The domain of realization of the L- phrase tone in American English

Jonathan Barnes, Stefanie Shattuck-Hufnagel**, Alejna Brugos* & Nanette Veilleux***, ***

*Linguistics
Boston University
Boston, MA 02215
jabarnes@bu.edu
abrugos@bu.edu

**Research Lab of Electronics
Massachusetts Institute of
Technology, USA
stef@speech.mit.edu

***Dept. of Computer Science
Simmons College
Boston, MA 02115
veilleux@simmons.edu

Abstract

The phonetic realization of intonational targets in the f0 contour is not always straightforwardly predicted by their affiliations in the segmental string, and the phrase tones of American English are a type of target for which several hypotheses about the domain of realization have been advanced. By varying the metrical structure of target words at the end of a phrase produced with the H* L- H% 'surprised dismay' contour, we determined that a) the right edge of the L-, signaled by the beginning of the rise for the H%, occurs close to the right edge of the phrase, b) the left edge of the L-, signaled by the end of the fall from the H*, stretches leftward to seek a prominent syllable, and c) there is significant variation in the resolution of the various factors that influence these two inflection point locations.

1. Introduction

Virtually all approaches to intonational phonology recognize a distinction between two types of pitch events: those signaling prominences, and those marking boundaries. In the current ToBI transcription standard (and associated theoretical frameworks), this distinction is captured by pitch accents on the one hand, and phrase or boundary tones on the other. Since Pierrehumbert 1980, many have recognized a distinction between two types of edge-marking tone: the phrase accent (hereafter T-), now held to be affiliated with the right edge of an Intermediate Phrase, and an analogous tone marking the right boundary of the Intonational Phrase, known as a boundary tone proper (hereafter T%). Despite their affiliation with phrase boundaries, phrase accents are often realized at a distance from the boundary. In an abundance of languages they have been documented migrating leftward from the right edge of the Intermediate Phrase. A prime example of this tendency is the long high plateau before the final rise in English questions such as the now emblematic *Does Manitowoc have a bowling alley?*, typically analyzed as an L* H- H% contour, with an anticipatory stretch of the H- all the way to the end of *Mánitowòc*. But what exactly conditions this migration? Pierrehumbert and Beckman 1988 characterize so-called secondary associations of phrase tones as the simultaneous realization of the tone in question both at the right boundary of the relevant Intermediate Phrase, and also somewhere else, for example at the boundary of the word bearing the phrase's final pitch accent in the example given above. Grice et al. 2000 support the general principle of secondary association, arguing however, contra Pierrehumbert 1980 and Pierrehumbert and Beckman 1988, that it is typically driven by attraction to metrically prominent phrase-internal syllables, rather than to phrase-internal constituent boundaries. They present a rich cross-linguistic typology of phrase accent behavior to identify several sometimes conflicting parameters governing the phrase accents' leftward march. These include the association with some material at the right edge of the Intermediate Phrase, and various attractions to, e.g., the rightmost stressed syllable in the phrase, the most prominent postnuclear stressed syllable, and the nuclear syllable itself. The interplay of these attractions results in a variety of contours. For English in particular, Grice et al. examine a number of fall-rise contours, e.g., a low valley formed between the nuclear syllable and the final postnuclear lexical stress in two contours (the traditional QFR and IFR), and attraction to the most prominent postnuclear stress in another (the traditional compound fall-plus-rise), but they acknowledge that these observations await laboratory confirmation. Gussenhoven 2000, on the other hand, analyzing the boundary tones of Roermond Dutch, observes some attraction to metrical prominence, but keeps it formally distinct from autonomous pressures motivating the phrase accents' leftward orientation. Instead, he motivates this pattern simply by the interaction of conflicting constraints mandating the occurrence of his boundary tones both as far right, and as far left, as is possible in the phrase.

This paper is an experimental study of the behavior of phrase accents in English as realized in a single contour, the emphatic fall-rise notably applied in Ladd 1996 to both *Sue?!* and *A driving instructor?!* By applying this contour, a H* L- H% in ToBI terms,[1] to a single (contextualized) noun phrase, while varying that phrase's metrical and constituent structure, we sought to survey the behavior of the English L- under circumstances diverse enough to identify the distinct pressures on its realization, and thereby to bring some measure of resolution to the conflicting views described above. The specific questions we address in the context of this contour are (1) whether English phrase accents are in fact attracted leftward to some metrically prominent position, and if so to

[1] We are that Ward and Hirschberg have worked extensively on a Fall-Rise contour which they analyze with an L*+H nuclear accent and a meaning involving incredulity (Ward and Hirschberg 1985, Hirschberg and Ward 1992). We were therefore particularly careful in our labelling, to be certain that the H component of the nuclear accent was in fact an H*. By this we mean that the rise associated with that H began well within the accented vowel. While there is always some ambiguity to such distinctions, we are certain that there were no clear instances of L*+H nuclear accents in our analysis set.

which (2) whether English phrase accents show any tendency to coincide with the boundary of the final pitch-accented word, and (3) whether English phrase accents exhibit anything like pure leftward and rightward orientation, regardless of what additional constraints they might also be obeying.

2. Methods

2.1. Stimuli: Because the experimental hypotheses concern the effect of metrical word structure on the alignment of turning points in the f0 track, the metrical structure of the target words was varied systematically. Eleven stimulus classes, only a few of which can be discussed in any detail here, were selected to control two aspects of metrical structure: the full-vowel vs. reduced-vowel status of the last two syllables of the final word, and the placement of main stress in the phrase-final and pre-final words. The eleven stimulus structures are summarized in Table 1.

Case	Metrical Pattern	Examples
1a	**S** w w (red)	Elizabeth, examiner, convertible
1b	**S** w w	Anthony, banality, tonality,
2a	**S** w s w (red)	rutabaga, minimizer, millimeter,
2b	**S** w s w	Abercrombie, palimony
3	**S** w S w w	apple harvester, eel examiner
4	**S** w S w w w	BARELY palatable, MOSTLY tolerable
5	**S** w w w	palatable, negligible
6	**S** w s w S w s w	cantalope enumerator, antelope exterminator
7	**S** w s w s w S w	cantalope enumeration, antelope extermination
8	**S** w w s w S w	panda enumeration
9	**S** w w w s w S w	elephant enumeration

Table 1: *Categories of metrical structures used in this study*. Bolded [**S**] denotes pitch-accented primary stress, plain [S] an unaccented primary stress, [s] a lexical secondary stress, and [w] an unstressed syllable. Where relevant, [w] (red) represents a reduced final vowel.

2.2. Experimental Design: These cases were designed to answer the following questions:

<u>Case 1</u>: As a baseline of sorts, in the absence of metrically strong postnuclear syllables, does the L- remain close to the phrase edge, creating a L- H% contour on the final syllable?
<u>Case 2</u>: Upon introduction of a metrically strong postnuclear syllable, does the L- retract to the penult, removing the rise from the final? Or does it rather *spread* to the penult, creating a short low valley from penult to final, leaving the rise in place? Alternatively, perhaps it is simply insensitive to the change in metrical structure? The (b) forms of both (1) and (2) were meant to determine whether the full-vowel vs. reduced-vowel status of the final syllable influences the effect of the postnuclear strong syllable in any way.
<u>Case 3</u>: Assuming that the L- displays attraction to a postnuclear metrical prominence (either through shift of the f0 valley or of the left edge thereof), does the distance of that syllable from the phrase boundary have an influence? In other words, is a tonal lapse at the phrase edge tolerated?
<u>Case 4</u>: As for (3), only with a longer lapse.
<u>Case 5</u>: Does extending the metrically weak postnuclear stretch found in (1) alter the behavior of the L-? In other words, how much of a tonal lapse between the nuclear pitch accent and the phrase accent is tolerated?
<u>Cases 6 and 7</u>: Given two metrically strong postnuclear syllables, which (if either) does the phrase accent target: the more prominent, the rightmost, or the leftmost?
<u>Cases 8 and 9</u>: As with (5), does a longer stretch of weak syllables between the nuclear pitch accent and the first postnuclear stress affect alignment behavior. (5) evaluates the effect of this lapse in the absence of postnuclear stresses. (8) and (9) evaluate this with such stresses in place.

For each target word, a short dialogue was constructed to maximize the chances of eliciting the grouping of the words of the target utterance into a single intonational phrase with a H* L- H% intonational contour (Table 2). The 62 stimuli were randomized; the subject read the sequence through twice, in opposite order, which took about 20 minutes.

2.3. Subjects: Nine volunteer speakers were recruited from an undergraduate class at Simmons College; all were between the ages of 18 and 22, were native speakers of American English, and had no self-reported history of hearing or speech problems. All were female.

S 1: Try to draw me a parabola. S 2: A parabola?!? I can barely draw a sine wave!	case 2a
S 1: That guy looks old enough to be the examiner. S 2: The examiner?!? He's our youngest student!	case 1a
S 1: That soup is barely palatable. S 2: BARELY palatable?!? It's deliciously palatable!	case 4

Table 2: *Examples of dialogues used to elicit the H*L-H% contour*. Speaker 2 (S2) is the subject, and the target words are contained in the first clause of each of her utterances. Speaker 1 (S1) parts were read by the experimenter.

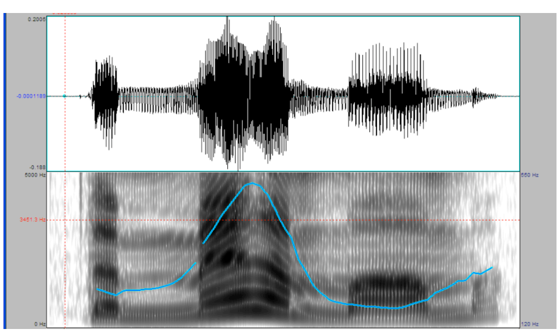
2.4. Elicitation method: Each speaker began by listening to a set of four dialogues that included the desired contour on a word that differed in its metrical structure from all of the target word classes (e.g., monosyllabic "warm" or disyllabic "Tiger"). The speaker then practiced producing the contour with this small set. When the experimenter was satisfied that the speaker could produce the target contour reliably, elicitation of the experimental utterances began. This overt training period was acceptable because the experimental hypothesis concerned not whether or when the speaker would produce the target contour, but rather how the low f0 region associated with the L- would be realized when the target contour was in fact produced, and because the subject heard no words with the target metrical patterns. Thus the aim of the training and practice periods was to ensure that the speaker grasped the contour and was able to produce it reliably in this experimental context for at least some introductory sentences.

2.5. Prosodic screening: In all, 1134 utterances were produced. The training and practice periods had the desired effect to some extent, in that most of the speakers produced the target H* L- H% contour (often realized as L+H* L- H%) on many of their utterances. However, there were also many utterances produced with different intonational contours, so the first step in analysis was to ToBI-label each of the utterances. A total of 631 utterances were determined to have a contour different from the intended target. Most often these could be characterized as either L* H- H% or L* H- L% (303 tokens). A further 261 tokens were perceptually somewhat similar to the target contour, but were realized with a sustained flat f0 after the fall from the pitch accent, i.e. without a final rise. (In some of these cases the f0 rise in the final syllable, associated with the H%, was very small, so a quantitative threshold of 1.5 semitones was established to resolve borderline cases that were ambiguous between a final rise and a flat final f0. This threshold was determined by perceptual agreement among 3 experienced labellers.) An additional 57 tokens produced with two pitch accents in the target utterance were also omitted.

2.6. Acoustic alignments: In order to determine how the metrical structure of the phrase-final word(s) affects the alignment of f0 turning points associated with the L-, it was necessary to label the location of these turning points within the syllables and segments of the target utterances. This labelling was carried out by four experienced ToBI labellers, i.e. the authors of this paper. For some of the utterances produced with the target intonation contour, the f0 track (computed by Praat) was too irregular (due to strong segmental effects, heavy glottalization, or vocal fry) to permit the alignment of the turning points to be labelled. The final number of utterances which could be labelled was 468, making up 41% of the total number of utterances collected.

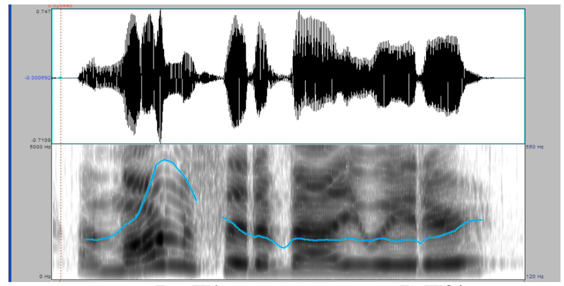
For these target utterances, there were a number of different shapes our contour could take, but in every case it was possible to identify both the end of the f0 fall, and the beginning of the following rise. In certain cases, the fall was complex, in such a way that an initial steep fall was followed by a shallower section. At this point, in some tokens the rise began immediately, while in others a short plateau intervened between the two, sometimes spanning several syllables. In Table 3, the first L denotes the end of the entire fall. The H denotes the beginning of the rise.

Figure 1: *Schematic of the three major f0 contours observed.*



L+ H*   L-H%
a   mi-   ni-   mum

Figure 2*: Example of a Case 1a exemplar*. The end of the steep fall occurs in the /m/, followed by a shallow descent and then rise in the final vowel.



L+ H*   L-H%
Marmoset eva- lu-   a- tor

Figure 3: *Example of a Case 6 exemplar***.** The end of the fall occurs at the onset of /va-/ and the rise begins after a flat plateau, at the /t/.

3. Discussion of Results

Results were coded as can been seen in Table 3, with L and H marking crucial F0 turning points.

Pre-ante-penult syllable	Ante-penult rime	Penult Onset C	Penult rime	Inter vo-calic C	Final syl rime	# tokens
A(b) **S**	er w	cr	om s	B	ie w	
			L	L	LH	17
			L	H		11
			LH			9
		L	L	H		4
				LH		2
				L	H	1
	L	L	H			1

Table 3: Summary of results for Case 2b, : 45 tokens of words of the form *Abercrombie*, (**S** w s w) were analyzable.

The major patterns observed for cases 1-5 are summarized in Table 4. From these codings, several primary conclusions can be drawn. First, the English L- we investigated is manifestly stress-seeking, in line with the predictions of Grice et al. This is seen in comparison of Cases 1a and 2a. For 1a (*Elizabeth)*, with a nuclear antepenult and a weak penult and final, a single pattern was predominant. In 53 of 57 analyzable tokens (nearly 93%), the fall from the pitch-accented syllable ended, and the rise to the H% began, within the rime of the final syllable. The L-, in other words, is edge-bound. In 2a (*rutabaga*—**S** w s w*)* however, something different happens. Here in every instance the fall was completed before the rime of the final syllable. In 40 out of 54 analyzable cases (74%), this turning point was clearly within the stressed penult. In the remaining 14 tokens, the end of the fall coincided with the ambiguously syllabifying consonant between the stressed penult and the final. The rise to H% was typically earlier in 2a words as well, beginning before the rime of the final syllable in 41 of 54 tokens (76%), and during the rhyme of the final only in 13 tokens (24%). We may conclude from this the following: in each case at least part of the L- occurs during the stressed penult. In most cases, both the end of the fall and the beginning of the rise occur before the final syllable altogether, while in a minority of tokens, the right edge (beginning of the rise) has stayed anchored to the final syllable, but the left edge (end of the fall) is found further leftward, creating a low plateau from the stressed penult to within the rhyme of the final syllable. Cases 1b and 2b show similar results. The repetition of the Case 1 pattern in Case 5 (*palatable*—**S** w w w), with 25 of 25 tokens showing both end of fall and beginning of rise exclusively within the rime of the weak final syllable, indicates that it is indeed metrical structure, and not, for example, phonetic distance from the nuclear syllable, that conditions the realization of the L-.

	end of fall	beg of rise
1a	final rime: 53/57 (93%)	final rime: 56/57 (98%)
1b	i) final rime: 37/51 (72.5%) ii) ambig C: 14/51 (27.5%)	final rime: 51/51 (100%)
2a	i) penult: 40/54 (74%) ii) ambig C: 14/54 (26%)	i) penult or ambig C: 41/54 (76%) ii) final rime: 13/54 (24%)
2b	i) penult: 41/45 (91%) ii) ambig C: 3/45 (7%)	i) penult or ambig C: 28/45 (62%) ii) final rime: 17/45 (38%)
3	i) antepenult: 30/30 (100%)	i) antepenult: 2/30 (6.7%) ii) ambig C: 5/30 (16.7%) iii) final rime: 23/30 (76.7%)
4	i) preantepenult: 26/27 (96%)	i) preantepenult: 1/27 (3.7%) ii) ambig C: 1/27 (3.7%) iii) final rime: 25/27 (92.6%)
5	final rime: 25/25 (100%)	final rime: 25/25 (100%)

Table 4: Most common f0 patterns for the various cases of metrical structure where H*L-H% was elicited.

Interestingly, this attraction to metrical prominences was far from unrestricted. Cases 3 and 4 show a surprising *absence* of the predominant tonal retraction pattern from Case 2. In 4, for example, (*BARELY palatable*—**S** w S w w), the occurrence of both end of fall and beginning of rise in the strong postnuclear syllable [*pal*] was documented just once in 27 analyzable tokens (4%). The predominant pattern, rather, was for the rise to remain within the rime of the final syllable, but for the end of the fall to coincide with the postnuclear stress, creating a low, flat stretch over four syllables. This was found in 21 of 27 cases (78%). Category 3 results were analogous, though with a shorter low stretch. The L- is still stress-seeking, but it achieves its goal through moving its left edge to an earlier syllable, leaving its right edge anchored, rather than by moving both together. The near-total absence of the Case 2 pattern from Cases 3 and 4 suggests a restriction on the distance that the right edge of an L- can stray from the boundary it signals. That it can retract one syllable, but apparently not two, potentially suggests a ban on tonal lapse, à la Zoll 2003, at phrase edges. Note that this is not predicted by a conflict between one constraint (as per Gussenhoven) mandating rightward alignment, and another mandating association to a stressed syllable. In this situation, if multiple association is not an option (as in 76% of Category 2a responses), then either the L- should retract fully from the weak final to the strong penult, antepenult, or preantepenult equally (assuming dominance of the licensing constraint), or the L- should fail to retract (assuming dominance of the alignment constraint). The asymmetrical distribution of the retraction pattern across our categories suggests that a *Lapse analysis may be superior (violated by retraction in 3 and 4, and not violated by retraction in 2).

The results from Categories 6, 7, 8, and 9 are more complex, and cannot be treated here in any detail. However, it should be noted that we found a certain amount of evidence for the pattern posited by Pierrehumbert and Beckman 1988, wherein the left edge of the phrase accent occurs not on a metrically prominent postnuclear syllable, but rather close to the right edge of the pitch-accented word itself.

4. Conclusions

These observations indicate that the realization of phrase tones in English is subject to a great deal of variation, and this variation increases with the complexity of the metrical contexts in which the contour is realized. Analysis of within-speaker variation and regularity is clearly a crucial next step in this project. We can conclude, nonetheless, as predicted by Grice et al., that the left edge of English L- is indeed stress-seeking. Further results not surveyed above suggest that this attraction may be sensitive to degree of stress, in that it prefers stronger prominences to weaker. There are also clear restrictions on how far from the phrase edge the right edge of the L- stretch can stray. This restriction looks very much like a ban on tonal lapse at the boundary. Lastly, there is also some evidence for attraction of the left edge of the L- to phrase-internal word boundaries. More generally, we can identify a variety of pressures which influence the domain of realization of English phrase accents, an interplay strongly suggesting an Optimality-Theoretic analysis as a promising future research direction.

5. References

Grice, M., Ladd, D.R., Arvaniti, A., 2000, On the Place of Phrase Accents in Intonational Phonology, Phonology 17, Cambridge University Press, 143-185.

Gussenhoven, C., 2000, 'The boundary tones are coming: On the non-peripheral realization of boundary tones', in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, M. B. Broe and J.B. Pierrehumbert (eds.), Cambridge: Cambridge University Press, pp. 132-151.

Hirschberg, J. and Ward, G. 1992. The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonational contour in English. *Journal of Phonetics* 20: 241-251.

Ladd, D.R., 1996, *Intonational Phonology,* Cambridge: Cambridge University Press.

Pierrehumbert, J.B., 1980, The phonetics and phonology of English intonation. Ph.D. dissertation, MIT.

Pierrehumbert, J.B. and Beckman, M., 1988, *Japanese Tone Structure.* Cambridge, MA: MIT Press.

Ward, G and Hirschberg, J. 1985. Implicating uncertainty: the pragmatics of fall-rise intonation. *Language* 61: 747-776.

Zoll, C. Optimal Tone Mapping. *Linguistic Inquiry* 34 (2): 225-268.