

WARNING

Concerning Copyright Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or reproduction.

One of three specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship or research.

If electronic transmission of reserve material is used for purposes in excess of what constitutes “fair use”, that user may be liable for copyright infringement.

This policy is in effect for the following document:

The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise–fall–rise intonation contour in English

Julia Hirschberg

AT&T Bell Laboratories, 2D-450, 600 Mountain Avenue, Murray Hill, NJ 07974, U.S.A.

Gregory Ward

Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60208, U.S.A.

Received 9th August 1990, and in revised form 17th September 1991

The present study investigated which prosodic variables were perceptually salient in the interpretation of a particular English intonational contour, rise–fall–rise ($L^* + H L H\%$ in Pierrehumbert's system of intonational description). A series of utterance pairs were recorded in which sentence and contour were held constant but context was varied to produce tokens of the "incredulity" and the "uncertainty" readings of the contour. From these utterances, stimuli were synthesized in which all possible combinations of four features (duration, amplitude, pitch range and spectral characteristics) from the "incredulity" reading were substituted for the corresponding features of the "uncertainty" reading. These features were selected based upon previous observations of consistent differences between productions of the two interpretations. Subjects listened to the utterances and selected one of the two meanings. ANOVA performed on the results indicated a large main effect for pitch range and a smaller one for spectral features in interpretation selection. Amplitude and duration showed no effect. It is suggested that the association of greater pitch range with incredulity can be accounted for by the previously noted tendency of listeners to associate larger pitch ranges with a greater degree of speaker involvement; conversely, the association of smaller pitch range with uncertainty can be explained as a consequence of the perception of lesser speaker involvement.

1. Introduction

Previous studies of intonational meaning have investigated the contribution of particular prosodic features to utterance interpretation. In particular, it has been shown that type of contour, in general, as well as pitch accent placement and intonational phrasing, in particular, all play a significant role in conveying different types of meaning (Lieberman & Sag, 1974; Ladd, 1980; Gussenhoven, 1983; Scherer,

Ladd & Liberman, 1984; Bolinger, 1989). Much less attention, however, has been paid to the interaction between such intonational and other prosodic features. In this paper, we investigate the relative contribution of F_0 (fundamental frequency) range, duration, amplitude and spectral characteristics to the interpretation of the rise-fall-rise intonational contour in English. This contour has been referred to in the literature as '2-4-3 contour (Pike, 1945); as a subtype of *Accent A* (Bolinger, 1958); as *tone 4* (Halliday, 1967); as *fall rise* (O'Connor & Arnold, 1961; Ladd, 1980; Cutler, 1977; Ward & Hirschberg, 1985); as *contrastive stress within contradiction contour* (Liberman & Sag, 1974); as *A-rise* (Bing, 1979) as well as *rise-fall-rise*. We will represent it as $L^* + HLH\%$, following Pierrehumbert's (Pierrehumbert, 1980) theory of English intonation. The contour is illustrated in Fig. 1, with the utterance *Eleven in the morning*.

In Pierrehumbert's system, intonational contours are represented as sequences of high (H) and low (L) tones in the F_0 contour. Contours have as their domain the *intonational phrase*, a level of phrasing composed of one or more *intermediate phrases* plus a *boundary tone*, which is either high (H%) or low (L%); this tone controls the pitch at the edge of the intonational phrase. An intermediate phrase consists of one or more *pitch accents* plus a *phrase accent*, also H or L, which controls the pitch from the last pitch accent to the end of the phrase. The utterance depicted in Fig. 1 consists of a single intonational phrase with an H% boundary tone; the intonational phrase itself consists of a single intermediate phrase with an L phrase accent.

For Pierrehumbert, there are six pitch accents in English: two simple accents, H* and L*, and four complex accents, L* + H, L + H*, H* + L, and H + L*. In complex accents, the tone marked with "*" is aligned with the stress of the accented word; so, L* + H is distinguished from L + H* by the way the tune is aligned with the text. The most prominent accent of an intermediate phrase is called the nuclear stress. In Fig. 1, the nuclear stress falls on *eleven*, which bears an L* + H accent. The $L^* + HLH\%$ contour, as illustrated in Fig. 1, thus comprises one or more L* + H pitch accents, an L phrase accent, and an H% boundary tone.

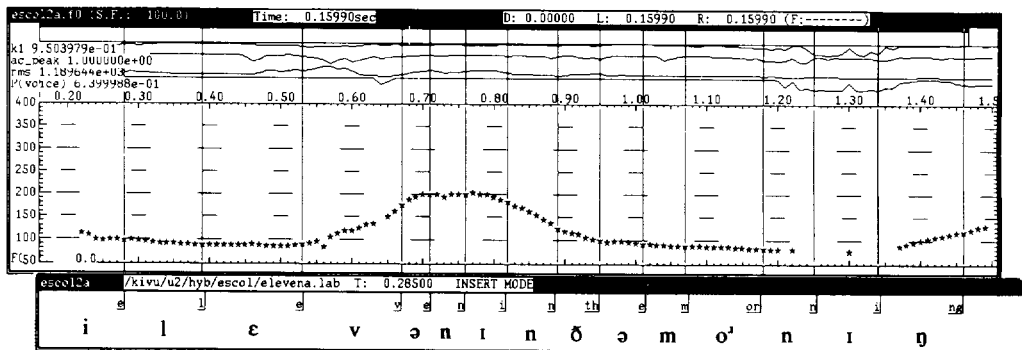


Figure 1. Fundamental frequency (Hz) for the $L^* + HLH\%$ contour produced on the phrase *Eleven in the morning*, with nuclear stress on *Eleven*. The vertical cursors mark segment boundaries, as determined from the waveform.

This contour would be appropriate in the following context:

(1) A: So, do you tend to come in pretty late then?

B: ELEVEN in the morning

$L^* + H$

$LH\%$

In earlier work (Ward & Hirschberg, 1985, 1988), we proposed a meaning for this contour which subsumed previous proposals and accommodated an even broader range of uses. In this paper, we present the results of an empirical study, testing how variation in duration, amplitude, pitch range and spectral characteristics affects subjects' interpretation of the $L^* + HLH\%$ contour. Some implications for the analysis of the meaning of intonational contours in general are then discussed. We begin by describing the possible interpretations of the contour.

1.1. The Meaning of $L^* + HLH\%$

Previous authors have proposed that the $L^* + HLH\%$ contour conveys: a statement or answer with reservation ("there's a 'but' about it") (Halliday, 1967); contrast (Lieberman & Sag, 1974); reservation or implied contrast (Bing, 1979); focus within a set (Ladd, 1980); selection of a variable from the background (Gussenhoven, 1983); and incompleteness or "something left unsaid" (Bolinger, 1986). In Ward & Hirschberg (1985, 1988), we noted some deficiencies in these accounts and proposed a new analysis of the contour, based upon a large corpus of naturally occurring data. We proposed that $L^* + HLH\%$ is employed to convey *lack of speaker commitment* to the appropriateness of a scale or scalar value evoked in the context.¹ This lack of commitment accommodates the two general interpretations of the $L^* + HLH\%$ contour discussed in the literature—the "uncertainty" interpretation (Halliday, 1967; Bing, 1979; Gussenhoven, 1983; Ward & Hirschberg, 1985; Bolinger, 1986), which we will indicate using "\ . . . /" around the text, and the "incredulity" interpretation (Lieberman, 1975; Pierrehumbert & Steele, 1987), which we will indicate with "! . . . !". For example, the phrase *eleven in the morning* uttered with $L^* + HLH\%$ in the context provided above in (1) appears to convey speaker uncertainty. In this context, the temporal scale is relevant, and B conveys, via $L^* + HLH\%$, his or her belief that a value (i.e. *eleven in the morning*) on that scale is inappropriate. That is, B conveys uncertainty about whether 11:00 a.m. amounts to "coming in late" for A. However, in (2), the same sentence uttered with $L^* + HLH\%$ appears to convey incredulity.

(2) A. I'd like you here tomorrow morning at eleven.

B. !ELEVEN in the morning!

Here, B conveys via $L^* + HLH\%$ his or her belief that 11:00 a.m. is too early to come in.

If $L^* + HLH\%$ can indeed be employed to induce either uncertainty or incredulity, what is it, then, that conveys one meaning over the other? While there do appear to be some relevant contextual factors, these turn out to be neither necessary nor sufficient in determining which meaning is to be understood.

¹ Scales are defined as *partially ordered sets*, following Hirschberg (1985). See Ward & Hirschberg (1985) for a discussion of the role of scales in the interpretation of $L^* + HLH\%$.

First, we note that, when speakers employ $L^* + HLH\%$ to convey uncertainty, it is usually directed at their *own* choice of some scale or scalar value. However, when the contour 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 $L^* + HLH\%$ conveys incredulity: when speakers use the contour to express incredulity, they generally express that incredulity about a value already evoked in the discourse. For example, notice the infelicity of B1's incredulity in (3); while B2's uncertainty is fine. (We employ “#” to denote pragmatic infelicity.)

- (3) 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 $L^* + HLH\%$ to items already evoked in the discourse, items do not need to be explicitly mentioned to be evoked. Consider (4):

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

Note that B2, with an uncertainty reading, is equally plausible—conveying that John and Mary are not really “calling it quits”. Thus, although context may favor one interpretation over another, it does not determine that interpretation.

So, we are still left with the problem of what hearers use to distinguish one meaning from the other. In Ward and Hirschberg (1988), we observed that there appeared to be consistent prosodic differences between utterances of $L^* + HLH\%$ produced to convey an uncertainty reading and those uttered to convey the incredulity reading—differences in spectral characteristics, duration, amplitude and pitch range. Tokens of the latter tended to be shorter and louder than tokens of the former, and tended to be uttered with a larger pitch range. We hypothesized that one or more of these factors might account for these differences in interpretation.

2. Methods

To determine the effect that duration, amplitude, spectral features and pitch range have on the interpretation $L^* + HLH\%$, we recorded a series of utterance pairs. In each pair, sentence and contour were held constant, but context was varied to favor one of the two readings. So, for the phrase *Eleven in the morning*, we elicited two natural utterances with the $L^* + HLH\%$ contour, one conveying uncertainty (appropriate in a context like (1)) and one conveying incredulity (appropriate in a context like (2)). The pitch track corresponding to B's reply in (1) was presented in Fig. 1 and that corresponding to B's reply in (2) is presented in Fig. 2.

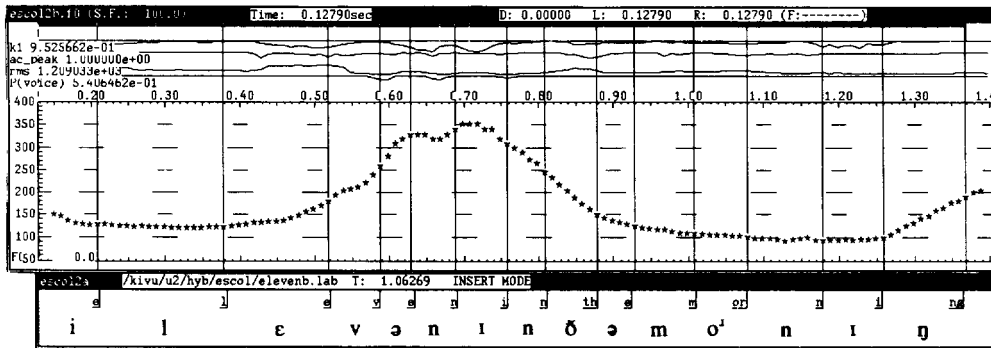


Figure 2. Fundamental frequency (Hz) for the L* + H L H% contour used to convey incredulity. As in Fig. 1, the phrase is *Eleven in the morning*, with nuclear stress on *Eleven*, and the vertical cursors mark segment boundaries.

As predicted, there are indeed differences between productions in duration, amplitude, spectral characteristics and pitch range. Specifically, *Eleven in the morning* uttered to convey uncertainty is somewhat longer than the same phrase uttered to convey incredulity, while the latter has a greater maximum amplitude and is uttered in a larger pitch range. A distinction in some aspects of voice quality is also apparent, although difficult to characterize. Impressionistically, utterances conveying incredulity appear to be uttered with a “tenser”, or “sharper” voice than do those conveying uncertainty. We focussed on the spectral characterization of these differences in this study, to separate these from other factors, such as intensity, which have also been associated with differences in voice quality, and which here are examined separately.

In all, eight pairs of utterances were recorded. The sentences recorded are listed in Table I. Note that the word uttered with the L* + H accent is indicated by small capitals.

Each token was checked by several independent observers to verify that the utterance conveyed the intended interpretation. Again, we noted consistent differences in each pair of utterances between productions that conveyed uncertainty and those that conveyed incredulity—differences in amplitude, duration, pitch range and spectral features. Data for each pair of stimuli are shown in Table II, with the uncertainty versions corresponding to the uncertainty stimulus and the incredulity versions, the incredulity.

F₀ peak was measured at the maximum of the “H” tone of the nuclear stress, which in each case bore the L* + H accent. Peak amplitude was measured at the amplitude

TABLE I. Stimuli used in perception study

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.

TABLE II. Acoustic features of stimulus pairs

Stimulus	Maximum F_0 (Hz)	Peak amplitude (RMS)	Difference in loudness (db)	Duration (ms)
1u	208	5584	1.88	1325
i	357	6930		1320
2u	278	4168	2.85	1469
i	333	5789		1358
3u	222	2928	1.24	2187
i	345	3376		1993
4u	270	4991	4.17	1571
i	345	8069		1358
5u	172	3381	2.92	1887
i	250	4733		1644
6u	172	4214	2.97	1996
i	222	5934		1974
7u	154	1274	0.49	1667
i	294	1348		1384
8u	182	1496	11.11	1596
i	286	5374		1569

maximum within the nuclear stressed accent. Note that, while in each case the duration of the uncertainty version exceeds that of the incredulity version, this difference in some pairs appears minimal.

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 from these utterances for a perception experiment. In the "hybridized" stimuli, features from the incredulity version of the sentence were systematically exchanged for the corresponding features of the uncertainty version, to produce stimuli with all possible combinations of amplitude, duration, spectral characteristics and F_0 range from the two original utterances, for a total of 16 stimuli synthesized for each original pair of utterances. (For instances of the same contour, we assume that switching F_0 contours is equivalent to switching pitch ranges. So, to substitute pitch ranges, we simply switched F_0 contours, after having pitch-tracked each natural utterance and checked to make sure that the contours were both instances of $L^* + H L H\%$.) That is, for each pair of original utterances, 16 new utterances were obtained by substituting the amplitude profile of the incredulity version for that of the uncertainty version, by substituting the amplitude and durational characteristics of the incredulity version for those of the uncertainty version, by substituting the amplitude, duration and spectral features of the incredulity version for those of the uncertainty version, and so on, for a total of 16 hybrid synthetic utterances.

To prepare these stimuli, each utterance was first digitized (using linear pulse code modulation, or PCM) at a sampling rate of 12 kHz and high pass filtered to remove low frequency noise. Next, amplitude and F_0 measurements were obtained for each digitized utterance using a cross-correlation pitch tracker and a sampling rate of 10 kHz to extract RMS amplitude and pitch. Alignment points for major F_0 excursions were marked by hand from the pitch tracks and PCM files for each stimulus, so that subsequent durational manipulations would preserve the original

contours. A 15-order autocorrelation linear predictive coding (LPC) analysis (with a 3 ms window) was then performed on each natural utterance.

New PCM files were created by first sampling RMS, timing, LPC coefficients and pitch information for each original utterance of each stimulus pair, and then synthesizing new F_0 and LPC files in which one or more of the sampled features of one utterance were exchanged for those of the other, with synthetic files produced by linear interpolation between exchanged sample points. New utterances were then resynthesized from these “hybrid” files using (pitch-synchronous) LPC re-synthesis and a sampling rate of 10 kHz.²

So, stimuli for the perception experiment consisted of synthesized utterances of all sixteen possible combinations of exchanges of the four features for each of the eight pairs of utterances, for a total of 128. Stimuli were blocked by sentence type and randomized within block and between blocks, so that one random hybrid token of each sentence type appeared in random order in each block. Subjects for the perception experiment were nine Bell Laboratories employees, all native speakers of English.³ They were instructed to listen (over headphones) to each utterance and to decide whether the speaker was trying to convey uncertainty or incredulity. To avoid biasing the judgments, no attempt was made to define these terms. The task was forced choice, “uncertainty” or “incredulity”; there were five invalid responses. Each subject judged each of the 128 utterances once, so the data were fully crossed. The task was self-paced; subjects could listen to a given stimulus as often as they liked before making a decision, although they were not allowed to reconsider earlier stimuli or responses. Each subject took between thirty and 45 minutes to complete the task.

3. Results

Figure 3 shows the general pattern of subject responses, pooling subjects and stimuli. The solid portion of each bar shows number of “incredulity” responses for stimuli with features taken from the incredulity original; the shaded portion shows number of “uncertainty” responses. In this figure, we see that stimuli to which subjects assigned the incredulity interpretation most frequently were those in which both F_0 range and spectral characteristics were from the “incredulity” utterance. Exchanging these features for the corresponding features of the uncertainty reading (either alone, or in combination with other features) appears to have been more effective than exchanging any other feature or set of features. While subjects varied in their general tendency to assign incredulity or uncertainty interpretations, the relationship of incredulity to uncertainty judgments for each subject in each category reflected the general pattern shown in Fig. 3. This pattern appeared to indicate that at least two factors were important in influencing subject judgments: F_0 and, to a lesser extent, spectral features. That is, taking the “uncertainty” stimuli as a base in each case, the substitution of the “incredulity” pitch range (via substitution

²We used software written by David Talkin for pitch extraction and LPC analysis. We also used Talkin’s WAVES software in the analysis of the original utterances. The hybridization and resynthesis software used was written by Julia Hirschberg, based on a hybridization program written by Mary Beckman and resynthesis software written by Mark Liberman. All software ran on SUN 3/110s.

³Ten subjects performed the task, but results from one subject were not scored, since the subject protested strongly on philosophical grounds that “uncertainty” and “incredibility” could not be distinguished. None of the other subjects found the task onerous.

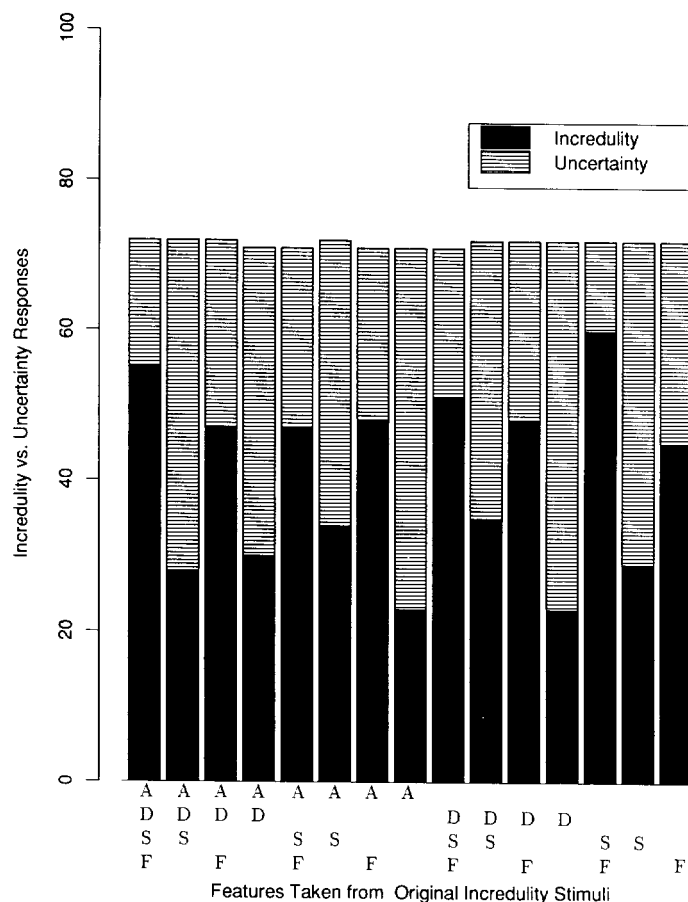


Figure 3. “Incredulity” vs. “uncertainty” responses to the 16 types of hybrid stimuli. Letters under the x-axis indicate features extracted from the stimulus originally produced with the incredulity meaning: A: amplitude, D: duration, S: spectral characteristics, F: F_0 . Responses pooled over all sentences and all subjects.

of F_0) seemed to have influenced subject judgments toward “incredulity” more than the substitution of any other of our four independent variables (amplitude, duration and spectral features), with substitution of spectral features a distant second.

To test this conclusion, we performed analysis of variance (ANOVA) tests on the data presented in Table III.⁴ In this analysis, spectral characteristics (S), amplitude (A), duration (D) and pitch range/ F_0 (F) were treated as within subject factors, with subject judgment the dependent variable. Initial tests showed a significant effect of sentence type (F -ratio = 146.3174, $df = 7$, $p < 0.001$); that is, certain sentences tended to elicit more “uncertainty” or “incredulity” judgments than others. However, we found no interaction of sentence type with any of the acoustic variables under consideration here. So, both subject and sentence type were treated

⁴For this analysis, we employed a repeated measures ANOVA algorithm (as described in Hays, 1981), implemented by Jan van Santen.

TABLE III. Results of analysis of variance

Source	SS	df	MS	F-ratio	p
S	1.9175	1	1.9175	8.2569	0.0042
D	0.0078	1	0.0078	0.0336	0.8545
A	0.0009	1	0.0009	0.0037	0.9513
F	26.5842	1	26.5842	114.4709	0.0000qq
S × D	0.0217	1	0.0217	0.0934	0.7599
S × A	0.4592	1	0.4592	1.9773	0.1600
S × F	0.0425	1	0.0425	0.1832	0.6688
D × A	0.0217	1	0.0217	0.0934	0.7599
D × F	0.0217	1	0.0217	0.0934	0.7599
A × F	0.1050	1	0.1050	0.4522	0.5014
S × D × A	0.0078	1	0.0078	0.0336	0.8545
S × D × F	0.0078	1	0.0078	0.0336	0.8545
S × A × F	0.0078	1	0.0078	0.0336	0.8545
D × A × F	0.0703	1	0.0703	0.3028	0.5823
S × D × A × F	1.4592	1	1.4592	6.2833	0.0123
Error	1.8579	8	0.2322		

S: spectral features; A: amplitude; D: duration; F: fundamental frequency range.

as random factors, and sentence data were collapsed over the sentence factor. Results are presented in Table III.

Line four of Table III confirms that indeed there is a large main effect for F_0 range (F) (F -ratio = 114.47; $df = 1$; $p < 0.001$) on subjects' interpretation of stimuli. The very high F -ratio for this factor should be noted. Line one shows a smaller main effect (F -ratio = 8.26; $df = 1$; $p < 0.005$) for spectral characteristics. Lines two and three of Table III indicate that there was no significant effect for either duration or amplitude. The remainder of the table shows results for two-, three- and four-way ANOVAs on the data, and reveals no significant interactions among any of the features, assuming the p -value for the four-way ANOVA (line 15) is agreed to be uninterpretable.

Thus, results of this study suggest that listeners were primarily influenced in differentiating between the uncertain and the incredulous interpretations of L* + H L H% by variation in pitch range. There was a lesser but still significant effect for variation of spectral characteristics as well—but none whatsoever for duration and amplitude, despite the consistent differences we found between these factors—as well as F_0 and spectral features—in the original pairs of utterances. Thus we would conclude that only two of observed types of difference were perceptually salient in listeners' interpretation of the contour as conveying uncertainty *vs.* incredulity.

Given our previous observation that the incredulity utterances exhibited consistently larger pitch ranges than the uncertainty utterances, we might wonder whether stimulus pairs in which the uncertainty and incredulity stimuli differed most in pitch range would be the ones in which pitch range substitutions had the largest effect. To test this hypothesis, we examined the scores for each stimulus pair separately. The three pairs with the largest pitch range differences (i.e., pairs 1, 3 and 5 in Table II, with differences ranging from 122.6 to 148.8 Hz) did rank as those in which F_0

substitution produced the greatest increase in (raw) incredulity judgment from stimuli in which the original F_0 was tested.

However, a simple linear model does not explain the results for the remaining pairs: For example, for pair 2—with a range difference of 55.5 Hz—substituting the higher range for the lower appears to have actually decreased the number of incredulity judgments. For all other pairs, substitution of the higher range did produce an increase (of at least one-third) in incredulity judgments. Furthermore, no special additional phonetic distinctions seem to separate pair 2 from the other stimuli; nor did subjects exhibit an unusual predilection for assigning the uncertainty interpretation to stimuli formed from this pair. Thus additional phonetic and semantic considerations do not appear relevant. While it might seem that the difference in pitch range observed for pair 2 (55.5 Hz) was insufficient to affect hearer judgments about the interpretation of $L^* + H L H\%$, in another pair (6), a difference of only 49.8 Hz in the substituted feature resulted in a 55% increase in the number of incredulity judgments. Furthermore, the change observed for pair 6 is explainable in terms of no other factor(s) but change in pitch range.

In sum, examination of four potential acoustic correlates of $L^* + H L H\%$ revealed that pitch range played the largest role in the interpretation of this contour, with larger pitch ranges inducing the incredulous interpretation and smaller pitch ranges, the uncertain interpretation. There was a smaller, though significant, effect of spectral characteristics. However, amplitude and duration appeared to play no significant role in subjects' interpretation of the contour.

4. Discussion

The characterization of the contributions that various intonational features make to utterance interpretation is a long-term goal of studies of intonational meaning. In this paper, we have investigated the influence of four prosodic features—duration, amplitude, spectral characteristics and F_0 range—on the interpretation of $L^* + H L H\%$. An empirical study of subject perceptions of utterances of this contour reveals that the primary factor distinguishing an incredulity interpretation from an uncertainty interpretation is pitch range. That is, when the (larger) pitch range of an incredulous utterance of $L^* + H L H\%$ was substituted for the (smaller) pitch range of an uncertainty utterance of the same sentence, subjects were significantly more likely to judge the utterance to be conveying incredulity than uncertainty. There was a lesser but still significant effect for spectral characteristics. While amplitude and duration for our stimulus utterances also differed consistently, these features did not prove perceptually salient in the experiment.

Given our findings, how can we explain why variation in pitch range distinguishes between the two interpretations of $L^* + H L H\%$? It has been claimed in the literature that one function of pitch range is to indicate degree of speaker involvement (Bolinger, 1986). That is, larger pitch ranges indicate a greater degree of involvement and smaller pitch ranges indicate a smaller degree. Intuitively, incredulity would appear to entail a greater involvement on the part of a speaker than would uncertainty. Recall that subjects interpreted stimuli with larger pitch ranges as conveying incredulity and those with smaller pitch ranges as conveying uncertainty. Thus, $L^* + H L H\%$ uttered with a relatively larger pitch range would be interpreted as incredulous, while $L^* + H L H\%$ uttered with a relatively smaller

pitch range would be interpreted as uncertain. Assuming then that pitch range does play a crucial role in the interpretation of $L^* + H L H\%$, we might wonder whether it plays a similar role in the interpretation of other contours. An examination of contours in this light will bring us closer to identifying the context-independent correlates of intonational meaning.

We thank Mary Beckman, Janet Pierrehumbert, David Talkin and Jan van Santen for useful comments, advice and generous sharing of software from the inception of this project. Thanks also to Bob Ladd, Toni Rietveld and an anonymous reviewer for comments on an earlier version of this paper.

References

- Bing, J. (1979) Aspects of English prosody. PhD thesis, University of Massachusetts at Amherst, Amherst.
- Bolinger, D. (1958) A theory of pitch accent in English, *Word*, **14**, 109–149.
- Bolinger, D. (1986) *Intonation and its parts: melody in spoken English*. Palo Alto: Stanford University Press.
- Bolinger, D. (1989) *Intonation and its uses: melody in grammar and discourse*. London: Edward Arnold.
- Cutler, A. (1977) The context-dependence of “intonational meanings”. In Papers of the fifth regional meeting of the Chicago Linguistic Society, Chicago, pp. 117–125.
- Gussenhoven, C. (1983) *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
- Halliday, M. A. K. (1967) Notes on transitivity and theme in English, part 2, *Journal of Linguistics*, **3**, 199–244.
- Hays, W. L. (1981) *Statistics*, 3rd edn. New York: Holt, Rinehart and Winston.
- Hirschberg, J. (1985) A theory of scalar implicature. PhD thesis, University of Pennsylvania.
- Ladd, D. R. (1980) *The structure of intonational meaning*. Bloomington: Indiana University Press.
- Lieberman, M. & Sag, I. A. (1974) Prosodic form and discourse function. In *Papers of the tenth regional meeting of the Chicago Linguistic Society*, Chicago, pp. 416–427.
- O'Connor, J. & Arnold, G. (1961) *Intonation of colloquial English*. London: Longman.
- Pierrehumbert, J. B. (1980) The phonology and phonetics of English intonation. PhD thesis, The Massachusetts Institute of Technology, Cambridge.
- Pierrehumbert, J. B. & Steele, S. (1987) How many rise–fall–rise contours? In *Proceedings of the eleventh international congress of phonetic sciences*, Vol. 3, pp. 145–147. Tallinn: Estonian Academy of Sciences.
- Pike, K. (1945) *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Scherer, K., Ladd, D. R. & Silverman, K. E. A. (1984) Vocal cues to speaker affect: Testing two models, *Journal of the Acoustical Society of America*, **76**, 1346–1356.
- Ward, G. & Hirschberg, J. (1985) Implicating uncertainty: The pragmatics of fall–rise intonation, *Language*, **61**, 747–776.
- Ward, G. & Hirschberg, J. (1988) Intonation and propositional attitude: The pragmatics of $L^* + H L H\%$. In *Proceedings of the fifth meeting of the eastern states conference on linguistics*, Philadelphia.