WARNING

Concerning Copyright Restrictions

INTONATION AND PROPOSITIONAL ATTITUDE:
THE PRAGMATICS OF L*+H L H%*

Gregory Ward   Julia Hirschberg
Northwestern University   AT&T Bell Laboratories

1 Introduction

Studies of intonational meaning have largely focussed on the contribution of individual intonational features such as contour, accent, phrasing, to utterance interpretation. However, the interaction among the various contributions of these features has been less studied. In this paper, we reexamine the contribution of one intonational contour, the **L*+H L H%** contour[Pie80],[1] illustrated in Figure 1. This contour would be appropriate in the following context:

(1) A: Would anybody in their right mind marry Manny?
     B: Anna may marry Manny.

We propose a meaning for this contour which subsumes previous proposals and accommodates an even broader range of uses. We then present preliminary results of empirical studies exploring the relationship between what the contour itself conveys and what is conveyed by other prosodic features, including duration, amplitude, pitch range, and voice quality.

2 The Meaning of L*+H L H%

Previous authors have proposed that the **L*+H L H%** contour conveys a statement or answer with reservation ("there's a 'but' about it") [Hal67]; contrast [LS74]; reservation or implied contrast [Bin79]; focus within a set [Lad80]; selection of a variable from the background [Gus83]; incompletion or 'up-in-the-airness' (Bolinger, p.c.).

    In [WH85], we noted some deficiencies in these accounts and proposed a new analysis based upon a large corpus of naturally occurring data. We found that **L*+H L H%** could be employed to convey speaker uncertainty about the appropriateness of some scale or scalar

---

[1] This contour is also known in the literature as °2-4-3 contour [Pik45]; as a subtype of Bolinger's Accent A [Bol58]; as tone 4 [Hal67]; as Bolinger's Accent B [Jac72]; as fall rise [OA61][Lad80][Cut77][WH85]; as contrastive stress within contradiction contour [LS74]; as A-rise [Bin79]; rise-fall-rise [Pie80].

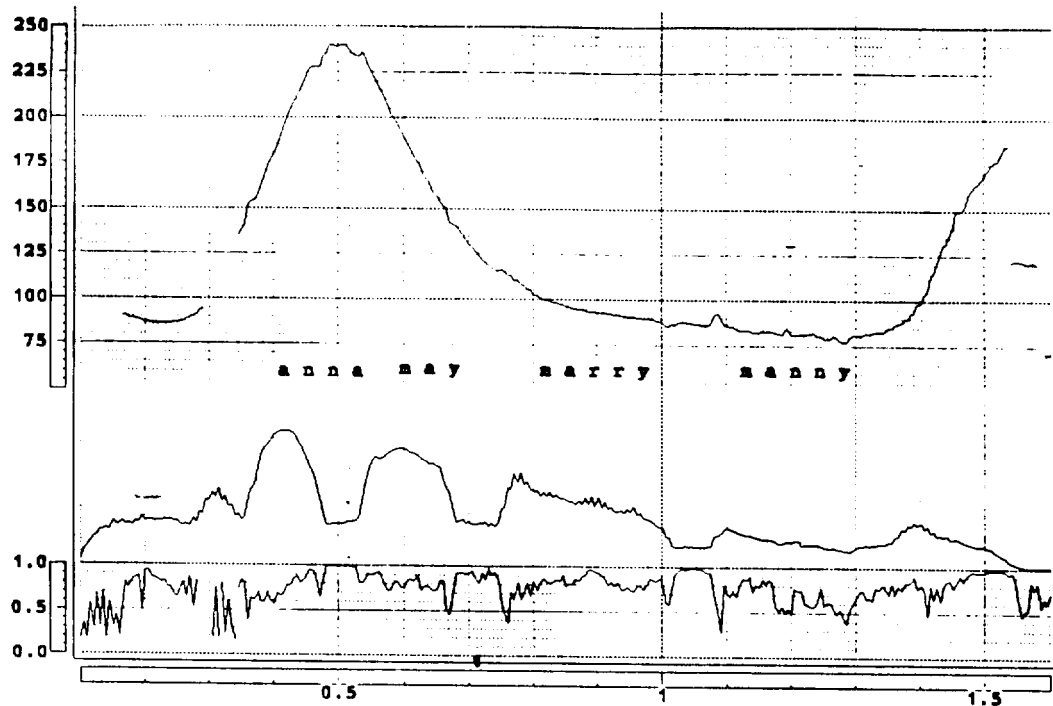Figure 1: The L*+H L H% Intonational Contour

value evoked in the context. Scales are defined as **partially ordered sets**, or **posets**, which are defined by a **partial ordering** R on some set. R must be **reflexive, antisymmetric,** and **transitive**, or alternatively, **irreflexive, asymmetric,** and **transitive.**[2]

This notion allows us to rank discourse referents as **values** on **scales**. The relationships that provide the basis for the felicitous use of **L*+H L H%** are just those that can be represented as partial ordering relations. A value on a scale may be associated with an entity, attribute, event, activity, time, or place – or with a set of such items. We can rank a property with respect to some entity which exhibits it via an *attribute-of* relation; an event, with other events, according to temporal precedence; elements or proper subsets of a set with respect to the set by an inclusion relation; and so on.

Given this definition of scales and scalar values, we claimed that a speaker may convey uncertainty about some scale or scalar value in three ways: First, a speaker may convey uncertainty about whether it is appropriate to evoke a scale at all in a given context. Second, a speaker may convey uncertainty about whether the scale evoked is an appropriate one in a given context. And, third, a speaker may convey uncertainty about whether the evoked scalar value is appropriate. So, for example, (1) illustrates this third type of uncertainty: B is conveying uncertainty about whether Anna is included by A in the set of sane persons.

---

[2] R is reflexive iff for all b1 ∈ B, b1 R b1. It is antisymmetric iff for all b1, b2 ∈ B, (b1 R b2 & b2 R b1) → b1 = b2. It is transitive iff for all b1, b2, b3 ∈ B, (b1 R b2 & b2 R b3) → b1 R b3. R is reflexive iff for all b1 ∈ B, b1 ¬R b1 and asymmetric iff for all b1, b2 ∈ B, b1 R b2 → b2 ¬R b1.

2.1 The Incredulity Interpretation

However, Liberman (p.c.) and Pierrehumbert & Steele[PS87] have noted that L*+H L H% can also convey incredulity. Consider (2):[3]

(2) A. I'd like you here tomorrow morning at eleven.
     B. !Eleven in the morning!

Here, L*+H L H% appears to convey incredulity on the part of the speaker about the proposed meeting time – *not* uncertainty. This clearly presents a problem for our previous analysis.

Three possibilities suggest themselves. The first is that we are dealing with two phonologically distinct contours, i.e., the contour Pierrehumbert & Steele and Liberman have associated with the incredulity reading is *not* the same contour that we have analyzed as conveying speaker uncertainty. If in fact we are discussing the same contour, then there are two additional possibilities: Either L*+H L H% represents a case of **intonational homophony**, i.e. one contour with two distinct interpretations; or, our earlier analysis of the meaning of this contour was overly restrictive in failing to capture this additional meaning – and thus requires modification.

2.2 Reconciling Incredulity with Uncertainty

To test the first of these possibilities, we recorded tokens of L*+H L H% in contexts which favor, first, incredulity and then, uncertainty. For example, in (3), *eleven in the morning* receives the uncertainty interpretation.

(3) A: Do you tend to come in pretty late then?
     B: \Eleven in the morning./

Recall that, in (2), the same string would be interpreted as incredulous. The pitch track corresponding to B's reply in (3) is presented in Figure 2 and that corresponding to B's reply in (2), in Figure 3. In this comparison, as in other pairs we tested, both pitch tracks turn out to represent instances of the same tune – L*+H L H%. However, there are other differences between these and other pairs in duration, amplitude, and pitch range. Specifically, utterances of L*+H L H% inducing an 'uncertainty' interpretation tended to be longer than tokens judged to convey 'incredulity', while the latter tended to be louder and uttered in a larger pitch range. Distinctions in voice quality are also apparent, though less easily measured. However, these features do not function as determinants of tune. So, the first possibility, that there are different tunes inducing incredulity and uncertainty, appears to be unfounded.

Given that we are dealing with a single contour, we must now determine how best to characterize its contribution to utterance interpretation. This presents a somewhat more difficult problem. That is, given two meanings associated with a single phenomenon, should that phenonemon be treated simply as ambiguous, or is there some abstraction from which both meanings can be derived? If this abstraction captures a significant generalization,

---

[3] We use ! ... ! to indicate an incredulous interpretation of L*+H L H% here, and we will employ \ ... / to indicate the uncertainty interpretation below.
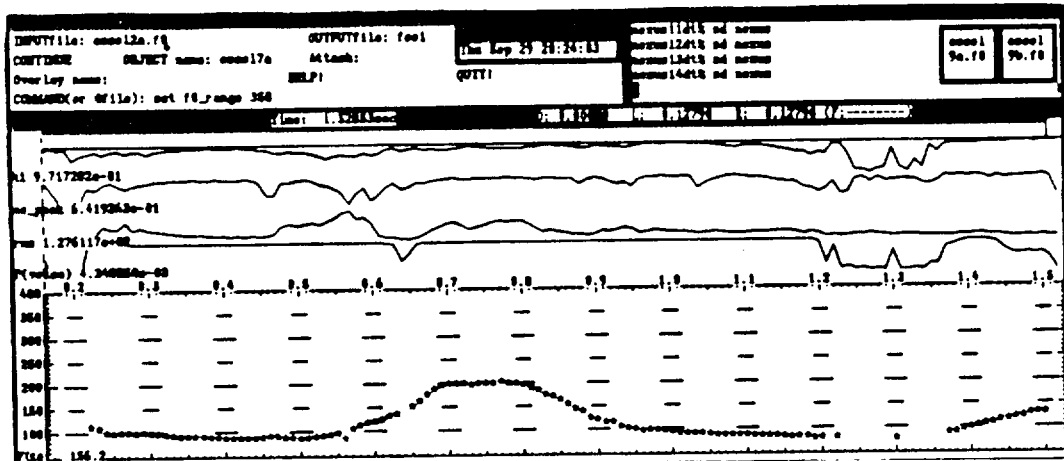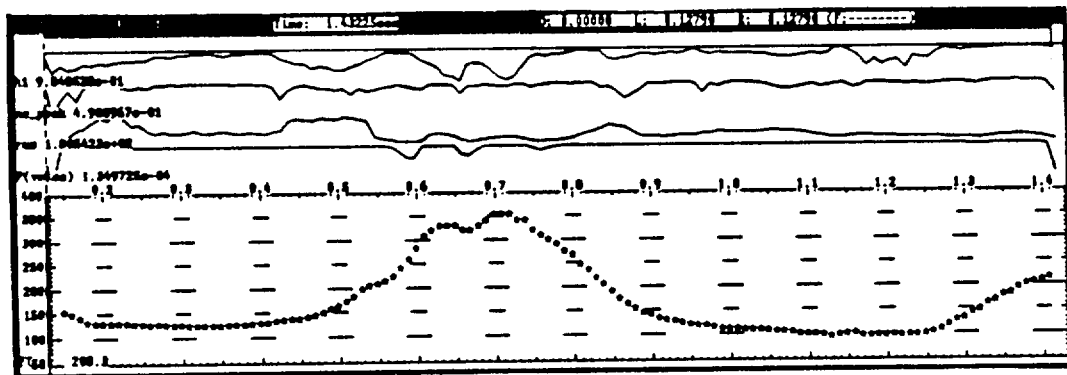
Figure 2: L*+H L H% Used to Convey Uncertainty

Figure 3: L*+H L H% Used to Convey Incredulity

then clearly it is to be preferred. In [WH86], we claimed that there exists just such a generalization to be captured in the case of L*+H L H%.

We proposed that a speaker's use of the L*+H L H% contour can convey *lack of speaker commitment to the appropriateness of a evoked scale or scalar value*. This analysis could accommodate both the notion that this contour is used to convey uncertainty, i.e., 'It's *not* the case that the speaker believes a scale or scalar is appropriate', and the notion that it can convey incredulity, i.e., 'It *is* the case that the speaker believes a scale or scalar is *inappropriate*'. The subsumption of incredulity and uncertainty by the abstraction 'lack of speaker commitment' can be explained as follows: For any speaker S and any scale or scalar x, there are four possibilities:

(4)   a. BEL(S, APPROPRIATE(x))
      b. BEL(S, ¬APPROPRIATE(x))
      c. ¬BEL(S, APPROPRIATE(x))
      d. ¬BEL(S, ¬APPROPRIATE(x))

We can say that S is uncommitted to the appropriateness of x whenever b ∨ (c & d) is true. Now, we can say that S is incredulous about the appropriateness of x just in case b

is true. And, we can say that S is uncertain about the appropriateness of x whenever (c & d) are true. So, lack of speaker commitment (b ∨ (c & d)) subsumes both incredulity and uncertainty.

In [WH85], we described three particular types of uncertainty that L*+H L H% can be used to convey, i.e. uncertainty about i) whether any scale at all is appropriate; ii) which of the possible scales is appropriate; and iii) which value on some particular scale is appropriate. This taxonomy can be extended to include cases in which the contour is used to convey incredulity.

An example of L*+H L H% involving incredulity about some value on a scale is illustrated above in (2), repeated below in (5):

(5) A. I'd like you here tomorrow morning at eleven.
     B. !Eleven in the morning!

Here, the temporal scale is relevant, and B conveys, via L*+H L H%, belief that a value on that scale, i.e. *eleven in the morning*, is inappropriate. Type I incredulity is exemplified by (6):

(6) B: Did you take out the garbage?
     A: Sort of.
     B: !Sort of!

In this exchange, B conveys incredulity about the fact that A has evoked a scale – where the scale evoked here is something like 'stages of a process of taking out the garbage'. For B, no scale is appropriate here. Note in fact that A may use an L*+H L H% contour also to convey uncertainty in this exchange, as in (7):

(7) B: Did you take out the garbage?
     A: \Sort of./
     B: !Sort of!

Here, B employs L*+H L H% to convey incredulity about the uncertainty conveyed by A's prior use of the contour. Finally, in (8), we illustrate Type II incredulity:

(8) A: I bet I know why Mary isn't dating John any more. He's ugly.
     B: !He's ugly!

In (8), B conveys that a scale of attractiveness is inappropriate.[4]

So, we see that there are three ways in which speakers can convey incredulity using L*+H L H%, corresponding to the three ways speakers can use the same contour to convey uncertainty.

---

[4] Of course, B's remark might also be interpreted as conveying type III incredulity: "!Ugly! I think John's quite handsome."

2.3 Distinguishing Incredulity from Uncertainty

If L*+H L H% can indeed be employed to induce two related but distinct meanings, the question then arises of how each meaning is is conveyed. While there do appear to be some relevant contextual factors, these turn out to be neither necessary nor sufficient in determining which meaning is understood.

First, we note that, when speakers employ L*+H L H% to convey uncertainty, it is usually directed at their *own* choice of some scale or scalar value. However, when it is used to convey incredulity, it is almost invariably directed toward another's choice of scale or scalar. This accords with another observation we have made about instances in which L*+H L H% conveys incredulity: Generally, in such cases, speakers express incredulity about a value already evoked in the discourse. For example, notice the infelicity of B1's incredulity in (9); while B2's uncertainty is fine.[5]

(9) A: Everybody had a good time.
    B1: #!Some people had a good time!
    B2: \Some people had a good time./
    B3: !Everybody had a good time!

Of course, it is difficult to imagine why B1 would want to convey that *some* is inappropriate in this context, since its appropriateness is nowhere in question – although the appropriateness of *everybody* could be (see B3). But B2 might plausibly wish to convey uncertainty about the appropriateness of a new value – *some* – which she herself has proposed. Despite the apparent restriction of an incredulous interpretation of L*+H L H% to items evoked in the discourse, items do not need to be explicitly mentioned to be evoked. Consider (10):

(10) A: I hear John and Mary are calling it quits.
    B1: !They're separating!
    B2: \They're separating./

And note that B2, with an 'uncertainty' reading, is equally plausible — conveying that John and Mary aren't really 'calling it quits'. Thus, although context may favor one interpretation over another, it does not determine that interpretation.

So, despite the fact that we can now explain how the two 'meanings' associated with L*+H L H% are related yet distinct, and that we can suggest some contextual features that may distinguish one use from the other, we are still left with the problem of what hearers use to distinguish one meaning from the other. In section 2.2 we noted that we had observed consistent phonetic differences between tokens uttered with the uncertainty reading and those uttered with the incredulity reading — differences in voice quality, duration, amplitude, and pitch range. The incredulity readings tended to be shorter and louder than the uncertainty readings, and to be uttered with a larger pitch range. We hypothesized that one or more of these factors might be responsible for the difference in conveyed meaning and decided to test this hypothesis by performing a perception experiment.

3 Empirical Investigation

To determine which of the phonetic distinctions were perceptually salient in leading hearers to differentiate between the two interpretations, we prepared a set of stimuli in which features

---

[5] We employ '#' to denote pragmatic infelicity.

Stimulus	Maximum Pitch (Hz)	Peak Amplitude (RMS)	Duration (Sec)
2a	208.3	5583.66	1.32460
b	357.1	6929.61	1.31950
4a	277.8	4167.985	1.46930
b	333.3	5788.562	1.35840
5a	222.2	2927.779	2.18650
b	344.8	3376.461	1.99270
6a	270.3	4991.326	1.57050
b	344.8	8069.461	1.35840
7a	172.4	3381.015	1.88730
b	250.0	4732.901	1.64390
8a	172.4	4213.625	1.99590
b	222.2	5934.369	1.97420
9a	153.8	1274.243	1.66680
b	294.1	1347.634	1.38380
10a	181.8	1496.344	1.59610
b	285.7	5374.462	1.56920

Table 1: Difference Between Stimulus Pairs

from one utterance of a given sentence conveying incredulity were exchanged for features of another conveying incredulity. For example, for the sentence 'The Smiths invited Veronica', we elicited two natural utterances with the L*+H L H% contour, one conveying uncertainty (appropriate in a context like (11a)) and one conveying incredulity (appropriate in a context like (11b)):

(11) a. Did the Smiths invite anyone interesting to their party?

b. I hear the Smiths have invited your worst enemy to the party.

We chose eight such pairs,[6] each similar to those shown in Figures 2 and 3 in displaying differences between 'uncertainty' and 'incredulity' versions in amplitude, duration, pitch range, and voice quality. Data for each pair of stimuli are shown in Table 1. In each case, the 'incredulity' versions were louder, shorter, and uttered in a higher pitch range than the 'uncertainty' versions, and voice quality in the former was more constricted than in the latter.

---

[6]The eight pairs used were (where nuclear stress is indicated by small capitals):

- ELEVEN in the morning.
- EMILY is willing to do it.
- George ordered BLUEBERRIES for dessert.
- EVELYN knows the answer.
- Bob's going out with ANNA.
- John's vacationing in VIENNA.
- Nine MILLION.
- Gary's buying a VOLVO.

From these pairs, we prepared resynthesized versions, in which one or more features of one version were replaced by those of the other.[7] Stimuli for the perception experiment consisted of resynthesized utterances of all possible combinations of exchanges of the four features, i.e.: a resynthesized version of the original 'uncertainty' utterance; the 'uncertainty' utterance with the amplitude (A) of the resynthesized 'incredulity' utterance and with all other features of the 'uncertainty' utterance unchanged; 'uncertainty' with the duration (D) of 'incredulity'; 'uncertainty' with the spectral features (S) of 'incredulity'; 'uncertainty' with the pitch range (F) of 'incredulity'; 'uncertainty' with A and D of 'incredulity'; 'uncertainty' with D and S of 'incredulity'; 'uncertainty' with S and F of 'incredulity'; and so on, for a total of sixteen variants per pair of original utterances.

Subjects were nine Bell Labs employees.[8] They were asked to listen (over headphones) to utterances played one at a time and to decide whether the speaker was trying to convey uncertainty or incredulity. The task was self-paced and subjects could listen to a single stimulus as often as they liked before making a decision, although they could not go back and change earlier responses.

3.1 Results

Overall results of the study indicate that listeners were most influenced by pitch range and spectral characteristics in differentiating between the 'uncertain' and the 'incredulous' interpretations of L*+H L H%. Figure 4 shows the general pattern of subject responses, pooling subjects and stimuli. (Shaded bars represent responses of 'incredulity'.[9] Letters under each bar indicate which feature(s) of the 'uncertainty' stimulus were switched for those of the 'incredulity' response).

From Figure 4 it appears that the major features distinguishing the two interpretations of L*+H L H% for these subjects was a combination of S and F. Exchanging both these features for those of the 'uncertainty' reading (either alone, or in combination with other features) was indeed more effective than switching any other feature or set of features. Comparing stimuli sets which differed only in whether both S and F were retained or both were exchanged for features of the 'incredulity' stimulus (e.g., comparing all stimuli in which ASF were exchanged to those in which A alone was exchanged, all in which DSF were exchanged to those in which D alone was exchanged, and so on) showed that over twice as many (2.089) of the former were judged to convey 'incredulity'. A two-tailed t-test of

---

[7] To switch pitch ranges, we simply switched f0 contours, after having pitch-tracked each natural utterance and checked to make sure the contours were of the same type, L*+H L H%. For instances of the same contour, we will assume that switching f0 contours is equivalent to switching pitch ranges. Characteristics of an utterance's amplitude, duration, spectral characteristics, and f0 contour were derived from pitch tracks (using a cross-correlation pitch tracker) and (autocorrelation) lpc analyses of each pair of natural utterances, using programs written by David Talkin. These substitutions were accomplished by means of a hybridization program adapted from one originally written by Mary Beckman. (Briefly, each feature for each utterance was sampled and new lpc analyses and pitch tracks produced by linear interpolation. Each new 'hybridized' utterance was then produced from some pairing of new or original lpc file with new or original file of f0 values, depending upon which features were to be substituted.) Resynthesis of hybridized results was performed via programs adapted from (pitch-synchronous) lpc resynthesis programs written by Mark Liberman. Analysis of stimuli was performed using Talkin's *WAVES* speech software.

[8] Ten subjects performed the task, but results from one were subsequently eliminated, since the subject protested strongly that 'uncertainty' and 'incredulity' could not be distinguished on philosophical grounds.

[9] While responses for individuals varied in terms of overall tendency to assign an 'incredulity' or 'uncertainty' interpretation to stimuli in general, the pattern of relative responses for each subject reflected the general pattern shown here.
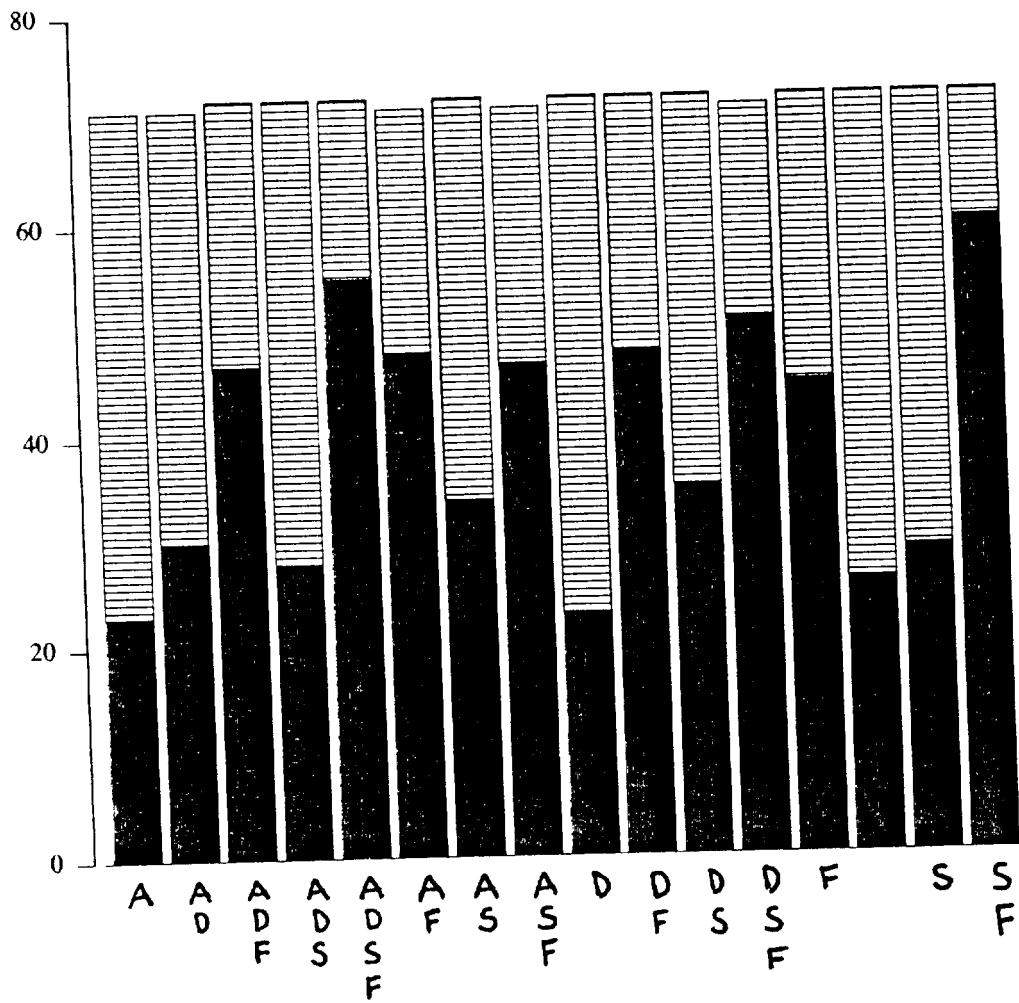
Figure 4: Results Over All Subjects and Stimuli

differences in number of 'incredulity' judgments by pair of stimulus sets was significant at the .002 level (t-stat=12.33, df=3).

However, when we compare stimuli in which SF were exchanged with corresponding stimuli in which S was not (e.g., ADSF exchanged vs. ADF exchanged, DSF vs. DF, and so on), it appears that it is F alone that induces the observed effect. There is no significant difference between judgments of stimuli in which both are exchanged and those in which F is exchanged while S is retained. When F exchanges are compared to F retentions with all other features constant, the ratio of 'incredulity' judgments for 'incredulity' F exchanges vs. 'uncertainty' F retentions is 1.76. A two-tailed t-test of differences by pairs of stimulus sets here was significance at the .001 level (t-stat=9.78, df=7). No other feature or combination of features produced a significant effect which could not be explained solely in terms of F.

Not surprisingly, stimulus pairs in which the 'uncertainty' and 'incredulity' stimuli differed most in pitch range were those in which pitch range substitutions had the largest effect. The three pairs with the largest differences (ranging from 122.6 to 148.8 Hz) also ranked as those in which F substitution produced the greatest increase in 'incredulity' judgment from stimuli in which the original F was tested. However, a simple linear model does not explain the results for the remaining pairs: For one (pair 4), with a range difference of 55.5 Hz, substitution of the higher range appears to have produced a negative change of 1 in 'incredulity' judgments. For all other pairs, change in F produced an increase of at least one-third in 'incredulity' judgments. Furthermore, no special additional phonetic distinctions seem to separate pair 4 from the other stimuli; nor did subjects exhibit an unusual predilection for judging stimuli formed from this pair 'uncertain'. Thus additional phonetic and semantic considerations do not appear useful. While it might appear that the difference in pitch range observed for pair 4 (55.5 Hz) was insufficient to affect hearer judgments about the interpretation of L*+H L H%, in another pair (8), a difference of only 49.8 Hz in the substituted feature produced a positive change of 20. Furthermore, the change observed for pair 8 is explanable in terms of no other factor(s) but substitution of F.[10]

In sum, examination of four potential acoustic correlates of L*+H L H% revealed that pitch range alone induced a reliable effect on interpretation. While each stimulus pair exhibited similar differences in amplitude, duration, and spectral characteristics, substitution of these features in our stimulus pairs resulted in no significant effect.

4 Discussion

The characterization of the sorts of contributions various intonational features make to utterance interpretation is a long-term goal of studies of intonational meaning. In this paper, we have proposed an extension of our previous analysis of the contribution the L*+H L H% contour makes to utterance interpretation, which accommodates both an 'uncertainty' and an 'incredulity' interpretation of the contour. Our empirical study of the phonetic variations that characterize the two interpretations suggests that the primary factor distinguishing the two perceptually is pitch range. When the pitch range of an 'incredulous' utterance of L*+H L H% was substituted for that of an 'uncertainty' utterance of the same sentence, subjects were significantly more likely to judge the utterance to be conveying 'incredulity' than 'uncertainty'. While amplitude, duration, and spectral characteristics for such utterances also differed in consistent ways, these features did not prove perceptually

---

[10]That is, no other feature substitution produces significant results when compared with that feature's (or set of features') retention, as in the comparisons described above.

salient in the experiment.

Assuming then that pitch range does play a crucial role in the interpretation of **L*+H L H%**, we might wonder whether pitch range plays a similar role in the interpretation of other contours. For example, does CONTINUATION RISE when uttered with a larger pitch range trigger a different meaning from continuation rise uttered with a smaller range? An examination of contours in this light will bring us closer to identifying the context-independent correlates of intonational meaning.

References

[Bin79]   Janet Bing. *Aspects of English Prosody*. PhD thesis, University of Massachusetts at Amherst, Amherst MA, 1979.

[Bol58]   Dwight Bolinger. A theory of pitch accent in english. *Word*, 14:109–149, 1958.

[Cut77]   Anne Cutler. The context-dependence of 'intonational meanings'. In *Papers of the Fifth Regional Meeting*, Chicago, 1977. Chicago Linguistic Society.

[Gus83]   Carlos Gussenhoven. *On the grammar and semantics of sentence accents*. Foris, Dordrecht, Neth., 1983.

[Hal67]   M. A. K. Halliday. Notes on transitivity and theme in english, part 2. *Journal of Linguistics*, 3:199–244, 1967.

[Jac72]   Ray S. Jackendoff. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge MA, 1972.

[Lad80]   D. Robert Ladd. *The Structure of Intonational Meaning*. Indiana University Press, Bloomington, Ind., 1980.

[LS74]   Mark Liberman and Ivan A. Sag. Prosodic form and discourse function. In *Papers of the Tenth Regional Meeting*, pages 416–427. Chicago Linguistic Society, 1974.

[OA61]   J.D. O'Connor and G.F. Arnold. *Intonation of colloquial English*. Longmans, London, 1961.

[Pie80]   Janet B. Pierrehumbert. *The Phonology and phonetics of English intonation*. PhD thesis, Massachusetts Institute of Technology, September 1980.

[Pik45]   Kenneth Pike. *The intonation of American English*. University of Michigan Press, Ann Arbor MI, 1945.

[PS87]   Janet B. Pierrehumbert and Shirley Steele. How many rise-fall-rise contours? In *Proceedings of the Eleventh Meeting*, Tallinn, 1987. International Congress of Phonetic Sciences.

[WH85]   Gregory Ward and Julia Hirschberg. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, 61(4), December 1985.

[WH86]   Gregory Ward and Julia Hirschberg. Reconciling uncertainty with incredulity: a unified account of the L*+H L H% intonational contour. Paper presented at the LSA Annual Meeting, New York, 1986.