

Chapter 15

The Pierrehumbert-Hirschberg Theory of Intonational Meaning Made Simple: Comments on Pierrehumbert and Hirschberg

Jerry R. Hobbs

Everything I know about intonation I learned from reading Pierrehumbert and Hirschberg's very fine paper. Therefore, for me to write a critique of their paper would be to climb far out on a very long and fragile limb. Nevertheless, I will do just that. The general outline of what I propose here is the same as Pierrehumbert and Hirschberg's—a compositional semantics for tunes in terms of what is mutually believed. But it seemed to me while reading their work that, by tidying up one's theoretical assumptions a bit and by modifying the compositional semantics of complex tones, one could give a somewhat more elegant version of their characterization of intonational meaning.

We will need three background principles. The first is this: Every morpheme in an utterance conveys a proposition. In the sentence

George likes pie.

the information conveyed can be represented

$$(\exists e, x, y) \text{Present}(e) \wedge \text{like}'(e, x, y) \wedge \text{George}(x) \wedge \text{pie}(y).$$

The word "George" conveys $\text{George}(x)$. The word "likes" conveys $\text{like}'(e, x, y) \wedge \text{Present}(e)$. The word "pie" conveys $\text{pie}(y)$ for some possible or nonspecific entity y . This principle may seem strange at first blush. What about a word like "the"? But the word "the" conveys a relation between two entities, a linguistic object (namely, the description expressed by the rest of the noun phrase) and an entity in the world, and it says about that entity that it is the most salient mutually known entity of that description. In fact, when we emphasize "the," it is just such a proposition that we are putting forth. Similar stories can be told about other morphemes that are not usually thought of as conveying propositions.

The second background principle concerns what discourse is all about. Imagine the space of all propositions. The speaker and hearer in the dis-

I have profited from discussions with Janet Pierrehumbert and Julia Hirschberg about this work. The research was funded by a gift from the System Development Foundation.

course each have subsets of this space as their own private beliefs. But in addition there is a large overlapping set of mutual beliefs. (There is also a penumbra around mutual beliefs, consisting of those propositions the speaker and hearer both believe but that the speaker doesn't believe the hearer believes, and propositions that the speaker believes both believe but that the hearer really doesn't believe, and so on, but these marginal cases will not concern us in these remarks.) In addition, there are all those propositions that are believed by neither the speaker nor the hearer, many of them because they are false. The prototypical utterance is a bid to turn some of the speaker's private beliefs into mutual beliefs. The utterance is anchored referentially in mutual belief and reaches out into the speaker's private beliefs. An utterance thus contains the given propositions that anchor it referentially and the new propositions that the speaker is informing the hearer of. Much of intonational meaning is concerned with keeping these different partitions of the space of propositions distinguished. For convenience, I will refer to the speaker's private beliefs that are expressed in an utterance as "new," to the relevant mutual beliefs as "given," and, with some violence to the complexity of the matter, to those things the speaker does not believe as "false."

The third background principle concerns discourse structure and will be discussed later.

Now we are in a position to look at a very simple characterization of Pierrehumbert and Hirschberg's theory of intonational meaning. A pitch accent on a morpheme means that the proposition conveyed by the morpheme is important for the correct interpretation of the utterance. When the pitch accent is realized by a H^* tone, the proposition is new. When it is realized by a L^* tone, it is *not* new. Note that I did not say "given"; it may also be not new because it is not believed at all. Thus, L^* is used in Pierrehumbert and Hirschberg's example (3) because the information should already be mutually believed, whereas in (4) it is used because it is false.

A simple tone can be given a prefix, resulting in a complex pitch accent of the form $X + Y^*$ where $X \neq Y$. This conveys a kind of correction or accommodation of what the speaker believes the hearer might incorrectly believe the status of the information to be. Thus, $L + H^*$ says something like, "you might think this information is not new, but it really is new," and $H + L^*$ says, "You might think this information is new, but it really is not new."

The simple tone can also be given a suffix, resulting in the complex pitch accents $L^* + H$ and $H^* + L$. In addition, intermediate phrases and intonational phrases can be given H or L suffixes. A H suffix signals incompleteness or open-endedness. A L suffix *does not* signal incompleteness or open-endedness. Note the scope of the negation here. It is not the case that a L suffix signals completeness; it merely fails to signal incompleteness. "Open-

ended" means something like, "What I've just conveyed by that morpheme or phrase requires further discussion before it is entered into mutual belief, or before its status with respect to mutual belief is agreed upon." I discuss below some of the ways this general characterization is instantiated in specific examples.

This is the whole theory.

Thus, I was completely convinced by Pierrehumbert and Hirschberg's account of the meaning of the H^* pitch accent. It seemed to me that their account of the meaning of the L^* pitch accent was also right, but could be made a bit cleaner. They say it means "salience without predication" and list a number of ways items can fail to predicate. In my account, there are just two ways something can fail to predicate, that is, fail to be proposed as new. It can be proposed as given or as false. The "given" accounts for their example (18):

- (18) Well, I'd like a Pavoni ...
 $L^* \quad L^* \quad L^* L H\%$

The "false" accounts for their example (17):

- (17) I was wrong
 $L^* H \quad L^* H H\%$

Pierrehumbert and Hirschberg note as a special case that cue phrases generally take a L pitch accent. But we need not see this as a special case. Cue phrases, like all morphemes, convey propositions. They convey propositions about the relations among segments of discourse. But if these relations in fact hold, the segments already stand in that relation, without the cue phrases being used. The cue phrases only help to emphasize relations that are already available in the text because of the information the segments convey. A cue phrase *could* convey a proposition that is either given or new, but usually the proposition will already be given implicitly by the content of the utterances. In fact, to use a H^* pitch accent would be to suggest that the hearer couldn't figure out the relation from the content and is hence frequently seen as overbearing. A similar story can be told for greetings and vocatives.

The biggest difference between my account and Pierrehumbert and Hirschberg's is in the treatment of complex pitch accents. They decompose them into a $L + H$ and a $H + L$ pattern, and the placement of the $*$ on either the H or the L . They come up with plausible interpretations for the $L + H$ and $H + L$ patterns but cannot come up with one for the $*$ -placement or for the compositionality of $L + H$ and $H + L$. They say, "... the meaning of each particular pitch accent may be derivable from the meanings of its constituent tones, plus some generalization about the interpretation of the star. However, beyond the observations just made, we are not able to

present such a decomposition as yet" (section 5.5). By contrast, my account is completely compositional. There are two basic pitch accents, L^* and H^* , and these can then have a prefix, a suffix, or no affix at all. The meanings of H^* and L^* are exactly the same in simple pitch accents and complex pitch accents; the meanings of the H and L prefixes are related to the meanings of H^* and L^* ; and the meanings of the H and L pitch accent suffixes are exactly the same as the meanings of the H and L phrasal and boundary tones. First I will examine some of Pierrehumbert and Hirschberg's examples to reinterpret them in light of my account. I will then present a broader comparison of the two accounts.

According to my account, the $H+L^*$ pitch accent should mean something like, "You might think this is new information, but it's actually not new." One way for it to be not new is for it to be given, so this pitch accent could be used where on the surface the information seems to be new, whereas in fact it is inferable from mutual knowledge and thus given. For instance, in Pierrehumbert and Hirschberg's example (45)

- (45) She's teething
 $H^* H+L^* H L\%$

the hearer might think the information is new since the utterance is a move in a disagreement, but the speaker is suggesting that the hearer knows or should have known this fact and its relevance.

The other way to be not new is to be false, so my theory predicts another reading for $H+L^*$ that Pierrehumbert and Hirschberg's would not predict. It is difficult to come up with examples of this use of $H+L^*$, because what it would convey is something like "This proposition seems like it would be new, but it is really false," and if it is false, why say it. One way to come up with an example is to find a case where the speaker is denying what the hearer has just said, and in particular, denying the validity of the terms the speaker has used. Thus, suppose someone has just suggested that the analysis I gave of a stretch of discourse was a Freudian account. As it happens, I think Freudian psychology is disreputable and unscientific. I reply,

- It's not a Freudian account
 $H+L^*$

- It's a cognitive account
 H^*

What is conveyed here is not that "Freudian" is not new because it is inferable, as Pierrehumbert and Hirschberg's account would have it, but that it is not new because it is ill-conceived and false.

According to Pierrehumbert and Hirschberg's account, $L+H^*$ conveys a contrast or a correction. According to my account, the H^* indicates that

the conveyed proposition is new, whereas the L prefix indicates that the hearer may have believed it to be not new. Our two accounts are consistent. There are two ways of being not new. The hearer could have believed the proposition false; this gives rise to the use of $L+H^*$ for correction. Pierrehumbert and Hirschberg's example (31) illustrates this:

- (31) It's even warm for December
 $L+H^* L H\%$

The hearer seems to believe that it is false that the month is December, and the information that it *is* December is apparently new.

Some examples of contrastive stress can be analyzed similarly. In

- John's a lawyer, but he's honest
 $L+H^*$

the hearer may have inferred from "John's a lawyer" is that John is dishonest. But this is false, and the speaker is conveying the true and new information that John is honest.

The second way of being not new is being given. Thus, $L+H^*$ could say that the speaker might believe the proposition is already mutually known, but it isn't because it contains new information. One way for this to happen is for a set of possible alternatives to be already given and for the proposition to convey the new information about the specific alternative selected. Pierrehumbert and Hirschberg's example (35) is an instance of this:

- (35) They want to know if we can come for dinner
 $L+H^* L H\%$

The hearer knew an invitation for *something* was in the offing. The new information is that it is for dinner.

Let us next examine the suffixes. The H suffix on the L^* pitch accent indicates that the proposition, although conveyed as given or false, is still open. It shouldn't be taken as relevant mutual knowledge until it can be considered further. All of Pierrehumbert and Hirschberg's examples (26) through (30) can be seen to exemplify this pattern.

- (26) He's a good badminton player
 $L^*+H L H\%$

It is mutually known that he's a good badminton player, but this is open in that the hearer might not have believed it relevant.

- (27) Sort of
 $L^*+H L H\%$

This is given in that the hearer should have known the speaker would carry out the garbage, but open in that it's not really an adequate answer.

- (28) Sort of
L*+H L H%

This is given in that the hearer had just said it, but open in that it is uninterpretable.

- (29) Leo
L*+H H L%

This is given in that the recalcitrant pet perhaps already knows it is wanted, but open in that the owner is not quite sure of this.

- (30) Jurgen's from Germany
L*+H H H%

Jurgen's existence is given, but the proposition is open because the speaker wonders why the hearer didn't think of it before and hence adds the H suffix to leave the matter open for discussion. To say the sentence with a simple L* pitch accent would be to contradict the hearer without leaving open the possibility that the hearer has an explanation for the oversight. Adding the H suffix mitigates the implied criticism.

My claims that the stressed information in examples (26), (27), and possibly (29) is given do not seem compelling to me. An alternate account might be this. A H suffix cannot be added to a H* pitch accent, for it could not be distinguished from a simple H*. The speaker thus has a choice of conveying the newness of the information or the openness. The openness is chosen, and this forces the pitch accent to be L*. Under this account, we might want to revise the theory to say that the L* does not signal "not new" but rather *fails* to signal "new." The L* pitch accent would then signal one of three things: given, false, or uncommitted.

According to my account, the H*+L pitch accent should say something like, "This is new information, and its truth and status are not open to question." One reason for it not to be open to question is that it is in fact inferable from mutual knowledge. This therefore explains compositionally Pierrehumbert and Hirschberg's observation that it indicates inferability. Thus, the pedantic tone in their (38)

- (38) Hit the hint key
H*+L H*+L L L%

indicates that the student should already be able to infer this information.

It doesn't seem that inferability should be the only reason an assertion is not open. Authority and expertise are two other possibilities. A sergeant might say to a private,

- Dig that hole
H*+L H*+L L L%

An obnoxious economist may tell a noneconomist in very positive tones,

- Inflation isn't why that happened
H*+L L L%

Pierrehumbert and Hirschberg argue that H+L means the proposition is inferable. This is the case in my account as well, most of the time, but for two independent reasons. H+L* usually says that the hearer might think the information is new, but since it is inferable, it is really given. H*+L says that this is new information and it is beyond dispute because it is inferable from mutual knowledge. However, I came up with examples of uses of each of these complex pitch accents where factors other than inferability justified their use.

According to Pierrehumbert and Hirschberg, L+H evokes a salient scale. Since the set {*false, true*} is a scale, this is too close to vacuous to be compelling. Their association of L*+H with uncertainty or incredulity can be seen to follow from the openness that the H suffix conveys. Their association of L+H* with contrast and correction follows from the corrective or accommodative function of the L prefix.

I should say that my account would be a bit stronger if I could argue that a pitch accent is composed of a basic pitch accent H* or L* with an optional prefix and an optional suffix. That would mean that there would be pitch accents of the form L+H*+L, conveying something like, "You might think this proposition is not new, but it is new, and that's that," and pitch accents of the form H+L*+H, conveying something like, "You might think this proposition is new, but it's not new, or is it?" But Pierrehumbert and Hirschberg tell me that these do not occur. My account would also be stronger if there were complex tones in which the affix were the same as the basic tone. The complex tone H*+H would convey, "This is new information, but we can discuss it." Such tones would be hard to distinguish from simple tones, however, so it's easy to see why they'd be of little use in communication.

In my account, suffixes also occur on intermediate phrases and intonational phrases, and there too a H means that the unit—in this case, the phrase—is open. The most common way for a phrase to be open is for it to satisfy Pierrehumbert and Hirschberg's description of the meaning of a H phrase accent: "the current phrase is to be taken as forming part of a larger composite interpretive unit with the following phrase" (section 6). Thus, in their example (55)

- (55) George ate chicken soup and got sick
H* H* H* H H* H* L L%

the first clause remains open, indicating that there is more to be said about the matter, and the second clause is taken to be strongly linked causally to

the first. In (56)

(56) George ate chicken soup and got sick
 H* H* H* L H* H* L L%

the causal link is not so strongly implicated. Pierrehumbert and Hirschberg say, "A L phrasal tone emphasizes the separation of the current phrase from a subsequent phrase" (section 6). This strikes me as a bit strong. I would say rather that it fails to emphasize the connection.

A H boundary tone in an intonational phrase also indicates openness. Pierrehumbert and Hirschberg's analysis of their examples (12) and (13) illustrates two ways for an intonational phrase to be open. In (12)

(12) My name is Mark Liberman
 H* H* L H%

the H boundary tone sounds strange if we take it to be suggesting that there is more to say about whether the speaker is Mark Liberman. But there is more that needs to be said concerning whether that fact links with the receptionist's expectations. In (13)

(13) I thought it was good
 H* H* H H%

the more that the speaker feels needs to be said is some validation of her opinion.

These examples, like yes-no questions, have an other-directed quality, but that is not what the H boundary tone signifies directly. Rather, it is derivative on the openness that is conveyed. There is more to be said, and it just happens that the hearer is the one who must say it. (It is an interesting question why ordinary *wh*-questions do not also take H boundary tones. One possibility is that incredulous *wh*-questions, which are more open than ordinary *wh*-questions, have preempted the H boundary tone.)

Boundary tones in connected discourse constitute just one more instance of the proposed meaning of suffixes. Pierrehumbert and Hirschberg say early in their paper, "An incorrect theory can make it difficult to establish interpretation, by grouping together contours that actually have disparate meanings or by drawing distinctions that have no meaning" (section 3.1). They apply this dictum to the theory of intonational phenomena to good effect, but unfortunately they do not apply it to the theory of discourse. The theory they adopt, that of Grosz and Sidner, is insufficiently rich in structural possibilities. It allows one to notice *some* of the discourse segments here and there and to talk about the obvious inclusion relations among them. But it recognizes only dominance and satisfaction-precedence relations among the intentions associated with the discourse segments. Using a theory this meager in its possibilities for discourse structure, all one

can say about the high boundary tone H% is that sometimes it means the current segment dominates the next segment, sometimes it means the next segment dominates the current one, and sometimes it means the current one satisfaction-precedes the next one. In other words, it doesn't tell one anything at all.

On the other hand, Pierrehumbert and Hirschberg could have appealed to a theory of discourse structure that focuses on discourse segment *boundaries* in a recursive fashion and asks for each boundary what relation spans that boundary and binds the two segments into parts of a single discourse. There is no shortage of such theories; see, for example, Grimes 1975, Longacre 1976, Hobbs 1978, 1985, Mann and Thompson 1986, or Polanyi 1986. If they had, they would have been in a position to characterize the discourse function of H boundary tones in a very elegant fashion.

Thus, the last of the three background principles we must assume is that discourse is typically structured in a hierarchical fashion. Individual utterances, or intonational phrases, are discourse segments, and when some coherence relation links two segments, the two together constitute a composed segment, which can then in turn be a constituent of a larger segment. In such a theory, Pierrehumbert and Hirschberg's example (59)

- (59) a. My new car manual is almost unreadable
 L L%
 b. It's quite annoying
 L H%
 c. I spent two hours figuring out how to use the jack.

has the structure shown in figure 15.1. Their example (60)

- (60) a. My new car manual is almost unreadable
 L H%
 b. It's quite annoying
 L L%
 c. I spent two hours figuring out how to use the jack.

has the structure shown in figure 15.2. The differing structures are signaled

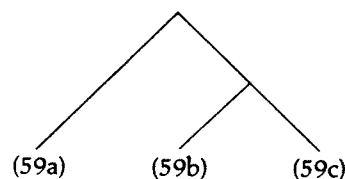


Figure 15.1
 Structure of Pierrehumbert and Hirschberg's example (59).

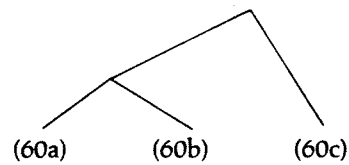


Figure 15.2
Structure of Pierrehumbert and Hirschberg's example (60).

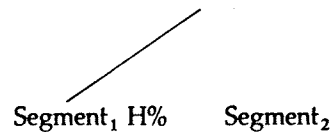


Figure 15.3
The meaning of a H boundary tone.

by the boundary tones, and the differing resolutions for the pronoun "it" fall out of these differing structures (Hobbs 1979).

Given this theory of discourse structure, the characterization of the discourse function of boundary tones can be stated quite succinctly and can furthermore be seen as just another example of signaling openness. A H boundary tone at the end of an intonational phrase can signal that the intonational phrase ends a segment that is a *nonfinal* subsegment of a larger segment (figure 15.3). A low boundary tone *doesn't* signal this.

As a final illustration, consider Pierrehumbert and Hirschberg's example (7):

- (7) The train leaves at seven
 H* H* H* L H%
 It'll be on track four
 H* H* L L%

The L phrase accent after the first sentence says that that bit of information is complete and can be believed as is, while the H boundary tone that follows it says that the sentence is not the last segment in the text but will be followed by more, related information.

Again, I feel that Pierrehumbert and Hirschberg's characterization of L boundary tones is a bit strong. They say that a L boundary tone indicates that the segment is to be interpreted with respect to what has come before, whereas I would say that it simply fails to indicate that the segment is to be interpreted with respect to what is to come. This explains the asymmetry they note at the end of their paper.

In summary, if the modifications I have suggested hold up, we see the English tune system as a simple, elegant structure for communicating the

status of the information being conveyed by the utterance. It consists of a mere three elements: a H*/L* choice to signal new or not new, a shift from a L or H to a H* or L* to indicate a kind of correction or accommodation to what the hearer might have believed the status to be, and a H suffix to indicate the status is still an open question. The power and, at the same time, the seeming complexity of the tune system arise from the abstract character of what the intonational elements signify, allowing them to be put to many uses. If all of this is so, then, with the suggested modifications, Pierrehumbert and Hirschberg's theory of intonational meaning is an impressive achievement indeed.

References

- Grimes, Joseph (1975). *The thread of discourse*. The Hague: Mouton.
 Hobbs, Jerry R. (1978). Why is discourse coherent? Technical Note 176, SRI International, Menlo Park, CA. Also in Fritz Neubauer, ed. (1983). *Coherence in natural-language texts*. Hamburg: Helmut Buske Verlag.
 Hobbs, Jerry R. (1979). Coherence and coreference. *Cognitive Science* 3, 67–90.
 Hobbs, Jerry R. (1985). On the coherence and structure of discourse. Report No. CSLI-85-37, Center for the Study of Language and Information, Stanford University, Stanford, CA.
 Longacre, Robert (1976). *An anatomy of speech notions*. Ghent: Peter de Ridder Press.
 Mann, William, and Sandra Thompson (1986). Relational propositions in discourse. *Discourse Processes* 9, 57–90.
 Polanyi, Livia (1986). The linguistic discourse model: Towards a formal theory of discourse structure. Technical Report 6409, Bolt Beranek and Newman, Inc., Cambridge, MA.

System Development Foundation Benchmark Series

Michael Brady, editor
Robotics Science, 1989

Max V. Mathews and John R. Pierce, editors
Current Directions in Computer Music Research, 1989

Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors
Intentions in Communication, 1990

Eric L. Schwartz, editor
Computational Neuroscience, 1990

Intentions in Communication

edited by Philip R. Cohen, Jerry Morgan, and
Martha E. Pollack

System Development Foundation Benchmark Series

1990

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England