1-1-1997

The Pragmatics of WH-Question Intonation in English

Christine Bartels
bartels@darkwing.uoregon.edu

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/pwpl/vol4/iss2/1
For more information, please contact libraryrepository@pobox.upenn.edu.
The Pragmatics of WH-Question Intonation in English
The Pragmatics of Wh-Question Intonation in English

Christine Bartels

1. Tonal Patterns in Questions

Any description and compositional phonological analysis of intonation contours must make certain assumptions as to what constitutes linguistically relevant contrasts in this domain—that is, a semantics. As Pierrehumbert & Hirschberg (1990) observe, “any theory of transcription must be viewed as provisional unless it is supported by considerations both of sound structure and of interpretation.”

However, the task of mapping sound into meaning is made difficult by the fact that a given intonation contour—a sequence of tones, or tune—can have very different connotations in different contexts. It is often suggested, therefore, that the contribution of tune, i.e., choice of tones, to utterance meaning in English is dependent on the discourse situation at utterance time and cannot be analytically reduced to constant semantico-pragmatic correlates of the relevant pitch movements; in short, that English pitch contours are polysemous.

By contrast, this paper argues, with Gussenhoven (1984) and Pierrehumbert & Hirschberg (1990), that tunes can be decomposed into tonal morphemes with invariant, abstract meanings/functions; all specific connotations are claimed to be pragmatic inferences derived from these abstract meanings in conjunction with contextual factors.

However, aside from different assumptions about the nature of the phonological and morphological building blocks involved, the model presented here differs from these earlier proposals for compositional models of tonal meaning in that the meanings of the tonal morphemes are drawn from a different domain. Both Gussenhoven and Pierrehumbert & Hirschberg find the denotata of the tonal

---

The content of this paper has been greatly influenced by discussions with Arthur Merin, whose Decision-Theoretic Semantics underlies the theory of tonal meaning presented here (see also Merin & Bartels 1997).
morphemes they identify—kinetic tones for Gussenhoven, level tones for Pierrehumbert & Hirschberg—in epistemic and discourse relations. They see the role of pitch accents as instructing the addressee on the joint epistemic status of the accented item, proposing, in Pierrehumbert & Hirschberg's words, that "speakers use tune to specify a particular relationship between the 'propositional content' realized in the intonational phrase over which the tune is employed and the mutual beliefs of participants in the current discourse." Thus Gussenhoven sees choice of nuclear accent tone (that is, the pitch movement associated with the sentence's main stress) as dependent on the 'manipulation' of the participants' shared cognitive 'background' that a speaker intends to effect: a fall means that the accented material is to be added to the background, a fall-rise that the material is being selected from the background, and a rise that its status is being tested. In Pierrehumbert and Hirschberg's level tone model, a high accent tone $[H^*]$ marks an item as 'new' to the common context, a low one $[L^*]$ as given. Phrasal tones instruct the hearer on interphrasal discourse dependencies: high tones $[H-, H%]$ indicate a connection to the subsequent prosodic phrase, whereas low tones $[L-, L%]$ indicate lack of dependency.

These semantics permit a plausible interpretation of tonal patterns in many contexts. However, in some cases they mispredict. For instance, while Pierrehumbert & Hirschberg would correctly ascribe 'given-ness' to Freudian (account) and 'new-ness' to cognitive (account) in (1),

(1)

It's not a FREUdian account - it's a COGnitive one.

$L^*$ $L-H\%$ $H^*$ $L-L\%$

corresponding to the contrast between $L^*$ and $H^*$, their account also predicts that in the alternative question (AQ) in (2) French is 'given' or 'not-new' and Flemish is 'new' to the discourse context.

(2)

Did the suspect speak FRENCH or FLEmish?

$L^*$ $H^*$ $L-L\%$

But this is not borne out by intuitions: the two disjuncts have the same status. Note also that permuting ‘French’ and ‘Flemish’ in (2) changes nothing intuitable regarding ‘given/newness’ but would force a switch in accent tones.

In other cases the discourse-epistemic semantics fail to make requisite predictions. Here the constraints on, and observed variations of, phrasal intonation in questions are a prime example, though not the only one. (Other unexplained phenomena are the tonal contrasts in adverbials such as always and usually noted by Allerton & Cruttenden (1978), or the association of obligatory genericity with rising intonation on indefinite sentence topics, to name only two.) The term ‘question intonation’ standardly refers to contours characterized by a final rise. Yet AQs such as (2) must obligatorily show a fall on the last disjunct. Yes-no questions (YNQs) such as those in (3a,b) may either rise or fall. (By corpus statistics, two thirds fall. Among rising questions, one can distinguish between high-rises and low-rises, as shown in (3a).)³

(3) I didn’t know John took a job all the way over in Tualatin.

   a. Does he have a CAR now? / Does he have a CAR now?
      H* H-H%
      L* H-H%

   b. Does he have a CAR now?
      H* L-L%

The same goes for wh-questions (WHQs) such as (4a,b), although these tend more strongly toward a falling pattern.

(4) A: I still have that mysterious backache. It simply won’t go away. I even went to see an orthopedic specialist yesterday.

³For simplicity, I'm going to ignore the possibility of fall-rises and other variants here and below.
a. E: And what did HE have to say?/ ... did HE have to say?
   H* H-H% L* H-H%

b. E: And what did HE have to say?
   H* L-L%

Only in their use as ‘echo questions’ (5a,b) do YNQs and WHQs consistently end with a rise:

(5)a. A: Did Amy get the summer job at the embassy?

   E: Did she get the job at the EMbassy? / ... at the EMbassy?
      H* H-H% L* H-H%
      (...Was that your question?)

b. A: Amy started her job at the embassy last week.

   E: She started

      her job WHERE last week? / ... WHERE last week?
      H* H-H% L* H-H%

Oddly, though, so-called ‘reference questions’ (Rando 1980) such as (6), which are superficially similar to echo questions in being ostensibly discourse linked, must always show a final fall.

(6) A: I just talked to him last night.

   E: You talked to WHO last night?
      H* L-L%

These tonal patterns have not yet received a satisfactory explanation.
2. A Decision-Theoretic Semantics for Intonation

2.1. Constructing a Domain of Denotata

The present account, based on a semantico-pragmatic model developed by Arthur Merin of the University of Stuttgart (Merin 1994, 1996; with precursors in Merin 1983, 1985), can make sense of the above observations in a straightforward way. It proposes fundamental sociopolitical relations governing the establishment, maintenance and negotiation of cooperation among potentially autonomous actors as the natural target domain for a semantics of intonation; within this model, negotiations regarding the discourse-epistemic status of propositions put forward by the participants might be seen as a special case.

Aside from covering a greater range of data, this approach also has the virtue of greater phylogenetic plausibility. Ohala (1983) suggests that high or rising tone is associated across species with ostensible submissiveness, i.e., low relative social power, and of low or falling tone with impositiveness, i.e., high relative social power—features usefully conveyed in a competitive Darwinian world. It is not obvious how these vocal gestures should have led to intonation as a gestural system involving discourse-epistemic denotata. By contrast, it is an uncontroversial assumption that humanity had to negotiate as it came into existence, and individual people have to start negotiating all too soon after coming into the world.

I can only give a brief sketch of Merin’s formal decision-theoretic model here. Cooperating actors in a minimal, i.e., biperson social situation—call them [Ego] and [Alter]—have to establish a ‘common ground’ (CG) of joint deontic-boulonnaic or epistemic commitments. (The default identification in the examples here is for Ego with the present speaker, and for Alter with the addressee.) To the extent of being autonomous, Ego and Alter are in need of persuasion. The paradigmatic question is always:

“Why (<expletive>) should I (do/believe that)’’

The need for persuasion implies that Ego’s and Alter’s preferences are formally inverse regarding points at issue. Indifference or consonance means, by definition, that there is no issue. Negotiations are in essence bargaining games (Nash 1953), i.e., social situations in which interests are neither wholly opposed nor wholly consonant,
promoting strategies of 'competitive cooperation'.

Negotiations on what becomes CG proceed by Elementary Social Acts (ESAs) consisting of Claims, Concessions, Denials, or Retractions (of a Claim). ESAs are transitions to (and from) negotiation states characterized by vectors of binary decision-theoretic parameters (Merin 1994). These parameters allocate ostensible agentrole [S]; preference [P] w.r.t. propositions under negotiation; dominance [D] w.r.t. balance of incentives/warrant; and initiator role [I] among Ego and Alter. For example, Ego's Claim for a proposition \( \Theta \) to become a mutually binding constraint—Ego's least marked act type—is formally characterized as \(<E,\Theta,E,E,E>\): speaker role, preference, dominance, and initiative (in that order) are all assigned to Ego. By contrast, Ego's Concession of \( \Theta \) is explicated as \(<E,\Theta,A,A,A>\): preference, dominance, and initiative are all Alter's. Similarly, for Ego's Denial of \( \Theta \) the settings are \(<E,\Theta,A,E,A>\); i.e., preference for \( \Theta \) and initiative are Alter's, but Ego is dominant. Whereas for Ego's Retraction of \( \Theta \) the settings are \(<E,\Theta,E,A,E>\); preference for \( \Theta \) and initiative are Ego's, but Alter is dominant. Other combinations of parameter settings for a given speaker and proposition are ruled out by a constraint setting \( P = 1 \) underlying this act typology: it is assumed that *hominis oeconomici* do not undertake counterpreferential initiatives. However, other act typologies, leading to explications of, e.g., Entreaty and Offer, can be obtained by relaxing or even inverting the constraint.

Once a claim is (vocally or tacitly) conceded, the propositional object \( \Theta \) becomes CG, i.e., a mutually binding constraint on future action—in particular, future discourse moves. If a claim is denied by Alter and retracted by Ego, its contradictory—not-\( \Theta \)—becomes CG. This concept of 'common ground' can be seen to subsume familiar notions in traditional, epistemically oriented models of discourse: e.g., Stalnaker's (1978) 'common context set', Gussenhoven's (1984) 'shared background', Clark & Marshall's (1981) 'mutual beliefs', and perhaps most directly, Hamblin's (1971) 'joint public commitment slate', or the combination of background assumptions and the evolving 'conversational record'.

In bargaining situations proper, extensions of the simple D and P parameters are determined in terms of cardinal (dis-)utilities. This, Merin argues, seems plausible for typical imperatives, e.g.:

"Give me your wallet! (Else be a casualty)"
but spurious for indicatives. Beliefs are, if anything, dispositions to action, not actions. In the present model, therefore, a measure of epistemic state change potential (Merin 1996), namely stochastic evidential relevance, instantiates the role of a utility. The expressed proposition is a more or less useful argument for or against an ulterior constraint on belief or action, i.e., an ulterior proposition at issue.

2.2. Intonational Morphemes

Intonational morphemes are assumed here to be kinetic tones: Falls, Rises, and some of their compounds, such as the so-called Fall-Rise. (For lack of time, compound tones will not be discussed here.) They thus involve combinations each of at least two of the phonemic tone units postulated in Pierrehumbert's now widely used (1980) model of English intonation: the accent tone associated with a stressed syllable and at least one subsequent phrasal tone. These kinetic morphemes denote, in the first and core instance, (re-)allocations of the [D]-parameter value—i.e., of the power of choice—regarding the instantiation of variables under negotiation. A Rise (L * H-/%) alienates choice to Alter, a Fall (H* L-/%) appropriates it. Variably defeasible default associations introduce preference ('scale') and initiative ('anaphoricity')-related aspects.

Unless there are more highly ranked variables under negotiation, in a typical discourse context tonally cued (re-)allocation of choice is likely to be interpreted with respect to propositional content: either with respect to propositions expressed by a whole sentence or clause or with respect to focus-identified subsentential items (usually syntactic constituents) that co-determine propositions. This is even more so the case for utterances presented in isolation—quasi-decontextualized 'citation forms' as approximated by some of the examples given here.

3. General Application to Questions

The decision-theoretic model offers the following account of the question data presented above for which the discourse-epistemic one fails to predict.

Variability of final pitch movement in YNQs and WHQs reflects the fact that Ego may foreground either of two choice-related aspects inherent to questions:
• in asking, Ego is alienating choice among alternatives (sets of possible worlds) to Alter, i.e., making a Concession;
• in demanding an answer from Alter, Ego is forcing Alter to commit himself to one mutually binding alternative, thus banning others from inclusion in the CG; Ego is thereby making a Claim, an attempt to restrict Alter’s future situational options.

Bolinger (1978a) already noted that questions oscillate between the force of requests and that of orders. As Merin (p.c.) puts it, the glass of situational options available to Alter is presented as half-full or half-empty by Ego. The claim, then, is that the dominant illocutionary force determines intonation: ostensibly concessive allocation of choice to Alter and ostensibly demand for commitment are conveyed through final rise (e.g., (3a,4a)) vs. final fall (e.g., (3b,4b)) respectively.

In AQs such as (2), rises on nonfinal disjuncts ostensibly concede to Alter the choice of whether the respective proposition is to become a mutually binding constraint, i.e., part of the CG. However, the last disjunct (Flemish) represents a proposition which must be added to the CG if none of the preceding ones have been. It is a demand (Claim) for the addressee to commit himself. Without this fall, the question as a whole would not convey that one and only one alternative must be chosen, and by inference, that the options are mutually exclusive and the list exhaustive; hence the obligatoriness of the final pitch movement.

In YNQs such as (3a,b), rising intonation conveys that the surface proposition is being posed for Alter to endorse or not; though one might say that logically, (at least) two alternatives are being offered, the covert one is not made salient. Falling intonation, by contrast, makes a YNQ akin to an alternative question in saliently evoking two mutually exclusive alternatives—the surface proposition and, most commonly, its negation. In other words, by ostensibly conveying a demand for Alter to restrict his options, i.e., to commit himself to the elimination of possible worlds that until then have still been “live options” from the point of view of the conversational record, the more peremptory falling intonation in itself serves to make this alternative set salient.

Note that falling YNQs are more suitably reported embedded under whether than are rising YNQs (Bolinger 1978b; Bartels 1997).
4. Wh-Questions

I mentioned that WHQs tend relatively more strongly than YNQs toward showing a final fall. Two factors—not unrelated—can be invoked to explain this tendency. One is the existential presuppositions inherent to WHQs according to most semantic analysts; e.g., in the case of (4a,b), the presupposition ‘He had something—one particular thing—to say’. To the extent that this presupposition is not yet perceived by the questioner to be part of CG, he can felicitously (nonvacuously) demand commitment to it from Alter by way of asking the WHQ. The presupposition represents an impositive lower bound on the epistemic commitment accompanying any direct, i.e., cooperative, answer and thus motivates the impositive intonation, even though choice of instantiation for the variable represented by the wh-expression is allocated to Alter (Bolinger 1982; Merin 1983).

However, this reasoning is not entirely convincing, in that most of the time, a speaker asking a WHQ does indeed assume that the relevant presupposition is part of the CG; if he did not, asking a YNO (e.g., in (4), “And, did he have anything to say?”) might often be a more appropriate strategy.

What holds more generally is that a WHQ is inherently impositive in that it always forces Alter to pick one and only one alternative from an explicitly or at least contextually restricted set of equally salient possibilities. In other words, even the most request-like WHQ demands of Alter that he renounce saliently evoked “live options” from the context. A speaker Ego still has a choice whether to foreground this Claim-like aspect of his question or whether to foreground instead the fact that he is, after all, offering Alter a choice among options: the glass can still be presented as half-empty or half-full. But by tendency, compared with YNQs, which by nature of their surface structure single out one alternative rather than evoking a set, WHQs will lean toward the intonation that signals restriction of Alter’s situational options.

4.1. Reference Questions

Evidence for this account of tonal meaning in WHQs are ‘reference questions’ such as (6) above (and (8) below), which are obligatorily falling. These have the express purpose, one might say, of producing previously evaded commitment from Alter to a specific extension of a designating expression, i.e., the wh-expression. They are
thus inherently, foremost Claims, imposing on Alter the demand to eliminate alternative options that he has intentionally or unintentionally preserved for himself through the vagueness—as perceived by Ego—of his original utterance.

While one could let the argument rest at that, it is worth noting that there is also the possibility of a more specifically linguistic line of reasoning here. Note that in reference questions, unlike other falling WHQs, the wh-expression must receive the nuclear accent; it is narrowly focused here. It has been argued by Berman (1990), Ginzberg (1992) and others that narrowly focused wh-expressions are always non-quantificational and specific in nature; one piece of evidence being that such wh-expressions scope over all other sentence constituents. Wide scope is generally taken as a criterion for specificity in NPs (see, e.g., Fodor and Sag 1982). Thus in (7), wide scope of an advisor over every student ('There is an x such that every y talked to x') is considered an indication of specificity in the indefinite.

(7) Every student talked to an advisor.

By the same token, the obligatory wide scope of the wh-phrase in the reference question in (8) points to this wh-phrase having specific reading ('There is an x such that most of you talked to every one of x's students today; who is x?

(8) A: Most of us talked to every one of his students today.

E: Most of you talked to every one of WHOSE students today?  

Enç (1991) defines specificity independent of scope possibilities as the property of being D-linked in Pesetsky's sense: a specific expression must stand in some sort of relationship to previously introduced referents, e.g., an inclusion relationship. This notion is compatible with Erteschik-Shir's (1986) view of wh-expressions in (non-echo) questions being "restrictively dominant," i.e., roughly, contrastively focused—asking for an entity to be picked from a salient set—when they bear the sole accent. By this criterion as well, the wh-expressions in (6) and (8) can be said to be specific.

What one might want to conclude, then, is that a reference
question such as (6) or (8) has a surface structure involving a specific placeholder ‘X’ rather than a variable ‘x’: ‘You talked to person X last night’, ‘Most of you talked to every one of X’s students today’, etc. In other words, a reference question represents a closed, assertable proposition that can be added as such to the CG. While Ego would have to grant, of course, given Alter’s previous utterance, that the corresponding existential presupposition, i.e., the open proposition involving a variable ‘x’, has already been committed to, he now ostensibly conveys a demand for Alter to newly commit himself to the epistemically stronger closed proposition not yet part of the CG (Bartels 1997). In this a reference question is no different from a declarative sentence, analyzed as an epistemic Claim in the present model.

5. Echo Questions

That leaves the case of echo questions, such as (5a,b). Why should they always rise? The pragmatic account laid out above predicts this tonal pattern, as follows: Ego cannot felicitously demand (claim) of Alter commitment to a proposition to which Alter has already made a commitment by his original utterance. Rather, in the case of an echo-YNQ Ego ostensibly offers Alter another choice to accept the posed sentence as an accurate echo of his original utterance or not. Only the posed sentence is made salient; alternative possibilities as to what Alter may have said originally are not. If one wishes to assume an implicit performance report frame at some level of linguistic structure, as given in (5a’), the echo utterance constitutes simply a special case of rising original YNQs.

(5a’) A: Did Amy get the summer job at the embassy?
   A’: Amy got the summer job at the embassy.

   E: [Did you ask] ‘Did she get the job at the EMBassy?’
      H*/L* H-H%

   E’: [Did you say] ‘She got the job at the EMBassy?’
      H*/L* H-H%
As to echo WHQs, the questioning Ego knows—and needs to convey that he knows—that Alter has already committed himself to a proposition containing a definite, referential expression in place of the wh-word. So even though such questions cannot be paraphrased as if-questions with an implicit performance report frame quoting the echo sentence with the wh-word in situ (cf. (5b')), the same general reasoning as with echo-YNQs applies.

(5b') A: Amy started her job at the embassy last week.
E#: Did you say 'she started her job WHERE last week'?
E': Where did you say she started her job last week?

Note that the wh-expression in echo questions must be assigned some of the same semantic properties as in reference questions: it is narrowly focused, nonquantificational according to Ginzburg and Berman, D-linked by Enç's definition, and takes widest scope, including scope over the implicit performative report frame, as illustrated by the E' paraphrase in (5b'). But even if one grants the wh-expression the relatively strong status of a specific placeholder here, the resulting proposition ('Amy started her job at place X last week') is still not stronger informationally than Alter's original statement. Any demand for commitment to the wh-based proposition would be vacuous, and thus, following Stalnaker (1978) and others, an infelicitous discourse move.

6. Extended Functions of Wh-Question

Intonation in Context

It was stated earlier that in the unmarked discourse context, tonally cued (re-)allocation of choice is likely to be (intended to be) interpreted with respect to propositional content; all of the examples presented so far were analyzed accordingly. However, in original, non-reference questions—questions that could be cast quite appropriately as either offers of choice or instructions for commitment given the current CG—it appears possible for discourse participants to forego this default interpretation in favor of reference to another salient, negotiable variable: the issue who of the participants is to take or maintain local or global control of discourse topic and development—control of the way in which the conversational record is
to be shaped.

Unfortunately there is no room here for presentation of extensive discourse fragments, and no statistically significant corpus analysis has yet been carried out. But consider your intuitive perception of the effect of final intonation on the addressee’s likelihood to assume initiative in discourse development in the following examples (loosely modeled after fragments cited in Selting 1991). In (9)-(11), the crucial WHQ shows a final fall:

(9) A: I always hate it when a class has only women in it.
E: Yeah, me too.
A: Just in general... But this term it’s really extreme.

E: *What do you study?
H*L-L%
A: Ah...sociology and music.
E: Hmm. I’m in speech pathology. Same thing there...
[dialogue continues with short alternating moves]

(10) A: [explaining about her difficulties with an ‘incomplete’]
and... and then I tried to explain this, repeatedly ... why I couldn’t make that time...

E: *Hmm... so who taught that course?
H* L-L%
A: George Bell is his name.
E: Oh, I know him. He came to our departmental potluck one time. Seemed a bit of an odd bird.
A: That’s what I thought.
(11) A: What are those scars you got there?
E: Oh those... those just look so bad because whoever did the stitches did a lousy job.

A: But what happened?
H* L-L%

E: Ah, I had a kind of accident in woodworking class in school...
[goes on the describe the event]

Now compare these with (12)-(14), in which the crucial WHQ shows a final rise.

(12) A: [describing why she quit her waitress job]
...and, I was exhausted, ... couldn't have done that much longer you see...

E: How long do they stay open at night?
H* H-H%

A: Oh, until one o'clock at least... Anyway, with school starting up again and Mom still needing me to help out on weekends...
[goes on in her description]

(13) A: I tell you, I was so upset with that woman...

E: Why?
L* H-H%

A: Because ... because of her political sheNANigans... like during the student senate election campaign...
[goes on to explain]
Selting (1991) states that falling intonation is common in WHQs conveying need for additional information on a given discourse topic or confirmation of an inference, whereas rising questions tend to move the discourse forward. I'd like to submit that the basic effect is better characterized interactively: utterance-final intonation in these questions—especially the more marked rising intonation where it occurs—signals on whose terms the cooperative development of the conversational record is to take place in subsequent moves. Falling WHQs as in (9)-(11) can indicate the questioner’s intent to assume control over the discourse; they are impositive—Claim-like in Merin’s sense. A cooperative addressee will oblige by trying to alleviate the need for information conveyed by the question—be it with a brief, single-clause response as in (9) or (10) or a more extensive description as in (11)—but he is no more likely than the questioner himself to then move the discourse forward to the next topic. Whereas rising questions like (12)-(14) are concessive in ostensibly leaving control of the discourse with the addressee: they, too, express a desire for a particular bit of information that a cooperative addressee will seek to satisfy, but they do not impose a new topic on the discourse and are often taken by the addressee as permission or invitation to elaborate further on the topic at hand, as illustrated in the examples.

7. Conclusion

We must conclude that in richer situational contexts, in which several variables are simultaneously under negotiation, tonal contours of WH-questions do not always pattern with whether or not the ad-
dressee is already committed to a relevant proposition. It is precisely such instances of lack of consistent ties to participant beliefs vis-à-vis propositional (or presuppositional) content that have led some discourse analysts to claim that intonation contours cannot be assigned invariant meanings or functions. However, given the sociopolitical domain of intonational meaning proposed here, the observed range of connotations in context can still plausibly be said to arise as pragmatic inferences from the basic interactive meanings of the respective tonal morphemes.

References


bartels@darkwing.uoregon.edu