EXCUSE ME, YOU'RE STANDING ON MY PENCIL --
THE NATURALNESS OF ESL DIALOGUES

DENNIS R. PRESTON
Eastern Michigan University

Though many recent proposals for the implementation of communicative language teaching make little or no mention of the dialog as a teaching device, it is not clear that this minimal example of native-speaker-like interaction is no longer important. In a recent work on communicative teaching strategies (Littlewood 1981), a number of suggestions depend on dialogs or parts of dialogs. The theoretical emphasis on discourse and the acquisition of discourse rules make exemplary dialogs a center of methodological consideration if not a major proposal for technical implementation. Withlowson (1978) bases nearly every argument concerning rhetorical and discourse facts (which he considers of primary importance to language teaching and learning) on invented dialogs of, presumably, native-speaker quality.

Perhaps more important, the use of dialogs to represent levels of discourse ("degrees of formality") seems to be a much-recommended technique. Paulston and Bruder (1976) in a number of sample lessons offer dialogs as effective means of displaying formal versus informal interactions. Those who claim that such dialogs are too restrictive still have the responsibility of providing students with the "pieces" of well-formed dialogs. Moreover, of course, the teacher presumably has some well-formed finished product in mind, and it is to that model which reference is made in judging the degree of success of the learner’s performance, no matter what degree of apparent error tolerance may be in effect.

For dialogs to be natural, of course, all areas of linguistic description must be satisfied, but there are three areas in particular which have been given considerable attention recently, and they are areas which have been singled out as particularly important by proponents of communicative teaching.

1 I ignore here the ambiguity of the "communicative" approach, I assume, no matter how contradictory, it has been influenced by at least discourse analysis, pragmatics and speech act theory, sociolinguistics and the ethnography of speaking, and psycholinguistics and learning theory trends more internal to the field of language learning.
In spite of these three areas of concern to recent proposals concerning the importance of communicative aspects of language teaching and their obvious relation to dialogues, this paper does not seek to justify in any way the use of dialogues in second and/or foreign language teaching and learning. It is obvious, I believe, that stretches of conversation ("dialogues") are of central concern to those who place communicative strategies first in language teaching and learning.

Whatever the specific methodological position or implementational function of the dialogue, how can its authenticity be judged? How can we be sure that what we give the students is an accurate representation of the language to be learned? There are clearly two means of analysis. The first, laborious but not doubt rewarding in the long run, would ask us to bring to bear all our ethnographic, speech act, and discourse analysis skills to bear on a sample or four six lines of speech. We would have to make sure that every act of sequencing, every representation of social ambience, and every functional representation was accurately and appropriately given. Such an analytic procedure has two shortcomings, though I would not hesitate to recommend such practice for prospective teachers.

1. The analytic sophistication we have in the areas necessary to carry out such an investigation is not yet sufficient.

2. The teachers, teacher-trainers, or textbook writers who would be most likely the ones to carry out such evaluations, at least at the practical level, do not have the training to employ the skills available.

The alternative is to rely on a global or synthetic response to the naturalness of dialogues, and this paper reports on one such attempt.

First off, I believe it is important to note that two potential sets of judges for the dialogues investigated here were rejected — linguists and language teachers.

Linguists are terrible judges of language at almost any level. Elsewhere I have noted that linguists have had so many sentences put to them for grammaticality judgment that their grammaticality recognizer are all burnt out, and, more seriously, that linguists are considerably more lenient than non-linguists in judging the acceptability of linguistic forms (Preston 1975).

To the extent that language teachers have been trained as linguists (no longer such a great concern), they may be disqualified on the same grounds, though I would add to their disqualification the following concern. The greatest complaints against the naturalness and authenticity of textbook language in general and dialogues in particular come from classroom foreign and, especially, second language teachers. Anyone who has spent time with such teachers will have heard over and over again "Who talks like that?" or "I wouldn't say anything like that if you paid me". It does not do any good, by the way, to suggest that teacher-modification of written dialogues to suit particular audiences...
or notions of accurate speech may be an easy cure; such discoveries of naturalness failures by a teacher is enough to send an otherwise valuable text to perdition. Perhaps some of this criticism is based, quite simply, on boredom. How many fresh examples of “In the Train Station” or “At a Cafe” can textbook authors provide! Even lines of authentic speech have taken on peculiar, teaching-dialog overtones for those long involved in instruction. I have spent a great deal of my professional life in ESL classrooms either as a teacher or observer, and, once, several years ago, when I heard a voice say “Excuse me, you’re standing on my pencil”, I was unable to respond in a normal manner. Fortunately the perpetrator was also an experienced ESL teacher, and as I started to laugh, he realized at once that his line was symbolic of lines we had taught so many times, realistic and appropriate as it was to the situation. (I was indeed, standing on his pencil!) That sort of unnatural reaction to natural language and the deep professional concern of many teachers for the practicality of dialogs make teachers as unreliable as linguists.

Teachers’ criticism of dialogs is matched to a certain extent in the literature. Here is one of the more polite ones:

In choosing a textbook, it is well to examine even apparently authentic dialogue material with the greatest of care. The dialogs may represent an interesting situation or relationship in the foreign culture, but be expressed in language which would never be used in such circumstances or among people of the type depicted. (Hill 1968, 275)

A little more strident than that is the following:

More importantly criticized from such information as has been discussed here under the term “register” is the faulty association of level and performance in large numbers of dialogs prepared for non-native speakers. The criticism here is not of the “dramatic” or “situational” lack in many dialogs though that is another area of common failure. Here I specifically refer to the fact that many situations which must be thought of as casual occurrences are vehicles for consultative or even formal English (Preston 1973, 14).

In my own case, at least, I felt I had done no more than look at language teaching dialogs through the eyes of an untrained drama critic, sophisticated only by classroom experience and linguistic training. How could such evaluations be more formally accomplished? Not all the linguist, not to mention pedagogical, concerns connected with dialogs can be attacked at once, for, as pointed out above, we lack, as yet, the linguistic sophistication for a thorough analysis of all matters which would determine the naturalness of a dialog.

Lacking a well-developed theoretical approach (and disqualifying teachers, and linguists), I decided to subject sixty published ESL dialogs to native speakers for judgments of their naturalness. Each dialog was given a brief intro-

duction providing the identity of the characters, their relationship, and the setting. The instructions given the judges were as follows:

This booklet contains sixty short dialogs in American English. Accompanying each is a rating scale of one through five. Please read each dialogue and assign each a rating. Use the following scale in indicating your judgment.

1. Completely natural; an accurate sample of American speech, suited to the participants and the setting indicated in the brief introduction.

2. A likely sample of American speech but perhaps a little off, not exactly suited to the people or the situation.

3. Undecided.

4. Not a particularly good sample of a likely American conversation; not especially well suited to the setting or the participants.

5. Completely inaccurate; such a conversation could not take place in American English.

The final sophistication of the data was hidden from the judges. According to a pre-judged analysis, the dialogs were classified as “formal”, “consultative”, or “casual”. Of course, there might have been some casual-consultative or consultative-formal overlap in the language itself, but the specifications of character, setting, and relationship should have made the intended registral identification rather precise. The dialogs were not labelled in any way, and they were presented in random order. The following samples illustrate three of the dialogs (one from each of the registers) exactly as they were presented to the judges:

1 A reader comes into a newspaper office and speaks to a man at the counter.
   A: I’m here to get some details about this classified ad.
   B: I’m sorry, but you’ll have to write to the advertiser.
   A: Why must I write to the advertiser.
   B: In order to get the information you want.

2 Two male college roommates are talking in their room.
   A: What’s the matter? What are you looking for?
   B: I have to pick Betty up in fifteen minutes and I can’t find my Dark Blue Tie.
   A: Have you looked in your drawer?
   B: I’ve looked everywhere—in my drawer, under my bed, on the dresser...

4. It might be argued that these brief introductions interfered with the judgment of the dialogs, but I felt that without some specification of situation and personal exactly those considerations of appropriateness thought to be so important in modern language teaching would not be brought to bear on these samples.

4. These terms are borrowed from the classification of English “styles” provided in Martin Jay’s The Five Clocks, 1961.
3 A college age man and woman who have just met are talking at a party.
A: Would you like to see some pictures?
B: Sure. Did you take them yourself?
A: Yes, I did. I've got a new camera.
B: These are really wonderful

Attached to each dialog was a rating scale, corresponding to the instructions given above, of the following shape:

| 1 | 2 | 3 | 4 | 5 |

As these particular items appeared in the questionnaire, they were 1, 2, and 3, respectively. In the following analyses, they are formal 6, casual 20, and consultative 19, respectively. The frozen and intimate registers, to use Jones' terms, were not represented since the former is seldom taught for oral production and the latter is, perhaps exclusively, learned through extensive language use in appropriate situations.

Four major conclusions may be drawn from the data:

1 At perhaps the most theoretical level, it is obvious that whatever cues are being responded to by the judges, there was considerable agreement in the ratings. Whatever global cues for dialog naturalness were being used by raters were being responded to with consistency. The standard deviation scores for individual dialogs show substantial similarity in responses. Of course, we cannot, except for individual items with especially high or low scores, hope to put together a profile of the clues used in evaluation. Suffice it to say, that the task, an evaluation of conversational naturalness or authenticity, does not seem to have been an unreasonable one.

2 Published dialogs (at least these ESL dialogs) are a lot better than I thought they were and better, I think, than most teachers would have admitted. It is true that these dialogs are far from the "worst" I have seen, but the selection of obviously bad dialogs would have prejudiced the results. I think the items included here are "typical." Most teachers, I think, would agree that these are not unlike the majority of published ESL dialogs.

On the other hand, perhaps I am too delighted by an overall rating of "2" for the entire sample (see Table IV). "2" indicates, to quote from the instructions, "a likely sample of American speech but perhaps a little off, not exactly suited to the people or situation." If we want to be real sticklers, we may demand that all our dialogs should earn "1." Nevertheless, I am surprised by this generally high rating for this set of conversations.

3 There is little difference in the averages for the different registral groups. The range (for men and women) is far greater among samples within one category than between any of the three registral levels. For example, at the casual level there is a range from 1.3 (#2) to 3.6 (#19) (see Table I), but the

| Table I: 10 male and 10 female responses to 20 casual ESL dialogs |
|-------------------|-----------|-----------|-----------|-----------|-----------|---------------|-----------|
| male responses   | MAV       | sd        | female responses | MAV       | sd        | TAV   |
| 1 2 3 4 5 6 7 8 9 10 |          |           | 1 2 3 4 5 6 7 8 9 10 |          |           |       |
| 1 | 4 | 2 | 1 | 3 | 5 | 6 | 7 | 8 | 9 | 10 | | 1.9 | | 1.7 |
| 2 | 2 | 1 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | 1.9 | | 1.7 |
| 3 | 5 | 4 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | 1.9 | | 1.7 |
| 4 | 3 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | |
| 5 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | 1.9 | | 1.7 |
| 6 | 7 | 8 | 9 | 10 | | | | | | | | |

| MAV = male average, FAV = female average, TAV = total average, sd = standard deviation |

| Table II: 10 male and 10 female responses to 20 consultative ESL dialogs |
|-------------------|-----------|-----------|-----------|-----------|---------------|-----------|
| male responses   | MAV       | sd        | female responses | MAV       | sd        | TAV   |
| 1 2 3 4 5 6 7 8 9 10 |          |           | 1 2 3 4 5 6 7 8 9 10 |          |           |       |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | 2.0 | | 1.9 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | |
| 5 | 6 | 7 | 8 | 9 | 10 | | | | | |
| 6 | 7 | 8 | 9 | 10 | | | | | | |
| 7 | 8 | 9 | 10 | | | | | | | |
| 8 | 9 | 10 | | | | | | | | |
| 9 | 10 | | | | | | | | | |
| 10 | | | | | | | | | | | |

Symbols as in Table I
(Table III), men 2.4, women 2.0. At every level men are less satisfied with the naturalness of the speech represented than women. Except at the formal level, the differences are not great enough to cause considerable comment, but two things should be noted. First, there is some research which suggests that women take less extreme positions in ratings regardless of the subject matter of the experiment. If “criticism” of the dialog may be taken to be a more extreme position, this helps explain the difference. A more subtle explanation suggests that the more radical difference at the formal level reflects the difference in men and women as regards standards in a speech community. If women adhere more to speech community standards, perhaps formal dialogs appear to be more natural to them, while less formal dialogs represent the more “natural” less standard-like register assumed more often by men.

In the long run these general tendencies may help us understand how such ratings may aid in the identification of more natural expression, but it is more likely that they will help us uncover the kinds of errors which produce the greatest failures in dialog construction.

With that end in mind, it is worth looking at some of the exceptionally high and low scorers to see if they give up any indication of their triggering exceptional scores.

In the casual register, the worst overall score was assigned to #19, 3.6 (Table I). What makes this item so bad?

Two college women who are close friends meet between classes on campus.

A: Hello there. I haven’t seen you for some time.
B: I’ve been terribly rushed this semester.
A: How are things going?
B: O.K., I suppose, but I don’t seem very well right now.

From the older, dialog critic’s point of view, I suppose the grammaticality of “I don’t seem very well” might be questionable, and perhaps that funny-sounding collocation is the reason for the low rating. On the other hand, I find more subtle errors here, errors more closely related to the kinds of communicative concerns discussed earlier.

For me, this dialog contains a sequencing error. The second time “A” speaks, her line is so much like a conventional greeting that I cannot help but feel she has somehow recycled the entire conversation back to the beginning. Even if there is no such simple rule as one which says you cannot greet a person twice

---

4 I do not mean at all to trivialize the important research that has been done on men and women’s speech. On the other hand, little has been done on responses to speech except in some of the classic sociolinguistic studies. In early every case, women were found to prefer “standard” forms. For an excellent example of this and some careful explanation see Milroy and Margrain (1978).
in the same conversation (or, a little more technically put, "open" a conversation twice), there are explicit rules which make "A"'s second line sequentially bad. For example, a rule of conversational sequencing suggests that relevancy is a dominating concern, and that relevancy is not generally determined by the situation but, in many cases, directly related to what has come before, often restricted to what has immediately come before. I understand "B"'s first line to be an adequate response to "A"'s first complaint that she hadn't seen her friend much this semester. "B"'s response is, in fact, an explanation, but it is the sort of explanation which already contains information of the sort sought in "A"'s second line. Briefly put, "terribly rushed" would be an appropriate response to "How are things going?" even if it were taken to be a sincere question rather than a formulaic greeting.

From a sequence point of view, then, either "A" has opened the same conversation twice (an incredibly easy rule to follow) or failed to respond relevantly to "B"'s previous remark. Though their rules of conversation which guide such sequencing are as yet poorly understood, it may be the case that just such failures to adhere to the gross rules outlined help cause the low ratings for this particular item. Sequencing, then, may be an important consideration in dialog naturalness.

Additionally, this dialog shows what I consider to be errors of appropriateness, perhaps induced by an inaccurate representation of the relationship between the participants. I do not find for some time and terribly items which characteristically belong to the casual level. I would have predicted such items as a while or really in their place. Compare the following, keeping in mind that the only objection directed against the earlier version at this point is the appropriateness of certain phrasing to the casual level which should have been indicated by the relationship between the participants outlined in the brief introduction.

A: Hi, I haven't seen you for a while.
B: I've really been rushed this semester.

Inappropriateness, in this case of registral level, may have been another factor in the low rating of this dialog.

Let's look now at an item which got especially high marks, casual #2, which earned an overall rating of 1.3.

Two men in their thirties who are neighbors and close friends meet in the street in front of their homes.
A: Well, Ted, how do you like your new car?
B: Don't talk to me about cars.
A: What's the matter?
B: I've already had a flat and the entire engine needs to be overhauled.
I should've taken your advice.

Men gave this dialog a 1.3 and women a 1.3; it is certainly idiomatic throughout, and, more interestingly, it contains a sentence which must not be interpreted literally in order for the dialog to make sense. "Don't talk to me about cars" (a negative imperative) is an invitation to "A" to do just that, inquire further about "B"'s new lemon. Notice that the sequence "Don't talk to me about..." serves this function while the superficially similar sequence "I don't want to talk about..." serves, in most cases, its literal function.

I would argue that part of the source of such high ratings for this dialog comes from the use of an indirect speech act. Though it has not been explicitly formulated, I suspect that there is some such conversational-speech act rule which says that the less formal a situation is, the more indirect speech acts there are.

If such careful proportions are the case, exposure to a variety of direct and indirect speech acts, presumably through practices and dialogues which contain such elements, will be extremely useful to second and foreign language learners.

In the worst dialog from this study and from the best, then, we have seen in operation all three of the areas of concern to recent linguistic theory which have had considerable influence on the communicative approach to language teaching and learning-discourse analysis, particularly sequencing; appropriateness, particularly the attention to level to formality; and speech acts, particularly the distinction between direct and indirect forms.

Of course the data presented here could be subjected to further particular analyses. Why, for example, do men and women differ so violently on formal #11 (1.3)? From this survey, however, it seems clear that a large number of published dialogues are perhaps not so bad as suspected. More importantly, considering the infancy of research methods in these areas of linguistics most important to communicative strategies, native speaker non-linguists do not constitute such a bad court of judgment for appeal in matters of language naturalness. At least, their responses are consistent, and particular analyses of high and low-scoring items suggest that such judges are sensitive precisely to those areas of analytic concern most recently influential in developing a methodological stance which emphasizes the importance of communication. Finally, perhaps, if nothing else, this paper will serve as a reminder that dialogues do constitute little units of communication and might be more thoroughly reviewed by communicationists in their search for techniques.

REFERENCES

LACT (Linguistic Agency, University of Trier).