PAPERS FROM THE ELEVENTH REGIONAL MEETING/ CHICAGO LINGUISTIC SOCIETY.

APRIL 18–20, 1975
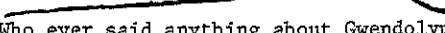
EDITED BY:

ROBIN E. GROSSMAN
L. JAMES SAN
TIMOTHY J. VANCE

Karttunen, Lauri. 1971. Some Observations on Factivity. Papers in Linguistics 4.55-69.

Kiparsky, Paul and Carol. 1970. Fact. Progress in Linguistics, ed. by Manfred Bierwish and Karl E. Heidolph. The Hague: Mouton.

Morgan, Jerry L. 1973. Presupposition and the Representation of Meaning: Prolegomena. Unpublished Ph.D. dissertation. Chicago: University of Chicago.

Rosenberg, Marc S. 1975. Counterfactives: A Pragmatic Analysis of Presupposition. Unpublished Ph.D. dissertation. Urbana: University of Illinois.

The Intonational Disambiguation of Indirect Speech Acts*

Ivan A. Sag
Mark Liberman

M.I.T.

0.   Sentences don't always mean what they say. In particular, they aren't always used to perform the speech act which their form makes them best fit for. By now it's a commonplace observation that sentences with the syntactic form of yes/no questions can be used to make requests, as in (1), or that sentences with the syntactic form of wh-questions can be used to make statements, as in (2), or suggestions, as in (3), and so forth:

(1) Would you stop hitting Gwendolyn?[1]

(2) Who ever said anything about Gwendolyn?

(3) Why don't you move to California?

    There is some disagreement as to whether a case like (2) is literally and conventionally a statement, or whether it is first and foremost a question, from the nature of which the hearer is expected to infer a statement. We will not address ourselves directly to this issue; the point of this paper is to discuss how the way in which such sentences are said, their intonation, is related to their intended usage.
    We began this investigation as a result of some remarks made by John Searle in a talk that he gave at the 1974 Linguistic Institute, on the topic of Indirect Speech Acts (which is his term for the non-question usages of cases like (1)-(3), as well as for other cases which we won't deal with in this paper). Searle suggested that the indirect versions of such sentences are intonationally different from the direct versions, in some way, and asked for help from linguists in pinning the differences down.
    From the beginning, we felt that there were real and characteristic differences. Although, for a given sentence, certain intonations are perfectly well suited to either type of usage, others are perfectly good in one case, but inappropriate in the other.
    For example, if two persons, say Aesop and Baldwin, were walking down the street, and Baldwin caught sight of a restaurant that was new to him, and wanted to know who was responsible for this new addition to their neighborhood, he might turn to Aesop and say:

(4)a Who opened the restaurant?

However, if our two characters were joint proprietors of the establish-
ment, and furthermore Aesop had accused Baldwin of having opened for
business on a day they were supposed to stay closed, Baldwin might deny
the accusation by saying:

(4)b Who opened the restaurant?

If in response to the accusation, Baldwin were to say (4)a, then what-
ever he might be doing, it would certainly not be the painless and ob-
vious denial accomplished by (4)b.

   In order to understand how and why these particular patterns
work in the way they do on these particular sentences, we spent consi-
derable time investigating the intonation of yes/no questions in general,
wh-questions in general, and other relevent categories. In this paper,
however, we will restrict our attention to the question raised by Searle,
even though our discussion will presuppose some more general ideas which
we don't have the space to develop here.[2] We will try

   I.   to show that there are ways to intonationally disambiguate
        direct vs. indirect usages of questions;
   II. to give a systematic description of the disambiguating factors;
   III. to explain why these factors should have the effect that
        they do.

1.1   What are the disambiguating intonations we're talking about?
First of all, there is a way of saying wh-questions which makes them
unambiguously "real" questions; its use in contexts suggesting Indirect
Speech Act usage is quite strange. Since a frequency/time graph for this
contour looks like an elongated tilde, we'll refer to this as the TILDE
contour. (4)a was an example of the TILDE contour, and (4)a was, as we
saw, quite inappropriate as a negative-implicating rhetorical question.
(5) is another example of the TILDE contour. Said in this way, (5) is
not appropriate as a suggestion:

   (5) Hey Baldwin, the climate here is really bad for you!
        I've got a suggestion—

        *Why don't you move to California?

   The TILDE contour has the following tonological characteristics:
1. High pitch on the wh-word (and associated elements). The utterance
    may begin on a mid pitch and rise to this high, or it may begin high.
2. A fall to low pitch on the next major stress of the utterance.
3. A rise in pitch on the last syllable of the utterance.
   To show that the initial high pitch is associated with the wh-
word, rather than simply occurring at the beginning of the utterance,
we consider the case of wh-questions with initial prepositions, which
sound strange with high pitch initially rather than on the wh-word, as
in (6).

(6) *With respect to which aspect of the problem are

                                          you most upset??

To show that the medial low pitch is associated with stress, rather than
lack of stress, we consider cases like those in (7):

(7)a Who opened the HUNGARIAN restaurant?

  b *Who opened the HUNGARIAN restaurant?

If we attempt to contrast the adjective Hungarian in the wh-question "who
opened the HUNGARIAN restaurant," said with a TILDE contour, we must main-
tain (or even reinforce) the low pitch on the stressed syllable, as in (7)a.
If we try to realize this stress in the more familiar mode, with a pitch
prominence, as in (7)b, the result is very strange.

   Why should this TILDE contour have the effect of forcing wh-ques-
tions to be "real" questions? We have two extremely speculative sugges-
tions to offer.

Speculation 1:
   We observe that the same contour is possible on greetings (as in
(8)a), assents (as in (8)b), and so forth.

(8)a Hello, Baldwin!   b That's a great idea!

A greeting said in this way is more jovial, expansive, and in general
just plain more of a greeting; likewise an assent is more assenting.
Perhaps the TILDE contour, in these cases, has the effect of "illocu-
tionary intensification," making an utterance of type X "more of an
X," thus making questions "more of a question."

Speculation 2:
   The TILDE contour on wh-questions is somewhat similar to what we
called the "contradiction contour" in an earlier paper.[3] An example of
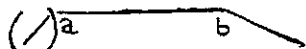the contradiction contour would be something like (9):

(9) Elephantiasis isn't incurable!

However, there are important tonological differences between the contra-
diction contour on declaratives and the TILDE contour on wh-questions.
In particular, the normal realization of stress inside the contradiction
contour will be precisely the familiar pattern[4] that is so strange in
examples of the TILDE contour like (7)b above. Also, as the contour in
(9) clearly shows, the initial high pitch of the contradiction contour
is not associated with the word stress of elephantiasis. Rather, it is

linked to the initial boundary of the utterance. This is not the case with the TILDE contour, which associates the initial high pitch with the stressed wh-word (whence the unacceptability of (6) above).

Despite these clear differences, it is interesting to note that both of these contours are quite literal-minded. For example, certain sentential matrices, such as "I'm afraid that..." can be parenthetical (or "adsentential"). Thus if one says, with almost any intonation in fact, "I'm afraid that you didn't pass the test", one may be talking not about one's fears, but about the test results. The contradiction contour, however, renders this sentence unambiguously a statement about one's state of mind.

Having apologized for the inconclusive nature of these remarks, let us turn to other ways to say wh-questions. One general class of wh-question intonations can be schematized as in (10), a shape which looks roughly like a hat.

(10)   Who opened the restaurant?

Hence, we will call this class the hat contour, for the sake of having a convenient name. These intonations may be characterized by:
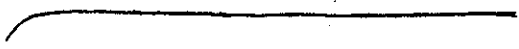1. A high pitch on the wh-word, optionally with a rise from an initial mid or low pitch.
2. The high pitch on the wh-word spreads to the right until a major stress is reached, and
3. After this first major stress, the pitch falls.
In the cases we will consider in this paper, this first major stress is also the last major stress, so that what happens in the more complex cases may safely be ignored.

The stress-linked fall (at b in the schema in (10)) may occur immediately after the wh-word, as in (11)a:

(11)a   Who opened the restaurant?

or it may be postponed until after quite a long stretch of high pitch:

(11)b   Who opened the new Hungarian restaurant on the

corner of Boylston street?

In general, almost any version of this "hat" contour may conceivably serve either for a real wh-question or for a negative-implicating rhetorical question, suggestion, etc. However, if the pitch at b in the schema in (10) is markedly higher than the pitch at a, then there is a fairly strong tendency for the utterance to be interpreted as something other than a real question, as in examples (2), (3), and (4)b above.
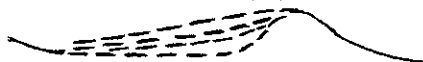
As we will see later on, this observation can be confirmed experimentally.

Such an observation is of some interest in and of itself; it helps to answer the question with which we began, "how are the intonations of indirect speech acts characteristically different from those of their direct counterparts?" However, we should not be satisfied to leave the matter there--the observation that intonation A has pragmatic effect B does not show even that there exists an entry A in the intonational lexicon, nor yet that if such an entry does exist, its "meaning" is B. The effect B may be achieved indirectly (e.g. by interaction with other elements of the utterance and with the context), the meaning being something rather different; and it is possible that the observed "intonation" may represent the interaction of several unrelated factors, being itself not a linguistic unit at all.[5]

As a matter of fact the "contour" observed in examples (2), (3), and (4)b above is subject at least to the conclusion that its connection with the effect of "indirectness" is an indirect one. It also may not be a significant linguistic unit--for example it may simply be a version of the "hat" contour, modified, say, by some process which links F_0 prominence to pragmatic importance.

However, for purposes of discussion, let's assume that the contour of these examples is a significant unit. To give an adequate descriptive characterization of this contour is beyond the scope of this present paper.[6] For now we will simply observe :
1. The presence of a low pitch near the beginning of the utterance, in association with a stressed syllable;
2. The presence of a high pitch associated with the last major stress of the utterance, with a subsequent abrupt fall;
3. Various ways of getting from the low pitch in 1. to the high pitch in 2. without any significant stretches of negative frequency/time slope.
These features are schematized in (12):

(12)   ...X........................Y..........
       (stress)   (stress)

Intonations with the characteristics schematized in (12) can be found on many different classes of utterance--if we consider a few such cases, certain generalizations about the use of this contour will emerge which can help us to understand its effect in the case of wh-questions.

Consider a simple declarative sentence, as in (13), said with the contour we've been discussing:

(13)   The blackboard is painted orange.

There are two kinds of circumstances in which this would be an appropriate way to say sentence (13):
1. where the speaker is expressing surprise--e.g. walks into the room,

observes the paint job, and exclaims "my god--the blackboard's painted orange!"

2. where the speaker is suggesting that the utterance is redundant or unnecessary--e.g. Question "what color is the blackboard?" Answer "I've told you a thousand times--" or "just open your eyes and look--the blackboard is painted orange!"

The reader is invited to work through a few examples to demonstrate that the appropriate contexts for such utterances are those that involve either surprise or redundancy, as opposed to, say, exasperation or vehemence.[7]

These twin (and in some ways rather opposite) implicatures of surprise and redundancy are also to be found in imperatives:

(14)a Go open the door! (I shouldn't even have to tell you...)

   b Well, butter my parsnips! (I never would've believed it.)

The same holds true of wh-questions:

(15)a (Albert walks into Sam's room, observes beautiful oriental carpet, and asks in amazement...)

      Where'd you get the rug?

   b (Sam and Albert have been discussing Sam's new rug. Albert suggests that Sam ought to get some matching runners. Sam asks well how can he find anything to match. Albert says...)

      Where'd you get the rug? (Obviously they would have more of the same stuff.)

Both (15)a and (15)b are real questions--the speaker doesn't know the answer, believes the hearer does, and wants to be told. But in (15)a the speaker is expressing surprise, while in (15)b he seems to be expressing the idea that his utterance is unnecessary.

One very common circumstance in which a wh-question is used to make a statement is in the kind of context mentioned in section 0 above, in which a statement or suggestion is made which the speaker regards as preposterous, and which (s)he attempts to dismiss by using a negative-implicating wh-question. In this kind of circumstance the wh-question is clearly felt to be redundant, the idea being that its answer is so obvious that the occasion for asking should never have arisen. Thus the observed association of Indirect Speech Act wh-questions with the contour schematized in (12) arises because of their commonly-found secondary characteristic of redundancy rather than the primary characteristic of indirectness.[8] If we consider some Indirect Speech Act wh-questions which lack the feature of redundancy, the use of the intonation in (12) is ludicrous:

(16)a (Albert grunts, throws down the thesis he's reading, grimaces, and says...)

      *When will I ever understand pseudoclefts?!

   b (Albert replaces the telephone in its cradle, walks to the window, and, looking out, says with a sigh...)

      *Who can imagine the treachery man is capable of?

1.2   Now for questions of the yes/no type. There are ways of saying sentences of yes/no question form which make them unambiguously real questions. For example, the sentence "would you stop hitting Gwendolyn?" can be said as in (17)a. On the other hand, it might be said as in (17)b, which disposes it in the direction of being an order or request.

(17)a Would you stop hitting Gwendolyn?

   b Would you stop hitting Gwendolyn?

In order to understand the nature of this distinction, it's necessary to say a few words about the intonation of yes/no questions in general. The familiar "rising" intonation characteristic of yes/no questions is differentiated from so-called "declarative intonation" in two ways. First, the major stresses in the declarative case are intonationally realized by a pitch accent which can be described as high-low or mid-high-low, while in the question cases, they are realized as low-high.[9] Thus in a yes/no question a strong, e.g. contrastive, stress near the beginning of the sentence, which subordinates stresses to the right, will result in a nearly level high pitch being carried along from the stressed syllable to the final syllable of the utterance. Thus in a question like (18)a, the result of the contrast on you is that everything from you to friends is relatively level in pitch, and quite high.

(18)a Are YOU one of Gwendolyn's friends?

   b So YOU'RE one of Gwendolyn's friends.

This is strikingly different from what happens in a similarly stressed declarative, in which a level low pitch is carried along to the end of the sentence, as in (18)b.

In addition to this difference in the mode of realization of stress, there is a second, more familiar distinction between declaratives and yes/no questions--the latter typically bear rising pitch on the terminal syl-

lable, while the former do not. This terminal-syllable rise is differentiated from the stress-realization rise (mentioned earlier) in two ways: first, it is not linked to stress, and second, it is not a rise from low pitch to high pitch, as in the stress-related rises in (17) and (18)a, but rather a rise from wherever the pitch happens to be (usually fairly high) to a still higher pitch. The presence or absence of this terminal rise is independent of the mode of stress realization.

That intonation that renders a yes/no question least likely to be interpreted as a real question, is a contour that involves low-high stress realization, as usual, but where there is no terminal rise, yielding a level high which spreads from the last major stress to the end of the utterance, as in (17)b. This case must be kept distinct from the case in which a yes/no question is said with declarative intonation, i.e. mid-high-low stress realization and terminal fall. Said in this way, the sentence could perfectly easily be either a real question or a request.

This far, what we've tried to do is to state the problem and to give a brief sketch of our hypotheses. In the remainder of the paper, we will report on the results of some preliminary attempts to justify our claims experimentally.

2.1   To corroborate our intuitions about the appropriateness of variously intoned utterances in various contexts, we recorded four subjects acting out various skits. We constructed these skits in pairs, so that it was clear that in one skit the question was being used as a real question, while in the other it was being used non-literally. For example:

(19)a Psychiatrist: Are you willing to make some sacrifices so
        that we can proceed with your therapy?
    Patient: I'm not sure.
    Psychiatrist: Well, for example, would you stop hitting
        Gwendolyn?

    b Teacher: You children are behaving very badly. Myra,
        please stop running around. And Lennie, I'm not going
        to tell you again--would you stop hitting Gwendolyn?

In (19)a it is clear that the sentence would you stop hitting Gwendolyn is being intended as a real question. The force of the same sentence in example (19)b is quite clearly indirect, i.e. that of a request or a command. Example (20) is another context pair:

(20)a Max: You know, Henry, the climate here is really bad for
        you. I've got a suggestion--why don't you move to Cali-
        fornia?

    b Max: Henry, I'm curious--why don't you move to California?
        Is it because you don't want to leave all your friends
        in Boston?

The underlined sentence is being used as a real question in (20)b, while in (20)a it is intended as a suggestion.

The first (though perhaps obvious) observation to be made about the intonations our four subjects used is simply that different speakers have different intonational styles. One speaker, for instance, never used the TILDE contour at all. Two others were quite fond of it. The systematic variations among these speakers' styles, however, will not concern us here.

A second observation is that just as we predicted, no speaker used the TILDE contour for suggestions or for rhetorical negative-implications. All occurrences of the TILDE contour were in real questions

Thirdly, it is clear that the negative-implicating questions, as well as the questions used as suggestions, involved less prominence on the wh-word, and relatively greater pitch prominence on the sentence-final stress, as in examples (2),(3),(4)a, and as predicted (given the nature of the contexts) by the ideas developed in section 1.1 above.[10]

We then conducted a pilot experiment, in which twelve subjects were (each) given six context pairs (like those in (19) and (20)), and were asked to listen to the underlined sentences intoned in various ways, and to indicate which of the paired contexts was more appropriate in each case. In this experiment, the stimuli were synthetic speech. The segmental phonology was identically synthesized, and duration was also kept constant. Only the fundamental frequency was varied.[11]

In addition to this forced-choice preference test, a second experiment conducted involved acceptability judgments. Informants were asked to judge whether or not an utterance (synthesized with some particular contour) was acceptable in some particular context. Both the stimuli and the contexts used were the same as in the preference test.

2.2   The results of these experiments lend considerable support to our intuitional observations about the TILDE contour. In the preference test, 31 out of a possible 36 judgments of the TILDE contour indicated a preference for the real question context. Likewise, in the acceptability test, 33 out of 35 judgments accepted the TILDE questions in real-question contexts, while 29 out of 33 rejected the TILDE questions as unacceptable in the nonreal-question contexts.

Even when we tried an utterance that we hypothesized would be fairly unlikely to be taken as a real question, i.e. "who ever said anything about Gwendolyn?" many subjects still judged it to be a real question when it bore the TILDE contour.[12]

In the case of the contour schematized in (12), our tests results although not as clearcut as the TILDE contour results, were very much in line with the ideas developed in section 1.1. In particular, such cases seem to retain their ability to function either directly or indirectly.

In the case of yes/no questions, the clearest results were those involving the underlined sentence in the contexts in (20). Here the presence of a marked terminal rise made all 12 subjects prefer the real-question interpretation, and made all but one find the request interpretation absolutely impossible. So here, as in the case of the TILDE contour, we find what seems to be effective disambiguation.

3.   In summary, in this paper we've argued that there are ways to dis ambiguate direct vs. indirect usages of questions in English by F_0 con-

tour alone, and that (relatively) naive subjects are reliably able to perceive this disambiguation. We have made some suggestions as to the nature of the disambiguating factors:
1. The TILDE contour forces wh-questions to be real questions.
2. The contour schematized in (12) sometimes suggests, but never guarantees, that an utterance is not a real question.
3. In yes/no questions, a prominent final rise forces the utterance to be a real question.

These conclusions suggest the following hypothesis: some intonation contours can "freeze" an utterance pragmatically, i.e. require a literal interpretation, but no intonation contour can force an indirect interpretation.

Whether this turns out to be true even for English, or true more generally, awaits further research. We hope we've given an indication of how such research can proceed.

FOOTNOTES

*We are extremely indebted to Jonathan Allen of the Electrical Engineering Dept. at M.I.T. for the use of his facilities and to Haskins Laboratories for the use of their facilities to synthesize the stimuli we used in our experiments. This research was supported in part by a grant from The National Science Foundation, and in part by grant 5 T01 HD00111 from The National Institutes of Health.

1.   The wavy lines in our examples represent fundamental frequency against time, with a certain margin of error because of typewriter spacing.

2.   For a more comprehensive treatment, see Liberman (in progress) and Sag (in progress).

3.   Liberman and Sag (1974).

4.   Ibid., 423ff.

5.   We feel that the relative lack of success of investigations of English intonation (compared to "conventional" phonological work) is due largely to the extreme difficulty of isolating the linguistic units involved. In the conventional domain of phonology, the question: "What are the linguistic segments (words, phones, etc.) that make up this example?" is relatively easy to answer, at least to a first approximation. In the area of intonation, this necessary first step seems to be a long one, although we are confident that it is a step in the right direction.

6.   See Liberman (in preparation) for some attempts.

7.   Thus it is appropriate to say, as a perfectly pleasant, unexasperated way to deliver a compliment:

(i)   That's the most beautiful hat I've ever seen!

It's inappropriate, however, for a flunky to come to the podium at the expected hour of the presidential press conference, and say solemnly:

(ii)
      Ladies and gentlemen, the U.S has invaded Abyssinia.

(unless, of course, he means to express surprise or shock)
      Conversely, it is appropriate to say (iii) as the most unvehement

(iii)   It's under the bed.

"who-cares-what's-the-difference" way imaginable of answering a question concerning the location of some irrelevant and uninteresting object.

8.   Another possibility is that the effect of redundancy is itself secondary. That is, perhaps this contour simply has a meaning of "surprise". Expressing surprise, under this account, would induce the effect that we have described as redundancy. For example, in the case of (15)b, the implicature: "This utterance is redundant, or unnecessary" could arise because the intonation of this utterance is expressing surprise at the very fact that it's necessary for the speaker to be asking such a question at all. Whatever the correct account may be, it will surely treat the redundancy cases and the surprise cases as a single phenomenon.

9.   For further discussion along these lines, see Goldsmith (1974a,1974b).

10.   Both these facts and our account of them are consistent with Larry Horn's observation that "destressing", when accompanied by phonological reduction of the wh-word and the following auxiliary, as in: " [waĩča] move to..." strongly favors a non-literal interpretation, here, that of a suggestion.

11.   We used synthetic, rather than natural speech in order to control for the effect of certain extremely poorly understood non-intonational factors such as "tone of voice".

12.   For a detailed discussion of the non-intonational factors that induce a negative-implicating interpretation, see Griffin (1974).

BIBLIOGRAPHY

Goldsmith, John (1974a), "English As A Tone Language". unpublished ditto,
_____ (1974b), "English As A Tone Language, Part II", unpublished ditto, M.I.T..

Griffin, Peg (1974), "Language Meaning and Reading". paper delivered at winter L.S.A. meetings, N.Y.,N.Y..

Liberman, Mark (in progress), M.I.T. Doctoral Dissertation.

_____ , and Ivan Sag (1974), "Prosodic Form and Discourse Function", in C.L.S. X.

Sag, Ivan A. (in progress), M.I.T. Doctoral Dissertation.