THE ROLE OF INTONATION IN SIGNIFYING SPEAKER INTENT

BRUCE FRASER

Boston University

Introduction. Even to the most casual observer it is obvious that a speaker usually, if not always, intends to convey a great deal more than he says. Recent work within the fields of linguistics, philosophy, sociolinguistics, anthropology, and adjacent areas has focused on this point, and we have now a growing body of speculation specifying how to account for the fact that we mean more than we actually say. See, for example, Fraser (1975), Gordon and Lakoff (1971), Searle (1969), Grice (1967), for some representative research. None of these scholars, however, has seriously addressed the following question: to what extent does the intonation of the utterance determine how the utterance is to be interpreted? Or, to invoke the title of this paper, what is the role of intonation in signifying speaker intent? The following represents a small, preliminary attempt to get at this question.

I am well aware that certain sentences are characteristically used in more than one way. For example, a speaker may utter It's getting late and intend to have it count as a simple report (say, during a discussion about whether we should go to the movies tonight), as a complaint (say, when waiting for the waiter to arrive to take the order prior to the theater), or as a request that we leave (say, if the speaker were my wife, whom I had brought to a party of linguists). But we are not aware of whether or not the speaker in one or more of these utterances of the same sentence signifies his intended message through the use of a particular intonation pattern.

To get some evidence which bears on the question of intonation and utterance-meaning, we designed an experiment to determine if native

English speakers could accurately judge the intended use of a series of utterances when these were removed from their conversational setting. We selected the following six test sentences, which have a relatively direct interpretation and an indirect interpretation as well.

		Direct	Indirect
(a)	Spend more time studying	Order	Suggestion
(b)	You must eat in that restaurant	Order	Recommendation
(c)	Could you do that before the operation	Question	Request
(d)	Can you lift your right arm	Question	Request
(e)	Shouldn't you be on your way to N. Y.	Question	Suggestion
(f)	Do you have to keep the lights on	Question	Suggestion

We hypothesized that while there was a class of possible intonation patterns consistent with the direct interpretation, whenever the speaker intended the indirect interpretation to be taken, he would signify this by utilizing a distinctive intonation pattern, not drawn from this class.

Method. Three male native speakers of English served as speakers of these test sentences. Each speaker was asked to consider each of the six sentences and to consider a context in which one of the interpretations of the sentences might be most appropriate. Then the speaker was asked to imagine himself in this situation, create whatever conversation was necessary to evoke the mood, and then utter the example sentence with the intended interpretation. The three speakers uttered each sentence twice, once for each intended interpretation, thereby providing a total of 36 utterances (three speakers x six sentences x two interpretations = 36). These 36 utterances (preceded by a statement of this number) were arranged at approximately five-second intervals on a tape in two groups of 18 utterances each. The same sentence was never placed in succession on the tape, though two of the speakers uttered successive utterances one time in each group. Each group contained some utterances of a sentence with the direct interpretation, some with the indirect interpretation.

The subjects were 22 white, female college students who had little or no linguistic training. The experiment was carried out in two parts.

Subjects were first given a questionnaire containing the six test sentences intermingled with seven additional examples. Each of the 13 sentences was arranged as illustrated by the following example:

Would you like a piece of cake
Example: Question 1 2 3 4 5 6 7 Offer

Subjects were instructed via written directions to indicate for each sentence how they thought the sentence might be used in an ordinary conversational situation. Selecting (1) in the foregoing example would signify that the sentence could be used only to make a question, not an offer, and vice-versa for selecting (7). Selecting (4) would signify that the sentence might be used equally well for both interpretations, depending on the context. Subjects were instructed to ignore any opinions about how a sentence is most often used. At no time was the term 'intonation' mentioned.

These questionnaires were collected and the subjects were given a second rating sheet. Subjects were informed that they would hear a series of 18 sentences spoken by several native English speakers and that each sentence was recorded when the speaker was using the sentence for one of the two purposes indicated on the rating sheet. The rating sheets did not contain the sentences, but only the utterance numbers and rating information, for example:

1. Question 1 2 3 4 5 6 7 Suggestion
2. Suggestion 1 2 3 4 5 6 7 Order
3. etc.

Subjects were to indicate, using the same criteria as in the earlier questionnaire, their opinion of how the sentence was intended. After finishing the first group of 18, the subjects were given a second rating sheet and completed the second group of 18 utterances.

Results. Subject judgments on the questionnaire sentences are presented in Table 1.

The numbers to the right of each example in Table 1 indicate the number of subjects who judged the sentence had only the first reading (those who selected (1) or (2) on the rating scale), followed by the number who judged the sentence could have both readings (those who selected (3), (4), or (5) on the rating scale), followed by those who judged the sentence could have only the second reading (those who selected (6) or (7) on the rating scale). For example, 14 subjects judged that (a) could have both interpretations, 5 judged that (b) could have only the Order reading, and 4 judged that sentence (e) could have only the Suggestion reading. Of the 22 subjects in this experiment, only one maintained her initial judgments about the use of these sentences; all others changed their opinions when they heard the actual recorded utterances. Changes in judgment were of two sorts: (1) moving from an initial judgment of only a single interpretation to the

TABLE 1.

(a) Spend more time studying	Ordr/Sugg	7/14/1
(b) You must eat in that restaurant	Ordr/Rec	5/ 7/10
(c) Could you do that before the operation	Ques/Sugg	5/12/5
(d) Can you lift your right arm	Ques/Req	8/12/2
(e) Shouldn't you be on your way to N. Y.	Ques/Sugg	4/14/4
(f) Do you have to keep the lights on	Ques/Sugg	5/10/7
(g) Why aren't you cleaning your room	Ques/Sugg	8/10/4
(h) Why not try that one	Ques/Sugg	0/10/12
(i) Can I see that	Ques/Req	4/10/8
(j) Will you get here by 8	Ques/Req	7/12/3
(k) Can't you try another	Ques/Req	3/15/4
(l) You could help me now	Stat/Req	2/10/10
(m) That would be wrong	Stat/Ques	16/ 4/2

could-be-either position; (2) a polarity switch to the other interpretation. In less than 1 percent of the 792 judgments (36 utterances x 22 subjects) was the polarity switch in the wrong direction (i. e. having first judged, for example, that (b) could be used only as an Order and then, when hearing it intended as an Order, judging it to be a Recommendation). On the other hand, more than half of the subjects changed their minds on two or more of the six sentences, usually on at least two of the six utterances of the sentence that they heard.

The subject judgments on the 36 utterances were analyzed to determine how consistent they were in recognizing the speaker intent. The original data was rearranged so that for each sentence, all the utterances with one of the intended interpretations were to receive a rating of (1), while the second interpretation was to receive a rating of (7). The subject judgments for each intended interpretation for each sentence (3 speakers x 1 sentence x 22 subjects) were then averaged. The difference in the two averages for each sentence was taken to be the degree of separation in recognizing speaker intent. The results are presented in Table 2.

TABLE 2.

Sentence	Difference in average rating
(a) Spend more time studying	3.4
(b) You must eat in that restaurant	2.5
(c) Could you do that before the operation	1.1
(d) Can you lift your right arm	0.43
(e) Shouldn't you be on your way to N. Y.	0.43
(f) Do you have to keep the lights on	0.26

In an effort to corroborate the judgments of the subjects that the utterances were indeed actually different, we analyzed the fundamental frequency (f_0) contours of half of the utterances.[1] This group of 18 analyses contained each of the six sentences three times (once by each speaker), with two of the three utterances with one intent, one with the other. The results corresponded to the subject judgments on two grounds.

First, the intonation contours of the three speakers for a sentence when uttered with one of the interpretations intended were roughly the same. Moreover, for a given sentence, the contours for the two intents differed markedly (both visually and in terms of relative frequency change) in at least one part, thereby appearing to signal the different speaker intent.

Second, the degree of intonation difference was much greater for the first two sentences ((a) and (b)) than for the remaining four ((c), (d), (e), and (f)). The frequency difference at the peak of the contour over studying and must was considerably more than the differences between the rising contour of the questions in opposition to the other interpretations of the interrogative forms in (c)-(f). But as I have indicated, subjects were most clear in their judgments of speaker intent for sentences (a) and (b), and much less clear in (c)-(f); thus, the degree of intonation difference also corresponds to the subject judgments.

Discussion. I want to stress what the preceding presentation should have made evident: this work must be taken only as a suggestion of where future, serious experimentation might take us. I think, however, there are several points that might be made.

First, the judgments on the questionnaires and the subsequent judgments on the utterances reveal once again that native speakers are not necessarily reliable informants. That many of the students when making the questionnaire judgments did not consider more than one intonation contour was evident from their comments subsequent to the presentation of the utterances. But the opposite side of the coin is that many linguists present their 'crucial' examples with a particular intonation pattern, thereby possibly prejudging the interpretation.

Second, this sort of experiment raises the question of whether the direct interpretation intonation contours form one equivalence class, with the indirect contours forming another. This may be too strong. For example (cf. Sag and Liberman 1975), it may be only that there is an intonation contour which can guarantee that the direct interpretation is understood. Or, it may be that although there are contours which clearly indicate either the direct or indirect interpretation, they are used only if confusion might result; the rest of the time the

contour is left to vary within broad limits. Moreover, although I have referred to these intonation contours as if they were clear, discrete, and well defined, this view is a great oversimplification.

Third, although both perceptually and acoustically the intonation seemed to vary systematically in terms of speaker intent, no control for timing, duration, or amplitude was made. Any one of these might have been responsible for the observed experimental results.

Fourth, we have no way of knowing the extent to which subject expectation influenced judgments. We tried to select the semantic content of each sentence to be equally compatible with each of the two interpretations, but no check was run on this. If, for instance, the example in (d) had been Can you resemble your pet rabbit, the meaning would have surely been taken to be that of a direct question, and a foolish one at that. In addition, it may be that a sentence such as (e) with a negative modal is nearly always used to convey a suggestion rather than a question; thus, even with a strong question intonation, subjects might have had to overcome a strong suggestion bias.

Finally, the use of staged utterances has a number of obvious disadvantages, the worst being the high likelihood that extraneous, nonconversational factors may be introduced and influence the experimental results.

NOTE

1. Thanks are due to John Allen and Doug O'Shaughnessy of MIT for making available their pitch extraction system as well as for assisting me in the actual analysis. Since the data was used as corroborative rather than primary data, and since there was only a single sampling from each speaker for a given utterance-meaning, the actual contours are not presented.

REFERENCES

Allen, J. and D. O'Shaughnessy. 1974. Fundamental frequency contours of auxiliary phrases in English. Ditto, MIT, Cambridge.

Fraser, B. 1975. Hedged performatives. In: Syntax and semantics III: Speech acts. Edited by J. Morgan and P. Coles. New York, Academic Press. 177-210.

Gordon, D. and G. Lakoff. 1971. Conversational postulates. In: Papers from the Seventh Regional Meeting. Chicago, Chicago Linguistic Society. 63-84.

Grice, P. 1967. Logic and conversation. Ditto, University of California, Berkeley.

Sag, I. and M. Liberman. 1975. The intonational disambiguation of indirect speech acts. In: Papers from the Eleventh Regional Meeting. Chicago, Chicago Linguistic Society. 487-497.

Searle, J. 1969. Indirect speech acts. In: Syntax and semantics III: Speech acts. Edited by J. Morgan and P. Coles. New York, Academic Press. 59-82.