Is that a real question?:
Final rises, final falls and discourse function in yes-no question intonation

Atissa Banuazizi and Cassandre Creswell
University of Pennsylvania

1 Introduction
How do intonation, syntactic form and discourse function map together to provide cues for interpretation of utterances? Clearly, there is no one-to-one correspondence between any of these three categories. Nor is it easy to determine clear-cut classifications of intonational contour type or of discourse function in the same way that one might be able to do for syntactic form. It is still worthwhile, however, to attempt to characterize certain principles that may guide listeners in their understanding of these utterances. In this study, we will approach these issues via the phenomenon of yes-no question intonation.

The most common final intonational contour for yes-no questions is a rising one. Different theories of intonation have attributed various discourse functions and intonational schematizations to this final rise. A small number of yes-no questions, however, end in a fall. Our study will characterize the discourse contexts that contribute to the realization of questions with a final fall. In contrast to previous analyses of the same intonational contour, we propose that discourse structure alone -- whether the question is somehow related to the following discourse -- cannot account for the usage of final rises and falls, but rather that how the speaker intends a question to be interpreted is a more important factor.

2 Previous studies
2.1 Final rises within a general theory of the meaning of intonational contours
The meaning or discourse function of final intonation patterns of yes-no questions has been explored in previous studies, both within a more general theory of what final intonation patterns signify and within studies specifically focusing on questions. With regard to the former we will look at Pierrehumbert and Hirschberg (1990) and Hobbs (1990).

Pierrehumbert and Hirschberg (henceforth P & H) make a general claim about the meaning of boundary tones. Boundary tones are the final tone of an intonational phrase: they can be either high or low phonologically, with some variation in their phonetic realization depending on the previous pitch and phrase accents. According to this model, a high boundary tone (H%) conveys "forward reference." A speaker uses a high boundary tone to convey that the addressee should "interpret an utterance with particular attention to subsequent utterances." They frame this claim within Grosz and Sidner's (1986) theory of discourse structure. According to P & H, an utterance ending in a H% signals that the intention underlying the utterance either dominates or is dominated by a subsequent utterance. A low boundary tone (L%) conveys that the utterance need not be interpreted with respect to subsequent utterances. It should be noted, however, that in the Grosz and Sidner model, within a coherent discourse without any interruptions, every non-final discourse segment is interpreted with respect to some later utterance, that is it either dominates or is dominated by some subsequent utterance. Under P & H's hypothesis, then, all non-discourse final utterances could have a H% tone. Such a hypothesis thus makes no prediction about when
speakers will use either a H% or L% tone.

Hobbs (1990) suggests a modification of P & H's claims about boundary tones by proposing to frame them within a different theory of discourse structure. Briefly, the type of theory he discusses is more articulated structurally; under such a theory, a discourse has a hierarchical structure where individual utterances are linked to constitute a larger segment which can in turn be linked together to form an even larger segment. Hobbs claims that under such a theory a H% will signify that an utterance is a nonfinal subsegment of a larger discourse segment, that is it will be followed by more closely related information. A L% will, in contrast, not signify anything of its own; rather, it will simply fail to signify that the segment should be interpreted relative to subsequent utterances. Hobbs offers an additional paraphrase of the meaning of H% as indicating "the status [of the utterance it ends] is still an open question."

2.2 Final rises and falls within a theory of felicity conditions

Other theories of the discourse function of final falls and rises are based on theories of sincerity conditions rather than on theories of discourse structure. Two of these, Hudson (1975) and Kenworthy (1978), build on the discussion of the intonation of many types of interrogative structures in Bolinger (1958). One aspect of Bolinger's discussion particularly relevant to the goals of our study is an analysis of conducive questions. These are utterances with the pragmatic force of questions but with almost any syntactic form "that show that a given answer is expected or desired." Bolinger points out quite justifiably that probably no one intonation pattern necessarily indicates conduciveness. He does suggest, however, that the strongest intonational correlates of conduciveness are a final, post-nuclear accent fall and a large degree of fall. An accent pattern (A) where the nuclear tone is followed by a "relatively steep drop" is likely to be conducive, particularly when two A accents in a row are used with the second lower than the first; Bolinger assigns an approximate meaning of "assertiveness" to this accent. An accent pattern (B) with a high nuclear tone followed by a continued rise is likely to not be conducive and is assigned the generalized meaning "incompleteness" or "connectedness". Finally, Bolinger discusses a third accent pattern (C), also with a final rise, but with a low nuclear tone; suggesting "anti-assertiveness" as a meaning for this pattern. The C accent is supposed to be somewhat conducive when combined with a preceding A, at least more so than an A-B sequence.

Hudson (1975) provides a formulation of the sincerity conditions that must hold for a question to be felicitous. He claims that to use an interrogative form the speaker must believe that the addressee knows "at least as well as he himself does" whether the proposition underlying the question is true. Hudson uses conducive to describe questions uttered when both the above condition holds and the speaker already believes she knows whether the proposition is true and anticipates that the addressee either agrees or disagrees with her. This use of conducive appears to be quite similar to Bolinger's use of the term. Hudson's claims about the intonation of questions also does not conflict with Bolinger's description. Hudson suggests that rising intonation shows that "the speaker DEFERS to the hearer with respect to the truth of the proposition." If a speaker already has a belief about whether the proposition is true, then this "deference" is ruled out, which would conflict with the use of a final rise when uttering a question expressing this proposition, according to Hudson. This then would explain the use of final falls with "exclamatory" interrogatives (Isn't that a pretty dress?); the speaker utters such a question already knowing whether the proposition is true.
An important aspect of Hudson's analysis is that the sincerity conditions on the use of yes-no questions do not require either that the speaker not know whether the proposition is true (as discussed above) or that the speaker desires the addressee to tell her whether the proposition is true (although in asking a genuine question both of these would hold.) Hudson instead derives these factors from the more general principles of social interaction, à la Grice's Cooperative Principle. In cases where the speaker already has a belief about whether the proposition is true and expects that the hearer does also, the speaker may use an interrogative form in order to get the hearer to consider the proposition--and whether it is true--because the speaker believes that such considerations are relevant to the current discourse context (for any number of possible reasons.)

Kenworthy's corpus study of question intonation is also based on claims about the sincerity conditions that must hold for use of questions. Her corpus was comprised of 200 questions collected in interviews, 25 of which were yes-no questions. Kenworthy claims that three sincerity conditions must hold for a speaker to ask a yes-no question felicitously. First, the speaker does not know whether the particular proposition denoted by the question is true or false; second, the speaker desires to know whether the proposition is true; and third, the speaker believes that the addressee does know whether the proposition is true. Intonation then correlates with whether the sincerity conditions hold. When all three hold, the yes-no question will end in a rising intonation (which can be realized in Scottish English as a high rise or a non-low fall); when the first condition does not hold, that is the speaker does know whether the proposition is true, then the question will end in a falling intonation, (a low-fall in Scottish English). In other words, a falling intonation will signify that the speaker has expectations about what the answer to the question will be. Kenworthy does not claim that the intonation will categorically vary with whether the sincerity condition holds. She merely suggests that "We can thus perhaps set up a correlation between a low falling intonation pattern and some modification or non-application of Sincerity Condition (I)." (italics ours)

A rough summary of the claims of these analyses is that falling intonation correlates with the speaker uttering a yes-no question under the condition that she already believes she knows whether the question's underlying proposition holds. Rising intonation then correlates with a lack of speaker knowledge about the truth of the proposition the question expresses.

2.3 Frequency of rise and fall intonation in questions
Fries (1964) argues against the existence of any question intonation per se in American English, pointing to a study of 2561 yes-no questions where 1580 had a falling intonation. He believes that this study contradicts previous suggestions that falling intonations on yes-no questions imply any special meaning apart from the question itself. Fries's corpus, however, consisted entirely of quiz shows, where panelists attempted to guess the occupations of contestants through the use of yes-no questions. Although Fries claims that these forms occurred within "the actual live function of language actively fulfilling its communicative function," the discourse setting can hardly be seen as a natural one. The purpose of asking the questions is specifically to confirm suspicions that the questioner already holds; indeed, the most strategic questions would be those where the questioner was already fairly confident of the answer. Kenworthy's point concerning the speaker's existing belief about the truth of a proposition would seem to apply in such cases.

Lee (1980) examined a similar corpus of radio quiz shows, focusing on the effects of frequency of occurrence of yes-no questions in stretches of speech. Lee
found that yes-no questions that occurred very close together were less likely to end on a fall (50 out of 199) than were yes-no questions that were widely separated by intervening material (20 out of 45). To begin with, we must note that these differences are not statistically significant ($p=0.2$). To account for his still fairly high percentage of final falls in both categories, however, we should also point out (1) the questions in his first category were drawn from the same sort of "twenty-questions" genre of quiz show as in Fries's study, and (2) in his second category, an unspecified number of those questions were actually asked by the quiz-show host -- that is to say, by someone who violates all three of Kenworthy's sincerity conditions.

2.4 Further considerations in describing and categorizing discourse function

Various problems arise in an attempt to categorize the discourse functions of utterances. Cameron et al. (1989) warn against facile identification of a recurrent syntactic or intonational form with some specific discourse function. In their discussion of this issue, which they label "the form and function problem," they suggest that for any given category of form, the communicative function is both variable and complex. Furthermore, individual utterances in discourse often perform more than one function simultaneously. Classifying these utterances according to discourse function is thus a problematic -- and, they argue, necessarily somewhat arbitrary -- task.

This variability comes in many forms. First, judgments about discourse function must necessarily be somewhat subjective. What sounds like a suggestion to one person may sound like a demand to another, and a request to a third -- even if the context of the utterance is held constant. In addition, a complete formulation or listing of categories is impossible. As Green (1989) points out, the function of even a single syntactic form may be limited only by the speaker's skill in exploiting the medium and influencing the addressee's beliefs about what the form is intended to mean.

Nonetheless, one might attempt to characterize discourse function in terms of speech act categories, e.g. promise, reminder, suggestion, criticism, assertion, directive (command, demand, plea, request), question (sincere, rhetorical, clarification), etc. (Green, 1989). Even if we could construct a fairly thorough list (that is, thorough enough to cover most of the discourse functions used in the situations we are examining in our study), a list is not necessarily the best, or even a very good, way to characterize such unwieldy phenomena. Lists can only classify phenomena in a finite number of categories; the relationships among the categories may be unclear and the boundaries between each category indistinct. More importantly, lists are not an explanatory tool. Without an accurate taxonomy of discourse function or meaning, our attempt to characterize the contribution or effect of intonational contours on speaker meaning would be difficult.

There are several possible approaches to a taxonomy of discourse meaning, many of which have been pursued in studies of intonational meaning. These include characterization by speaker attitude (politeness, judiciousness) or emotion (anger or joy); propositional attitude (belief, uncertainty); speech act (statement, request); and information status of propositions and referents. [See P & H (1990) for references.] P & H themselves in their compositional approach to intonational melodies utilize the concepts of mutual belief and intention. According to them, the speaker uses intonational tune to convey "how S[peaker] intends that H[earer] interpret an intonational phrase with respect to (1) what H already believes to be
mutually believed and (2) what S intends to make mutually believed as a result of subsequent utterances."

The use of a mutual belief framework or sincerity conditions may be a theoretically more appealing way to approach discourse meaning; for example, it avoids the problems of lists. It is not clear, however, that either of these approaches are entirely applicable for our purposes. Our data are drawn from informal phone conversations between participants who do not know each other. Most of the talk is informative. The speakers are exchanging information not with the intention of getting each other to do anything with or act upon the information, but getting their conversational partner to accept the information as part of the mutual beliefs and the conversational record. Although one would still expect these conversations to be in accordance with a mutual-belief based approach, the discourses may have a very different pattern of information exchange than a dialogue where the participants have a goal of getting the speaker to do some task beyond just maintaining a congenial relationship and continuing the conversation.

3 Present corpus study
3.1 Data
We used, courtesy of the Linguistic Data Consortium, the SWITCHBOARD corpus, a body of recordings of telephone conversations between strangers given a preassigned topic to discuss. The corpus was collected by Texas Instruments to be used in the development of speech recognition systems. This corpus has been transcribed orthographically. The transcriptions have been coded for a variety of linguistic information -- syntactically, phonetically, and by discourse function.

3.2 Coding for syntactic form
In order to search for yes-no syntactic questions, we used a version of SWITCHBOARD coded under a modified version of the Dialog Act Markup in Several Layers (DAMSL) coding scheme (Jurafsky et al., 1997). DAMSL characterizes discourse function within the frame of Speech Act Theory (Core and Allen, 1997; Jurafsky et. al. 1997; Allen and Core, 1997). This scheme was developed as way to annotate dialogs for discourse functions systematically in order to use the annotated texts both in the training of computer-automated systems for dialog analysis and speech recognition and in linguistic research. The scheme allows utterances to be marked with multiple communicative functions. The broad categories of functions that DAMSL uses include Forward Communicative Functions (yes-no questions, statements of opinion, thanking, apology), Backward Communicative Functions (acknowledgment of answer, appreciation, summary-reformulation), and Utterance Features (whether the utterance is about the communicative process itself or the task at hand.) The type of structure annotated with the DAMSL scheme is referred to as shallow discourse structure; it is a model of discourse at the level of speech act types and how conversants might expect conversational units to pair together. (For example, a question is responded to with an answer, but a statement of opinion might be responded to with agreement.) This shallow level of discourse structure contrasts with deep discourse structure, which models things like plans and intentions of the participants or of the attentional focus of the discourse.

The discourse coding system presents certain problems for this study. Ideally, we would search the corpus for instances of syntactic yes-no questions, code them for post-nuclear accent intonational contour, and then compare their
discourse contexts. The classification for yes-no questions (qy), however, does not include every instance of a yes-no syntactic form. The categories of tag questions, rhetorical questions, "or"-questions (i.e. Did you go or did you just stay home?), and back-channel questions (i.e. Is that right?), all of which may include yes-no-questions are separately coded. Additionally, not everything coded as qy is in fact a yes-no question. Some forms are declaratives (i.e. no subject-auxiliary inversion), wh-questions, or single words (e.g. Right? Really?). We excluded such deviantly coded utterances. We extended our study to include utterances from the codings backchannel (bh) and rhetorical question (qh) for purposes of comparison. The validity of our findings rests on the assumption that the discourse code qy marks the great majority of all syntactic yes-no questions in the corpus.

3.3 Coding for intonation
We coded final post-nuclear accent intonational contours as either falls or rises. These two categories, however, encompass a variety of different contours in other theories; here we will provide a discussion of the rough equivalents in two theories.

In P & H (1990) final contours are essentially the equivalent of the concatenation of the phrase accent and the boundary tone. There are two choices for both and so four possible combinations, L-L%, H-L%, L-H%, and H-H%. Our categories of fall and rise are almost entirely correlated with the boundary tones, L% and H%. Final intonations that would most likely be categorized as H% by P & H were classified as rises, L% as falls. One potential difference is a pattern of H phrase accent and L boundary tone. When this pattern is combined with a L pitch accent (L*-H-L%), the intonational contour will end in a high, but level plateau. A contour like this was most likely characterized by us as "Level" and excluded from the discourse study (about 1% of all tokens looked at), but potentially may have been characterized as a Rise, particularly because it ends at a pitch higher than the nuclear accent.

A very different approach to the analysis of intonational contours is the "British school," (Armstrong and Ward, 1952; Crystal, 1969; Halliday, 1967). The specific study we refer to here as a comparison is O'Connor and Arnold (1961). There are no boundary or final intonational contours per se in O'Connor and Arnold's intonational system. Instead, final rises and falls are grouped in with the nuclear accent as "nuclear tones." There are six different nuclear tones, High Rise, Low Rise, Fall-Rise and High Fall, Low Fall, and Rise-Fall. These will correspond quite closely to what we categorize as Rise and Fall. O'Connor and Arnold also provide a categorization for intonational tunes as a whole, dividing these into ten Tone Groups. For each Tone Group they discuss the meaning, essentially framed in terms of speaker attitude, when combined with a category of syntactic form. Clearly, any discussion we have of meaning will not encompass as many fine-grained categories as that of O'Connor and Arnold's method of classification.

3.4 Results
We coded 3789 qy-type utterances in Switchboard; 249 were realized as final falls. We then examined the discourse contexts of the yes-no questions with final falls and compared them with the contexts of a comparable number of randomly selected final rise questions.

In addition, we performed two similar, but much smaller studies of backchannel function questions and rhetorical questions. We examined a portion
of the 1,019 utterances labeled "bh" for "backchannel questions" (e.g. *Is that right? Did it really?*) in the discourse-coded Switchboard material, and coded them as either rises or falls. Only complete yes-or-no syntactic questions were coded; other one word or declarative backchannel utterances were disregarded. Of the 100 looked at, 41 were falls and 59 were rises. When, however, we examined the qy-coded questions, we found that a large number of them (e.g. the selfsame *Is that right?*) would more accurately have been coded as backchannels themselves. We found 25 more backchannel falls and 5 more backchannel rises, bringing our total percentages to 51% falls and 49% rises.

From the 565 utterances labeled "qh" for "rhetorical question," we likewise coded only complete yes-or-no syntactic questions (excluding wh-questions, declaratives, choices, etc.). Of the 81 remaining, 38 were falls and 53 were rises. We also found a number of questions among the qy-coded utterances that we felt to be clearly rhetorical: 7 falls and 4 rises. These brought our total percentages to approximately 44% falls and 56% rises. There is not, evidently, a significant difference between backchannel and rhetorical questions with respect to final rise vs. fall; both show a roughly equal distribution between the two.

It is thus necessary to revise our original estimation of the number of final falls among qy utterances, given that not all appear to have correctly coded for discourse function. As explained above, we reclassified a total of 32 qy falls as either rhetorical question or backchannel, removing them from our tally. We further assumed, based on the number of rhetorical and backchannel questions among our random sample of qy rises, that there had been a 3.65% rate of error in coding. From an original count of 1961 total qy rises, we thus estimated a probably more accurate total of 1889. Having already discarded the utterances ending in a level pitch, as well as the specially annotated utterances detailed above, this leaves us with 2106 actual qy utterances, 217 of which are falls. These results are summarized in Table 1:

<table>
<thead>
<tr>
<th>Discourse Coding</th>
<th>Total Number</th>
<th>Rises (%)</th>
<th>Falls (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>qy (yes-no question)</td>
<td>2106</td>
<td>89.7</td>
<td>10.3</td>
</tr>
<tr>
<td>qh (rhetorical question)</td>
<td>102</td>
<td>55.9</td>
<td>44.1</td>
</tr>
<tr>
<td>bh (backchannel question)</td>
<td>130</td>
<td>49.2</td>
<td>50.8</td>
</tr>
</tbody>
</table>

### 4 Analysis

#### 4.1 Discourse function of rises and falls

We examined the discourse contexts of the 217 yes-no questions with final falls and compared them with the contexts of a comparable number of randomly selected final rise questions, and coded each one as genuine or non-genuine. To determine a definition for a “genuine question,” we tried to articulate a standard set of conditions to be fulfilled by a question. We began, following Hudson and Kenworthy, by postulating three preliminary sincerity conditions for a genuine question:

- The speaker S must not already know, or believe she knows, the answer to the question she is asking the hearer H.
- S must desire a response from H.
• S must expect that H can provide a response.

4.11 What counts as a non-genuine question
We found several categories of syntactic yes-no question that we felt violated one or another of these conditions:

**Expected response is something other than yes/no.** Certain questions, although phrased as yes-no questions, might better be understood as wh-questions. An example would be "Do you have any particular things that are priorities on your list?" If H responded with a simple "yes," he would be considered to be remiss in his contribution to the conversation. H understands the actual question to be "What is a priority on your list?": S does not actually desire to know whether the proposition expressed by her question is true or false.

**Introduction of the assigned topic.** We limited this category to utterances found in the openings of conversations, where the speakers used question form to state the assigned conversational topic, e.g. "The Middle East crisis, should we be there?". This case is somewhat similar to the above in that S does not expect H to provide a simple yes/no answer, but rather is attempting to initiate a more extended discussion.

**Introduction of new topics within the conversation.** This category is similar to the previous one, except that it can occur anywhere within the conversation. S may wish for H to provide an extended answer, or often S will use her question as a springboard for making further statements of her own.

**Requests couched in question form.** A question such as "Can you tell me anything else about the house?" is not inquiring as to H's capability to do so; rather, it is a request for further information about the house. Similarly, "Can I shut my radio off?" is a request to be allowed to temporarily remove oneself from the conversation. The truth of the proposition is not in question here.

**A particular response is clearly expected.** Although S has no way of knowing for sure what the answer to the question will be, she assumes that H will supply a particular response. Obviously, it is somewhat problematic for us to attempt to read the speakers' minds. Evaluating S's expectations about H's response is complicated by the constraints of relevancy under which speakers and hearers interact. Following Grice's Cooperative Principle entails making only relevant contributions to a conversation. A relevant question is presumably one with a relevant answer. As such, if S asks a question, she assumes that the answer will be relevant. As a hearer then, one assumes that speakers just do not ask questions "out of the blue" about which they have no expectations at all. Furthermore, given that there are only two possible responses to a yes-no question, it is rare that S will be completely unable to hazard a guess as to the probable answer. Given these complicating factors, we have limited this category to situations where it was, in our subjective opinion, reasonably obvious to anyone following the conversation or with basic knowledge of the world what the answer might be.

**Discourse markers.** An example is "Do you know what I mean?". S does not necessarily require any particular response from H, although H may provide one as a form of backchanneling.

**Conversation management.** Questions such as "Do you want to start?" or "Do you think we've talked long enough?" serve as ways to initiate, close, or otherwise promote the smooth running of the conversation. They may function as requests or as suggestions specifically related to the conversation's progress. As
with discourse markers and indirect speech acts such as requests, the truth of the proposition is taken for granted.

Quoted or reported questions, or questions in citation form. S may introduce a question with the clause "The question is," thus providing a citation form of the question that she does not necessarily intend H to answer. Questions may be similarly embedded in statements when S quotes or paraphrases questions that others have posed, or that she herself has uttered on another occasion. In such cases, S has no expectations of H providing a response.

Questions followed immediately by another clause. In such a case, S does not give H the opportunity to respond immediately; S defers her expectation of a response from H until such time as she has completed her own utterance. Lists of questions also fall into this category.

Correction question. These are questions where H has apparently misunderstood the S the first time she asked. The repeated question may have falling intonation, particularly if there is an intonational focus on a word that has been misunderstood.

(1) (Context: They have been discussing H's pets, including her fish.)
S: Well, did you all catch it, or uh,
H: Uh-huh, we have a cat, Stripper.
S: Uh-huh. No, did, no, did you CATCH the fish?
(SWB file: 2877_1244_1259)

The illocutionary force of a correction question is not the same as when the question was originally asked. Here S is not asking H about a proposition so much as informing H about what she has just said.

Jokes. One example of this comes from the end of long conversation about a controversial topic, where S asks, “Well, have we come up with a solution yet, Salina?” As both S and H are well aware that the answer is "no," this is clearly not a genuine question.

Information couched in question form. We often see questions of the form Did you know that X? In many such cases, S's purpose is to inform H of the proposition X, not to determine the extent to which H is up-to-date. S may also offer suggestions to H in question form. In one such case, H has explained to S the difficulties of finding a good indoor pool in the area, and S responds with the question "Have you tried the pool at the Spring Creek Fitness Center?" Although S does not know the answer to this question, her primary object is to inform H about the facility.

Questions to self. In cases where the addressee of the question is actually S herself (e.g. if S is unsure of a title of a movie that she is describing to H, and interjects "Is that X?"), S does not expect a response from H.

Among questions with a final rise, we found that 57.8% were genuine questions that fulfilled all the sincerity conditions. Among questions with a final fall, however, we found only 12.0% were actually genuine questions. These findings are summarized in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>No. of Genuine Questions</th>
<th>Percent Genuine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rise</td>
<td>237</td>
<td>137</td>
<td>57.8</td>
</tr>
<tr>
<td>Fall</td>
<td>217</td>
<td>26</td>
<td>12.0</td>
</tr>
</tbody>
</table>
This small percentage of final falls used as genuine questions that remains does not greatly disturb us. It is not, of course necessary to account for every instance of final falls because intonation is inherently somewhat variable. However, we certainly ought at least to attempt to explain the fact that nearly 42% of our rises were not, according to our definition, genuine questions: should not more of these, then, have been expressed as final falls? We discuss this issue further below.

5 Explaining falls and rises
The myriad of uses of final falls with yes-no questions and the differences between the uses of rises and falls appear to fit well within an explanatory framework of sincerity conditions. The patterns in our data indicate that rises are associated with questions that fulfill all our previously stated sincerity conditions. In contrast, speakers use falling intonation for yes-no questions where sincerity conditions do not hold.

These sincerity conditions can be articulated within a framework of mutual beliefs. If a rising boundary tone signals that the status of the truth or falsity of the proposition corresponding to the question asked is still undecided, as Hobbs suggested, then this could potentially explain the fact that rises are used when S has no strong expectations about what the answer to the question might be. Falling intonation would then fail to signal this lack of expectation. Of course, even within this framework, certain uses of questions might not be well accounted for, particularly those used as suggestions, requests, and conversation management techniques. Utterances with these illocutionary forces are not used to directly manipulate the common ground of mutual beliefs, but rather to affect or coordinate with H's actions (which must take place via a change in mutual beliefs (Clark and Carlson, 1982; Clark, 1996) or to inform H of S's intentions.

A discourse structure framework as suggested by Hobbs and P & H would appear to work less well, particularly when tested on actual data. If "status as an open question" is construed more narrowly within a theory of discourse segmentation, it should mean that a rising intonation should signify that a subsequent utterance is somehow related to the question within a larger discourse segment. This claim is in direct conflict with the frequent use of questions with falling intonation to begin a new topic for discussion. Here, clearly, the question is to be grouped with subsequent utterances -- something with which falling intonation should have no particular association, according to Hobbs and P & H.

5.1 Explaining final falls in backchannels and rhetorical questions
How can we account for the enormously higher rate of final falls in the syntactically similar categories of backchannel questions and rhetorical questions? Backchannel questions violate the first two of our sincerity conditions; when S asks, "Isn't that something?" she fully believes that it is, in fact, something, and does not require H to answer the question for her. Rhetorical questions similarly fail to fulfill these conditions. The answers to rhetorical questions are, in the speaker's mind, either perfectly obvious or perfectly obviously unknowable -- in neither case would S require or expect an answer.

5.2 Rises with non-genuine questions
One explanation for our seeming overfrequency of rises in non-genuine questions may lie in the politeness strategies used by the speakers. If a rising intonation correlates with S uttering a question that denotes a proposition that S
believes is not mutually believed, then sincere questions are more likely to be uttered with a rising intonation. If this is in fact the correlation that holds, (lack of mutual belief = rising intonation) then speakers could exploit this correlation as a politeness strategy. Even if S expects she knows whether or not a certain proposition holds and believes that H knows S knows, she could still use a rising intonation simply as a way to offer the addressee more options for interpretation. Offering options, even sham options, is one strategy of maintaining interpersonal relations when performing a face-threatening act (Lakoff, 1973; Brown and Levinson, 1987). For example, using a question like Did you know X? to inform H of X may be more polite than informing H of X directly, because S can use the former whether or not H knows X already. Informing H of something H already knows is face-threatening in that it implies that S thinks H needs to be reminded of the proposition. By using a rising intonation, S offers H an opportunity to "weigh in" on the truth or falsity of the proposition, even though this may in fact be a fake option if the truth value of the proposition (or S's belief about the proposition's truth value) is already mutually believed. So, rises are a strategy for making a question less threatening and more polite by pretending that H's viewpoint still matters. Because falls are both marked as intonation for yes-no questions and because they are ambiguous about whether the proposition is already mutually believed, they do not lend this same optionality to the interpretation of the utterance. However, because the use of a question is itself a way to offer options, a question, even with a falling intonation, can still be a more polite way of discussing or introducing a proposition about which S already has beliefs (Creswell, 1993).

Rising intonation on a yes-no question may not always be polite, however. If the status of a proposition's truth value should not be questioned and S nonetheless uses a rising intonation, it may be interpreted by H as an indication of S's doubt. For example, using a low nuclear accent and a steep rise on a normal backchannel question makes it seem as if rather than affirming what H has just said, S believes otherwise, as in the following example:

(2) A: It couldn't have been Johnny vandalizing the neighborhood. He's been in school every day this week.
B: Has he?

6 Conclusions
Our data clearly indicate that final rises are the overwhelmingly favored intonational variant for yes-no questions, despite some earlier claims to the contrary. We have also tried to show, however, that there is a certain regularity to speakers' use of final falls. Discourse function does correlate -- if not categorically, then at least somewhat predictably -- with final contour.

We hope that even though our investigation makes claims only about a single syntactic form, our claims about the discourse functions of final rises and falls could be evaluated, using the same methods, with respect to other syntactic forms. We do not mean to imply by this that final rises and falls will carry the same conventionalized meaning across all other syntactic forms. The caveat of Cameron et al. discussed above still applies: rigidly identifying a particular discourse function with a specific intonational pattern glosses over the complexities and ambiguities inherent in actual human interaction. It also ignores the enormous variability of the intonation of individual speakers.

In addition, our methods could perhaps be applied cross-dialectically and cross-linguistically. Although studies based on introspection or much smaller
corpora have made valuable theoretical contributions to the understanding of these phenomena, we feel strongly that it is through the systematic investigation of large bodies of data drawn from natural conversations that we can best understand the complex patterns that actually occur.

Notes

1 We thank Kazuaki Maeda for extensive computer assistance throughout this project.
2 Kenworthy refers to questions as denoting the proposition corresponding to the true answer; we take yes-no question as denoting the set containing the corresponding true and false propositions (Karttunen, 1977); but we will often use Kenworthy's phrasing as a convenient shorthand.
3 We found only very few instances of such an intonational contour in our corpus, and chose to exclude them from our study.
4 Some utterances require further annotation, for example, if they seem to end in a level pitch. In addition, some utterances receive special coding for the following reasons:
   • Incorrectly coded as yes-no questions. Some utterances are in fact wh-questions or declaratives. We also eliminated all questions where the syntactic status was ambiguous due to a contracted or dropped auxiliary.
   • Incomplete utterances. We eliminated unfinished utterances from our study. These include instances where speakers break off or are interrupted midway through an utterance.
   • Utterances cut off by the software program. Because the alignment of the transcription and the sound files is not exact, some utterances are clipped prematurely.
   • Choices. Two yes-no questions connected by "or" are somewhat problematic for the purposes of this study. Often these have the intonation of coordinated declarative clauses; the first conjunct ends in a rise and the second in a fall. The final fall in the second conjunct may be more dependent on its syntactic structure rather than its discourse context. In addition, although syntactically these are yes-no questions, their discourse function is inherently tied to the fact that they present a choice of response to the addressee other than "yes" or "no". In order to avoid these issues, we eliminated these forms from our investigation.
   • Rises ending with or. Often speakers follow a yes-no question with "or" even though the second conjunct is not presented. These questions have the final rise of normal questions followed by the "or" at a much lower pitch. We excluded these forms from the investigation of discourse context also.

4 That is, "some relevant completion of the incomplete thought represented by her utterance is relevant." (Sperber and Wilson, 1987.)
5 They will, nonetheless, have an effect on what is mutually believed. For instance, if S asks "Can I run a minute and turn my radio off?" the mutual belief that S is exiting the conversation for a moment will be added to the common ground.
6 A steep fall, interestingly, will produce a similar effect, with the difference that it suggests that B has knowledge of Johnny's actual absence from school.

References


Core, Mark and Allen, James. 1997. Coding dialogs with the DAMSL annotation scheme. In AAAI Fall symposium on communicative action in humans and machines.


Kenworthy, Joanne. 1978. The intonation of questions in one variety of Scottish English. Lingua 44, 267-82.

Lakoff, Robin. 1973. The logic of politeness; or minding your P's and Q's. Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, Chicago, IL.


