hold up its end of a conversation.

Key words: discourse modeling, human/computer interaction, natural language interfaces, recipient design.

1. Conversations with Computers

Why is it that natural language has yet to become a widely used modality of human/computer interaction? Visionaries seem to have no difficulty imagining a future where we'll be able to talk to software applications – or even computer *agents* – in plain English (Laurel, 1990). And yet the only exposure large numbers of users have had to such interfaces has been through limited question answering systems and keyword interfaces to adventure games. There is pessimism as to whether there will ever be a useful natural language interface technology, particularly among those who advocate direct manipulation and desktop interfaces (Schneiderman, 1981, 1982). On the other hand, there is optimism among those computational linguists who have made significant progress in formalizing the structure and semantics of sentences and mapping them onto database query or command languages. Despite significant progress, there is still the stark reality of human language *use*. That is, people don't always speak in grammatical sentences. They are often indirect; they don't say what they mean. They rely on the unspoken knowledge and presuppositions they share with their conversational partners. For these reasons, pessimists argue that natural language is inherently ambiguous (we might wonder how people have managed all these years without being native speakers of predicate logic). It has been argued that natural

User Modeling and User-Adapted Interaction 1: 67–86, 1991.

© 1991 Kluwer Academic Publishers. Printed in The Netherlands.

language understanding is just too hard a problem to solve in any robust, formal, domain-independent way. This may be so. But a robust, formal, domain-independent solution may also be unnecessary. The question is: can natural language be a useful interface technology *despite* all these unsolved problems?

Consider first how people manage to make themselves understood. To speakers and addressees, a conversation can seem as concrete as a desktop. This is because people in conversation build mental models of what they are discussing. They can then share and manipulate and refer to elements of these models. In addition, language use is opportunistic. People are able to adapt to conversational partners who differ from themselves, and they have well-developed strategies for collaborating with others to establish that they've been understood. How do these strategies work? Can any of them be carried over from human/human communication to human/computer interaction?

2. Adapting to Addressees

When people talk to each other, they tailor their utterances to their partners (Clark and Murphy, 1982; Bell, 1984; Milroy, 1987); this is what is meant by *recipient design* or *audience design*. People can accommodate a wide range of partners: friends, strangers, the hard of hearing, disembodied voices on the telephone, readers who will come along after they are gone, foreigners, children, and even dogs. This flexibility in designing utterances for a particular addressee has been documented even among the very young. For instance, five year olds have been observed to use more simple language and a different pitch range when talking to three year olds than they do talking to adults (Shatz and Gelman, 1973). Likewise, we should expect people to talk differently to a computer than to another person, depending on the kind of language input they think the computer can handle.

On the other hand, some of the mishaps people experience in their dealings with computers may result from carrying over some of their expectations from human conversations into dialogues with computers (Brennan, 1990). For instance, polite and informal responses from a computer system have been known to shape commands from users to be more polite and informal (Richards and Underwood, 1984). Take the following example of a Stanford undergraduate who was trying to communicate with Socrates, an on-line library program:

Socrates: PLEASE SELECT COMMAND MODE
 Student: Please find an author named Octavia Butler.
 Socrates: INVALID FOLIO COMMAND: PLEASE

Here, user friendliness in the form of gratuitous politeness backfired when the student tried to reciprocate. Note that the error message was successful;

User: *Who has a terminal?*

System: *DAN FLICKINGER HAS A TERMINAL.*

Fig. 1. Adjacency Pair from the HPNL Program (Brennan, 1988)

afterward the student was able to repair the problem. But the problem came up in the first place because the student expected the system to *use* only words it could *understand*. We generally expect this from our conversational partners, unless they are reading from a Berlitz phrasebook.

A natural language computer interface is an unusual kind of linguistic partner. A user must figure out what its limitations are and then design utterances with these limitations in mind. Limitations may be due to gaps in the system's linguistic coverage or else to unexpectedly narrow coverage of the application's vocabulary and domain. But often these limitations are well hidden from the user, who consequently cannot employ the powerful strategies available in a human/*human* conversation for establishing meaning.

In conversation, contiguous turns are likely to be relevant to one another; whatever one partner does after another takes a turn may be interpretable as a response to that turn. This relevance between pairs of conversational turns has been expressed by considering them as *adjacency pairs* (Schegloff and Sacks, 1973; Sacks et al., 1974). For instance, a question/answer sequence is an adjacency pair, with the question turn as the first part and the answer turn as the second part. An adjacency pair is not a fixed ritual, but a sequence of turns relevant to the same conversational purpose. While a question is not *always* followed by an answer, it *often* is, and when it isn't, it is likely to be followed by another relevant turn such as a request for clarification or a side sequence.

In conversation, people adapt to one another. For instance, they establish local conventions for referring expressions (Clark and Wilkes-Gibbs, 1989; Isaacs and Clark, 1987; Garrod and Anderson, 1987) and become more similar in the language they use over the course of a conversation (Isaacs and Clark, 1987; Milroy, 1987). Answers tend to contain syntactic constructions and lexical items which are repeated from the questions they follow (Levitt and Kelter, 1982). Such *lexical entrainment* is embodied in some question-answering computer applications, as illustrated in Figure 1.

In human language use, entrainment may happen for a variety of reasons: because of lexical and syntactic priming (Bock, 1986), or else because the conversational partners have negotiated a common model (Garrod and Anderson, 1986) or a single perspective (Schober and Clark, 1989). In any event, using the same words in a conversational context is one way partners can know and convey that they are talking about the same thing. In addi-

User: *Who has a terminal?*

System: *Shall I do the following? Create a report showing the full name and the manager and the equipment from the forms on which the equipment includes 'TERMINAL'?*"

Fig. 2. Dialogue from Q&A (Hendrix, 1985)

tion, if a computer dialogue partner were to keep track of how a user refers to objects in the domain, in order to use the same referring expressions itself and to expect similar references from the user in the future, then the search space for potential referents would be reduced for both dialogue partners.

Not all question-answering systems mark relevant adjacent turns in this way. Figure 2 shows the kind of response provided by a commercially available natural language interface to a database. Such paraphrases are not particularly conversational. To understand them, the user must translate her query into procedural or logical language (Brennan, 1990; Carbonell and Hayes, 1987). Paraphrases should be available upon request, if the user needs reassurance or more detail about how the system interpreted a particular query. However, the response style in Figure 2 also requires explicit confirmation after every query. Confirmation is appropriate when the system's interpretation is incomplete or ambiguous, or when the consequences of a particular interpretation are potentially destructive. But requiring this extra work from the user after every turn is tedious.

Every contribution to a conversation must be understood and accepted before the participants go on; this follows from the *collaborative view* (Clark and Wilkes-Gibbs, 1986; Clark and Murphy, 1987; Isaacs and Clark, 1987; Clark and Schaefer, 1989; Clark and Brennan, 1990). But an acceptance turn need not be explicit - one can *implicitly* accept what a conversational partner says simply by going on with the next relevant turn. Implicitly or explicitly, people seek and provide evidence of their understanding, step by step; this is the process of *grounding* (Clark and Brennan, 1990). Grounding includes back-channels, monitoring and displaying attentiveness (as indicated by eye contact), and all of the other things people do to check on one another's understanding. Whether two people are having an argument, moving a piano, or dancing a tango, each monitors what the other is doing. They use all the evidence they can get, including linguistic evidence, visual evidence, and tactile feedback. Grounding is an important part of any coordinated activity between two partners, however different they may be from one another.

Direct manipulation interfaces handle the grounding problem well, at least for simple tasks that can be concretely represented; a user can get back continual evidence about how the system has "understood" an action. However, in other kinds of human/computer dialogues, there are fewer

opportunities for grounding than in human/human communication, since interfaces typically do not provide information simultaneously in parallel channels, and since people may not expect a system to take any responsibility for the success of the interaction (and many systems are incapable of doing so). People learn to look for indirect evidence that their query or command has been understood (for examples of indirection in Unix, see Norman (1981). Most of the responsibility for grounding a human/computer dialogue typically falls on the user. Some features of a computer partner may facilitate grounding, for instance: the FTP file transfer program that keeps us posted as the bits go by, the run bars at the bottom of a LISP process that provide reassurance that our program hasn't crashed yet, and the flashing red light or whirring sounds from a disk drive that reveal the good news or the bad news about what a program is doing to a file system. But features such as these are usually not part of any premeditated architecture for supporting human/computer dialogue. Such features arise all too infrequently, and when they do, it is too often due to serendipity or mutation, as opposed to good design.

3. What do you say to a natural language interface?

Consider the view that people approach using language to a computer in light of their experiences using language with other people. This view suggests that choice of words and syntax may emerge from strategies such as grounding, expectations about adjacency, and adaptations such as entrainment. An alternative hypothesis is that people may simply use a predictably restricted subset of natural language to a computer partner, as has been suggested by Guindon et al. (1987) and others. Whether this is true has practical consequences for research and development; which of the hard problems should computational linguists concentrate their efforts on? For instance, are pronouns really a crucial area of coverage, or will people tend not to use them to a computer partner? In one experiment which simulated a natural language interface to a statistics package, people used very few pronouns (Guindon et al., 1987). To understand whether this result really means that there is an assumption of poor shared context (as claimed in Guindon et al., 1987) or whether the task just didn't lend itself to using pronouns, we must examine people's behavior in a task where co-reference between turns is possible. That is, pronoun use with a computer partner should be compared to that with a human partner, in situations where the task underway is the same.

The collaborative model would predict that any restricted subset of natural language used with a computer partner will not be a robust subset across situations. That is, the text a user is likely to type to a computer partner cannot be represented by a predictable set of defaults. Word and wording

choices will depend instead on what people are trying to do, their model of their partner's abilities, what they have negotiated so far, whether they understand their partner's last response, and whether they think their partner understood theirs. How do these factors influence a user to choose particular words in a particular situation? Understanding these choices should help us design natural language interfaces (and other kinds of interfaces as well) that present themselves honestly as the limited conversational partners they are. It may be possible for a natural language interface to provide invisible constraints that prevent a user from falling off the edge, for instance, by generating its responses to use only those grammatical constructions and lexical items that the system itself will be able to interpret.

Considering the conversational strategies people use with one another then leads us to the following predictions. First, answers that show lexical and syntactic parallelism with queries should be better than short answers or ones which do not preserve parallelism. Second, short answers should beget short queries and complete sentence answers should beget complete sentence queries. Third, if general expectations about connectedness between turns in a discourse is the main factor governing pronoun use, rather than simply a tendency to use fewer pronouns with computer partners, there should be just as many pronouns used with computer partners as with human ones. Finally, people should be able to treat a dialogue with a computer partner differently than one with a human partner; for instance, they shouldn't expect much *social context* with a computer partner. However there is still a *dialogue context*, and people should be able to use some of the same strategies in adapting to a computer partner as they would with a human partner, over the course of a dialogue. The design of an utterance should depend on at least two factors: 1) the model a person has in mind of her conversational partner at the outset, and 2) how this model changes over the course of the conversation, as an adaptation to to the kinds of responses her partner makes.

4. Method

To examine how people formulate utterances to a natural language interface, I designed an experiment which varied the type of partner – *human* or *computer* – and also the style of responses, in the context of a database query task. There were three styles of responses: 1) A *short* response to the query "What is Aida's profession?" would be "engineer"; 2) a *sentence* response would be "Aida's profession is engineer" and 3) a *lexical change* response would be "Aida's job is engineer". The second response style, a complete sentence that closely paralleled the words and syntax of the query, was based on the strategy embodied in HPNL, a natural language interface at Hewlett-Packard Labs (Brennan, 1988). The third style, lexical change,

was an attempt to explore the effects of canned messages or responses generated from logical form without heed to the particular lexical items used in the adjacent query. The experiment design is summarized in Table I.

Since recovery from errors was beyond the scope of this study, I chose a *Wizard of Oz* technique (Guindon et al., 1987) that allowed me to avoid errors by simulating the natural language interface in the computer partner condition.

TABLE I
Experiment Design

		Style		
		Short	Sentence	Lex change
Partner	Human	n = 6	n = 6	n = 6
	Computer	n = 6	n = 6	n = 6

4.1. SUBJECTS

Subjects were 44 Stanford students, most of whom had little or no experience in using computers except for word processing and electronic mail. They volunteered for the experiment in exchange for \$5 or as part of their participation in an introductory psychology class.

4.2. MATERIALS

Materials included a small database of six fictitious people and their attributes, such as nationality, profession, hobby, etc. The database was displayed as an array printed on a piece of paper with six names filled in across the top. Missing were two out of six values for each of the attributes. The attributes were unlabeled, but their categories were obvious from the values that were filled in. Subjects had to generate their own labels for the information they requested. The database array and the instructions for both partner conditions are reproduced in Appendix A.

Two computer conferencing programs were written for the experiment, one for each partner condition. They were identical except for the prompts printed on the subject's screen. Subjects assigned to human partners saw the prompts " $=>$ " when it was their turn to type and "Partner:" preceding their partner's response, while those with computer partners saw "User:" and "System:". The experimenter saw the same prompts in both partner conditions: " $->$ " and "Subject:". These conferencing programs were similar to Unix's *write* program; typed characters were not sent to a partner's screen

until the typist hit carriage return. The text generated in each session was automatically saved into a file.

The conversations were conducted over two terminals connected to a Hewlett-Packard 9000 series 300 workstation, with the subject in one room and the experimenter in another.

4.3. PROCEDURE

Subjects were run individually. Just before a session began, an experimental confederate randomly assigned the subject to a partner condition and a style condition. The confederate then started up the appropriate conferencing program and informed the experimenter which style to use. I played the role of the experimenter or "wizard". In an attempt to avoid being biased to act particularly "human" or particularly "computer-like", I was blind to which partner condition a subject had been assigned to. It was important that responses to the subject's questions be consistent across both partner conditions. In my role as wizard, I had a copy of the database chart with all the values filled in and a well-rehearsed set of rules to use in generating the answers for each style condition.

The confederate greeted the subjects and told them that they were about to perform an information-gathering task. Half of the subjects were told that they would be using a natural language system; the other half were told that they would be using the terminal to communicate with another person who was networked to the same computer and who could look up the information in a booklet. The instructions for both partner conditions were identical except for the part necessary to establish the cover stories. Each subject was given the database chart and instructed to find the missing information, one item at a time.

As wizard, I tried always to understand the literal content of what a subject typed (including ellipses, pronouns, typos, and ungrammatical constructions). I tried to answer only what was explicitly asked, and I provided the information from one cell of the database chart at time. I was careful not to produce responses with typos or misspellings. I typed a response quickly, checked it for errors, and then sent it to the subject's terminal. I avoided using sentence-initial capitalization or sentence-final punctuation. For generating the responses, I used the rules in Table II.

It was evident that the wizard needed a policy for getting the task started and responding to subjects who ventured beyond the task or felt playful. So some "canned" responses were used: starting the conversation with "is anyone there?"; greeting the partner with "hello," and then, "what do you want to know?" after the subject had an opportunity to greet in return; answering only the first part of a compound request and then typing "-what else?"; "what do you want to know?" for extremely general requests, "what was

TABLE II
Rules for Generating Responses

Style	Response
1. Short	A noun phrase
2. Sentence	Use request as a template and fill in the answer. Begin any yes/no questions with "Yes" or "No". Avoid the phrase "kind of" (in answering "What kind of...").
3. Lex change	Give sentence answers, except make one lexical change. If necessary, this may involve a syntax change also, but try not to. Keep track of lexical changes and use the same word consistently.

that again?" for anything unintelligible, "I can't answer that" for questions outside of the task domain (such as "who is this?"); "you're welcome" after "thank you" and "goodbye" or "bye" after any closing remark; and "no comment" when all else failed (as it did when one person asked "*perhaps we could go out some time?*").

There were 22 blank cells on the database query form and it took subjects an average of 27 turns (including opening and closing turns) to gather the information. After completing the database query task, subjects filled out a paper and pencil questionnaire. They answered questions concerning their experience using computers and their experience during the experiment. While a subject was doing this part of the task, the confederate informed me whether I had been playing the role of human or computer.

Finally, we interviewed subjects at length to determine what they thought the experiment was about and whether they really believed what they had been told about their partner. Then we debriefed them thoroughly. After a subject left, the confederate and I went over the questionnaire and our notes and then decided whether we thought the subject had believed the cover story. On that basis, we had to discard the transcripts of six unbelievers. We also discarded one because he failed to do the task, and another because of equipment failure. We ran the experiment until there were six believers per cell and analyzed only the transcripts from those 36 subjects.

4.4. RESULTS AND DISCUSSION

Are people as likely to use third-person pronouns (such as he, she, him, her, his, it) with a natural language interface as with another person? They are. There was no difference in the number of third-person pronouns between the two partner types ($F(1,30) = .165, p = .69$). Each transcript included 0–25 third-person pronouns, with a mean of 9.3 pronouns ($sd = 8.2$) used with computer partners and 8.3 pronouns ($sd = 7.1$) used with human partners. 93% of these pronouns referred to individuals and entities on the database chart that had been mentioned in a previous sentence; 85% of these pronouns referred to individuals or entities mentioned in a previous turn (some turns contained more than one sentence). A closer look at third-person pronouns by partner type showed that 97% of the pronouns used to computer partners were intersentential, which did not differ significantly from the 89% of intersentential pronouns used to human partners. The pronouns used to computer partners were actually more likely to co-refer with noun phrases in previous turns than were the pronouns used to human partners (95% to 74%, $\chi^2(1) = 8.7, p < .005$). This is evidence that people expected connectedness between sentences and turns, regardless of whether they believed they were talking to a computer or another person. Such a result fails to support Guindon's (Guindon et al., 1987) finding, that people tend not to use pronouns to a natural language computer interface because of a general assumption of poor shared context.

On the other hand, first-person and second-person pronouns (such as *I*, *you*, and *me*) were much more likely to be used with human partners (an average of 5.3 times) than with computer ones (an average of 0.6 times, $F(1,30) = 13.04, p < .002$). These pronouns serve a very different purpose than do third person pronouns, since they do not get their interpretations from the linguistic context supplied by previous turns. These pronouns are often meta-conversational, e.g. "can you tell me what Aida does for a living?". They often appear in indirect queries, which are typically more polite. They also acknowledge the social context of the conversation by explicitly specifying the typist and the addressee. Not surprisingly, this social context was mostly absent with a computer partner.

Context bears also on the use of ellipses, e.g. "what about Ellen's?" People used ellipses an average of 4.2 times in dialogues with a human partner, to only 1.2 times with a computer partner ($F(1,30) = 4.838, p < .04$). So it seems likely that people expect some but not all aspects of context to be shared with computers.

To what extent did people adapt by modifying their utterances over the course of the conversation to be more like the responses of their partners? There are several ways that one could adapt to a partner; one way is by using or not using grammatically complete sentences. Each turn was coded

as either a sentence or a non-sentence (which included all ungrammatical input such as key words, telegraphic input and phrases). Change over time was determined by comparing the percentage of sentential turns for the first half of each dialogue to the percentage for its second half, for both the short and sentence style conditions ($N = 24$). Across response styles, there was no significant difference in the percentage of complete sentence turns in the first half of the dialogues (nor was one expected), but there *was* a difference in the second half, in the direction predicted. Regardless of whether a partner was supposed to be human or computer, the partner's complete sentence answers evoked more complete sentence questions in the second half of the dialogue than in the first half, and short answers evoked more short (phrasal or keyword) questions in the second half than in the first ($F(1,20) = 6.6$, $p < .02$). This interaction between response style and dialogue half is shown in Table III.

TABLE III
Syntactic Entrainment: Percentage of Sentence Turns

		First half of dialogue	Second half of dialogue
Response style	Short	76.07	65.62
	Sentence	72.54	82.89

People used different initial strategies in forming queries, depending on who they thought their partners were. The first query in each dialogue was always a complete sentence with human partners, whereas with computer partners, half the time the first query was a phrase or key words. This difference may reflect not only people's initial models of what kind of linguistic input the computer could handle, but also a tendency to be polite to human partners. However, by the last half of each dialogue, the mean percentage of complete sentences was no different across both kinds of partners, and was affected only by whether the response style was short or sentential. These results support the prediction that the design of utterances is shaped both by the initial model of the partner and also by the partner's responses, and that people can use responses to tune their models of a partner. Exactly what kinds of models people bring to human/computer dialogues is beyond the scope of this study.

Recall that *grounding* is the process by which speakers and addressees establish that they've understood one another, by seeking and providing evidence of their understanding. Among the ways people can provide evidence is with a verbatim repetition of portion of a partner's utterance, a relevant next turn, or an explicit acknowledgment (Clark and Schaefer, 1989; Clark and Brennan, 1991). I counted the number of turns that contained explicit

acknowledgments, such as the *OK* in “*OK*, now what kind of vehicle does Takis drive?” and the *Nice pet!* in the sequence, “Takis owns a pit bull”, “*Nice pet!* What about Megumi?”

Turns directed at human partners contained significantly more acknowledgments than those directed at computer partners ($F(1,30) = 21.02$, $p < .001$). The response style also made a difference: both the lexical change and short answer conditions showed more explicit acknowledgments than the sentence style condition ($F(2,30) = 7.68$, $p < .002$). One possibility is that the sentence answers, which very closely paralleled the queries, provided enough *implicit* evidence (in the form of repetition) to ground answers with queries, so that more explicit grounding was unnecessary. An alternative possibility (which cannot be distinguished from the first one on the basis of these data) is that sentence answers were somewhat monotonous, which may have caused subjects to stick more closely to the task and give fewer spontaneous acceptances. The interaction between partner and style was significant as well ($F(2,30) = 5.71$, $p < .01$). That is, it was when the partner was assumed to be human that the number of acknowledgments people gave depended on the partner’s response style. The most acknowledgments occurred in the lexical change style with a human partner; recall that this was the condition where the human partner responded by using different terms than the subject had used. This may reflect an increased effort to ground a conversation that seems to be going strangely. The mean number of explicit acknowledgments per conversation by partner and response style is in Table IV.

TABLE IV
Mean Number of Explicit Acknowledgments per Conversation

		Response style		
		Short	Sentence	Lex change
Partner	Human	3.3	1.17	8.00
	Computer	0.00	0.67	1.00

5. Conclusion

This research bears on some practical issues which have consequences for natural language interfaces. How does one partner adapt to another over the course of a dialogue? Is this adaptation different with a human partner than with a human partner? How can users be constrained to type input that a system can handle? How should responses be articulated? Is there

an expectation of connectedness across sentences? When users restrict their input, are their choices predictable as a kind of default behavior, or do their choices depend on the discourse context?

When using natural language, people treat computer partners differently than they do human partners, as well they should. Nevertheless, expectations about conversational interaction still play a significant role in human/computer dialogue. This experiment demonstrates that tailoring utterances for a partner is affected both by the initial model of the partner and by the partner's subsequent responses. In general, people were very sensitive to response style. The six whose dialogues were excluded because they didn't believe the cover story were *not* in the computer partner condition, as I had expected would be the case. Every one of them was in the human partner condition - that is, they refused to believe that their partner was indeed human. Apparently, these subjects detected that the responses were rule-based. In contrast, none of the subjects in the computer partner condition (the deception condition) had any difficulty whatsoever believing their partner was a computer that could interpret natural language. Judging informally from the post-task interviews, these subjects seemed to enjoy the experiment more.

As a dialogue proceeded, people adapted to their partners by designing queries that were more similar to their partners' responses. Complete sentence answers evoked complete sentence queries, and short phrasal answers evoked non-sentence queries. This form of adaptation or *syntactic entrainment* may seem a little surprising in the computer partner condition, since there should have been no practical or social reason to adapt. Recall that nearly every query was accepted and understood by the computer partner, and people always got some kind of contingent response. So this adaptation may not be an intentional strategy for trying to type only what the computer would understand. It may be a syntactic priming effect or it may be a well-practiced way of indicating to a partner that a model or perspective has been accepted.

This study has several implications for natural language interfaces. First, representing dialogue context in the form of a discourse model is essential. The turns taken by the user and the system should be treated as connected, contingent sequences and not as disembodied grammatical sentences composed of strings of characters; simply translating a string into a logical formula does not a conversation make. People expect connectedness, whether the partner is a computer or another person. Second, while there are many unsolved problems facing those who develop natural language interfaces - knowledge representation, commonsense reasoning, and plan recognition, to name only a few - these problems need not be entirely solved before natural language can be used as a computer interface technology. In addition to using the linguistic structure of an utterance, a natural language inter-

face should exploit the interactive, collaborative nature of language use, to increase the odds of successful human/computer dialogue. Third, a system should present itself (and its limitations) in a way that is informative about what kind of conversational partner it is. The response style of the system can be used to subtly constrain and ground the queries the user types. Many natural language parsers can handle only grammatical sentences as input; however, even users who are aware of this sometimes type strings which are not well-formed (Carbonell, 1979). So if a system cannot interpret partial input, answers to queries should be delivered as grammatical sentences; then people can use their natural conversational strategies to adapt to this strange partner and formulate utterances that it can handle.

Future work on natural language interfaces should support both the user's and the system's ability to negotiate conversational repairs, an important and unavoidable part of language use, especially with a strange partner. While enabling systems to adapt to users is an important long-range goal in making natural language a viable interface technology and in the invention of "intelligent" interfaces, such interfaces should also exploit the natural ability of human beings to adapt to their communicative partners. This ability to adapt is what makes the whole human/computer enterprise possible in the first place.

6. Acknowledgements

This work was supported by a National Science Foundation predoctoral fellowship and by Hewlett-Packard Labs. Thanks to Steve Lowder for software assistance. I also thank Herb Clark, Arvind Joshi, Diana Roberts, Bridget Bly, the Stanford Language Users Group (SLUGs), Hewlett-Packard's Human-Computer Interaction Department, and all those who attended the Second International Workshop on User Modeling.

Appendix A: Database query task

Each subject in the experiment was randomly assigned to a partner type and a style condition, and saw one of the two sets of instructions below. All subjects saw the same database array. Those in the computer partner condition were given these instructions:

Please use the natural language computer program provided and discover the missing information represented by each blank cell on this form. The program will interpret what you type and then respond to you in English. It is capable of looking things up one at a time in its database. Thus, please go through the form systematically, one item at a time.

TABLE V

	Igor	Aida	Takis	Ellen	Fred	Megumi
	Russian		Greek	American	Australian	
	actor		veterinarian	psychologist		surgeon
		MSEE	DVM		BA journalism	MD
		MIT	Cornell	Stanford		Penn
	stamp collecting		gardening		piano	painting
		mansion		condo	ranch house	loft
	silver motorcycle	white sedan		green jeep	black pickup truck	
	goldfish	7 cats		none	horse	
	2 daughters		0	0	1 daughter, 1 son	
		married 10 years	single	married 1 year		divorced 2 years
		Jon	nobody		Anne	nobody

Do not take any notes while you do this part of the task. In part 2 of the task you will be asked to make some judgments about the people represented on this form. When you finish part 1 of the task, please find the Experimenter.

Subjects in the human partner condition were given these instructions:

Please work with the other person and discover the missing information represented by each blank cell on this form. Your partner, who is at a terminal in a room nearby, does *not* have a copy of this form, and he or she has to look things up one at a time in a book. Thus, please go through the form systematically, one item at a time.

Do not take any notes while you do this part of the task. In part 2 of the task you will be asked to make some judgments about the people

represented on this form. When you finish part 1 of the task, please find the Experimenter.

All subjects were given the same database array (Table V).

Appendix B: Example dialogue

Computer partner, sentence responses

System: Is anyone there?

User:

System: hello

User: hello

System: what do you want to know?

User: igor's degree

System: Igor's degree is a BA, in fine arts

User: igor's school

System: Igor's school was the Sorbonne

User: igor house

System: Igor's house is a commune

User: igor marry

System: what was that again?

User: is igor married

System: no, Igor is not married

User: aido nationality

System: Aida's nationality is Brazilian

User: aido profession

System: Aida's profession is engineer

User: aida hobby

System: Aida's hobby is hang gliding

User: aida children

System: Aida's children are 1 daughter and 3 sons

User: takis house

System: Takis's house is a trailer

User: takis car

System: Takis's car is a purple van

User: takis pet

System: Takis's pet is a pit bull

User: ellen degree

System: Ellen's degree is a PhD, in psychology

User: what hobbies does ellen do

System: Ellen does weight lifting

User: who is ellen married to

System: Ellen is married to Charles

User: what does fred do
System: Fred is a joke writer
User: tell me more about fred
System: what do you want to know?
User: school
System: Fred's school was Syracuse
User: married? how long?
System: Fred is married, 5 years
User: how about megumi?
System: Megumi is divorced, 2 years
User: what nationality is he or she
System: Megumi's nationality is Japanese
User: what does he drive or she drive
System: Megumi drives a red convertible
User: any pets?
System: yes, Megumi has 1 pet
User: what kind of animal is it?
System: It's a tortoise
User: does megumi have any children?
System: yes, Megumi has 1 child
User: is it a son or a daughter?
System: It's a daughter
User: thanks mister computer
System: you're welcome
User: good bye
System: bye

Appendix C: Example dialogue

Human partner, short responses

Partner: Is anyone there?
=>
Partner: hello
=> howdy, are you ready
Partner: what do you want to know?
=> ok, i'm not too sure what you've got there,
but what i have here is a list of people and
i'm going to start by asking you:
what do you know about Igor?
Partner: what do you want to know?
=> what does Igor have a degree in?
Partner: BA, in fine arts

=> From what school?

Partner: the Sorbonne

=> Where does Igor live?

Partner: in a commune

=> What is his marital status?

Partner: divorced, 3 weeks

=> Who is Igor's friend?

Partner: I can't answer that

=> Why not?

Partner: no comment

=> What do you mean : No comment?

What I need to know is a name of a person associated with Igor.

Partner: nobody

=> Thank you

Partner: you're welcome

=> Next person is Aida. What is Aida's nationality

Partner: Brazilian

=> What is her occupation

Partner: engineer

=> Hobbies?

Partner: hang gliding

=> any kids

Partner: yes

=> how many

Partner: 3

=> sons or daughters

Partner: 1 daughter, 2 sons

=> Next person is takis. Where does takis live?

Partner: in a trailer

=> what kind of car does takis drive?

Partner: purple van

=> i hope i don't have to remember all of this

Partner: no comment

=> Does takis have any pets

Partner: yes

=> what type and how many

-i know i'm only supposed to ask one question at a time, just trying to save some time

Partner: a pit bull

=> next person is Ellen. What degree

Partner: PhD, in psychology

=> hobbies
 Partner: weight lifting
 => name of person associated with Ellen
 Partner: Charles
 => next person-Fred. Fred's occupation?
 Partner: joke writer
 => what school did he attend
 Partner: Syracuse
 => marital status
 Partner: married, 5 years
 => last person is Megumi. What is Megumi's nationality
 Partner: Japanese
 => what vehicle does he drive
 Partner: red convertible
 => pets and how many
 Partner: a tortoise
 => just one
 Partner:
 => just one tortoise
 Partner: one tortoise
 => and last, any kids
 Partner: yes
 => sons or daughters
 Partner: 1 daughter
 => thanks, that's all i need to know
 Partner: you're welcome

References

- Bell, A.: 1984, 'Language Style as Audience Design'. *Language in Society* 13, 145-204.
- Bock, J. K.: 1986, 'Syntactic Persistence in Language Production'. *Cognitive Psychology* 18(3), 355-387.
- Brennan, S. E.: 1988, 'The Multimedia Articulation of Answers in a Natural Language Database Query System'. In: *Proc., Second Conference on Applied Natural Language Processing*. Association of Computational Linguistics, Austin, TX, pp. 1-8.
- Brennan, S. E.: 1990, 'Conversation as Direct Manipulation: An Iconoclastic View'. In: B. Laurel (ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, MA.
- Carbonell, J. G.: 1979, 'Towards a Self-Extending Parser'. In: *Proc., 17th Annual Meeting of the ACL*. Association of Computational Linguistics, pp. 3-7.
- Carbonell, J. G. and P. J. Hayes: 1987, 'Robust Parsing Using Multiple Construction Specific Strategies'. In: L. Bolc (ed.), *Natural Language Parsing Systems*. Springer-Verlag, New York.
- Clark, H. H. and S. E. Brennan: 1991, 'Grounding in Communication'. In: J. Levine, L. B. Resnick, and S. D. Behrend (eds.), *Shared Cognition: Thinking as Social Practice*. APA Books, Washington D.C.

- Clark, H. H. and G. L. Murphy: 1982, 'Audience Design in Meaning and Reference'. In: J. F. LeNy and W. Kintsch (eds.), *Language and Comprehension*. North-Holland Publishing Company, Amsterdam.
- Clark, H. H. and E. F. Schaefer: 1987, 'Collaborating on Contributions to Conversations'. *Language and Cognitive Processes* 2, 1-23.
- Clark, H. H. and E. F. Schaefer: 1989, 'Contributing to Discourse'. *Cognitive Science* 13, 259-294.
- Clark, H. H. and D. Wilkes-Gibbs: 1986, 'Referring as a Collaborative Process'. *Cognition* 22, 1-39.
- Garrod, S. and A. Anderson: 1987, 'Saying What You Mean in Dialogue: A Study in Conceptual and Semantic Co-ordination'. *Cognition* 27, 181-218.
- Guindon, R., K. Shulderberg, and J. Conner: 1987, 'Grammatical and Ungrammatical Structures in User-Adviser Dialogues: Evidence for Sufficiency of Restricted Languages in Natural Language Interfaces to Advisory Systems'. In: *Proc., 25th Annual Meeting of the ACL*, Association of Computational Linguistics, Stanford, CA, pp. 41-44.
- Hendrix, G.: 1985, *Q&A. Software*, Symantec.
- Isaacs, E. and H. H. Clark: 1987, 'Reference in Conversation Between Experts and Novices'. *Journal of Experimental Psychology: General* 116, 26-37.
- Laurel, B.: 1990, 'Interface Agents: Metaphors With Character'. In: B. Laurel (ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, MA.
- Levinson, S. C.: 1983, *Pragmatics*. Cambridge University Press, Cambridge.
- Levelt, W. J. M. and S. Kelter: 1982, 'Surface Form and Memory in Question Answering'. *Cognitive Psychology* 14(1), 78-106.
- Milroy, L.: 1987, 'Style-Shifting and Code-Switching'. In: *Observing and Analyzing Natural Language*. Blackwell, New York, pp. 171-198.
- Norman, D. A.: 1981, 'The Trouble with Unix'. *Datamation* 27, 139-150.
- Richards, M. A. and K. M. Underwood: 1984, 'How Should People and Computers Speak to One Another?'. In: *Proceedings, Interact '84: First IFIP Conference on 'human-Computer Interaction'*. London, International Federation for Information Processing, pp. 33-36.
- Schneiderman, B.: 1981, 'A Note on Human Factors Issues of Natural Language Interaction with Database Systems'. *Information Systems* 6(2), 125-129.
- Schneiderman, B.: 1982, 'The Future of Interactive Systems and the Emergence of Direct Manipulation'. *Behavior and Information Technology* 1, 237-256.
- Schober, M. F. and H. H. Clark: 1989, 'Understanding by Addressees and Overhearers'. *Cognitive Psychology* 21(2), 211-232.
- Shatz, M. and R. Gelman: 1973, 'The Development of Communication Skills: Modifications in the Speech of Young Children as a Function of Listener'. *Monographs of the Society for Research in Child Development* 38, 5:1-37.
- Schegloff, E. and H. Sacks: 1973, 'Opening Up Closings'. *Semiotica* 8, 289-327.
- Sacks, H., E. Schegloff, and G. Jefferson: 1974, 'A Simplest Systematics for the Organization of Turn-Taking in Conversation'. *Language* 50, 696-735.