UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Language Differences in Face-to-Face and Keyboard-to-Keyboard Tutoring Sessions

Permalink

https://escholarship.org/uc/item/7dh9q229

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 13(0)

Authors

Seu, Jai
Chang, Ru-Charn
Li, Jun
et al.

Publication Date

1991

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Language Differences in Face-to-Face and Keyboard-to-Keyboard Tutoring Sessions*

Jai Seu and **Ru-Charn Chang** and **Jun Li** and **Martha Evens**
Computer Science Department
Illinois Institute of Technology
Chicago, IL 60616
csevens@harpo.iit.edu

Joel Michael and **Allen Rovick**
Department of Physiology
Rush Medical College
Chicago, IL 60612

Abstract

Face-to-face and keyboard-to-keyboard tutoring sessions were recorded and analyzed as a first step in building a machine tutor that can understand and generate natural language dialogue. There were striking differences between these modes of interaction. The number of turns in an hour-long session dropped and so did the length of the sentences, although the number of sentences per turn stayed roughly the same. Students contributed 37% of the words in the face-to-face sessions; their share dropped to 25% in the keyboard sessions. Sentence structure is simpler in the keyboard sessions. The tutors ask more questions in the keyboard sessions; they explain this as a deliberate strategy to keep the dialogue going. The tutors also use a much wider range of expressions of acknowledgement in the keyboard sessions in a deliberate attempt to communicate verbally the kind of encouragement that is often expressed nonverbally in a face-to-face situation.

Introduction

This paper compares aspects of the language used in face-to-face and keyboard-to-keyboard tutoring sessions. These sessions were recorded as part of a project to design and develop a machine tutor that can generate tutorial dialogues. The tutor is intended to help first-year medical students understand the complex negative feedback system that regulates blood pressure (Kim et al. 1989).

The first step in our research was the recording of human tutoring sessions and the analysis of the transcripts of those sessions with regard to both language and content. The language analysis involves the study of vocabulary, syntax, and discourse for use in the understanding of ill-formed input (Lee et al. 1990; Seu et al. 1991) and the generation of responses (Zhang et al. 1990). The work described here formed part of the first phase in the language analysis.

Studies of keyboard-to-keyboard sessions in the literature had led us to expect that the language in keyboard sessions would be terser and more elliptical (Thompson 1980; Grosz 1978). It was. Many of our ideas about what phenomena to look for come from Grosz or from Grishman & Kittredge (1986). From Grosz (1978) we expected frequent use of "OK" with multiple meanings. In fact, it was our most frequent single word sentence and it seems to mean "I understand," "You are on the right track," and "Get on with it" in about equal proportions. The use of "you" also parallels that described by Grosz, but "I" is used more often as an abbreviation for "increase" than as a personal pronoun. When it appears as a personal pronoun, it is often used in pseudo-questions by students, such as "I don't understand about contractility."

Our tutors responded to Fox' (1988b) description of repair situations in tutoring by trying very hard to avoid ambiguity, even though they do not seem to be as verbose as the tutors described in (Fox 1990).

Capturing the Tutoring Sessions

The data used in this study comes from two face-to-face sessions and eight keyboard-to-keyboard sessions. Each session lasted approximately one hour. The tutor in every case was either Joel Michael or Allen Rovick, faculty members in the Department of Physiology at Rush Medical College. The students were first-year medical students at Rush, who were learning this material as part of their first-year curriculum. The students volunteered to take part in our research and were aware that we were recording these sessions.

Each tutoring session begins with a description of a procedure that perturbs the system; then the student is asked to predict how certain important parameters will change. The procedures used in these examples were

*This work was supported by the Cognitive Science Program, Office of Naval Research under Grant No. N00014-89-J-1952, Grant Authority Identification Number NR4422554, to Illinois Institute of Technology. The content does not reflect the position or policy of the government and no official endorsement should be inferred.

haemorrhage (loss of 0.5 or more liters of blood) and a malfunctioning pacemaker. The student is asked to predict what will happen to these parameters at three different stages: at the direct response stage immediately after the perturbation takes effect, at the reflex response stage when the baroreceptor reflex provides negative feedback, and in the steady state.

The data on face-to-face tutoring sessions comes from two hour-long sessions both conducted by Joel Michael with two different first-year medical students at the time when they were studying this material in class.

The keyboard-to-keyboard sessions were captured by a program we wrote for the purpose. To run this program (called STS) requires two PC's or compatibles and a Hayes smart modem or an equivalent and a telephone line from one site to another. With this program we are able to capture 60-90 turns an hour; the limitation seems to be the typing abilities of the student and the tutor and not the equipment. We would be happy to give the STS program and the data to anyone who would like to use it.

In the first version of the program we had a formal turn-release mechanism, but it was possible for both tutor and student to type at the same time. This resulted in several unintelligible passages of intermingled characters. In the sessions (K1–K8) where this occurred we edited these passages, where we could reconstruct the words, into sequential turns. The interruption is marked TI or SI, depending on whether the tutor or the student performed the interruption.

In the second version of the program we made it impossible for the tutor and the student to type at the same time. Barbara Fox' work suggests that these interruptions play an important part in face-to-face sessions, so in the third version of the program we buffer what the interrupter types and display it on the recipient's screen. It appears in the transcript after the turn that it interrupted. We have now captured 24 keyboard-to-keyboard tutoring sessions, each one hour in length.

An important first step in our research was identifying each item in the data. Jai Seu wrote a program that numbers each sentence or fragment in a way that identifies the session, who is typing, the turn, and the sentence within the turn. For example, the number K24-ST-057-01 indicates that this sentence comes from Keyboard session 24, that the student is typing, that this is turn 57, and that it is the first sentence in that turn. The second component may be TU (tutor), TI (tutor interrupting the student), ST (student), or SI (student interrupting the tutor). The data described here comes from the first eight keyboard sessions.

Differences in Session Length and Sentence Length

Both the face-to-face and keyboard-to-keyboard sessions were roughly an hour long, but many more words were exchanged in an hour of face-to-face tutoring than in a keyboard session. This is consistent with results reported by Gamoran (1991) using somewhat different experimental conditions. She found three times as much tutoring in face-to-face sessions as in remote sessions.

The two face-to-face sessions contained a total of 6,410 words divided into 489 turns. The eight keyboard-to-keyboard sessions had a total of 7,303 words divided into 699 turns. Thus, the number of words per turn dropped from an average of 13.1 in the face-to-face sessions to 10.1 in the keyboard-to-keyboard sessions.

The students were affected by this drop much more strongly than the tutors, with the result that the tutors controlled the floor much more in the keyboard sessions than in the face-to-face sessions. In the face-to-face sessions the tutor spoke 4,094 (62.8%) of the 6,410 words and the student spoke 2,386 (37.2%). In the keyboard sessions the tutor typed 5,452 (74.7%) of the 7,303 words while the student typed only 1,851 (25.3%).

The sentence length dropped for both the tutor and the student. The average sentence length in the face-to- face sessions was six and a half words; this average dropped to five words in the keyboard sessions. The distribution in sentence lengths is shown in more detail in Tables 1, 2, and 3. The number of sentences per turn stayed roughly constant, however. There were 1.86 sentences per turn in the face-to-face sessions and 1.74 sentences per turn in the keyboard sessions.

Table 1. Average Sentence Length

	Face-to-Face	Keyboard-to-Keyboard
Tutor	8	6
Student	5	4
Combined	6.5	5

Table 2. Distribution of Sentence Lengths

	Face-to-Face		Keyboard-to-Keyboard	
	N	Perc	N	Perc
Tutor				
1 to 3 words	187	37.6%	266	32.7%
4 to 10 words	155	31.1%	355	43.6%
above 10 words	155	31.1%	193	23.7%
total	497		814	
Student				
1 to 3 words	186	45.1%	273	67.7%
4 to 10 words	155	37.6%	69	17.1%
above 10 words	71	17.2%	61	15.1%
total	412		403	
Combined				
1 to 3 words	373	41.0%	539	44.2%
4 to 10 words	310	34.1%	424	34.8%
above 10 words	226	24.9%	254	20.9%
total	909		1217	

Table 3. Average Number of Sentences in a Given Length Category per Session

	Face-to-Face	Keyboard-to-Keyboard
Tutor		
1 to 3 words	93.5	33.2
4 to 10 words	77.5	44.3
above 10 words	77.5	24.1
total	244.5	102.0
Student		
1 to 3 words	93.0	34.1
4 to 10 words	77.5	8.6
above 10 words	35.5	7.6
total	206.0	50.0
Combined		
1 to 3 words	186.5	67.3
4 to 10 words	155.0	53.0
above 10 words	113.0	31.8
total	455.0	602.0

Differences in Sentence Complexity

With the difference in sentence length came a marked drop in sentence complexity. We counted the numbers of subordinate clauses. The frequency of sentential complements and relative clauses was cut in half. The frequency of adverbial clauses (mostly "if" and "when" clauses) also decreased but not as markedly. We assume that this happened because adverbial clauses are used to express the underlying causal relationships that are the focus of the tutoring effort. The numbers of subordinate clauses and the percentages obtained by dividing these numbers by the number of sentences are shown in Table 4.

Table 4. Subordinate Clauses

	Face-to-Face		Keyboard-to-Keyboard	
Clause Type	#(Cl)	Perc	#(Cl)	Perc
Sent.Comp.	96	10.6%	57	4.7%
Adverbial	81	8.9%	86	7.1%
Relative	48	5.3%	32	2.6%
#(Sent)	909		1217	

Although the tutor use of subordinate clauses stayed roughly the same over the different keyboard sessions, the students varied a great deal, so we have given the data for individual students in Table 5.

Table 5. Student Use of Subordinate Clauses in the Keyboard Sessions

(S1-S8 are the eight different students involved. Clause Types (CT) are Sentential Complement (SC) Adverbial (Ad), and Relative (Rl).)

CT	S1	S2	S3	S4	S5	S6	S7	S8	Stu Tot	Tut Tot
SC	1	4	0	3	3	0	0	0	11	46
Ad	0	3	1	12	8	0	0	0	24	72
Rl	0	0	0	2	0	0	0	0	2	30

Differences in Vocabulary

We expected to find a wider range of vocabulary items in the face-to-face sessions. We were wrong. There are 740 types (different individual words) in the face-to-face sessions and 915 types in the keyboard sessions. There were 6410 tokens (total words) in the face-to-face sessions, so the type-to-token ratio is 740/6410 = 0.115. There were 7303 tokens in the keyboard sessions; the keyboard type-to-token ratio is 915/7303 = 0.125. Details of the type distribution are shown in Table 6.

Table 6. Distribution of Word Types

	Face-to-Face		Keyboard-to-Keyboard	
Total Used	740		915	
Used Once	432	58.4%	573	62.6%
More Often	308	41.6%	342	37.4%
Tutor	611	62.4%	770	66.6%
Student	368	37.6%	386	33.4%
Both	239		241	

Our reading and conversations with Bozena Thompson had led the computer scientists among us to expect phatics (particularly expressions of rage and frustration). Our colleagues at Rush predicted that no phatics would appear and they were largely right. The only phatics that we found were "Well" and "Oh." They predict that Rush students will not express rage and frustration when communicating with our machine tutor either. We are eager to find out. The distributiion of "Well" and "Oh" is shown in Table 7. It does not seem surprising that these expressions appear more often in spoken language than in typed language. We cannot explain why the tutors did not stop using them as completely as the students did.

Table 7. Distribution of Phatics

	Face-to-Face		Keyboard-to-Keyboard	
	Well	Oh	Well	Oh
Tutor	18	2	14	2
Student	49	22	3	0

Differences in Discourse Strategy

We have only begun to explore this dimension but there are certain fairly obvious differences between the two media here. The tutors asked more questions in the keyboard sessions (217) than in the face-to-face sessions (129). They explain this as a conscious attempt to get students to respond.

The tutors actively sought verbal means to express approval and enthusiasm in the keyboard-to-keyboard sessions, information that they express nonverbally in face-to-face seessions. Thus, the words "absolutely," "super," and "excellent" appear four or more times in the keyboard sessions and not at all in the face-to-face sessions. They also softened disconfirmations (cf. Fox 1988a) with "not exactly," or "not quite."

Our tutors responded to the description of repair situations in (Fox 1988b) with a determination to avoid such situations. Examination of repair situations arising in three talk-aloud sessions not included in this study suggested that the most common source of repair problems was "How" questions. They have made a deliberate effort to replace these questions with more specific questions beginning "By what mechanism ... " or "In which direction ... " For whatever reason we have too few repair situations in our current sessions to make any interesting generalizations about them.

The tutors expressed a feeling of loss of control after the keyboard sessions. We suspect that this feeling may explain why the tutors uttered only ten imperative sentences out of the 909 sentences in the face-to-face sessions (3 in one and 7 in the other), but typed in 90 imperative sentences (out of 1217) in the keyboard-to-keyboard sessions. The students produced no imperative sentences in either medium.

Future Research

Clearly we need to collect and analyze more data. We have already recorded six more face-to-face tutoring sessions, but they have not yet been transcribed. We are making plans to record eight more. We actually have twenty-four keyboard-to-keyboard sessions in three groups of eight; we used the first eight in these analyses because the protocols involved are closer to those used in the face-to-face sessions.

The numbers of subordinate clauses of different types were recorded by hand. These numbers are probably not quite accurate. We are constantly expanding our parser to handle ill-formed input and we hope in a few months to be able to parse all the sentences produced by both students and tutors. At that point we will be able to undertake a more sophisticated analysis of syntactic differences and carry it out more accurately.

We have already written software to count differences in vocabulary between the student and the tutor. We want to examine other differences in language usage. We are particularly interested in studying how student language changes over a series of sessions and in finding out whether it becomes more like that used by the tutor. We need to record a series of sessions with the same student carrying out a number of different procedures. This project should also contribute to the design of the student modeler.

Our tutors reacted to (Fox 1988b and 1990) with a deliberate and largely successful attempt to avoid repair situations in the keyboard-to-keyboard sessions. We are still trying to figure out how they accomplished this, because we do not know how to make the machine tutor recognize the need for repair or perform it.

Summary

We have recorded human tutoring sessions carried out both face-to-face and keyboard-to-keyboard in the process of gathering information about language, tutoring strategies and modeling techniques for a machine tutor. In carrying out the language analysis we discovered some major differences in language use.

Predictably, the number of turns and the sentence length dropped, but the students were affected much more than the tutors, so that the percentage of words contributed by the student dropped from 37% to 25%. As sentences grew shorter there were fewer subordinate clauses of all kinds, but there were still a fairly large number of adverbial clauses that express the basic causal relationships in the underlying domain.

The language used by the tutors shows evidence of several strategies for overcoming the lack of nonverbal communication. The tutors ask more questions and use more imperatives in an apparent effort to prompt student response. They also provide more verbal encouragement.

References

Fox, B. 1988a. Cognitive and Interactional Aspects of Correction in Tutoring. Technical Report #88–2, University of Colorado.

Fox, B. 1988b. Repair as a Factor in Interface Design. Technical Report #88–4, University of Colorado.

Fox, B. 1990. Final Report to the Office of Naval Research. University of Colorado.

Gamoran, M. 1991. The Tutor Behind the Curtain: Face-to-Face and Remote Tutoring. Paper presented at the Annual Meeting of the American Educational Research Association.

Grishman, R., and Kittredge, R. 1986. *Analyzing Language in Restricted Domains: Sublanguage Description and Processing.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Grosz, B. 1978. Discourse Analysis. In Donald Walker ed., *Understanding Spoken Language*, New York, New York: Elsevier-North Holland. 138–176.

Kim, N., Evens, M., Michael, J.A., Rovick, A.A. 1989. CIRCSIM-TUTOR: An Intelligent Tutoring System for Circulatory Physiology. In Mauer, H. ed., *Computer Assisted Learning, Proceedings of the International Conference on Computer Assisted Learning*. 254–266. Berlin: Springer-Verlag.

Lee, Y.H., Evens, M., Michael, J.A., Rovick, A.A, 1990. IFIHS: Ill-Formed Input Handling System. In Proceedings of Second Midwest Artificial Intelligence and Cognitive Science Society Conference. 93–97. Carbondale, Illinois: Computer Science Dept., Southern Illinois Univ.

Seu, J.H., Evens, M., Michael, J.A., Rovick, A.A. 1991. Understanding Ill-Formed Input to an Intelligent Tutoring System in an LFG Framework. In Proceedings of the Third Midwest Artificial Intelligence and Cognitive Science Society Conference. 36–40. Carbondale, Illinois: Computer Science Dept., Southern Illinois Univ.

Thompson, B.H. 1980. Linguistic Analysis of Natural Language Communication with Computers. In Proceedings of COLING 80, Tokyo, Japan. np.

Zhang, Y., Evens, M., Michael, J.A., and Rovick, A.A. 1990. Extending a Knowledge Base to Support Explanations. In Proceedings of the Third IEEE Conference on Computer-Based Medical Systems. 259–266. :IEEE.