Repair as Factor in Interface Design

Barbara A. Fox

Technical Report #88-4

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION University of Colorado 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Personnel and Training Research Programs Office of Naval Research (Code 1142PT)
6c. ADDRESS (City, State, and ZIP Code) Campus Box 345 Boulder, CO 80309-0345	7b. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, VA 22217-5000
8a. NAME OF FUNDING / SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Repair as a Factor in Interface Design

12. PERSONAL AUTHOR(S)
Barbara A. Fox

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	88, 01, 19	14

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

One of the little understood and equally little studied aspects of natural language interface design lies in the area of turn-taking between user and system. In many systems, the user is allowed a single command at a time while the system is allowed a large section of, let's say, canned text; in others, the student can be interrupted immediately after producing some kind of error, while the student is given no facilities for interrupting the system at all. These various solutions to the turn-taking problem appear to be unprincipled and ad hoc, designed to satisfy the particular needs of a particular system, at least as those needs are understood by the designer of the system.

In this paper I would like to examine, in a principled fashion, the turn-taking system required for a certain kind of system--namely, Intelligent Tutoring Systems. (cont.)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

19. Abstract (continued)

The findings reported here strongly suggest that the need for repair of one's own and one's interlocutor's utterances critically affects the kind of turn-taking mechanism that will be viable in a tutoring system. I thus propose an unusual hypothesis: repair is an essential factor in natural language interface design.

Repair as Factor in Interface Design

Barbara A. Fox
Dept. of Linguistics and
Institute of Cognitive Science
University of Colorado, Boulder

DRAFT: comments welcome

1 Introduction

One of the little understood and equally little studied aspects of natural language interface design lies in the area of turn-taking between user and system.[1] In many systems, the user is allowed a single command at a time while the system is allowed a large section of, let's say, canned text; in others, the student can be interrupted immediately after producing some kind of error, while the student is given no facilities for interrupting the system at all. These various solutions to the turn-taking problem appear to be unprincipled and ad hoc, designed to satisfy the particular needs of a particular system, at least as those needs are understood by the designer of the system.

In this paper I would like to examine, in a principled fashion, the turn-taking system required for a certain kind of system--namely, Intelligent Tutoring Systems. The findings reported here strongly suggest that the need for **repair** of one's own and one's interlocutor's utterances critically affects the kind of turn-taking mechanism that will be viable in a tutoring system. I thus propose an unusual hypothesis: repair is an essential factor in natural language interface design.

[1] To give an idea of how little studied this area is, I found no mention of this topic in the interface design literature. See Coombs and Alty (1980); Gaines and Shaw (1986); Baldwin and Siklossy (1977); Miller and Thomas (1977); Fitter (1979); Kidd and Cooper (1985); Edmonds (1982); the issue of IJMMS edited by Rissland; Hayes and Reddy (1983); and Webber and Finin (1985).

2 Background

The present paper is based on data collected for a project on human tutorial dialogue. The design of the project was quite simple. Tutors were located through various science departments at the University of Colorado and students were located by advertising in the local student newspaper. Each tutor-student pair was brought together for a total of three hours, spread out over three different sessions. The first session for each pair was a normal face-to-face session, in which the tutor and student met in a small lab room and were video-taped. They were given no instructions as to the content or structure of these interactions. The second session for each pair was somewhat different: the subjects were put in separate rooms and were asked to communicate via terminals hooked up to a VAX 780. Each was told to operate the terminal in the following way: they could each enter up to 20 lines of text, and when they were done with their turn, or had reached 20 lines, they could send their text to the other person. The other person, after a relatively short delay, then received the text, displayed on their terminal screen. At that point, the person who had just received the text could begin to work on their own turn, sending it when it was done (or already up to 20 lines), and the whole process would repeat. There was no way for either person to retrieve a text that had already been sent or to fix something that they had stated incorrectly in the already-sent text: each participant was paralyzed until the other had sent back their text. The following diagram illustrates the turn-taking system at work here:

This particular turn-taking mechanism was designed specifically to test the effect of one kind of turn-taking mechanism on the process of tutoring.

3 Results

The most striking finding of this study was completely unexpected; it manifested itself in numerous complaints from tutors and students that they had made a mistake somewhere in their text, not noticed it until after they had sent the text off, and had been unable to fix it until **after** the other person had replied, thus making it possible for the error to thoroughly confuse the hearer--clearly a dangerous prospect in a tutoring situation. This was the one universally agreed upon design flaw of the system. As one student put it on the evaluation sheet: "She [the tutor] could comment that I was taking a wrong turn while I was taking it."

It is clear from these findings that any kind of communication model, including a computer interface, must have built into it a reasonable and effective mechanism for fixing trouble in the communication when such trouble arises. Exactly how such a mechanism should **function** is the topic of the remainder of the paper.

4 Turn-taking and Repair in Face-to-Face Tutoring

The face-to-face tutoring sessions we video-taped exhibited a much different turn-taking system from the one described above for the terminal-to-terminal session. In face-to-face tutoring, as in everyday conversation, turns are allocated by a much more flexible system: speaker A can speak again even if there has been no overt response from speaker B; and either participant can interrupt the other--in the middle of a sentence or during a longer turn--to repair a misunderstanding (or can overtly show that no repair is needed, for example by saying "mhm," "yes," "right," etc.). One possible mechanism of turn-taking which could generate this kind of flexibility is presented below.

We need to first establish that there are units out of which turns can be constructed. These units have been referred to as turn constructional units (TCU's), and can be single lexical items (yes), phrases, clauses or sentences. According to the turn-taking system, each speaker is at first alloted one of these TCU's. The end of such a unit constitutes a place where speaker-change could occur;

that is, at this point another person could begin talking. The end of a TCU is thus a transition relevance place (TRP), since it is a place at which a transition from one speaker to another can (but need not) occur.

The following turn-taking rules, which are based on these concepts, were originally proposed for everyday conversation. It is clear from the data collected for this project that they hold for face-to-face tutoring as well (the rules are taken format from Levinson (1983), which is based on Sacks, Schegloff and Jefferson (1974)):

Rule 1----applies initially at the first TRP of any turn

(a) If the current speaker selects a next speaker in current turn, then current speaker must stop speaking, and that next speaker must speak next, transition occurring at the first TRP after next-speaker selection.
(b) If current speaker does not select next speaker, then any (other) party may self-select, first speaker gaining rights to the next turn.
(c) if current speaker has not selected a next speaker, and no other party self-selects under option (b), then current speaker may (but need not) continue.

Rule 2---applies at all subsequent TRP's

When Rule 1(c) has been applied by the current speaker, then at the next TRP Rules 1(a)-(c) apply, and recursively at the next TRP, until speaker change is effected.

In order to allow for interruption, I postulate the following additional rule:

A non-speaker may elect to start speaking before the current speaker has reached a TRP.

This system also accounts for a number of other interesting characteristics of the face-to-face tutoring data collected. For example, one of the striking characteristics of these data is the variability of turn length: in some cases, tutors pursued extremely long, multi-unit turns, while in others they constructed very short turns; similarly, in somewhat rarer cases students produced long turns, while other times they produced short, one-unit turns. Longer tutor turns seemed to correlate with

something more of a teaching style, but even in a highly interactive, non-teaching session, one can find an occasional long turn from the tutor. Examples illustrating this variability are given below.

(S=student, T=tutor)

S: And uh (0.6) And okay, so, (0.5) uhm (1.3) and I have a value of three, and they want to know what the tangent is.So, I have one over cosine of theta equals three. (0.8) And I have the sine of theta over cosine of theta (1.0) u:hm. (0.8) .hh Okay, so I guesss- I somehow have to: (0.8) tangent of theta is going to be: (0.4) sine of theta over cosine of theta. (2.0) One over cosine of theta

T: You know **their ratio**
S: Yeah.
T: But you don't know exactly what they are.
S: What they are, yeah.
T: So=
S: =Okay. So this is **the basic.**
T: Yeah.

T: The outside derivatives get one over this inside function, times the inside derivative. .hh So .hh this actually gives us a real powerful integration tool. Suprisingly powerful. .hh Because now if we have ( ) some function like this (1.2) and we want to differentiate it, .hh (.) the natural log disappears in the differentiation process.You take your inside derivative, (0.5) three ex squared plus six, divided by the inside function, (.) by the chain rule, .hh and then get a rational function.

Clearly, a turn-taking system like the one presented above is needed to produce such flexibility in turn length, with no single party distributing turns (see below). Since it is not the case that all tutor

(or student) turns are long, and since it is also not the case that all possibly lengthy turns are allowed to go to completion (in many cases they are interrupted by the other participant), then it must be that length of turn is freely negotiated between speaker and hearer, as it is in everyday conversation, and not pre-determined by a rigid on-off kind of mechanism. The only kind of mechanism that could produce such flexibility in turn structure is one that allows a) speakers to speak again, even if the interlocutor has not in the meantime taken a turn, b) participants to repair a previous utterance, even if it looked as though that turn was already complete, c) participants to speak before the other's turn is complete (this allows the hearer to interrupt the speaker in the middle of a sentence or at the ends of units in a long turn).

An additional prominent characteristic of the face-to-face data is the lack of tutor control over turn allocation. This lack of control contrasts with the typical classroom setting, for example, in that in the classroom the teacher has control over who speaks, for how long, and when. Turn-taking in the classroom is modulated in the following way: the teacher asks a question, the students then raise their hands (forming a pool of possible next speakers), and the teacher selects one person to speak immediately next. After that person is finished speaking, the turn bounces back to the teacher, rather than, say, to one of the other students (this system is described in some detail in Schegloff, 1985). In the tutoring data, no such turn allocation exhibits itself.

Most importantly, this turn-taking system provides for a fundamental process of tutoring: correction. This point can be nicely illustrated by contrasting the possibilities for correction in face-to-face tutoring with the possibilities for correction in the terminal-to-terminal tutoring I described above. The terminal-to-terminal interface allowed for multi-unit turns (in fact encouraged them) but did not allow for two critical elements: (1) the opportunity for the hearer to indicate understanding or lack of understanding at the end of every unit; and (2) the opportunity for the speaker to initiate correction on his/her own turn after it was sent. The problems with this method of turn-taking are immediately apparent. Let's say the tutor constructs a 20-line text with an error (which goes unnoticed) in the

first line. The student receives this text and can't understand past the first line, since the error completely confuses the explanation the tutor meant to be giving. The student then sends a repair initiator to the tutor, who then must reconstruct in her mind what the error could have been, and how it affected the rest of the text. The tutor must then not only repair the error but also reproduce the other 19 lines of text she had already sent. If this happened with every turn, the efficiency of the communication would drop to close to zero and very little learning would take place. In fact, the participants quickly learned to take relatively short turns when possible; that is, they tried to manipulate the interface they were given to make it match more closely the turn-taking system of natural conversation.

In contrast, the turn-taking system of face-to-face tutoring enables speakers to produce multi-unit turns while at the same time allowing the hearer to initiate repair at the end of every unit and allowing the speaker to initiate repair, even after the end of the multi-unit turn, if there is some sign--which could be as little as silence--that the hearer is having difficulty understanding. How is this accomplished?

The fundamental principle underlying this system is that each speaker is alloted one unit (TCU) per turn; at the end of such a unit, someone else can begin to speak. But, we know that even in natural conversation people take sometimes very long turns--e.g., stories, jokes, reports. Such multi-unit turns, however, regularly display to the hearer at the outset that something long is in store, for example, "guess what," "you'll never guess what happened to me today," "did you hear the one about..." So we know that, even given a system that allocates one unit per turn, there are ways to secure multi-unit turns.

But, securing a multi-unit turn in the course of a tutoring session does not mean, as one might expect, that one person speaks and the other is completely silent. Even though a speaker may have secured a multi-unit turn, the hearer will produce things like "uh huh," "mhm," at predictable points

within the multi-unit turn.[2] In fact, these so-called backchannel utterances are positioned at the end of TCU's and signal that, given the end of a new TCU, the hearer is having no difficulty understanding or agreeing with what the speaker has said. In other words, even in a multi-unit turn, there are places--namely the end of each TCU within the larger turn--where the hearer can initiate repair ("huh," "what," "I don't get it," and the like). In the following passage, S, the hearer of a multi-unit turn, indicates at many TCU boundaries that she is understanding and agreeing, and not about to initiate correction or repair.

T: Let's say you have a triangle, (1.2) and this angle is
theta, and this is ar.=
[
S: Aha
S: =Right.=
T: =Okay? (0.7) This'd be the wy. Wy equals ar sine theta.
And ex, equals ar cosine theta.
[
S: ar cosine theta. Right.
T: And- and I use that all the time, it has to be something
I've used like (0.4) over and over and over.=
S: =Mhm.

Without a facility for producing such "backchannel" comments, the student can become completely confused: there would be no opportunity for the student (or the tutor, in the case of a long student turn) to clear up misunderstandings as they occur, so misunderstandings could accumulate, until the point when repair becomes extremely difficult and time-consuming (see Anderson et al (1986) for a similar point). If the level of understanding and agreement is constantly being displayed by both parties, then there is little opportunity for serious misunderstandings to arise.

Furthermore, the mechanism of tutoring turn-taking provides a facility for interruption, as a device for repair. That is, there are instances in conversation where one participant starts talking before the other participant has finished a complete unit, usually to initiate or perform some kind of repair.

---

[2]This seems to contradict Hayes and Reddy's notion of implict confirmation, wherein no overt indication of comprehension is necessary.

While this facility is not captured well by the rules given above, it is clear from looking at the tutoring data that certain kinds of interruption are essential for maintaining mutual comprehension. In the following passage, for example, the tutor interruption repairs the now obvious misunderstanding which has grown because the two have been looking at different problems.

S: And (0.5) the magnitude would be:
(13.5)
S: The equations I have f- f- for this is uhm, force over charge>
(1.6)
S: But (1.5) I don't know if force would be (//)
T: Okay, nah, nah you've got another one (0.9) You have one that says down there, energy stored in capacitor.
S: O:h.
(3.9)
S: Energy stored in capac- but- don't they want
(3.5)
S: So this is electric field? (0.9) This is e-
T: Oh, excuse me, excuse me, I was looking at this one, and you're still working on this.

In the following passage, the tutor interrupts the student to prevent a potential difficulty.

S: Equal to one half. And then the mass. (0.8) I need the mass of an electron.
T: Mhm
(1.8)
S: And that's in my book. (1.7) And th//en
T: e-e-e- what units are you going to put that in. This is the main thing I'm worried about.

In the following case, the student interrupts the tutor to clarify the signifance of something the tutor has suggested:

T: Does that make sense to you, or not // yet.
S: Well to- not- I mean, (no sense really).=
T: Okay. Well (0.4) just start (0.6) like for example=
S: =It seems like that's just you know the- the- back of the problems

Both aspects of the design of turn-taking in tutoring function to enhance smooth, efficient mutual

understanding, by allowing misunderstandings to be clarified as soon as they come up (in some cases, even before there are overt symptoms of misunderstanding), while both participants can remember what is likely to need repair, without having to repeat long portions of previous turns.

In addition to providing for on-line repair, this turn-taking system also allows the speaker to speak again without waiting for the hearer to complete a turn first. For example, in the following passages the speaker talks again even though the hearer has not completed a turn in between:

T: And you probably have some way of remembering that already.
(1.3) [student stares blankly at tutor]
T: Maybe not>

T: So one electron volt is the charge of an electron,
which is, one point six times ten to the minus nineteen
joules.
(1.7)
T: So I don't know, do you have to write that number down?
(1.1) [student stares blankly at tutor]
T: Can you write that down on your piece of paper?
S: Sure

In these passages, the tutor expects a confirmation of or answer to her first line; the student, however, displays signs of not being able to provide the response, so the tutor speaks again.

It seems to me not at all odd, then, given the tremendous benefits of this particular turn-taking system, that people engaged in face-to-face tutoring choose to utilize this system. Tutoring represents enough of a challenge to minimize misunderstandings without the system within which it is embedded aggravating the dangers.

My suggestion, then, is that any natural language interface designed for a tutoring environment incorporate the facilities for turn-taking described above.

5 Further Refinements

There has been at least one tutoring system developed (see Anderson et al (1986)) which has built into it a turn-taking system which in one respect matches the suggestions I have given for designing tutoring interfaces. In this system, the tutor can--in fact must--interrupt the student as soon as s/he has produced an error (for example, typed in an incorrect answer to a question).

In this system, then, at least one of the parties has the ability to initiate repair close to the source of the trouble.[3] The problem I see with this system is how it has defined close: it positions the correction immediately following the errorful word. While I think this approach represents an interesting attempt at modeling repair, it does not incorporate what is now known about repair in tutoring.[4]In order, then, to further refine my suggestions for interface design, I will devote the remainder of the paper to exploring where, given an interface which allows us to position repair relatively freely, a tutor's correction of a student error should be positioned.

5.1 Other-Correction

It has been known for some time that other-correction, that is correction of the speaker of the error by the hearer, is avoided in naturally-occurring conversation (Schegloff, Jefferson and Sacks (1977)). For example, let's say speaker A says something that contains some kind of error. Speaker B does not rush in as soon as the error is produced to fix it; s/he regularly withholds correction until after the error, in fact until after the turn containing the error is complete. If speaker A still has not fixed the problem (the turn appears to be over and there is even a small bit of silence growing), B may finally say something to suggest there was an error and that A should fix it. Note in the following passage, for example, how B withholds correction until it is clear that A is not going to perform it on his own; B then initiates correction and lets A actually perform the fix:

---

[3]Of course, this system does not allow for the possibility that the student might need to repair something generated by the tutor.

[4]Anderson et al do propose as an alternative method of waiting until the student has finished a complete answer before starting correction. This alternative is fully in keeping with the suggestions below.

A: Hey the first time they stopped me from selling cigarettes
was this morning.
(1.0)
B: From selling cigarettes?
A: From buying cigarettes.

Now, in examining face-to-face tutoring sessions, it became clear that this same preference for self-initiation of correction and for self-correction holds for tutoring as well. Even though correction would seem to be an inherent part of tutoring, tutors and students manage to negotiate it in such a way that students are given the opportunity to a) correct their own mistakes, and failing that, b) at least initiate correction that the tutor can then accomplish, and failing that, c) the tutor can initiate (and potentially also accomplish) correction. As in everyday conversation, the tutor withholds correcting a student until the student has passed opportunites to initiate correction. In the first passage below, for example, the student makes a mistake but is given the opportunity to correct it, which he in fact does, successfully. Similarly, in the second passage, the student makes a mistake and the tutor gives the student a chance to self-correct; in this example, however, the student fails to take this opportunity and the tutor performs other-correction.

S: Ey: is minus one, bee is zero.
(*)
S: No, bee is one.
T: All right.

S: Ey plus cee equals zero.
T: Right, so that tells // you
S: Ey equals cee.
(0.5)
T: Minus cee.

T: Do you know the difference between velocity and speed?
(2.4)
S: Velocity is the distance times time>
(1.7)
S: Speed is uhm
(2.9)
T: No, distance is- velocity isn't distance times time

In other words, tutors do not behave the way Anderson et al's system currently behaves: they do not immediately correct an error as soon as it is produced.[5] They wait until slightly after the student's errorful turn appears to be complete, and then, if the student still shows no signs of correcting his/her error, they suggest that some correction may be in order.

6 Conclusions

Based on the findings of the present study, I would like to suggest the following changes to ITS interface design.

1. Turn-taking should not be an on-off option. That is, participants should not have their ability to speak turned off until the other participant has finished a turn. The interface must allow for each person to participate as they see fit.

2. As part of the preceding condition, it is especially important that, even when one person is producing a long turn, the other have the ability to show understanding, initiate repair, disagree, and so on, at the end of every unit.

3. The turn-taking mechanism must provide flexibility in turn length.

4. Correction of the student, or initiation of such correction, should be withheld until the student has had an opportunity to self-correct, or initiate self-correction.

It is clear that future ITS natural language interface designs must make room for principled changes in mechanisms for turn-taking and repair. Such changes should make for much smoother, more effective tutoring systems.

[5] In fact, they rarely overtly correct a student's error at all.

REFERENCES

Anderson, J., Boyle, C., Corbett, A., and M. Lewis. 1986. Cognitive modelling and intelligent tutoring. Report for ONR.

Baldwin, J. and L. Siklossy. 1977. An unobtrusive computer monitor for multi-step problem solving. International Journal of Man-Machine Studies 9, 349-362.

Coombs, M. and J. Alty. 1980. Face-to-face guidance of university computer users--II. International Journal of Man-Machine Studies 12, 407-429.

Edmonds, E. 1982. The man-computer interface: a note on concepts and design. International Journal of Man-Machine Studies 16, 231-236.

Fitter, M. 1979. Towards more "natural" interactive systems. International Journal of Man-Machine Studies 11, 339-350.

Gaines, B. and M. Shaw. 1986. Foundations of dialog engineering. International Journal of Man-Machine Studies 24, 101-123.

Hayes, P. and D.R. Reddy. 1983. Steps toward graceful interaction in spoken and written man-machine communication. International Journal of Man-Machine Studies 19, 231-284.

Kidd, A. and M. Cooper. 1985. Man-machine interface issues in the construction and use of an expert system. International Journal of Man-Machine Studies 22, 91-102.

Levinson, S. 1983. Pragmatics. Cambridge: Cambridge University Press.

Miller, L. and J. Thomas, Jr. 1977. Behavioral issues in the use of interactive systems. International Journal of Man-Machine Studies 9, 509-536.

Sacks, H., Schegloff, E. and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking. Language 50, 696-735.

Schegloff, E. 1985. "Micro and macro sociology." MS.

Schegloff, E., Jefferson, G., and H. Sacks. 1977. The preference for self-correction in the organization of repair. Language 53, 361-382.

Webber, B. and T. Finin. 1985. In response: next steps in natural language interaction. In ?

I. Notational Conventions in Transcripts

The following notational conventions are used in the transcripts.

//	point at which current utterance is overlapped by the next utterance produced by another speaker.
(0.0)	Numbers enclosed in parentheses indicate length of silence.
Underlining	indicates stressed syllables
:	lengthened syllable
-	Glottal stop cutting off a word
=	indicates a relationship between two utterances in which there is not the usual beat of silence between them.
?	rising intonation
(( ))	non-linguistic action
( )	unintelligible stretch
hh	audible outbreath
'hh	audible inbreath
(hh)	laughter within a word