**Robust Learning Environments:
The Issue of Canned Text**

Barbara A. Fox

Technical Report #88-5

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 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
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)*
Robust Learning Environments: The Issue of Canned Text

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 ______		

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)*

There is a growing sense among people in the area of cognition and instruction that computer tutoring systems offer a much needed alternative learning environment for our current school system. Such totoring systems, in an ideal form, would provide one-on-one instruction, context-grounded apprenticeship-style learning, and greater tolerance of individual learning styles (see Collins, Brown and Newman, 1986). These features could greatly enhance what has become for the most part a mass production of socialized, but uneducated, adults.

Intelligent tutoring systems as they now exist do not encompass the features listed above. Inasmuch as they are pre-programmed and incapable of making spontaneous judgments and decisions, they do not tolerate individual differences; they do not yet model apprenticeship modes of learning, and they offer one-on-one (continued)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION 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

Robust Learning Environments:
The Issue of Canned Text

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

1 Introduction

There is a growing sense among people in the area of cognition and instruction that computer tutoring systems offer a much needed alternative learning environment for our current school system. Such tutoring systems, in an ideal form, would provide one-on-one instruction, context-grounded apprenticeship-style learning, and greater tolerance of individual learning styles (see Collins, Brown and Newman, 1986). These features could greatly enhance what has become for the most part a mass production of socialized, but uneducated, adults.

Intelligent tutoring systems as they now exist do not encompass the features listed above. Inasmuch as they are pre-programmed and incapable of making spontaneous judgments and decisions, they do not tolerate individual differences; they do not yet model apprenticeship modes of learning, and they offer one-on-one instruction only in the sense that one person can work at a computer without others present, not in the sense that the instruction is tailored to the individual learner. In order to realize the goals stated above, then, a great deal of work remains to be done to better understand apprenticeship and tutoring as they occur between humans. In particular, one pressing issue to be addressed is: How will understanding the characteristics of human tutoring lead us to re-vision machine tutoring?

The present paper attempts to address this issue in a preliminary way by exploring two essential aspects of human tutoring comunication with an eye to a feature of many current ICAI systems--canned text. It is the conclusion of my work on human tutoring dialogue that canned text limits the functionality of ICAI systems beyond reason and inhibits the growth of the kind of system Collins et al hope to inspire. In particular, I hope to demonstrate here that canned text is not capable of handling the two tutoring phenomenon examined and therefore prevents effective tutoring from taking place.

1.1 The role of canned text in ITS

While there are a variety of projects concerned with text generation (e.g. Mann et al's Penman project, Bates et al's ILIAD), most of the interactive systems in intelligent tutoring systems operate with canned text to communicate with students. The decision to use canned text rather than naturally generated language is, I believe, a purely pragmatic one, having to do with the limitations on our current understanding of text generation, in particular with the limitations on our understanding of dialogue. I assume that if sophisticated text generation facilities are ever available, researchers in ITS will make use of them in tutoring systems.

Canned text is thus apparently a necessary evil to be lived with until natural language generation is considerably more sophisticated than it is today (see Mann et al, 1981, for a state-of-the-art report on text generation)--not something that is likely to happen in the next few decades.

Nonetheless, even though there seems to be no alternative to canned text at this moment, I think it is crucial for researchers in ITS to realize the full limitations of canned text and to decide if it is wise to wait for the implementation of text generation systems before solving the problems of canned text.

1.2 Background

The work presented here is part of a larger study of human-human tutorial dialogue. The study as a whole attempts to characterize, in extreme detail, the characteristics of tutorial interaction, how that interaction is changed when the participants are not face-to-face, and how student attitudes towards computers affect their interaction with them.

The data for the study were gathered by bringing together tutors and students from on campus. Graduate students from chemistry, physics, math and computer science were located and asked to participate as tutors in the study. All were recommended by their departments and all had worked as tutors before. They were paid for their work, in keeping with normal tutoring arrangements. Students were located by advertising in the student newspaper for people interested in being tutored as part of a research project. Each student-tutor pair met for an hour of face-to-face tutoring in a small office, and they were video-taped. The pairs also met for an additional 2 hours each, when they were asked to communicate using computer terminals hooked to a VAX 780. They were given no instructions regarding any aspect of the tutoring, either structure or content. The students represent a fairly wide spread of expertise, from nearly complete novice to intermediate.

1.3 The Transcripts

Detailed transcripts of each face-to-face session were made, following the conventions of Sacks, Schegloff and Jefferson (1974). In order to help the reader recreate as much as possible the sound of each fragment used below, I provide here a brief description of each of the major transcription conventions used. Readers familiar with this method of transcription can skip to section 2.

A double slash (//) indicates the place at which a speaker's utterance is overlapped by talk from another speaker.

M. No. They're a//ll thin.
C. They're not

Thus with this notation we can see that C's utterance starts after the a in M's all.

An utterance which has more than one double slash in it is overlapped at more than one place, and the utterances which do the overlapping are given in sequential order after the overlapped utterance.

G. they're all Keegans like the ones around Greensprings
they're all kind'v, // bout five five, five si//x,
M. They're all from around Greensprin//gs.
C. Yeh

Here, M's utterance overlaps G's starting at bout, and C's utterance overlaps G's starting with the x in six. Notice that C's utterance also overlaps the very end of M's.

A left-hand bracket at the beginning of two lines indicates that the two utterances begin simultaneously.

M. Yeh.
[
C. Lo:ng time ago it reminds me

M and C begin talking simultaneously.

The equals sign (=) indicates latching, that is, the next speaker begins without the usual "beat" of silence after the current speaker finishes talking. In this case there is an equals sign at the end of the current speaker's utterance and another equals sign at the beginning of the next speaker's utterance. If two speakers simultaneously latch onto a preceding utterance (that is, they begin talking simultaneously), this is indicated in the transcript with a left-hand bracket preceded by an equals sign.

(R) (h)hh (h)uh (h)uh (h)uh!=
(S) hhh(h) H(h)m
=[
K. Which la:mpost?

Here S and K simultaneously latch onto R's laughter.

Numbers given in parentheses indicate elapsed silence, measured in tenths of seconds. Single parentheses with a dot between them represent a silence that is less then a tenth of a second but still longer than the usual beat of silence.

Certain facts about the production of the talk are given through the orthographic symbols used. Punctuation is used to suggest intonation; underlining indicates stress. A colon after a letter means that the sound represented by that letter is somewhat lengthened; a series of colons means that the sound is increasingly lengthened.

The letter h within parentheses indicates "explosive aspiration," and usually means some type of laughter is being produced. A series of h's preceded by a dot represents an inbreath (where number of h's is meant to correspond to the length of the inbreath), while the same series preceded by nothing represents exhaling.

Questionable transcriptions are enclosed within single parentheses; the transcribers thereby indicate that the exact form of the utterance is not clear. Speaker's initials given in single parentheses means that there is some question about the speaker's identity. Single parentheses with capitalized words enclosed --e.g.,(CLEARS THROAT)--represent non-transcribed material (i.e., noise which is non-linguistic).

These are the major transcription conventions which will be used in the data fragments. For a more detailed guide to CA notational conventions see Sacks, Schegloff and Jefferson (1974).

2 Two Faces of Canned Text: Repair and Internalization Routines

While canned text displays limitations in a variety of tasks, it is theoretically incapable of handling a range of discourse phenomena. For the purposes of the present paper, I have chosen to look at just two of these discourse tasks: repair and internalization routines.

2.1 Repair

According to the conversational analysis literature, repair is the process in conversation whereby some source of misunderstanding (or non-understanding) is fixed (Schegloff, Jefferson and Sacks, 1977). Although to my knowledge there are no systems currently which offer the functionality of repair, the importance of repair, and other discourse processes like it, for the future of user interfaces is made clear in the following quote from John Seely Brown:

> A dialogue involves constant conversational repairs between
> two people. When someone doesn't understand what I've just said,

> I must try to diagnose not only what, but why, he didn't understand and then accordingly repair what I said. It is basically an adjusting process that goes on between the two of us communicating....We have been misled into thinking that natural language, per se, is so powerful. Instead, I think it is the dialogue process that is so powerful, e.g. the notion of conversational repairs that occur between two people. If we can understand this process and how to capture it in man-machine communication, we will have made a major breakthrough on the perceived friendliness of machines.

Given this strong endorsement from one of the field's leading theoreticians, repair stands as a natural area to be explored for its compatibility with canned text.

For the purposes of the present paper I will focus most of my attention on one type of repair, known technically as third position repair (Schegloff, Jefferson and Sacks, 1977). Third position repair can be illustrated schematically as follows:

1. A: utterance
2. B: response
3. A: repair initiator
4. B: new response

Speaker A produces some utterance, which B then responds to. A can tell from B's response that B misunderstood A's utterance and uses the next turn--third position--to initiate repair on the utterance at line 1, so that B can be led to the intended interpretation.

Not all repair involves third position repair; at the end of section 2.1 I examine a more common, but somewhat less interesting, form of repair.

In order to proceed with this analysis, it is first necessary to establish, at least to a first approximation, what the subprocesses are which make up the phenomenon of third position repair. Consider the following passage from a physics tutoring session, in which there are two instances of initiating repair on the part of the tutor.

(1)

T: [READING PROBLEM] What is the speed of a three
hundred and fifty ee vee electron.Okay.
(1.4)
T: So the main thing he:re, I mean, when you look
at that, what is electron volts, what kind of a,
what are we talking about.
(1.5)
S: Isn't it- the charge of an electron times?
(0.9)
T: Right,//but what is it-what is that. Is it-=
S: The voltage?
S: =It's smaller
(0.2)
T: No- okay, I'm n- I'm a I'm a I'm not asking a specific
enough question. U:hm, (1.1) Uh (0.3) is this units
of length?
(0.9)
S: Oh, not it's uhm
(1.9)
S: It's voltage, isn't it?

The tutor in this passage first asks a question regarding the units expressed by electron volts. The student responds with a formula (<u>charge of an electron times the voltage</u>). The tutor asks the question again; this time the student responds with a description (<u>it's smaller</u>--presumably than joules). The tutor asks a third time, this time with a completely reformulated version, and the student is finally able to answer the question at the right level of analysis (although not with the right answer).

What did the tutor do in order to accomplish these repairs? In order to perform third position repair the tutor must engage in a set of extremely complex subprocesses (not meant to be sequential):

1. The tutor must be able to interpret the utterance, or set of utterances, requiring repair.

2. The tutor must be able to recognize that the interpretation displayed by that set of utterances requires repair.

3. The tutor must be able to see what in the preceding discourse could have led to the interpretation requiring repair and how it could have led to it.

4. The tutor must be able to re-articulate the source of trouble such that the new formulation takes into account the path which led to the misinterpretation (and eliminates it as a viable interpretation).

For example, the tutor in this passage must see what kind of an answer "electron volts is the charge of an electron times the voltage" is. It is, of course, a formula for how to calculate electron volts. It arises from the formulation of the question "What is electron volts, what kind of a what are we talking about" by the possible interpretation of _is_ as 'equals', a common interpretation in physics problems. The tutor recognizes that as not the kind of answer she was looking for and sees that it will not serve as a resource for helping the student solve the problem at hand. She rephrases the question as "but what is it- _what is_ that," using _that_ to refer to the equation just proffered by the student (and equations are not equal to something else). The student answers this new formulation with a description, "it's smaller," which is not the kind of description the tutor is looking for. The tutor reformulates the question once more, this time not in the form which gave rise to the previous answers. This time she has designed her question to specifically rule out the various interpretations of _what is X_; she accomplishes this by naming a member of the class from which the right answer should be chosen but which is itself obviously not the right answer (the obvious incorrectness of her choice is critical--otherwise the question would have been heard as asking a simple yes/no question rather than guiding the student to the right level of analysis). The student finally produces an answer at the right level (although it is not the answer the tutor was getting at).

I think it is safe to say that no system currently in use could manage these four processes. In particular, the last two steps--tracing the path which led to the misinterpretation and reformulating the trouble source--seem beyond the capabilities of current systems. These processes involve not an abstract domain of knowledge--such as physics or chemistry--but **rather** a history of the preceding discourse, the ability to find alternative interpretations, based on another's utterances, of one's own verbal behavior **and** the ability to re-design one's own utterances to rule out undesired interpretations and guide the hearer towards the desired interpretation. These are extremely sophisticated processes, which demand an extensive familiarity with possible relationships between syntactic/semantic structure and discourse structure--natural language understanding in the very deepest sense--and an entirely different kind of cognitive flexibility than computers have displayed up to now.

But what of canned text? Would it be possible, given advances in other areas of knowledge representation, natural language understanding, etc., to perform repair with canned text?

For the sake of brevity, I will not explore here the possibilities of the first two subprocesss of repair. While they are extremely important in repair and represent highly complex behaviors, they are somewhat less implicated in the canned text issue; I have therefore chosen to save them for another occasion.

Let me start, then, by looking at the third subprocess of repair--seeing how a prior utterance could have led to the misunderstanding observed. This step is critical inasmuch as understanding where the difficulty is coming from enables the tutor to repair the difficulty effectively. But, while this step is extremely important in formulating an effective repair initiator, it requires an understanding of the kinds of interpretations which can arise from certain kinds of utterances, as well as an understanding of syntactic and lexical ambiguities; it would therefore be extremely difficult to model in a computer system.[1]

It is quite simple to show that everyday interaction requires processing of this sort at high speeds and in real-time. Consider the following passage, for example, M (mother) is able to see (a) that her son has misinterpreted her, (b) what his misinterpretation was, (c) how it could have followed from her utterance, and (d) how to initiate repair so that he eventually produces the desired kind of response.[2]

(2)

M: Do you know who's going to be at that meeting?

C: Who.

M: I don't kno:w.

C: Oh::. Probably Mrs. McOwen ('n detsa) and probably Mrs.
Cadry and some of the teachers.

R originally hears his mother's utterance as a pre- announcement; that is, she is going to **tell him**

[1] For example, since any object can be characterized in an unlimited number of ways (Schegloff, 1972), a phrase such as What is X has an unlimited set of answers, as in the following invented example:

What is milk?

A white liquid
It has the following molecular structure:
A food
A dairy product
It's made from cows
An ingredient in cakes
The bane of every child's existence
An allergen
.
.
.

Each of these answers to what is milk? calls forth its own response.

[2] For reasons beyond the scope of this paper, this passage can also be treated as an instance of fourth position repair. See Fox, 1987.

who is going to be at the meeting. The appropriate response to a pre-announcement is No or I don't know, or to repeat the question word from the pre-announcement, R's choice for his utterance at line 2. At the third line the mother reveals two things: (a) that she understands that R heard her utterance as a pre-announcement, in other words, that she knows the list of people and she is going to tell him, and (b) that his interpretation is incorrect--she does not know, and hence could not be about to announce it.

Repair might work then something like this for passage 2. It is part of every speaker's knowledge that an utterance of the form do you know + clause {wh-} has two easily accessible interpretations: pre-announcement and information-seeking question. Furthermore, it is part of every speaker's knowledge that who is not an answer to an information-seeking question with a wh- word in it.

So when the mother hears who as the response to her question, she knows that it is the sort of response which follows a pre-announcement; it is not the sort of thing that follows an information-seeking question, since it neither provides information nor overtly claims to lack the requested information. She thus hears that R has not understood her utterance as an information-seeking question. Her best source from which to infer exactly what he has understood is his own response. The repetition of a wh- word (with falling intonation) in a response to a question has only one interpretation--it is a gatekeeper after a pre-announcement.[3] So the mother can infer from this piece of information alone that R heard her to be making a pre-announcement.

---

[3]In this context, a wh- word with falling intonation is not a repair initiator:

(invented examples)

A: Do you know where you're going?
B: Where.

A: Can you show me which to choose?
B: Which.

These awful examples further illustrate the context-dependence of interpretation. In other contexts, a wh- word can serve to initiate repair.

She can now go back to her original utterance and see how R heard it as a pre-announcement; it is the do you know + clause {wh} structure which is ambiguous. Her third position response is carefully designed to not repeat any of this structure. She responds by directly negating the central proposition underlying R's interpretation, namely, she knows who.

Judgments regarding the origin of misunderstanding come from knowledge about other high-frequency interpretations for an utterance, about what kinds of responses are appropriate for what kinds of utterances, and as we saw in passage 1, knowledge about syntactic and lexical ambiguities. Clearly, then, there is a great deal of knowledge involved in understanding how a misinterpretation arises, as well as how to fix it.

Furthermore, in tutoring it is critical that the tutor be able to distinguish an incorrect answer to the right question from a misunderstanding of the question. For example, if the tutor asks "what's farads" and the student says "resistance," then it would not be appropriate for the tutor to say "yes, but what is that"; correction procedures would be in order.

How would an intelligent tutoring system manage third position repair, with special regard to the third subprocess (we can assume for our purposes that the system can negotiate the first two subprocesses)?

As far as I can tell, there seem to be 3 ways of handling the third subprocess of repair. The first, and most desirable, involves building an enormous natural language understanding module which would be able to look at an utterance like charge of an electron times the voltage, understand it as an equation, go back to the tutor's original utterance, see that is can be heard as asking for an equation when coupled with a unit that could be one side of an equation, and provide instructions for how to rephrase the question so that that particular ambiguity could be avoided. Building a system with these capabilities should be one of the tasks for the next 50 (or 100 or 200) years for researchers in natural language processing; it is not a viable solution to the problem now. Moreover, even if there were such a system, it would only be wasted on a machine which used canned text on the production end, since there would be no way to use the information derived from the comprehension system to tailor the output of the system, except to proliferate the number of canned text responses.

The second alternative is to have one stored response for each misinterpretation. In this case, the system obviously does not need to know how the student came to the misinterpretation-- it will use its one stored response no matter what. No special mechanism is needed. Chances are this may work some of the time in a kind of half-hearted way and may work some of the time not at all.

The third alternative is to complicate the system by adding to the collection of responses to misinterpretations. In this case, the system could look for pre-determined misinterpretations based on the possible ambiguities of the tutor's utterance. Each source of ambiguity would be connected to a range of responses which display having fallen prey to those ambiguities. Each of these would in turn be connected to a set of repair initiating utterances. This option requires that the programmer guess at all the possible ambiguities and their related misinterpretations and connect those to specific pieces of canned repair text.

This solution is at once too much work and not enough. It is too much work because it requires a blind search for ambiguity, without regard to the context in which the text will occur. It is well known (Schegloff, 1976) that what look like syntactic and lexical ambiguities outside of any context turn out in a particular context not to be ambiguous at all.

And yet it is not enough work because the effort is almost certain, with its necessary emphasis on context independent judgments of ambiguity, to miss just those kinds of ambiguities which will arise in context. In other words, this method will provide for many ambiguities which will never arise but will not provide for many of those that will.

In addition, the programming work required for this method would be tremendous. Let's say, based on the sessions I gathered, that there are an average of 500 tutor utterances for a one hour session (not counting utterances like mhm, right, yeah); now let's say, to be extremely conservative, that each utterance displays two sources of ambiguity (outside of any context), leading to two different misinterpretations. Each of these misinterpretations would have to have its own repair initiating text stored with it. We would thus need 2 stored pieces of text for every tutor utterance--roughly 1000 pieces of canned text *just to handle the possibility of one kind of repair*--these utterances would have no other function in the system; they would not even necessarily be useful in managing other kinds of repair.

Now let us suppose that we have a system which has roughly 1000 pieces of canned text specifically for third position repair and which is capable of figuring out roughly which piece of canned text to use at which time. Is this enough to satisfy the fourth process?

In human tutoring, the tutor reformulates the source of trouble so that the specific misunderstanding it responds to is excluded as an interpretation. The repair is extremely sensitive to the context of its production, and shows this sensitivity in its very design.

For example, consider the following passage. It is very close to passage 1 in character, but **differs** from it in important details. It follows passage 1 in the transcript.

(3)

T: [READING PROBLEM] How strong is the electric field between the plates
of a twenty microfarad.Now what's microfarads

S: Ten to the negative six
(0.7)
T: Okay, so what is a farad, what units is that.
(0.6)
S: Wait- oh what's a farad equal?
[
T: Yeah.
(0.2)
T: Like no, I uh- are they talking about length? That's

what I want (to know).=

S: =Oh. (0.6) Uhm, the farad is the capacitance=

T: =Good.

The student responds to the first version of the question with a value rather than an equation--a microfarad is ten to the minus six farads. Notice the tutor's second attempt in comparison to the second attempt of passage 1: first, it picks up on the implied "farad" in the student's answer and uses farad rather than microfarad (notice that there was no corresponding switch to a higher level of units in the earlier passage); second, it stresses is rather than what, presumably because of this very shift in units--the student has identified the relevant set of units but has not said what they are used to measure; and third, the tutor starts off with okay, so instead of right, but.

The student takes the stress on is to mean that the tutor is looking for an equation (as in passage 1). At this point, the tutor steers the student to the right level of analysis, again using length as an obviously incorrect but right level of analysis answer. Notice too that the tutor does not preceed her last formulation with "I'm not asking a specific enough question"; she assumes that the student will recognize the situation they are in and will recognize the tutor's use of length as the wrong answer but the right level of analysis.

In order to test the context specificity of the repair process, let me take the liberty of making up a passage based on passages 1 and 3 above, with the tutor questions patterned on those of passage 3 and student utterances patterned on those of passage 1.

(4)

(Altered example)

T: Now what's electron volts.

S: Charge of an electron times voltage?

T: Okay, so what is electron volts?

S: It's smaller

T: Like no, are they talking about length?

Even though the two passages on which this blend is based are extremely similar, their parts cannot be interchanged. While the result is not completely bizarre, it is pragmatically odd. The tutor's question at line 3 does not respond to the student's answer at line 2 at all: it does not take in to account that the student has just offered an equation, so that stressing is, in conjunction with mentioning a set of units (electron volts) which could be one side of an equation, is more misleading than it is helpful; it makes use of a device, okay, so, which says that the student has gotten one step closer to solving the problem and now should see the next step, which in this case is not true; it does not refer in any way to anything brought in by the student's answer, and so does not bring the student an inch closer to being able to understand (a) what her misinterpretation was and (b) what the intended interpretation is. In addition, a tutor is not likely to use a somewhat unusual technique, like asking an obviously incorrect question, for the first time without prefacing it in some way. The tutor's question at line 4 is therefore unnecessarily jolting and likely to be misunderstood.

Tutor utterances which initiate repair are thus intensely sensitive to the exact context of utterance, including the history of the interaction to that point--e.g. what the tutor and student have talked about, whether the tutor has engaged in similar repairs earlier in the interaction. There is no way that canned text--given that it is fabricated outside of any context and must be usable in all contexts--could reach this level of sensitivity. Repair as human tutors do it is thus theoretically impossible for systems using canned text.

The question is, then, given these findings, could canned text ever be even reasonably effective in accomplishing repair?

Before initiating repair, the tutor knows what kind of answer is appropriate for her question, knows the kind of misinterpretation the student has made, and chooses a path to lead from one to the other, so that the tutor uses the misinterpretation to get to the intended interpretation. For example, it is quite obvious that the response to "one point six times ten to the minus nineteen" cannot be the same in 5 below as in 6, since the intended answers are different in the two cases, and therefore the tutor must be "steering" in different directions.

(invented examples)

(5)

[Looking for the answer: charge times voltage]

T: What is electron volts

S: Energy.

T: Right, but what does it equal.

S: One point six times ten to the minus nineteen joules.

T: Yes, but how is it calculated?

S: charge of an electron times the voltage

T: Right.

(6)

[Looking for the answer: energy]

T: What is electron volts.

S: One point six times ten to the minus nineteen joules.

T: Yes, but how is it calculated?

S: ?

So the tutor must keep in mind the kind of answer which would be the appropriate kind of answer.

Furthermore, the tutor must design a new version on the basis of what the misinterpretation was, so that the misinterpretation can be used as leverage to manouever the student to the right kind of answer.

(invented examples)

(7)

[looking for the answer: energy]

T: What is electron volts
S: One point six times ten to the minus nineteen joules.
T: Right, but what are joules.
S: Energy.
T: Right.

(8)

T: What is electron volts.
S: Charge times voltage
T: Right, but what are joules.
S: ?

It is essential to notice here, too, that the reformulation responds to the surface form of the utterance as well as to the basic underlying misinterpretation. For example, the following invented example displays the same misinterpretation as example (2) above, but with a different surface form for R's first utterance. The rest of the passage remains the same.

(9)

(Altered example)

M: Do you know who's going to be at that meeting?
R: No.
M: I don't know.
R: Oh::.

While No and Who are both legimiate responses to pre-announcements (Schegloff, class notes), and thus share the same misunderstanding of line 1, they must engender different responses. Since no does not ask a question, it is not appropriate to respond to it with an answer. We could replace Who with any of its usual alternatives--I don't know, I give up--and the result would be the same--unacceptable discourse. The surface formulation of the utterance displaying the misunderstanding plays a large role in determining what is an acceptable and useful repair initiator.

The same thing holds for repair in tutoring discourse. Consider the following altered example:

T: Now, what's microfarads.
S: Isn't it ten to the negative six?
T: Okay, so what is a farad.

Although intuitions about invented examples are not entirely uniform across individuals, it is my sense that line 3 of this passage is not appropriate given the form of the student's utterance at line 2.

Above I suggested that a canned text system would need somewhere near 1000 pieces of stored test just for third position repair. Now, if we take into account the fact that repair utterances respond to the surface form of the misinterpreting utterance, the number expands towards infinity, since the number of ways a student could phrase something is huge, if not infinite.

Now, it is my opinion that either 1000 or upwards of infinity stored pieces of canned repair text is enormously greater than any canned text system could manage. A knowledge of discourse structure and usual interpretations would help reduce the number, but one of the reasons for using canned text to begin with is that we don't have enough conscious knowledge of these matters ourselves to be of much use to a computer. My conclusion, then, is that handling the process of third position repair seems beyond the reasonable (and very likely theoretical) limits of canned text.

If we extend our view to include other kinds of repair than third position repair, the problems multiply. Consider the following case, a more common kind of repair.

The student has brought with her a sample exam, presumably made up by her instructor as a study guide for the up-coming exam. At one point during the session, the tutor looks through the sample exam, and the following exchange takes place.

(10)

T: (LOOKING AT A SAMPLE EXAM) So how long do you have this?
(0.6)
T: To do this?
(0.3)
S: To do what.

T: This test.

S: .hh (0.4) uhm (0.3) two hours.

In the first three lines, the tutor appears to be asking how long the student has to do the sample exam (which the tutor is holding as she utters lines 1 and 3). But this question does not make sense-- a sample exam is not done in a limited amount of time the way a real exam is. Furthermore, the tutor knows that the student has her real exam the next day, so from that she could presumably deduce how long the student has to work through the problems on the sample exam. There is thus clearly a confusion between the sample exam and the real exam it stands for, hence the initiation of repair by the student. The student designs her repair initiator so that she indicates what she has understood and what she hasn't--with this formulation she indicates she does not understand the referent of this. The tutor responds by providing a more explicit version of the NP targeted by the student's repair initiator.

The surface form of the student's repair initiator greatly affects the form the tutor's repair will take. For example, the following uses of this test are not appropriate:
(Invented examples)

(11)

T: How long do you have to do this?
S: Huh?
T: This test.

(12)

T: How long do you have to do this?
S: How long?

T: This test.

In order to help the student, the tutor must know exactly what piece of the original utterance to repeat or elaborate on. Huh? would normally get a repeat of the entire utterance while How long?, as a check of hearing, would get a confirmation (yes, I did say how long). The variety of operations which could be performed on just one utterance would be quite large, greater than the number of words in the utterance, and each of these would need a separate response from the tutor:

(Invented examples)

(13)

T: How long do you have to do this?
S: How>
T: Long.

(14)

T: How long do you have to do this?
S: How long?
T: Yeah.

(15)

T: How long do you have to do this?
S: How long what?
T: Do you have to do this.

(16)

T: How long do you have to do this?
S: Who?
T: You.

(17)

T: How long do you have to do this?
S: How long do I what?
T: Have to do this.

etc.

For a system using canned text, there would seem to be two options for dealing with this problem.

The first would be, for every utterance, store the response to every possible repair initiator. This effort would take at least 10 responses for the average utterance, times roughly 500, which would give 5000 more stored pieces of repair text per session. The second option would be to have a separate subroutine which could be called to perform the operations necessary to produce the appropriate repair.

In some cases such a solution would work. For example, it is nearly always appropriate to respond to huh? by repeating the whole utterance. That would be easily enough taken care of by a separate module. In general, anything that requires repetition could be easily handled; in response to something like how long to do what? the system could repeat the previous utterance verbatim starting with the word after do. But repair initiators which call for some elaboration of the trouble source--as in example (10)--cause terminal difficulties for this solution, on two counts: (a) it is impossible to distinguish, without intonation cues, some uses of elaboration initiators from repeat initiators; and (2) appropriate elaboration cannot be done as an operation on an existing utterance, since it requires more information. Example (18) below illustrates the first difficulty:

(18)

(invented example)

A: I want you to go to the store tomorrow.
B: When
A: Tomorrow
B: No I mean when tomorrow
A: Oh, in the morning

B's when was taken to be asking for simple repetition when in fact it asked for elaboration. These two uses would usually be distinguished by intonation contours (request for repetition would be done with high rising intonation, while request for elaboration would be done with falling intonation). Since intonation is not available over a terminal, except by means of punctuation (which would be problematic, since standard American English punctuation does **not** correspond to intonation), these items represent a source of ambiguity.

But more importantly, these items can request something that was not in the original utterance. Where then would the requested information come from? For example (10) above, what part of the system would have known to produce this test?

As far as I can tell, there would have to be a module of the system which could go back to the knowledge representation and code the desired piece of information into natural language. But that, of course, requires natural language generation capabilities, which we, in this imaginary system, do not have.

This kind of repair, then, cannot be handled by means of canned text.

In fact, it seems that repair in general is beyond the capabilities of canned text, even in the context of a system that has highly advanced natural language understanding facilities. Creating 1000 or 2000 repairing utterances to be stored away would fix the trouble in a proscribed set of cases but would leave a much larger set in a bigger muddle than they were to start with; it is also an excessive, brute-force solution to a problem which deserves careful and insightful analysis.

2.2 Internalization Routines

One of the key tasks set before a learner at the outset of the learning cycle is to make the transition from assisted learner to solo problem-solver.[4] This transition is facilitated by what Collins et al (1986) have called an "internalized guide," which is constituted in part by the learners perceptions of aspects of the instructor's performance.[5] The learner uses this internalization of the instructor as a resource during practice in a process known as "reflection" (Collins et al, 1986)--"the process that underlies

[4]It is certainly possible for individuals to teach themselves certain skills. But notice that this is a situation to be remarked on: the learner in such a case will very often say "I taught myself." In contrast, one rarely says "Oh, someone taught me that."

[5]My own experience as a learner suggests that the aspects incorporated are not only the functional, problem-oriented ones; characteristic mannerisms, tone of voice, phrasing, even style of dress, can all be part of this internalization.

the ability of learners to compare their own performance, at both micro and macro levels, to the performance of an expert" (p.4). Internalization is thus a critical process in learning.

In human tutoring, tutors seems to be intensely concerned with the internalization process, to the extent that all of the tutors in my study engaged in formulating for the students repeatable procedures which the students could carry out when solving problems alone. I think of these as something like socialization routines (Ochs and Schieffelin, 1983), since they serve the function of having a less than fully competent member of a community internalize the knowledge of a more competent member, so that the former can perform competently even when the latter (or someone like her) is not present. I have therefore chosen to call these internalization routines, since they are prompts for how to behave in the future, when the tutor is not present.

The kinds of procedures offered cover a wide range of contingencies: what to do if you start to panic in an exam, how to ask yourself questions about a problem before working it through, how to write things on an exam to maximize the number of points you'll get, how to choose a method of solving a problem, and others. They deal with real-life situations (such as how to do well on an exam, what to do if you panic, etc.) and problem-oriented situations (how to set up a problem). One thing they all have in common is they are fitted to the history of the particular interaction within which they are situated.

The following passage is completely typical of this phenomenon.

(19)

T: What is the speed of a three hundred and fifty ee vee
electron.Okay.
(1.2)
T: So the main thing he:re, (.) I mean when you look at that,
what is electron volts, what kind of a

S: It//'s ahm

T: What are we talking about.

.

.

T: here you got kyu going through one volt. And what are these,

these are>

(0.9)

T: Those aren't lengths, so what are they=

S: =That's- (.) the work?

(0.3)

T: Work or energy.

.

.

T: So they want you to relate speed to energy.

.

.

T: What- equation do you know, that is an energy equation,

that has the velocity, of (0.7) the particle in it.

.

.

T: so:, when you look at this, (.) you know you got

this three hundred electron volts, and you go- and you

always go, oh my God what (0.4) what is an electron volt

.

T: and then if you can any way fool around with the units,

to figure out well what is it I'm talking about=

S: =a//ha

T: you know, and you go well its kyu times vee: and you

look over (at) your equation and you go okay well

that's (0.2) work and energy are the same thing.

S: Aha

T: And then you go, energy, okay, well what energy

d- equation do I know that has velocity in it.

In this procedure, the tutor labels for the student the steps which she has just led her through to solve a problem the student had been unable to even start by herself. It is at once a summary of what they have just done and a procedure for how to do something like it in the future. In some cases the tutor repeats nearly verbatim what she had said in the course of solving the problem with the student. The student now knows the procedure in two forms: she knows it experientially from having worked through the problem in this way; and she knows it verbally, from having heard the tutor's description of what they've just done. The verbal routine serves at least two functions: it provides the student with appropriate labels for her behavior; and it gives the student something to follow when she has to work a similar problem by herself.

The important point about this passage for the purposes of the present paper is that the tutor doesn't provide just any homily to end the problem with; she chooses a procedure which has been critical in solving the problem and words the procedure in such a way that the connections to the preceding problem-working efforts are obvious. Now, since the steps the tutor and student take cannot be planned in advance--they depend entirely upon the needs of the individual student--then the tutor's final formulation of the procedure cannot be planned in advance. Therefore procedural routines of this sort cannot be stored as canned text.

Another example is given below.

(20)

(the student has been working a problem)

S: Is that-
(0.2)
T: This is- // this is how I would- show it.

S: Okay.

.

.

T: You don't plug in- n:umbers, .hh until you've f:ound
what you're solving for.

.

T: And that way- (0.7) It doesn't matter so much in-
in (0.5) problems like this, b//ut when you get really

messy,

S: Right

S: Mhm>

T: It's easier to leave the numbers out (.) at first.
(0.5)

S: Right

T: And put 'em in later.

The tutor here has shown the student how to solve a problem without using specific values until the last step (the student had been trying the solve the problem by putting values in right away). After she demonstrates the technique, the tutor provides a little take-away procedure for handling (in general) values versus symbols. Notice that this particular procedure is invoked because of its relevance to what has just preceded it.

One last example is given below.

(21)

T: And then the other thing is is- y- the thing to do
is .hh like (0.5) if I'm grading these exams if
someone writes down the right equations, okay

S: Right

T: With all the right things. And then makes the mistake
on plugging the numbers in, that not that bad of a mi//stake,
right? .hh Right. So what you want to do is make sure=

S: It's not as bad as, right

T: =you write down the equation and then (.) plug in
the numbers.

S: Okay.

T: But be sure to write down the equation. Just don't

(0.6) go off putting numbers down.

Occasionally the tutor will provide procedural routines of this type even when the trouble addressed by the routine has not yet arisen in the particular interaction. It appears that tutors can anticipate, based on their own experience and their experience tutoring other students, that certain kinds of trouble will come up once the student tries to work problems solo. But these cases are distinguished from the others by virtue of explicit reference to someone other than this particular student's past or habitual difficulty (could be the tutor's):

T: This is- this is certainly the hard part.=

S: =Yeah, ge- breaking it dow//n is- is definitely the hard part.

T: .hhh

T: Yeah, and- what I recommend doing- is once you think you
have the numbers, (0.9) plug them in and check it,
(0.7) because I usually have to (0.4) go back- like
two or three times, to catch all my little mistakes.

3 Conclusion

Human tutoring is a highly complex, interactive process, wherein each next utterance is carefully fitted not only to the previous utterance but to the entire history of the discourse. This context-dependence is nowhere more apparent than in repair and internalization routines; here we see the constant self-adjustment of tutor and student that characterizes all successful interaction. If ICAI systems are to move in the direction suggested by Collins et al, they will have to accomodate the context-sensitivity of human interaction. In particular, they will have to overcome the limitations of canned text, which inhibits effective communication and successful tutoring.

REFERENCES

Collins, A., Brown, J.S., and S. Newman. 1986. "Cognitive apprenticeship: teaching the craft of reading, writing, and mathematics." BBN report no. 6459.

Fox, B. 1987. "Interactional reconstruction in real-time language processing." Cognitive Science.

Mann, W., Bates, M., Grosz, B., McDonald, D., McKeown, and W. Swartout. 1981. "Text generation: the state of the art and the literature." ISI report RR-81-101.

Ochs, E. and B. Schieffelin. 1983. Acquiring Conversational Competence. Boston: Routledge and Kegan.

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

Schegloff, E. 1976. "Some questions and ambiguities in conversation." Pragmatics Microfiche.

----------. 1972. "Notes on a conversational practice: formulating place." In D. Sudnow (ed.) Studies in Social Interaction. New York: Free Press.

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