JOURNAL ARTICLE

Analysis of Intonation: the Case of MAE_ToBI

Carlos Gussenhoven

Radboud Universiteit Nijmegen, Netherlands
c.gussenhoven@let.ru.nl

Annotation systems for intonation contours are ideally based on a well-motivated phonological analysis of the language in question, such that instances of indecision are restricted to uncertainties over what intonational structure the speaker has used, rather than over the choice of label in situations where no suitably distinctive label is available or more than one suitable label is available. This contribution inventorizes a number of cases of overanalysis and underanalysis in MAE_ToBI and argues that they are in large part due to the decision by Pierrehumbert (1980) to analyze a rising-falling accent as a rising pitch accent (L+H*) followed by a L-tone from a different source (an 'on-ramp' analysis). It is shown how the opposite choice, a falling pitch accent preceded by a L-tone from a different source (an 'off-ramp' analysis), avoids most of these problems. Results from a perception experiment testing MAE_ToBI's prediction of intonational boundaries show that steep falls do not always signal a boundary. The inclusion of a tritonal prenuclear pitch accent, which explains the absence of an intonational boundary after a steep fall followed by a gradual rise, can readily be accommodated in the 'off-ramp' analysis, but not in MAE_ToBI.

1 Introduction

Faced with the need to identify the phonological elements in a single rising-falling accent peak in an otherwise low-pitched intonation contour, an analyst has three options, all of which have been adopted for West Germanic (**Figure 1**). First, the rise could be the pitch accent and the fall a transition between H and a following L-tone. This option was taken by Pierrehumbert (1980), as L + H*, and was inspired by Bruce (1977). In his description of Central Swedish, lexically contrastive pitch accents occur in the stressed syllable of the word, while a focus-marking H-tone, functionally equivalent to the pitch accents of English, is sequenced after the lexical pitch accent of the last word in the focus constituent. In broad-focus sentences, this focus marking H-tone occurs after the last lexical word and thus before the final boundary L-tone.[1] Although Pierrehumbert (2000, p. 20) does not make this equation when she lays out her indebtedness to Bruce (1977), it is plausible that, despite the difference in functionality of Bruce's (1977) 'sentence accent' and the 'phrasal accent' of Pierrehumbert (1980), the combination of pitch accent and phrase accent was transferred to English nuclear melodies. This ultimately led to boundary tones of an intermediate phrase (L- or H-) in the analysis of Mainstream American English (MAE) known as MAE_ToBI (Beckman & Pierrehumbert, 1986; Beckman et al., 2005; Silverman et al., 1992), whose development from Pierrehumbert (1980) is charted in Ladd (2008, ch. 3).

---

[1] Riad (2014, p. 254) analyses the Central Swedish lexical tone distinction privatively, such that the intonational tones only appear after the lexical tone in the case of Accent 2.
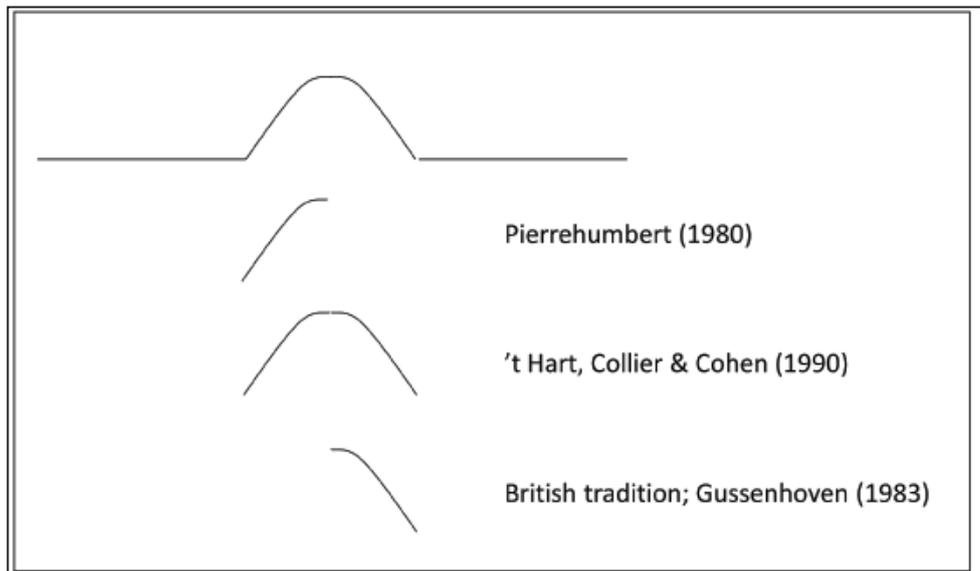
Figure 1: Three analyses of an accent-lending pitch accent.

The second option is to take the entire rise-fall as the pitch accent. An analysis that came close was proposed by 't Hart, Collier, and Cohen (1990), whose model used constantly changing line segments as primitives. Their analysis took both the rise and the fall to be pitch accents ('accent-lending pitch movements') and included a convention whereby a syllable may be marked as accented by more than one accent-lending movement, thus making the accent-lending property of any movements that are added to the first in the same syllable vacuous.[2] Goldsmith (1980) and Leben (1976) assumed a tritonal MH*L, whereby the M was deletable in Goldsmith's analysis and insertable in Leben's, thus taking positions intermediate between the second and third. The third option was unequivocally adopted by Palmer (1922), who set the stage for the term '(High) Fall' for this pitch accent in the British tradition of intonation analysis (Ladd, 1980, ch. 1). This interpretation was taken over in subsequent descriptions of English intonation as well as in the autosegmental analysis of Gussenhoven (1983). Apart from the discussion of the position of M by Leben (1997) and Goldsmith (1980), none of the three options was argued for in a comparison with the other two by any of the authors concerned.

Ideally, a phonological account provides a unique transcription for any expression in some language, while there is a unique expression that will correspond to any legitimate transcription. That is, there are no unusable transcriptions (no overanalysis), and there aren't any expressions for which no transcription is available (no underanalysis; 'expression' here refers to an intonation pattern, abstracted away from its morpho-syntactic content). A further property of a successful analysis is its predictive power. Analytical decisions about the tone structure may imply a particular prosodic constituent structure. For instance, a decision to transcribe a steeply falling pitch movement as resulting from a H* followed by a phrasal L-tone in English brings the prediction with it that falling pitch movements are phrase-final.

The purpose of this contribution is to show that examples of underanalysis, overanalysis, and incorrect boundary prediction can be found in MAE_ToBI. Section 2 restates the grammar of MAE_ToBI, including the conventions governing the phonetic implementation (2.1), and considers what that analysis would look like under an off-ramp view (2.2).

---

[2] 't Hart et al. (1990) replaced the accent-lending falling movement ('A') with a movement that had been used to describe the transition between a high-ending and a low-beginning IP ('B'), thus opting for an on-ramp analysis in the summary of their work.

Section 3 discusses a number of cases of underanalysis in MAE_ToBI, while section 4 does the same for cases of overanalysis. Section 5 identifies a boundary prediction and reports a perception experiment whose results indicate the incorrectness of that prediction. Section 6 reviews empirical evidence presented earlier that bears on the choice between an on-ramp and an off-ramp analysis. Section 7 finally summarizes the off-ramp grammar and discusses the implications of our findings.

2 MAE_ToBI and its off-ramp alternative

2.1. The MAE_ToBI grammar

MAE_ToBI uses four tone paradigms. Addressing them from early to late, there is an optional initial boundary tone of the intonational phrase (IP), five pitch accents to be used for accented syllables, two final boundary tones of the intermediate phrase (ip), and two final boundary tones of the IP. These are listed in (1). In addition, optional downstep applies to any H-tone other than H% (notated !H), provided another H-tone precedes in the IP. In (2), five phonetic implementation conventions applicable to (1) are listed.

(1)   a.   Initial IP-boundary: %H (optional)
         b.   Pitch accents: H*, L*, L+H*, L*+H, H+!H*
         c.   Final ip-boundary: H-, L-
         d.   Final IP-boundary: H%, L%

(2)   a.   The F_0 between adjacent targets is obtained by linear interpolation, except for targets of T-, which are 'spread' between the pitch accent on the left and the boundary on the right.
         b.   H% after H- is upstepped to extra high.
         c.   L% after H- is upstepped to the value of H-.
         d.   H*, trailing H, and H- are optionally downstepped relative to a preceding H.
         e.   The pitch between adjacent H*'s sags.

Convention (2c) has a special position. A phonetic implementation rule will not categorically assimilate a tone to another tone, but rather raise or lower a tone's target such that its identity is detectable in the signal. However, (2c) leaves no trace of L% and thus effectively turns the phonetic implementation into a mechanism for deleting tones. Ignoring this point, we can calculate the number of two-accent IP-contours by multiplying 2 initial boundary conditions (optional %H) by 5 (prenuclear pitch accents) by 5 (nuclear pitch accents) by 2 (phrase tones) by 2 (IP-tones) = 200 contours. To these, we should add the downstepped contours. If there is no initial %H, 64 contours will have a H in both pitch accents ($4 \times 4 \times 2 \times 2$), and an additional 16 will have a single pitch accent with H followed by H- ($1 \times 4 \times 2 \times 2$). When %H is used, only $1 \times 1 \times 1 \times 2 = 2$ will *not* have a following H tone. This puts the total number of downstepped contours at $80 + 98 = 178$, making for a total of 378 two-accent contours.

2.2 An off-ramp alternative

Pierrehumbert's (1980) decision to analyze a rising-falling accent-lending contour as a L+H* pitch accent followed by an extraneous L-tone was referred to as an 'on-ramp analysis' in Gussenhoven (2004: 127). An 'off-ramp' analysis will assume a H*+L pitch accent preceded by an extraneous L-tone. A crucial difference between the MAE_ToBI on-ramp analysis and an off-ramp analysis lies in the number of targets that are needed after the nuclear pitch accent. After a pitch peak, two further targets may occur in English, a low

	MAE_ToBI		MAE_ToBI (overt tones only)	Off-ramp alternative
1	H* H- H%		H* H%	H* H%
2	H* L- H%		H* L-H%	H*L H%
3	H* H- L%		H*	H*
4	H* L- L%		H* L%	H*L L%
5	L* H- H%		L* H-H%	L*H H%
6	L* L- H%		L* H%	L* H%
7	L* H- L%		L* H-	L*H
8	L* L- L%		L*	L*
9				H*L
10				H* L%
11				L* L%
12	L*+H L- L%		L*+H L%	L*H L%

Table 1: Representations of nuclear contours in MAE_ToBI (column 1) with graphic phonetic implementations, after Pierrehumbert 1980 (column 2). Column 3 repeats the representations without tones that have no overt target. Column 4 gives representations in an off-ramp analysis without phrase tones and with optional IP-boundary tones.

target followed by a high target at the IP-end; while after a low valley, there can follow a mid target and high target at the IP-end. To represent these post-peak and post-valley targets, MAE_ToBI provides T- and T%. Most contours, however, have only a single such target. **Table 1** lists the eight MAE_ToBI contours with single-tone H* and L* pitch accents in the first eight rows. It shows two post-nuclear targets for contours 2 and 5, the other six having a single overt target after T* (contours 1, 3, 4, 6, 7, and 8). If we simply leave out tones with abstract targets, we produce the representations in column 3.

Still concentrating on contours 1 to 8, column 5 presents the off-ramp analysis, in which the MAE_ToBI ip-boundary tones have been absorbed as trailing tones in the pitch accent, except for contour 4, which has a trailing L in column 5, but no corresponding L- in column 4. These off-ramp versions amount to a system with four pitch accents (H*, H*L, L*, L*H) and an IP-optional boundary tone. Spelling out the 12 representations by combining these four pitch accents with the three boundary conditions H%, L%, and Ø (no tone) yields four further contours. The representation H*L L% for contour 4 contrasts with contour 9, the 'half-completed fall', and contour 10, the 'High Level-Slump', and 12, the delayed fall, to which we turn in Section 4.4. A discussion of contour 11 appears in Section 3.2.

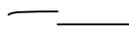## 3 Underanalysis in MAE_ToBI
3.1 Contours ending in mid pitch

Pierrehumbert (1980, p. 88) discussed the contrast between mid-ending (3) (cf. Pierre-humbert's Figure 6.4) and (4), noting that her analysis had a single representation for them. She argues that contour (4) is a 'chanted' version of (3) and that chanted speech is an orthogonal variable not requiring a separate tonal representation. MAE_ToBI notates them as H* !H- L%. Against this view, Hayes and Lahiri (1991) showed that the English vocative chant requires a representation which accounts for the neutralization of vowel quantity contrast in IP-final syllables, causing *Je-en!* and *Ja-ane*! to be prosodi-cally identical. Moreover, !H- crucially requires a syllabic association to a post-accentual

stressed syllable, since its phonetic alignment is with -*nath*- in (3) rather than with either the preceding or following unstressed syllable (Ladd, 1978; Liberman, 1975). Example (3) could be a tentative suggestion (Crystal, 1969, p. 147; Gibbon, 1976, p. 135; Gussenhoven, 1983, p. 40; Uldall, 1961) or be used to chide someone. These effects are quite different from that of (4). That is, the contrast between (3) and (4) represents a genuine case of underanalysis in MAE_ToBI.

(3)   Abernathy!

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx3.

(4)   Abernathy!

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx4.

The off-ramp analysis provides H*L Ø, contour 9, for (3), which contrasts with the rapid final fall, contour (4). By assuming that trailing L has mid-low pitch, while L% is pronounced at fully low pitch, the two L-tones in contour 4 acquire overt tonal targets. Also, the mid-low ending of (3) is explained by the pronunciation of trailing L at a point near the IP-boundary.

  Vocative chants require an additional pitch accent, notated H*+H in Gussenhoven (2004, ch. 15). It is given in **Table 2**, as H*H, where also the falling-rising vocative chant (Gussenhoven, 1983, p. 41, with reference to Pierrehumbert, 1980) and the low falling vocative chant (Gussenhoven, 2004, p. 315) are included, and accounted for by the addition of H% and L%, respectively. After the extension of the off-ramp grammar with this H*H pitch accent, contour 13 takes care of (4). In addition, we generate representations for two further vocative chants.

  There is in fact a third mid-ending contour, for which MAE_ToBI would equally have to use H* !H-L%. Contour 10 is part of class of contours ending in a fall to mid after a high stretch beginning after the accented syllable. We will return to these contours in Section 4.4.

3.2 Scathing intonation

The second case of a missed contrast concerns contour 8, L* L- L%, the 'scathing' contour, as it was called by Alex Monaghan in a now defunct Linguist List message. It is an echo-statement, typically used as a repetition of a listener's earlier utterance, used to express disparagement and disbelief. Gussenhoven (2004, p. 301) claimed that there are two 'scathing intonations'. One remains level from the low-pitch accented syllable onwards, which has the force of *Here we go again!*, a 'routine' meaning identified by Ladd (1978),

	MAE_ToBI		MAE_ToBI (overt tones only)	Off-ramp alternative
13	H* !H- L%	▬▬	H* !H-	H*H
14		▬╱		H*H H%
15		▬▬		H*H L%

Table 2: Representation of the vocative chant in MAE_ToBI (column 2) with graphic phonetic implementations after Pierrehumbert (1980) and representations for the mid-falling, falling-rising, and low vocative chants in the off-ramp analysis.
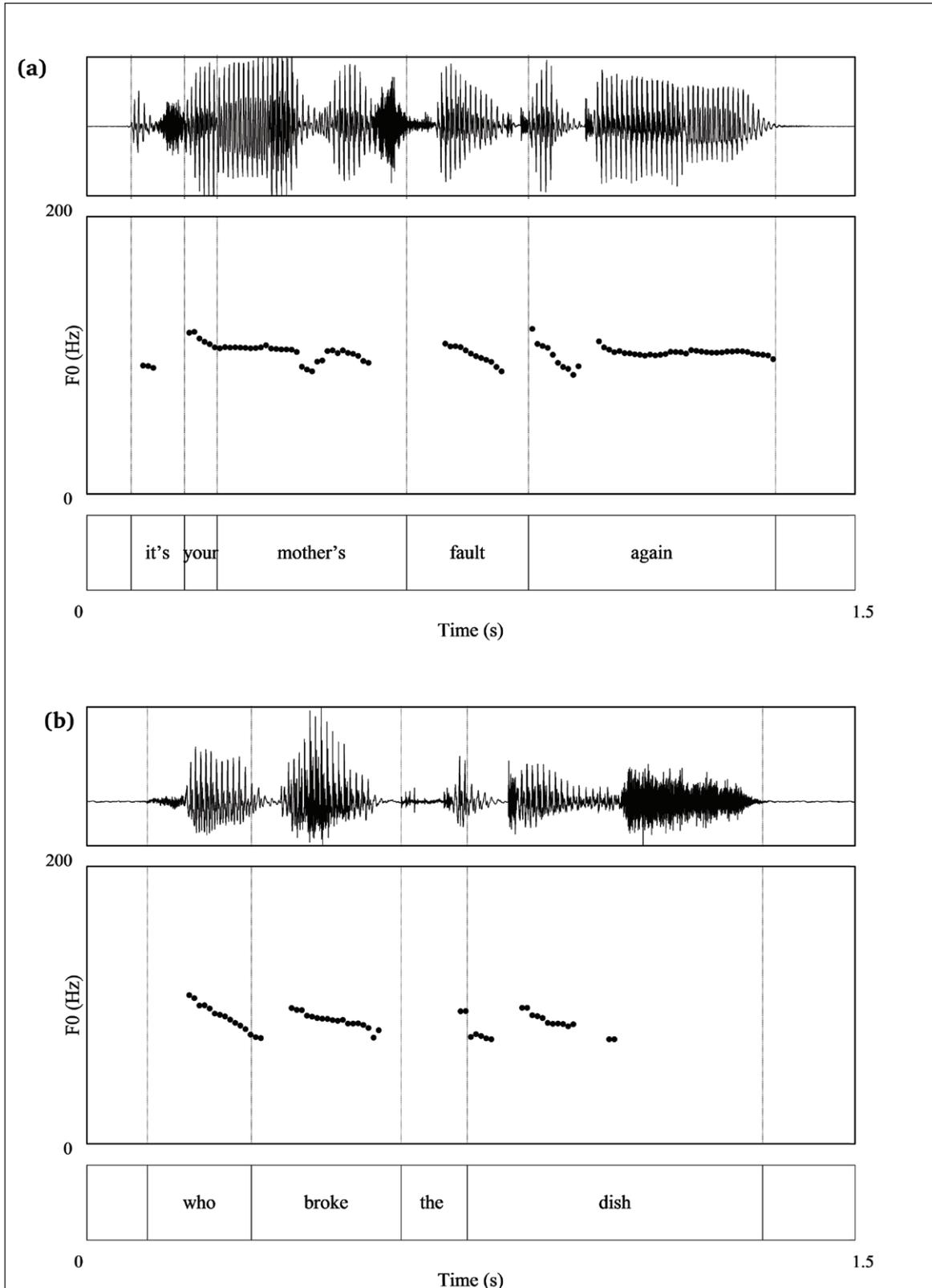
Figure 2: Two 'scathing' contours, a low level contour on *It's your MOTHers fault again* (panel a, male GBE speaker) and the low falling contour on *WHO broke the dish!* (panel b, speaker CG). From Gussenhoven (2004). This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav2a and http://dx.doi.org/10.5334/labphon.30.wav2b.

shown in panel (a) of **Figure 2**.[3] The other contour descends somewhat within a low register. It may express a stronger degree of mockery, as in panel (b), but has other uses too, as in Pierrehumbert (1980, Figure 4.19), where it is used on *damn* after H* on *God* in *God damn it!* The off-ramp analysis transcribes these as L* Ø and L* L%, respectively.

3.3 H+!H*, but no H+L*

The third case of a missed contrast was pointed out to me by Bruce Hayes with reference to Pierrehumbert (1980) (personal communication, July 1991) and concerns the contrast between downstepped !H* and L* after a preceding high syllable. MAE_ToBI provides H + !H* to cover the first case, but since there is no H + L*, it cannot describe the second.[4] Grice (1995) independently treated this distinction, exemplified by her with (5) and (6), pointing out that these contours required the adoption of a generally applicable leading H, which is prefixed to either L* or H*. In (5), the accented syllable *-ma-* is fully low-pitched, due to L*, while that in (6) is mid-pitched, as for a downstepped !H*. Illustrative contours are presented in **Figure 3**. Possibly, the slowly rising pitch towards H% in the contour in panel (a) serves an enhancement of L*. The contrast was included in the analysis of German by Grice, Baumann, and Benzmüller (2005).

(5)   A: The onions are over there, but I can't see any tomatoes.

      B: The toMAtoes haven't arrived yet?

           L*

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx5.

(6)   A: We've got all the stuff now, haven't we?

      B: The toMAtoes haven't arrived yet!

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx6.

Following Grice (1995), the off-ramp analysis assumes a prefix H, here notated in italic font to separate it from the base pitch accent (*H*H*L, *H*L*, etc.). Strikingly, H* is invariably downstepped after the pre-accentual peak (Grice, 1995, p. 202). The generalization that arises from the pronunciation of this pitch accent and of the vocative chant (Section 3.1) is that within pitch accents downstep is obligatory. Under an assumption of 'P(itch) A(ccent)-internal downstep' (Gussenhoven, 2004, p. 301), the inclusion of H* and its leading H in the same pitch accent renders the downstep inevitable, quite as in the case of H*H, the vocative chant. When trigger H and target H* are not contained within a

---

[3] A reviewer pointed out that definitions of 'core' or 'abstract' (Cruttenden, 1997) meanings for intonational morphemes may be problematic and that factors like politeness, social distance, and physical distance need to be considered for a better understanding of intonational meaning. For instance, Bórras-Comes et al. (2015) show that compared to non-chanted vocatives, the vocative chant of Central Catalan is favoured by larger physical distance and speaker superiority. These factors may well also apply to the English vocative chant.

[4] Pierrehumbert (1980) had H + L* instead of MAE_ToBI's H + !H*, with a convention that L* in H + L* was realized as !H*.

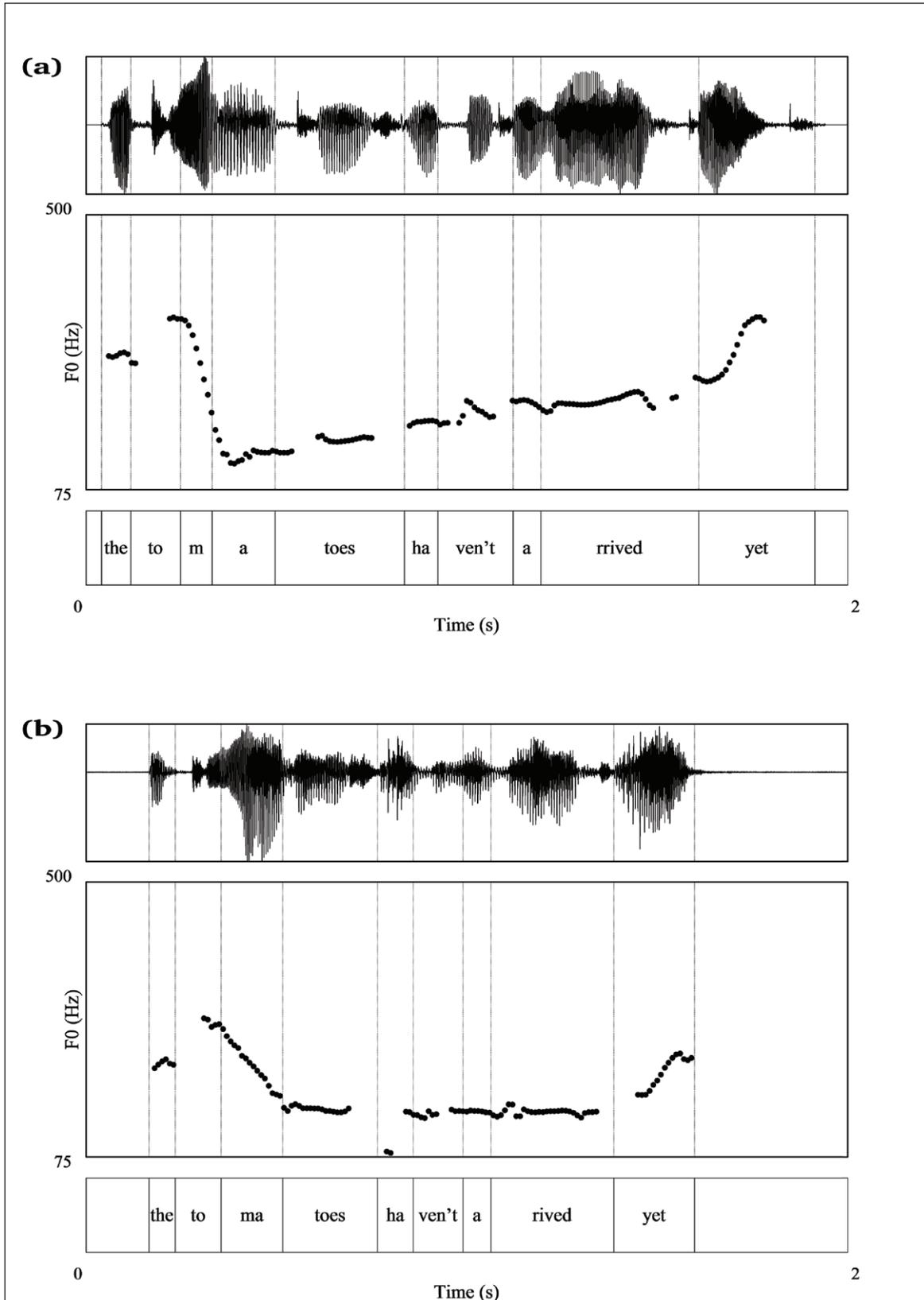**Figure 3:** A L* target (panel a, female GBE speaker) and a downstepped !H* target (panel b, female Mid-Western MAE speaker) with leading H's on *tomatoes* in *The tomatoes haven't arrived yet*. The contour in panel (a) is an echo question, that in panel (b) can be used as a statement. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav3a and http://dx.doi.org/10.5334/labphon.30.wav3b.

pitch accent, downstep is optional, but while any H-tone can be the trigger, only H* can be downstepped. Against the background of these generalizations in the off-ramp analysis, the postulation of downstep of the phrasal tone in the chanted call in MAE_ToBI now looks arbitrary, since in other contexts no contrastive downstep on H- is in evidence. For instance, there has been no demonstration that H* L- H% (high-low-high) is categorically distinct from H* !H- H% (high-mid-high).

3.4 Virtual vs. real leading H

My fourth case has not been discussed before, as far as I am aware. To describe high level pitch between a high and a downstepped high pitch accent, ToBI uses a prenuclear H* which is followed by H + !H*, where the high stretch between the pitch accents is described as an interpolation between H* and leading H. An example of this contour (cf. 't Hart, Collier, & Cohen's [1990] 'flat hat') is shown in **Figure 4**, panel (a). The general descending profile is a common, though not a necessary feature of this contour. The MAE_ToBI analysis implies that there is no transcription available for the same contour with an upstepped high pitch on the syllable before the second accented syllable, as in the contour in panel (b). This contrast seems quite categorical, with a distinct note of liveliness in contour (b) which is absent in contour (a).

To account for the difference between the contours in **Figure 4**, we must assume that the pronunciation of the target of prenuclear H* continues until just before the first tone in the nuclear pitch accent. In fact, contours 3 and 8 already made it clear that tonal targets are continued rightwards if there is no further tonal target in the IP: without any following tones, a string-final H* is realized as high level pitch until end of the IP, while string-final L* in the same position produces low level pitch. Similarly, a trailing tone is continued when string-final, as in contours 7 and 13. This 'continuation' of tones appears to apply generally to any English morpheme-final tone. In addition to the situation before a toneless boundary, there are three inter-morphemic stretches in which this continuation occurs:

(i)   from a boundary tone to a pitch accent;
(ii)   between pitch accents;
(iii) from a pitch accent to a boundary tone.

MAE_ToBI presents the continuation of tonal targets as an anomaly, applicable only to the phrase tone, i.e., the equivalent of context (iii). The most widely discussed case here is that of L- between H* and H%, which forms a 'floodplain', in the terminology of Lickley et al. (2005), but the same is true for mid-level stretches in the MAE_ToBI L* H- H% contour. From the off-ramp perspective, these anomalies disappear as part of the generalization that unspecified inter-morphemic stretches are filled with the tone on the left. Thus, prenuclear H* in (7) continues its pronunciation from *-ron-* onwards, until preparations need to be made for the pronunciation of the downstepped target of !H*. To account for this continued pronunciation, Gussenhoven (2000) introduced the concept of double alignment. Alignment with other phonological constituents quite generally determines the location of tonal targets (cf. McCarthy & Prince, 1993). It is expressed as a coincidence of the edges of two constituents, such as when a prefix is said to align its left edge with the left edge of the word it attaches to. Thus, an initial boundary tone aligns its left edge with the left edge of the IP, a final boundary tone aligns its right edge with the right edge of the IP, a leading H aligns its right edge with the left edge of the following T*, and an associated tone aligns with an edge of the accented syllable rime (cf. Pierrehumbert, 1993).
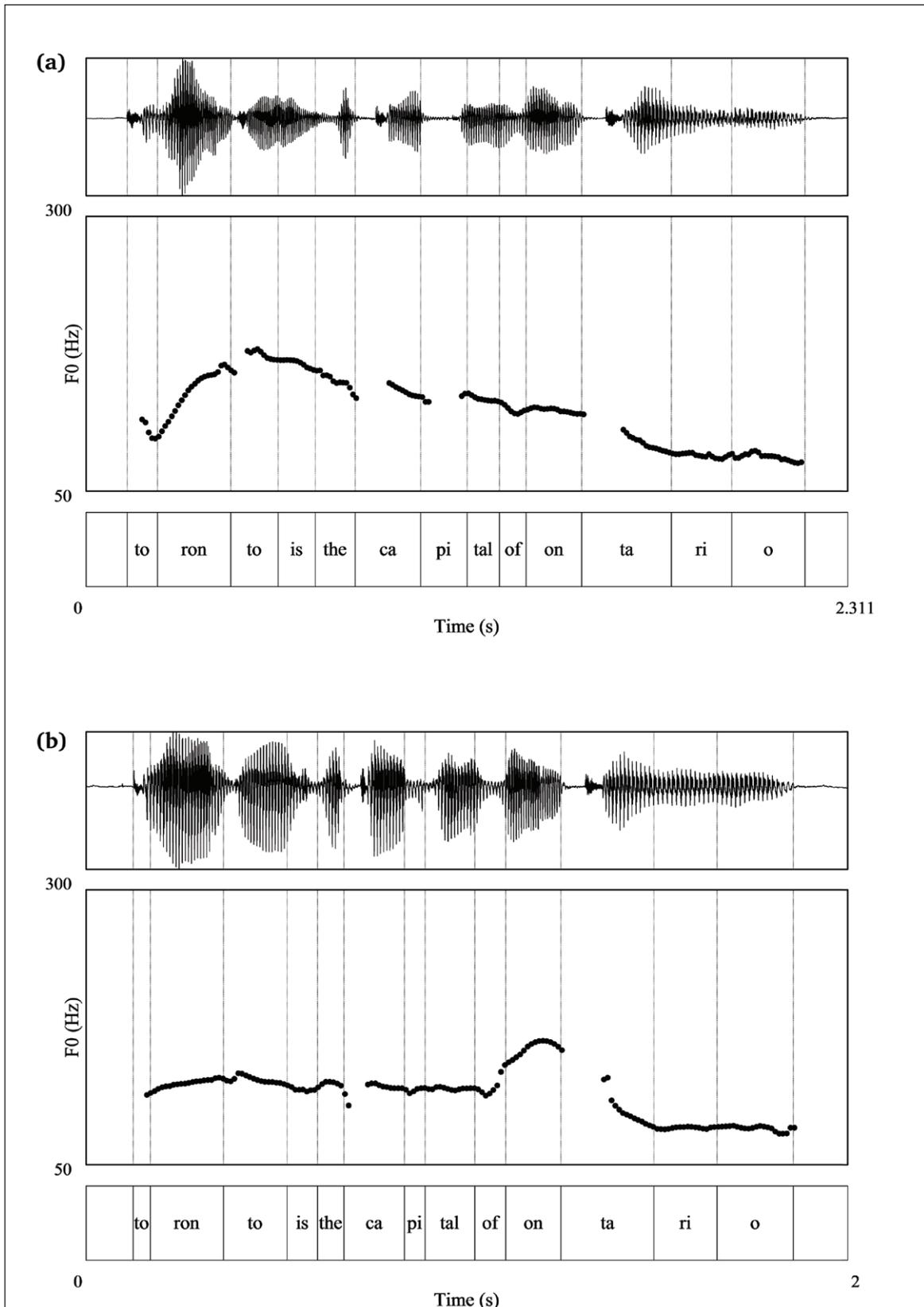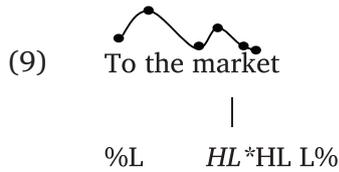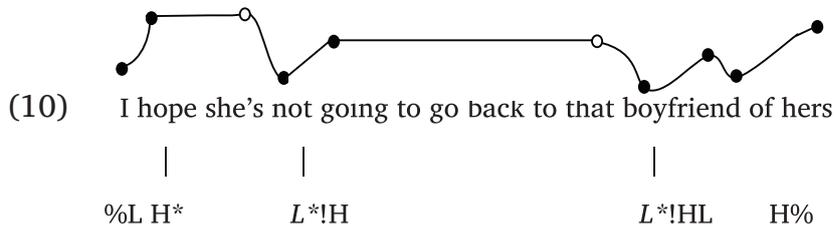
Figure 4: A descending 'flat hat' contour without (panel a) and a 'flat hat' contour with a raised peak on the syllable before the second accented syllable (panel b). Male Can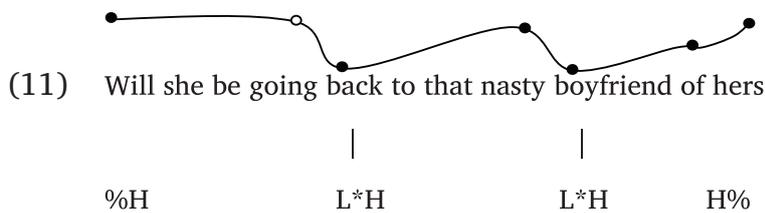adian English speaker. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav4a and http://dx.doi.org/10.5334/labphon.30.wav4b.

Unspecified space between targets is covered by an interpolation between them in MAE_ToBI, following Pierrehumbert (1980). Double alignment means that the left-hand target additionally acquires a right-hand target, since the tone is both left-aligned and right-aligned, the latter being shown as empty bullets (van de Ven & Gussenhoven, 2011). In (8), leading H now defines a contour distinct from (7), one with raised pitch on the syllable immediately before the nuclear accent. A contour like (8) is reported for *Now you're CURving to the RIGHT* in **Figure 1** in Shattuck-Hufnagel et al. (2004), where I interpret the mid target on *CUR-* to be a realization of H* and *the* to be the location of leading H. In their small corpus, 39% of two-peak contours had an intervening peak on an unstressed syllable, many of which are likely to be further examples.

(7)

But Toronto is the capital of Ontario

H*   !H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx7.

(8)

But Toronto is the capital of Ontario

H*   *H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx8.

3.5 Prefix L* and L*+H

Contour 12 in **Table 1** raises two issues in the intonational phonology of English, corresponding to two contour classes which have a low-pitched accented syllable followed by a rising-falling contour, *viz.* 'delayed' contours and contours ending in a 'slump'. The MAE_ToBI representation belongs to the first class. It was characterized as having 'scoop' by Vanderslice & Pierson (1967) with reference to Hawaiian English. For American English, Vanderslice (1972, p. 1053) notes that scoop, which corresponds to Ladd's 'scooped' or 'delayed peak' contours (Ladd, 1980, 2008) and my own [Delay] (Gussenhoven, 1983), 'delays the upward pitch obtrusion associated with an accented syllable'. Semantically, it has been characterized as having an intensifying (O'Connor & Arnold, 1973, p. 78; Tench, 1996, p. 126, among others) or dominating effect (Brazil, 1985, p. 129), or as expressing that the speaker is impressed (Wells, 2006, pp. 218, 221). These scooped or delayed contours can be captured by a prefix L*-tone, to be inserted to the left of H* (Gussenhoven, 2004, p. 307). Prefixal *L**, notated in italic font, associates with the accented syllable, dislodging following H*, whose asterisk is now left out. It may combine with prefix-H. Following our discussion of Obligatory PA-internal downstep, the presence of leading H in the pitch accent implies downstep on the F_0 peak due to underlying H*, located on *market* in (9). That is, no contrast between a downstepped and non-downstepped second peak in (9) is expected (Gussenhoven, 2004, p. 321). In (9), there is low pitch on *To,* high pitch on *the,* late rising pitch on *mar-,* and falling pitch on *-ket.*

(9)   To the market

%L   *HL*HL L%*

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx9.

The question arises then whether the existence of the simplex pitch accent L*H in the off-ramp analysis by the side of a prefix-L* attaching to H* represents a case of overanalysis, i.e., whether L*H is equivalent to L*H. There are two arguments for considering them to be contrasting representations. In L*H, H has the status of a dislodged H*-tone, which retains the properties of H*. This means, first, that it is not treated as the last tone of a pitch accent, which would require it to align with the next pitch accent, like H in monomorphemic L*H, but rather will continue its pronunciation until the next pitch accent, creating high level pitch. Second, downstep targets H*-tones, predicting that L*-prefixed H-tones (i.e., underlying H*-tones), but not trailing H-tones, can be contrastively downstepped. So while L*H has a counterpart L*!H, there should be no L*!H. Example (10) illustrates a prenuclear L*!H, in which the H-tone creates mid level pitch, before two occurrences of L*!HL. This contour is predicted to contrast with a non-downstepped version. In contradistinction to (10), contour (11) has two occurrences of L*H in prenuclear position, predicting that the pitch between *back* and *boy* is a slow rise, and also that the pitch on -*sty* of *nasty* is not contrastively mid or high.
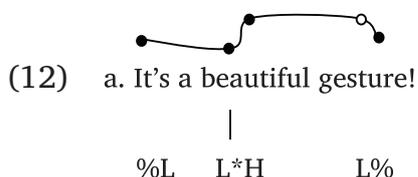
(10)   I hope she's not going to go back to that boyfriend of hers

%L H* L*!H L*!HL   H%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx10.

(11)   Will she be going back to that nasty boyfriend of hers

%H   L*H   L*H   H%

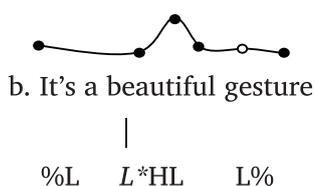This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx11.

An empirical argument may be based on meaning. An eye-tracking study in fact suggested that L*HL is associated with newness, just like H*L, while L*H is associated with givenness (Chen et al., 2007). Yet, the above claims evidently require more empirical research before it can be decided whether the off-ramp analysis is here running into a case of overanalysis or whether we here have another case of underanalysis in MAE_ToBI.

   The second class of rise-fall contours end in a 'slump', a truncated type of final fall, which is characteristic of Northern British English contours, variants of which

are surveyed in Cruttenden (1997, ch. 5). Nolan and Grabe (1997) pointed out that Pierrehumbert's convention of using H- L% to mean mid pitch (MAE_ToBI's !H-L%, see convention (2c)) makes it impossible to use H- L% to describe the slump. This type of contour, to be sure, has not been included in descriptions of MAE, and MAE_ToBI cannot be criticized for failing to provide a representation for the Rise-Level-Slump of Northern Irish English on which Nolan and Grabe (1997) base their case. Indeed, Mayo et al. (1997) abandon clause (2c) so as to use H-L% for the 'slump' in Glaswegian English. Yet, it may be argued that a phonology of a complex intonation system like that for English may generate contours that do not occur in all varieties. As noted by Pierrehumbert (2000, p. 27), 'nothing like the full set generated by [MAE_ToBI] has ever been documented'. A large proportion of a grammar's legitimate contours may never be encountered in anyone's lifetime, any more than will be the majority of morphosyntactic structures generated by some simple mini-grammar of English. Such non-occurrence may well be interpreted as absence from the grammar, provided it takes the form of a stochastic algorithm (Dainora, 2006). Either way, varieties are likely to differ in the frequency with which certain structures are used for certain pragmatic functions (Grabe & Post, 2004; Ritchart & Arvaniti, 2014), while there will also be cases of absolute non-use (see also Cole & Shattuck-Hufnagel, 2016). Wells (2006, p. 245 fn 8), for instance, notes that the second edition of O'Connor and Arnold (1973) was the first British English course book that awarded the Fall-rise (H*L H%) full treatment as a neutral polar question contour. Earlier, it had not been reported for questions in GBE, and in MAE it is apparently (still?) not used in that function. Or again, I have found it hard to elicit H* H% contours from GBE speakers, who tend to produce L*H H% instead, and the speaker of the contour in **Figure 3**, panel (b), associated it with GBE, while having no problem producing it. Be this as it may, the off-ramp analysis readily provides representations for slumped contours by providing L% after pitch accents like H* and L*H, as shown in (12a), which contrasts with (12b) of the standard languages. The off-ramp analysis offers contour 10 in **Table 1** for a high-beginning equivalent of (12a), a contour which has not been reported even for northern British English. Arguably, therefore, we are here dealing with a systematic case of overgeneration. However, there is a difference between this case and the cases of overanalysis in MAE_ToBI to be discussed in Section 4. The MAE_ToBI cases concern putative contrasts that one would not expect to turn up in any variety of English, while contour 10 in **Table 1**, being clearly distinct from other contours, might.

(12)   a. It's a beautiful gesture!

%L   L*H   L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx12a.

b. It's a beautiful gesture

%L   *L*HL*   L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx12b.

4 Overanalysis in MAE_ToBI
4.1 Prenuclear L*+H

Overanalysis may arise from sequences of H-tones, one or both of which are unstarred. Some of these are given in (13), where the transcriptions to the right of the arrow would not appear to describe a different contour from that on the left.

(13)   Ambiguity of analysis I
            a. L*+H   H+!H*   ⇔ L* H+!H*
            b. L*+H   H*   ⇔ L*   H*
            c. L*+H H- H%   ⇔ L* H- H%

Figure 5 presents F_0 contours on *toRONto is the capital of onTArio* for cases (13a) in panels (a) and (b) and for cases (13b) in panels (c) and (d). The contours in panels (a) and (c) might at first sight be transcribed as on the left of the arrow, while those in panels (b) and (d), in which the mid sections have been resynthesized, might be expected to be transcribed with the symbols to the right of the arrow. However, the original and the resynthesized contours are not easily interpretable as representing different intonations. Chapter 5 in Pierrehumbert (1980) discusses these complications of the analysis and characterizes the contours in panels (b) and (d) as 'impossible'. Like case (11c), these ambiguities are an inevitable consequence of her analysis.

   We begin by observing that the absence of a L+H* pitch accent in the off-ramp analysis forces it to interpret inter-accentual slow falls as instances of H*L and inter-accentual slow rises as instances of L*H. This is shown graphically in (14) for prenuclear H*L. The low target before the nuclear F_0 peak is described by aligning the trailing L rightwards, thus moving its target to a point just before the target of the next tone. The space between the targets of H* and L is filled by an interpolation. In MAE_ToBI, which lacks H*+L but has L+H*, the slow fall is an interpolation between a prenuclear H* and a leading L of the next pitch accent.

(14)

H*   L+H*   L-   *MAE_ToBI*

H*L   H*L   *Off-ramp*
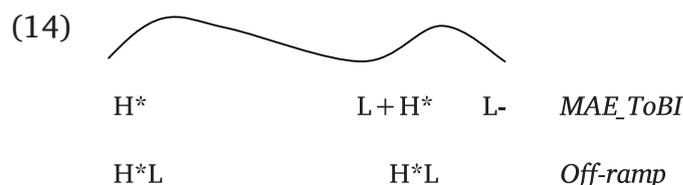
The off-ramp view thus suggests two things. One is that trailing tones of *prenuclear* pitch accents are aligned rightmost, i.e., with the left edge of the first tone of the next pitch accent. The trailing L of the *nuclear* pitch accent is aligned leftmost, i.e., defines a rapid fall. The second implication is that linear interpolations are restricted to tones *within the same pitch accent.* This intra-morphemic linear interpolation between tones thus contrasts with the inter-morphemic continuation of tonal targets over stretches of speech between pitch accents and boundary tones (see Section 3.4). By having stretchable interpolations between tones in pre-nuclear pitch accents, we guarantee that H* and H*L will be distinct in all positions and all contexts, as will L* and L*H.[5] The slow rises in panels (a) and (b) of

---

[5] Until Gussenhoven et al. (1999–2003), tonally unspecified stretches of speech between pitch accents were filled by interpolations instead of double alignment. Grice et al. (2009) assumed linear interpolation in their critique of my off-ramp analyses of (2004, 2005), which however include 'continuation'. The term 'spreading' was avoided because of its implication of tonal association. The relevance of interpolation shapes for the perception of the alignment of pitch targets was demonstrated by Barnes, Veilleux, Brugos, and Shattuck-Hufnagel (2010, 2012), who introduced a Tonal Center of Gravity (TCG) to predict perceived alignment. This paper recognizes this view of pitch movements, but it is not directly relevant for the discussion of the more coarse-grained alignments of tones in this article.
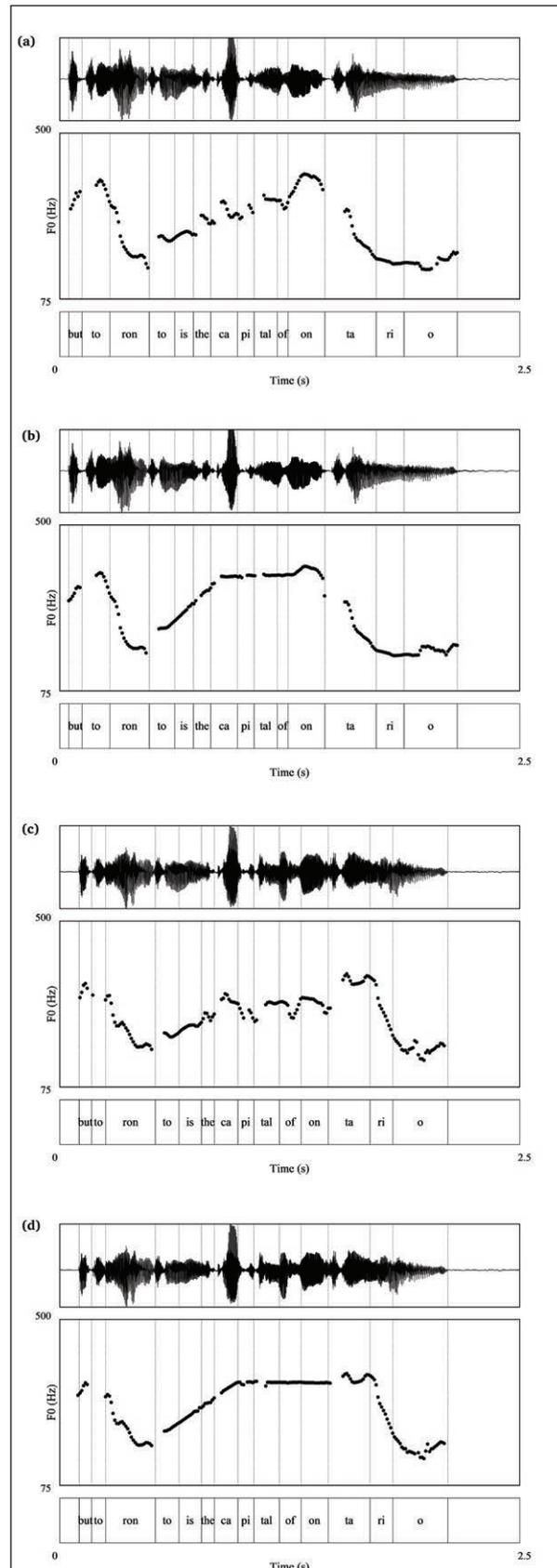
Figure 5: A pre-nuclear L*-beginning rise in *But ToRONto is the capital of onTArio* and a resynthesized version with an accelerated early part of the rising movement before H+!H* (panels a, b) and before (non-downstepped) H* (panels c, d). Female GBE speaker. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav5a, http://dx.doi.org/10.5334/labphon.30.wav5b, http://dx.doi.org/10.5334/labphon.30.wav5c, and http://dx.doi.org/10.5334/labphon.30.wav5d

Figure 5, therefore, are described by a pre-nuclear L*H followed by *H*H*, whereby contour (b) is a less felicitous realization of that representation. Likewise, those in panels (c) and (d) are described by L*H before H*. The overgeneration in (11c) similarly disappears, both of them corresponding to L*H H%.

Like MAE_ToBI, which has an optional %H, the off-ramp analysis offers two choices at the initial boundary of the IP, notated as %L for low-to-mid pitch and %H for high pitch. The contours in **Figure 5** have %H. **Figure 6** shows rising prenuclear stretches after %L. Since MAE_ToBI uses leading H in H+!H* merely to provide a high target in the syllable before !H*, as we saw in Section 4.1 above, it is unclear which of the four available transcriptions (L* H*, L*+H H*, L* H+!H*, or L+H H+!H*) describes which contour in **Figure 6**.

4.2 L*+H vs. H*

In (15), a source of ambiguity is given which has been widely commented on (Arvaniti, 2016; Gussenhoven, 2004, p. 319; Ladd, 2008, p. 96, fn 3; Pitrelli et al., 1994), in particular for occurrences on the first syllable of the IP, case (15a). In (15), the accolade represents the initial IP boundary.

(15) Ambiguity of analysis II
    a. { … L+H*   ⇔ { … H*
    b. { H* … L+H* ⇔ { H*… H* (assuming 'sagging')
    c. { L+H*   ⇔ { H*

In cases (15a, b), an unaccented syllable precedes a H*-bearing syllable with leading L. For (15a), the issue is whether the initial unaccented syllables are low enough to warrant the choice of the leading L, in a situation in which low or mid pitch is predicted anyway, given the absence of an initial %H. If the prediction for L+H* is that of a later peak, as suggested by Steedman (1991), the analysis would imply a three-way peak timing contrast: early peak (H*), later peak (L+H*), and very late or delayed peak (L*+H). This is not, however, a claim that has been explicitly made, as far as I know. Steedman (1991, p. 273) consistently uses H* in combination with L-L% and L+H* in combination with L- H%, noting that the H* is somewhat later in the second type of contour than in the first, thus interpreting the difference as allophonic. The examples in the MAE_ToBI manual (Beckman & Ayers, 1994) suggest that for L+H* there is low pitch preceding the peak *in addition to* a high pitched peak, resulting in a wider rising flank than for plain H*. In panels (a) and (b) of **Figure 7**, realizations of *MariANNa won it* are given (from Beckman & Ayers, 1994; creak occurs on [nɪt]). Under this interpretation, the question arises how pronunciations of L* H-H% are to be transcribed that similarly differ in pitch range. The pair of contours in panels (c) and (d) are arguably analyzable as %H L* H-H%, for the narrower range one in panel (c), and as %H L*+H H- H%, for the wider-range pronunciation in panel (d), thus providing a use for the putative contrast in (15a). This approach would however leave yet further pitch range differences unaccounted for, like two heights for the beginning of the fall in H+!H* L-L%. Understandably, no such more general representation of pitch range variation has been included in MAE_ToBI, which views pitch range differences other than downstep as orthogonal to the symbolic transcription (cf. Bolinger, 1951; Ladd, 2008, p. 36, sec. 5.2). The putative contrast in (15a), therefore, is an anomalous feature of the analysis, if the interpretation is in terms of pitch range.
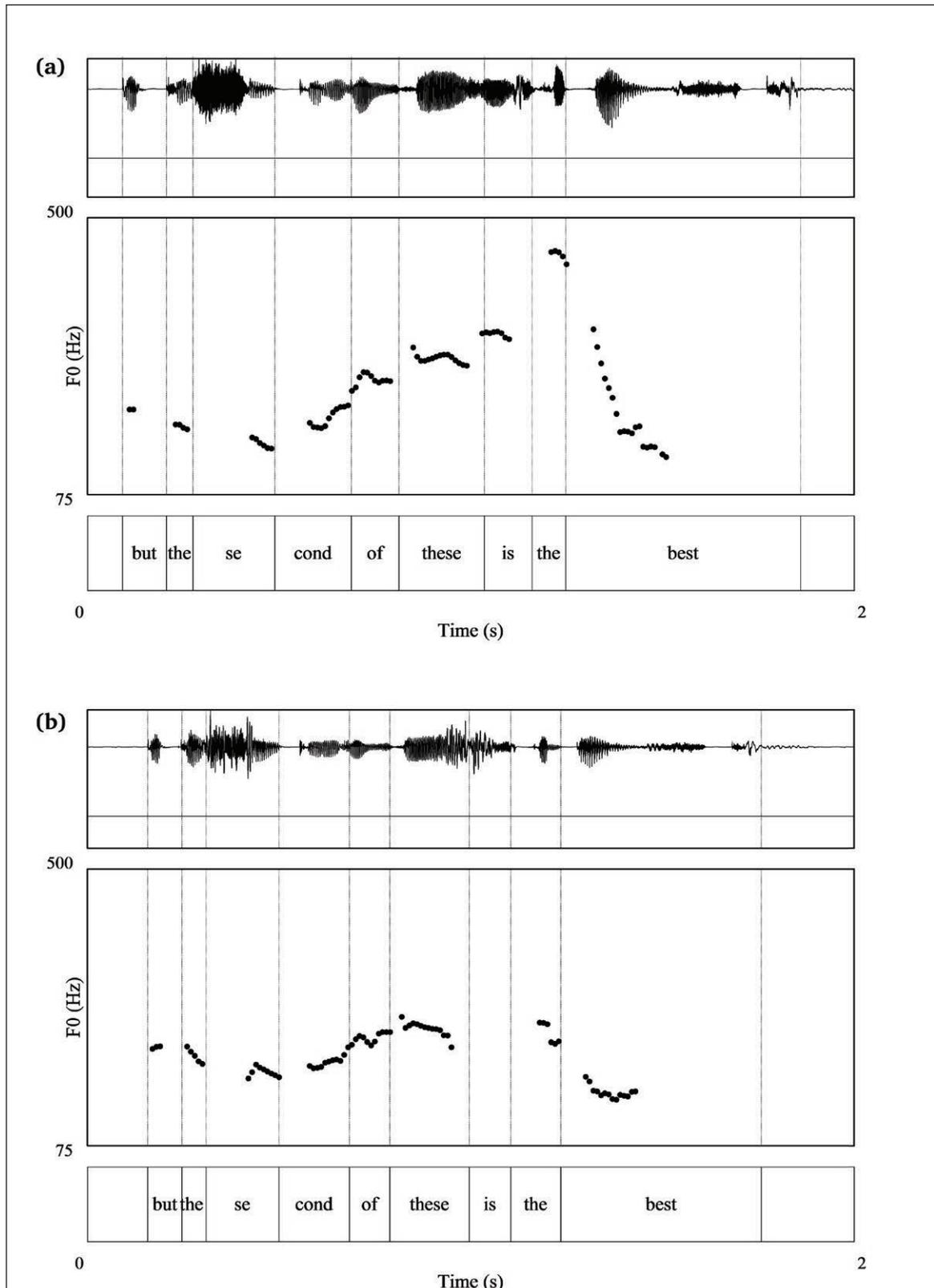
Figure 6: Two slow rises from the pre-nuclear accent in *But the SECond of these is the BEST* to a nuclear !H* without leading H (panel a) and *H*H* (i.e., with leading H, panel b). Female speaker of GBE. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav6a and http://dx.doi.org/10.5334/labphon.30.wav6b.
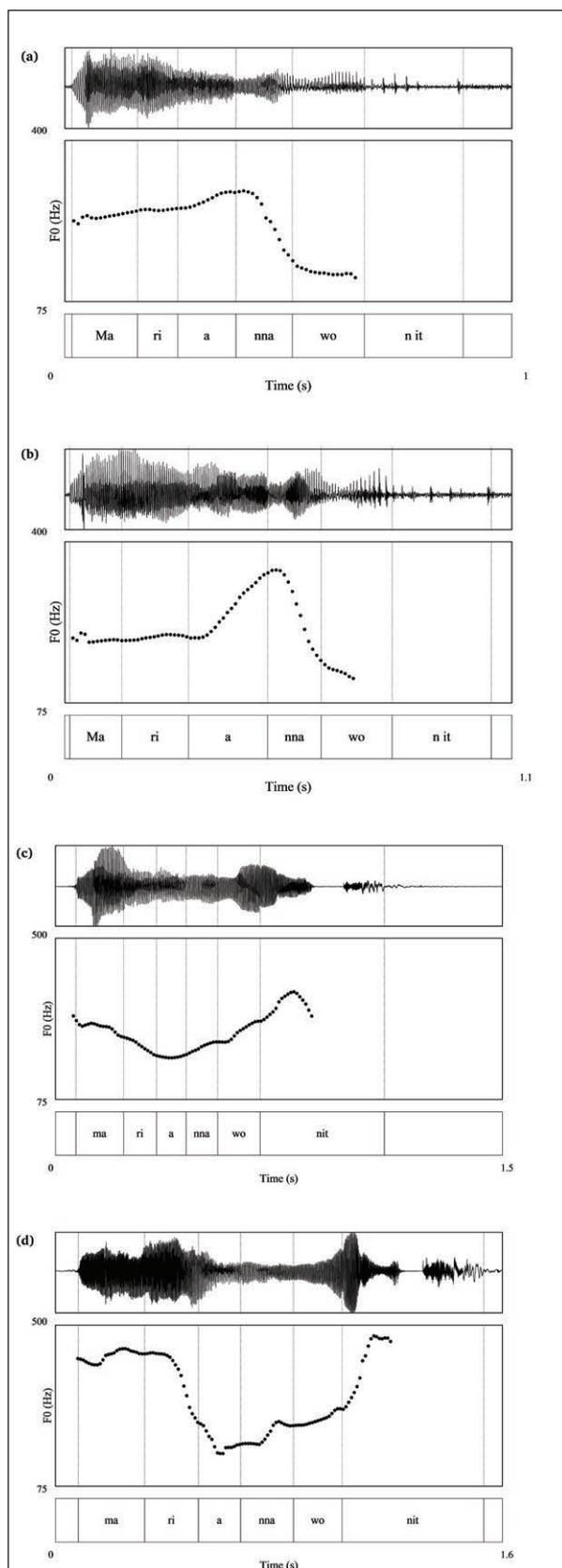
Figure 7: *Marianna won it* with H* L-L% (panel a) and L+H* L-L% (panel b) (female MAE speaker, from Beckman & Ayers, 1994) and *Manianna won it?* with neutral (panel c) and wide pitch range (panel d) pronunciations of %H L* H- H%. Female GBE speaker. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav7a, http://dx.doi.org/10.5334/labphon.30.wav7b, http://dx.doi.org/10.5334/labphon.30.wav7c, and http://dx.doi.org/10.5334/labphon.30.wav7d.

The evaluation of case (15b) depends on the assumptions made for the shape of the interpolation between H*-tones in the contour to the left of the arrow and the realization of the leading L-tone. For the first aspect, Pierrehumbert (1980) argued for a sagging transition, instead of a level interpolation, as would be predicted by double alignment assumed in the off-ramp analysis. The second aspect concerns realization of the leading L of L+H* in the left-hand transcription. If sagging is assumed and the realization of L in L+H* is low-pitched, the prediction is that contour (c) in **Figure 8** is a realization of H* L+H*, while contour (b) is the realization of H* H*. The two contours do not, however, appear to represent different intonations, while both are distinct from contour (a).

The need for a deep valley for the nuclear L+H* after a prenuclear H* was questioned by Ladd and Schepman (2003), who showed that its depth varied with the distance between the H*-tones. While arguing on the basis of the location of the low target that it in fact derives from a L-tone, they point out that there is no contrast between different depths and thus no contrast between contours (b) and (c), and thus no contrast between L+H* and H* under an assumption of sagging. More realistically, targets of leading tones are implemented by gradient realization rules creating undershoot in Grice (1995, pp. 226–228), in the spirit of Chen & Xu's (2006) weak targets, whose realization has less priority than a target of T*, say. Abandoning the requirement of a low realization of leading L as well as the convention of sagging interpolations would enable MAE_ToBI to correctly describe the difference between contours (b) and (c) on the one hand and contour (a) on the other. Without L+H* (or LH*, as I would have notated it), the off-ramp analysis

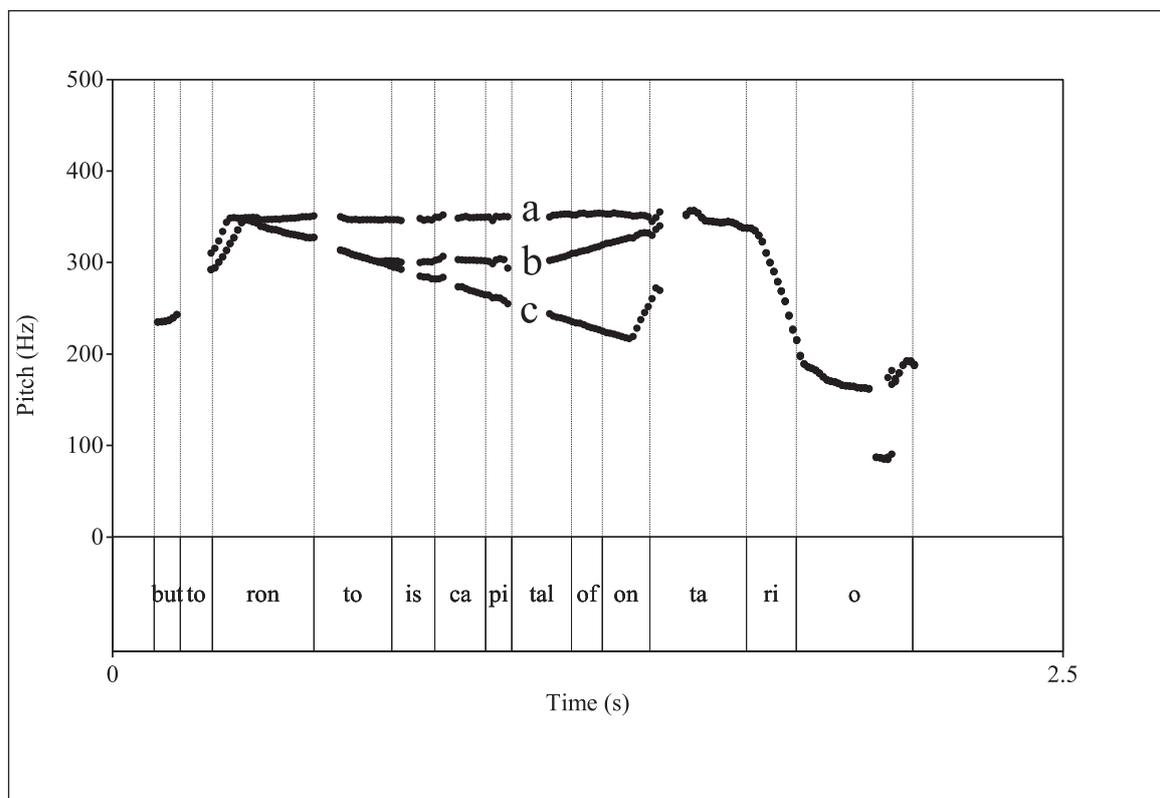**Figure 8:** *ToRONto is the capital of onTArio* with MAE_ToBI H* . . . H* and level pitch (top), H* … H* and sagging pitch (middle), and H* . . . L+H* (bottom), resynthesized versions from a source utterance by a female GBE speaker. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav8a, http://dx.doi.org/10.5334/labphon.30.wav8b, and http://dx.doi.org/10.5334/labphon.30.wav8c.

cannot run into this ambiguity between H* H* and H* L+H*. If the pitch is high level, we have a case of H* H*, while a slow fall is described by H*L H*.

Neither can there be any ambiguity between L+H* and H* in the case of an IP-initial accented syllable in the off-ramp analysis (15c). With only H* available as a transcription (abstracting away from the option of a trailing L and downstep on H*) and with %L and %H as initial boundary tones, there are two ways in which preceding pitch may contrast, mid/low pitched (%L) or high pitched (H%), phonetically realized in the onset and early section of the rime. Compare this with the four transcriptions that are available in MAE_ToBI, H*, L+H*, %H H*, and %H L+H*. If unaccented syllables occur before the pitch accent (15a), we may include a leading H before nuclear T* in the off-ramp analysis and before H* in MAE_ToBI. Four transcriptions are now produced in the off-ramp analysis and six in MAE_ToBI. Off-ramp leading H is pronounced higher than a preceding H, including %H, while following H* is obligatorily downstepped (see Section 3.2). For the first three syllables of *The tomatoes*, the four off-ramp options are therefore %L H* or low-low-high, %L *H*H* or low, extra high, downstepped high, %H H* or high, high, high, and %H *H*H* or high, extra high, and downstepped high. MAE_ToBI's six patterns have not explicitly been described.

5 An incorrect boundary prediction
5.1 Boundaries in MAE_ToBI

MAE_ToBI specifies two prosodic boundaries. First, a T- without a following T% indicates the end of an intermediate phrase (ip), while any T-T% combination additionally indicates the end of an intonational phrase (IP).[6] As observed by Ladd (2008, p. 107), MAE_ToBI would appear to predict boundaries where there are none. As a result of the absence of any falling (H*+L or H+L*) pitch accents, a sharpish accent-lending fall minimally predicts an ip-boundary, since that fall can only be described by H* followed by an ip-boundary tone, L-. The incorrectness of this implication is suggested by a contour type that is not often discussed in the literature on English (but see Cruttenden, 1997, pp. 59, 76; Gussenhoven, 1983, p. 35; Ladd, 2008, p. 107, where his (3) can be interpreted in this way), although it figures prominently in the description of Dutch, which has a similar intonation system to English (Collier & 't Hart 1980; 't Hart et al., 1990, p. 116). Panel (a) in **Figure 9** gives the F_0 and speech waveform of an English example. It contrasts with the contour in panel (b), which would appear to have an IP-boundary after *finance committee* (Gussenhoven, 2004, p. 305; see also panels (c) and (d)). Section 4 reports an experiment that was designed to decide whether a medial boundary exists in contour (a).

5.2 A perception experiment

Adverbs like *honestly* and *oddly* can modify adjectives, predicates, and clauses. Only in the third case are they obligatorily separated from the clause by an intonational boundary. This is illustrated in (16a), which minimally contrasts with (16b), where *honestly* modifies a predicate.

---

[6] MAE_ToBI has two ways of indicating prosodic boundaries. In addition to the tonal one, there are the 'break indices', whereby digits 0 and 1 indicate a clitic group boundary and a word boundary, respectively, and 3 and 4 an ip and an IP boundary, respectively. Digit 2 is there for cases where the tonal transcription implies boundaries that are not perceived to be as strong or as weak as the transcriber's tonal transcription is felt to imply. The distinction between 0 and 1 falls outside our interest. The information about perceived outliers in pause durations, the only non-redundant information in the break indices, is not relevant to our research question.
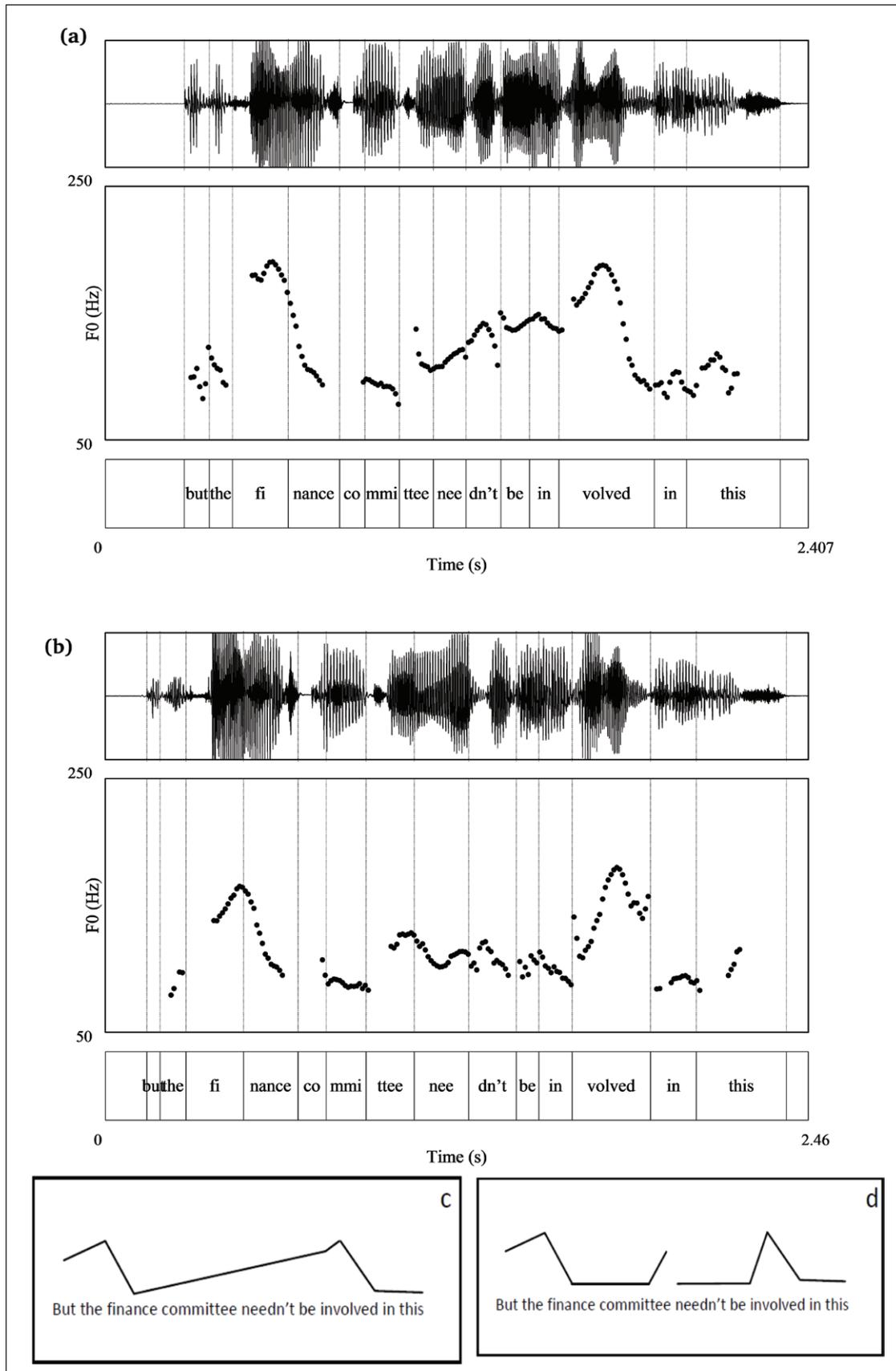
Figure 9: A sharp pre-nuclear fall on *fi-* in *But the FINANCE committee needn't be INVOLVED in this* (panel a) and a contour with an IP-boundary after *finance committee* (panel b). Stylized versions are given in panels (c) and (d). Male GBE speaker. This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav9a and http://dx.doi.org/10.5334/labphon.30.wav9b.

(16)   a. {She treated him}{honestly}'I am honest when telling you that she treated him'
        b. {She treated him honestly} 'The way she treated him was honest'

In order to find evidence for the assumption that the interpretation of sentence-final English adverbs depends on the presence of an intonational boundary before the adverb, more specifically that contour (a) of **Figure 9** does not have an internal intonational boundary, a semantic judgement task was used in which native speakers of English identified one of two meanings of string-identical sentences of the kind illustrated in (16) which had been provided with a number of artificial F_0 contours.

5.2.1 Method
Four minimal sentence pairs with string-ambiguous adverbials were composed (17).

(17)   a. She TREATED the poor man(,) HONESTLY
        b. I THOUGHT she responded(,) ODDLY
        c. He NEVER acted(,) STRANGELY
        d. He DEALT with the woman(,) HONESTLY

The eight sentences were recorded by a female native speaker of MAE in her thirties from Portland, Oregon. By judiciously cutting and pasting sections in the stretch of the waveform before the adverbial, one durational hybrid of each pair of utterances was created, using the software Praat (Boersma & Weenink, 1992–2009). The single-IP versions were used as the source utterance in the case of (17a, c) and the split-IP ones in the case of (17b, d). Appendix C gives the durations of the sections in the original speech files for two sentences with and without boundary whose averaged durations were created in the hybridized source files. By using these as source utterances for F_0 manipulation, we neutralized the effect of any durational marking of the IP-boundary in the original recordings. With the help of the resynthesis program in Praat (Boersma & Weenink, 1992–2009), we then superimposed 12 declining F_0 contours on each of these four sound files, with F_0 values which are representative of the speaker's original utterances (see **Table 3** for these values; unmarked turning points have the same values as equivalent points with F_0-labels). The twelve contours come in two sets of six, as shown in the six cells of **Table 3**. In order to increase the variation in the stimuli, one set had a low-pitched syllable before the first pitch accent (*She, I, He, He*), while these syllables had high pitch in the other set, as indicated by the interrupted sections, phonologically equivalent to initial %H. The
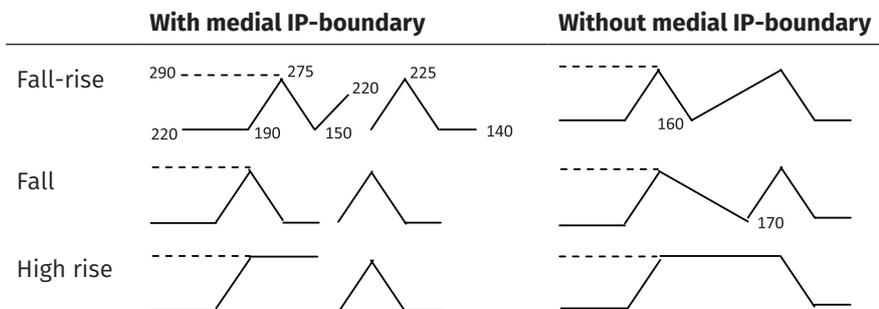

Table 3: Schematic representations of double and single IPs for three two-accent contours with 'Fall-rise', 'Fall', and 'High rise' pitch accents for the first accent and falling pitch accents on the second accent, with F_0 values (Hz) for turning points as used in the artificial contours. The interrupted line indicates versions of the contour with initial %H.

crucial comparison is that between the contours in the two cells of the row labelled 'Fall-rise', which reproduce the contrast in **Figure 9**. As a baseline, we included a contour with an accent-lending fall which *does* signal an intonational boundary in other descriptions of English, as shown in the row labeled 'Fall'. The pitch after the first F_0 peak continues low; its counterpart without an intonational boundary is taken to have a slow fall between the accent peaks. As a further control, the contours given under 'High rise' were included. Here, the fall just before the rise towards the second peak is also taken to predict an intonational boundary. The counterpart without the boundary has high level pitch between the accent peaks. It is stressed that the F_0 manipulations were applied to only four sound-files, one for each sentence, and that any effects are therefore based on F_0 differences only. The interrupted contour sections correspond to the implied IP-boundary.

5.2.2. Procedure

Contours were exhaustively paired within each set of six contours for each of the four source files, excluding pairings of identical contours. This gave two sets of 30 pairs, one with low and one with high beginnings. In order to avoid an unmanageably large set of stimuli, which would arise if we had included 30 (pairings) × 2 (sets) × 4 (sentences) = 240 stimulus pairs, we composed two sets of 30 stimuli, one with initial low F_0 selected from sentences (17b, d) and one with initial high F_0 selected from sentences (17a, c) (see Appendix A). The inclusion of all four source files was intended to avoid fatigue and boredom among the participants. Two test versions were prepared with counterbalanced orders of these 60 stimulus pairs, augmented with four filler pairs inserted at the beginning. Moreover, the members of the stimulus pairs occurred in reversed order in the two test versions.[7]

Seventeen native speakers of American English, approximately equally divided over male and female genders, participated in this semantic identification task. Fifteen participants were recruited from the student population of the Linguistics Department of UC Berkeley, while two were staff members in similar departments in the UK and the Netherlands. Each stimulus pair was presented once, with a latency of 800 ms after a warning signal. The interval between the members of each pair was 800 ms, while 5 seconds elapsed between each pair and the warning signal for the next pair. The participants, 8 of whom did one test version and 9 the other, were asked to identify which of the two members in each pair corresponded best with the interpretation of the sentence-final adverb as a predicate modifier (Version A) or a sentence modifier (Version B; see Appendix B for these instructions). They gave their judgements on a 3-point scale, labelled '1' (for the first member), '0' (for no preference) and '2' (for the second member).

5.2.3. Results

The 1, 0, 2 score values were converted to −1, 0, +1 (version A) and +1, 0, −1 (version B), respectively, so that a higher score represents a higher degree of predicate adverb interpretation of the adverb. A RM Anova on the scores pooled over source files was performed with Initial Boundary Tone, Medial Boundary, and First Pitch Accent as factors. It only showed significant main effects for Medial Boundary ($F_{2,16} = 424,254$; $p < 0.0001$) and First Pitch Accent ($F_{1.621,16} = 134,797$; $p < 0.0001$; Huynh-Feldt corrected). Since there was no effect of the F_0 of the contour beginning, scores were averaged over low-pitched and high-pitched initial syllables and displayed in **Figure 10**. Post-hoc pairwise

---

[7] A typing error caused one of the 30 pairs to be missing, for which aggregate scores were imputed, and another to appear twice, for which scores were averaged. For details, see Appendix A.
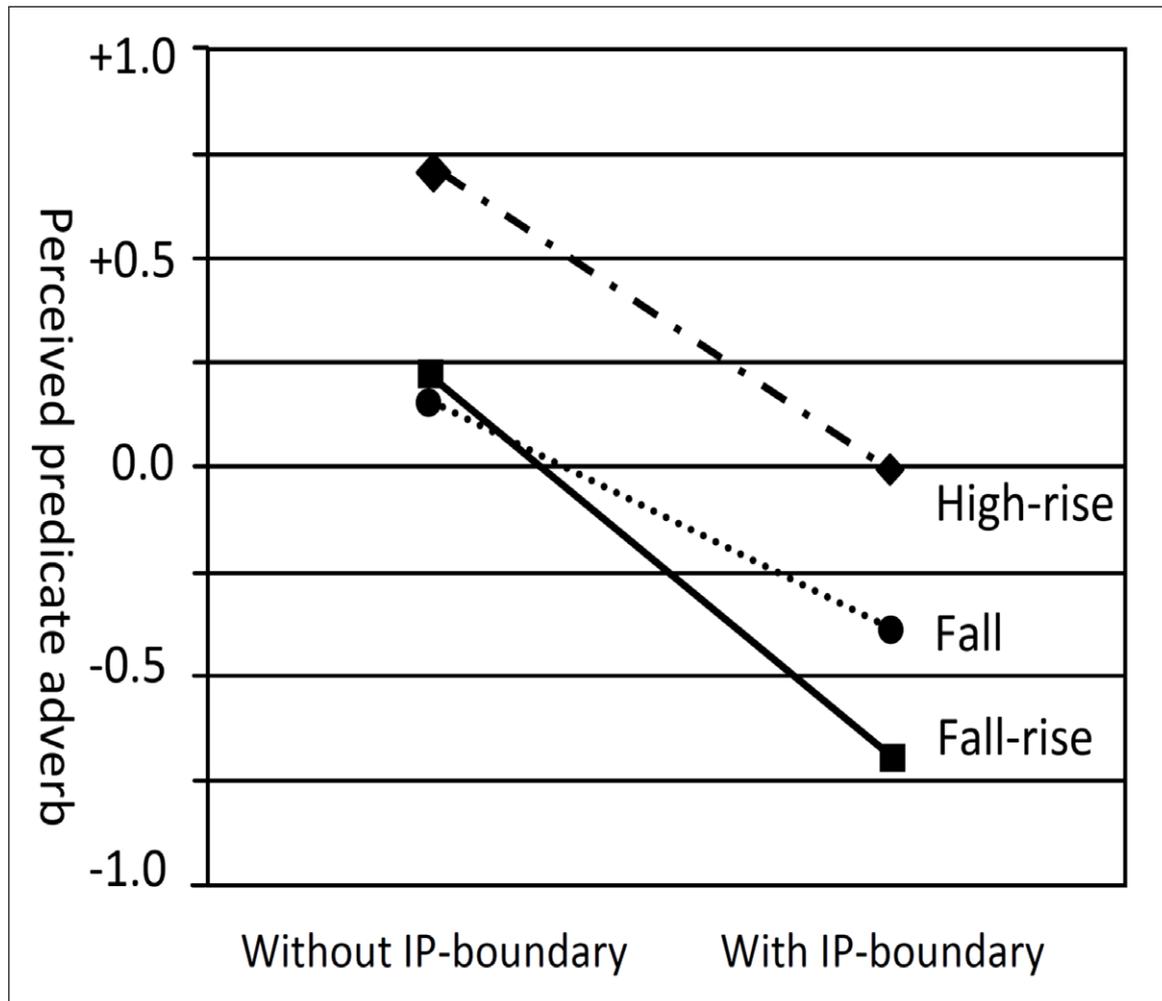
Figure 10: Perceived scores for predicate adverb, aggregated over two test versions and high and low beginning stimuli for three contours with (right) and without (left) a medial IP- boundary.

comparisons show that the High-rise pitch accent attracted significantly higher scores for the interpretation as a predicate adverb than both the Fall ($p < 0.01$) and the Fall-rise ($p < 0.001$).

5.2.4. Discussion

The results confirm the interpretation of the contours in the second column in **Table 3** as having no medial intonational boundary. Crucially, the Fall-rise contour in the column 'With medial IP-boundary' is interpreted to differ from the Fall-rise contour in the second in the same way as do the single-IP and two-IP versions of the Fall and High-rise contours. There is therefore no motivation for a transcription of the first Fall-rise contour with L- after the first pitch accent.

Three additional points are made. First, the finding that the High-rise contours are more readily interpreted as lacking an intonational boundary than either the Fall-rise or Fall contours is attributed to the low phonetic salience of the F_0 features separating the two pitch accents. In the contour without medial boundary, the pitch continues level from one peak to the next, *modulo* the declination, and for the contour with the medial IP-boundary, it is only the falling-rising pitch movement just before the adverb which can be held responsible for the perceptual effect of the intonational boundary. Second, it is striking that this subtle phonetic feature has the same interpretation effect as the more substantial phonetic differences between the two contours for the pre-boundary Fall-rise and the Fall. The fact that there is no interaction between Medial Boundary and First Pitch

Accent means that the effect sizes of the medial boundary do not vary across the three contour types. There is therefore no evidence in these data for two intonational prosodic constituents, like the intermediate phrase in the case of the right-hand Rise and Fall contours, and the intonational phrase in the case of the right-hand Fall-rise contour. Thirdly, the absence of any effect of %H was to be expected, as it has no role to play in signalling an upcoming boundary.

These results replicate those obtained in Gussenhoven (2008) for Dutch. In that experiment, participants indicated their interpretation of three ambiguous words on a 5-point scale, which had a modal adverb at one end and a predicative adjective at the other. There were three such words, one example being *vast*. As a modal adverb it means 'surely', as in *Ze zit VAST op de SNELweg* 'She must surely be on the motorway', while the predicative adjective means 'stuck', giving 'She has got stuck on the motorway'. If the pitch accent on the target word, here *VAST*, is identical to that on the VP (here *zit op de SNELweg*) and there is an IP-boundary between them, a pattern arises that is referred to as 'tone concord' by Wells (2006, p. 85) and which uniquely gives the interpretation of predicative adjective. However, in the interpretation as a modal adverb, there is no IP-boundary. Ignoring details, those results were the same as those reported here for English.

5.2.5. The interpretation of the prenuclear fall-rise

According to the exposition so far, neither MAE_ToBI nor the off-ramp analysis can account for the results for the Fall-rise contours. In the off-ramp analysis, a pre-nuclear fall is described as H*L, but this would rather give a slow fall, not a sharp fall plus a slow rise. It is reasonable to assume that a historical reinterpretation of {%L H*L H%} {%L H*L L%} as a single IP retained the salient medial H% at the expense of medial %L. If this H-tone is reinterpreted as the final tone in a tritonal prenuclear pitch accent, as in {%L H*LH H*L L%}, the realization with H in rightmost position follows from the grammar. It will locate the target of the final trailing tone just before the target of the next H*, and interpolate to it from the target of preceding L (cf. Cruttenden, 1997, p. 76). This contour is presented by O'Connor & Arnold (1973), here given as (18), though analyzed there as a contour containing an IP boundary. **Figure 11** gives the pitch track of their recorded example, overlaid with a resynthesized version, which to my ear sounds the same. The actual phrasing of this contour is somewhat ambiguous due to the long duration of the final syllable of *Paris*, which suggests a pronunciation with two IPs.

(18)   (The food in) \/Paris was su \perb

The analysis of (16a, b) in the off-ramp view is shown in (19a, b). As observed above, the IP-final H% of (19a) ends up as a third tone in the prenuclear pitch accent in (19b), which aligns rightmost, as usual. The initial %L in the second IP is deleted in the restructured form. An unexpected confirmation of the analysis in (19b) for Dutch, where the same contours exist, is provided by 't Hart et al. (1990), who reported an accelerated rise following the slow rise, occurring just before the second accented syllable, which they labeled '5' (see panel (c) in **Figure 9**). Similarly, Steedman (2014) discusses this contour in terms of how the theme is signaled, placing the intonational boundary between the theme *Anna will marry* (pronounced L + H* LH%) and the rheme *Manny* (pronounced H* LL%, his example (10)). It is tempting to interpret the two consecutive high targets in these descriptions as reflecting the targets of prenuclear trailing H and nuclear H*, respectively.

MAE_ToBI cannot easily account for this contour. A newly introduced prenuclear H* + L would have the arbitrary property of requiring a nuclear pitch accent beginning with a H-tone, to make sure there is a slow rise from the prenuclear accented syllable. This

Figure 11: F_0 track (black speckles) with superimposed smoothed contour (grey speckles) of *Paris was superb* (speech file from O'Connor & Arnold, 1973). This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wav11a and http://dx.doi.org/10.5334/labphon.30.wav11b.

measure would however not account for the wider facts, since pre-nuclear H*LH may also appear before pitch accents beginning with L*, in which case there would be no H-tone to explain the slow rise (Gussenhoven, 1983, p. 63). The alternative decision to introduce a pre-nuclear H*+L+H would have the disadvantage of requiring a unique timing policy for the final H tone, in order to prevent it from being realized immediately after the pitch fall described by H*+L. In other words, while the off-ramp analysis can naturally incorporate a pre-nuclear H*LH, the on-ramp analysis cannot.

(19)   a. She treated the poor man,   honestly

           %L   H*L   H% %L H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx19a.

       b. She treated the poor man honestly

           %L H*LH   H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx19b.

The contours labeled 'Fall' have the representations in (20a, b), those labeled 'Rise' are given in (21a, b). As will be clear, the a-examples all have the same phonological boundary, a prediction that was supported by the results of the perception experiment.

(20)   a. She treated the poor man,   honestly

     %L   H*L   L% %L H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx20a.

    b. She treated the poor man honestly

    %L   H*L   H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx20b.

(21)   a. She treated the poor man,   honestly

    %L   H*   % %L H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx21a.

    b. She treated the poor man honestly

    %L   H*   H*L L%

This audio content is available at: http://dx.doi.org/10.5334/labphon.30.wavEx21b.

6 Other empirical evidence

The identification of the falling section of an F_0-peak as a pitch accent would appear to avoid the cases of underanalysis and overanalysis by MAE_ToBI which were discussed in Sections 3, 4, and 5. Two findings have been presented that more specifically support the off-ramp view. First, Dilley et al. (2005) show that there is a low correlation between the timings of the first valley and the peak in F_0 rise-falls, suggesting that the targets of L and H* do not obey a constant interval, as suggested by the MAE_ToBI L + H* pitch accent, but are timed independently with reference to the segmental string. Conversely, Barnes et al. (2010) show that the target of the L-tone *after* H* is located with reference to the target of H*, and not with reference to any following segmental landmark, which does not support the MAE_ToBI analysis of the fall as being composed of H* followed by a heteromorphemic phrase tone. The latter result was also obtained for a number of varieties of continental West Germanic (Peters et al., 2015). These two sets of findings are just as would be expected under an off-ramp view, in which the rise is defined by heteromorphemic tones and the fall by tautomorphemic tones. In addition to these alignment facts, there are pitch

span effects for Dutch that appear to confirm the off-ramp view. Chen (2011) measured the pitch span of rises and falls of accentual pitch peaks on the S of SVO sentences in elicited adult speech. In about half the data, the S was contextually focused, while in the remainder it was topic, the O being focused. When dividing the data up into utterances in which the pitch after H* continued at a high level and utterances in which the pitch sloped down from the peak, she found that the H* + level contours differed significantly in the span of the rise towards H* as a function of the focus structure, rise spans being wider because of a lower end point. However, the rises in the H* + fall were not significantly different in the two focus conditions; rather, it was the fall that had a significantly wider pitch span, because it ended lower in the focus condition. These results do not match the on-ramp analysis, which would describe both H*-peaks as consisting of a pitch accent that represents the rise, L + H*. By contrast, the off-ramp analysis analyzes the H* + level as H* (preceded by a %L boundary tone), while the H* + fall is analyzed as H*L. Focus in Dutch can thus coherently be described as causing a raising of H* and a hyperarticulation of the fall represented by H*L.

Lastly, it is reiterated here that the results of Gussenhoven & Rietveld (1991) favoured the off-ramp analysis of Gussenhoven (1983) over the analysis in Pierrehumbert (1980). The two sets of 210 differences in terms of phonological elements among 15 nuclear melodies as expressed in those two theories showed a modest correlation of $r = 0.38$, meaning that the theories made very different predictions about the degree of similarity between pairs of nuclear melodies. Semantic differences obtained from a perception experiment with auditory stimuli representing those same pairs of nuclear melodies correlated fairly well with the off-ramp theory ($r = 0.57$), while no significant correlation was found between the Pierrehumbert data and the perception data.

7 Summary and conclusion

The off-ramp intonation grammar derived above and earlier provided in Gussenhoven (2004, p. 313)[8] is summarized in (22), with the conventions in (23).[9]

---

[8] My first attempt at an autosegmental analysis was based on Goldsmith (1980) and a familiarity with 't Hart & Collier (1980) and O'Connor & Arnold (1973), none of which explicitly featured boundary tones (cf. Ladd, 1983). My 1983 description of English was based on three pitch accents that could undergo modifications, much as in Ladd (1978, 1983), and treated the effects of boundary tones as the phonetic realizations of the pitch accents. In that description, I assumed that trailing tones of nuclear pitch accents 'spread' (i.e., 'continue' in the terminology used in this paper), but recoiled from assuming that final tones of prenuclear pitch accents do so (1983, p. 72), instead opting for a Tone Linking Rule which deleted trailing tones of pre-nuclear pitch accents. The generalized notion of continuation was originally formulated for Dutch (Gussenhoven et al., 1999–2003). An English version appeared as chapter 15 in Gussenhoven (2004). Unlike the wider formulation there, which maintained the 1983 proposal of a modification [DELAY] for both L*-initial and H*-initial pitch accents, (23) allows the affixation of the L*-prefix to H* only. This agrees with Cruttenden (1986, p. 123).

[9] The number of contours (22) generates is larger than that for MAE_ToBI, and *ceteris paribus* cases of underanalysis should be rarer, while the risk of overanalysis might be expected to be higher. Prefix L*, which may be attached to H*, H*L, and H*H, puts the number of nuclear contours at 2 (H-Prefix) × 8 (5 + 3 L*-prefixed pitch accents) × 3 (IP-endings) or 48 nuclear melodies. With 2 IP-beginnings and 5 prenuclear pitch accents, this gives 480 two-accent contours. Here, my assumption is that prenuclear scooped contours typically imply nuclear scooped contours, so that L*-prefixation is not counted separately for prenuclear accents. Because downstep is obligatory on H* after leading H on H*, I assume there are no additional downstepped versions of contours with leading H. Among %L-beginning contours, four pre-nuclear pitch accents have a H-tone (H*, H*L, H*LH, L*H), while six nuclear pitch accents have a targetable H* (i.e., not preceded by a leading H) (H*, H*L, H* + H, L* = H*, L* = H*L, and L* = H* + H), i.e., 4 × 6 × 3 (IP-endings) or 72 downstepped contours, 144 if H% beginning ones are included. Among the %H-beginning contours which have one H* in either prenuclear or nuclear position, there are 18 with L* in prenuclear position before a nuclear pitch accent with H* but without leading H, and 18 with L* or L*H in nuclear position with a prenuclear pitch accent containing H*, adding another 36 downstepped contours, or 660 in all.

(22) $\left\{\begin{array}{c}\%H \\ \%L\end{array}\right\} \left\{\begin{array}{c}(L*)H*(L(H)) \\ L*(H)\end{array}\right\} (H) \left\{\begin{array}{c}(L*)H*(L) \\ L*(H) \\ (L*)H*+H\end{array}\right\} \left\{\begin{array}{c}H\% \\ L\% \\ \emptyset\end{array}\right\}$

(23)   I.   a. The last trailing tone of a prenuclear pitch accent aligns rightmost.
         b. Other trailing tones align leftmost.
      II.   a. Within a pitch accent, interpolations are linear.
         b. Otherwise, unspecified speech is governed by the leftmost tone.
     III.   a. Within a pitch accent, downstep of H after H is obligatory.
         b. Otherwise, downstep of H* is optionally triggered by a preceding H.

Significantly, the conventions in (23) refer to pitch accents, as opposed to similar tone sequences belonging to different morphemes. The off-ramp analysis thus brings out the phonological and morphological relevance of this concept, making its tones distinct from otherwise identical sequences of tones. This sensitivity to the morphemic structure strengthens the case of the off-ramp analysis, because reference to morphological structure is a routine feature of phonological generalizations across languages.

It was suggested that a historical accident, Pierrehumbert's (1980) adoption of an equivalent of the focus-marking H from Bruce's (1977) tonal phonology of Central Swedish, lay behind her decision to analyze an accent-marking rising-falling pitch configuration in Mainstream American English as a rising pitch accent L+H* followed by a low tone from some other source, instead of a falling pitch accent H*+L preceded by a low tone from some other source. This on-ramp analysis led to a number of questionable properties of her analysis, many of which were inherited by a widely used transcription system for the language, MAE_ToBI. First, it created the need for two further tones after a nuclear pitch accent, later leading to the introduction of a tonally marked prosodic constituent, the intermediate phrase, by the side of the higher-ranking intonational phrase (Beckman & Pierrehumbert, 1986). Since no other analysis of a West Germanic language had earlier seen the need for that constituent,[10] the MAE_ToBI intonational phrasing analysis is unique among those many analyses of West Germanic intonation. Other assumptions which were in part generated by the on-ramp view and which were questioned here include the sagging of pitch between H*-targets, instead of sustained high pitch; the use of leading H in H+H* to ensure continued high pitch preceding downstepped !H*-targets, which usurps a general function of leading H to describe pre-accentual peaks; downstepped H-tones other than !H*, instead of downstep of H* only; linear interpolations between pitch accents, instead of a continuation of the left-hand tonal target; and the equation of L*-prefixed ('scooped' or 'delayed') contours with rising pitch accents.

Section 2 gave a summary statement of the MAE_ToBI analysis. Section 3 inventorized cases of underanalysis, the absence of a transcription for some contour, and Section 4 did the same for cases of overanalysis, the existence of more than one transcription for some contour. Section 5 presented perception data that suggest that MAE_ToBI's prediction of an intonational boundary is false in the case of a prenuclear sharp fall which is followed

---

[10] This is not to say that no phrasal distinctions have been made in earlier descriptions. 'Subordinated' tone groups have been claimed to account for IPs with reduced pitch range (Crystal, 1969, p. 244), while Trim (1959) pleads for a continued use of double bar (||) and single bar (|) boundary markers to imply freedom of tonal dependence across the double bar. In modern terms, these correspond to utterance and IP boundaries, respectively. Pierrehumbert (1980), among others, incorporated utterance-final unaccented IPs in her analysis, following Bing's (1979) 'O-domains'. None of these phrasal distinctions concern MAE_ToBI's ip, however.

by a gradual rise to the next accented syllable. Those data also revealed a lack of evidence for a two-tier intonational phrasing structure.

Throughout the discussion, it was shown how the opposite choice, the identification of a falling pitch accent in the accent-lending rise-fall (an off-ramp analysis), avoids the disadvantages of the MAE_ToBI analysis. The off-ramp analysis was similarly a historical accident, since it tacitly continued the off-ramp view of the British tradition (Gussenhoven, 1983, 2004). It shares with MAE_ToBI the incorrect prediction of a phrase break as described in Section 5. In the off-ramp case, this is because any trailing L-tone in a pre-nuclear pitch accent will be realized late, creating a slow fall rather than a slow rise. However, it was argued that the introduction of a tritonal pre-nuclear pitch accent H*LH, which was claimed to have resulted from a phonological change triggered by phrasal restructuring, fits neatly into the tone grammar that was independently yielded by the off-ramp view. In addition, two potential cases of overanalysis were identified for the off-ramp analysis. One concerned the occurrence of a L*H pitch accent by the side of a L*-prefixed set of nuclear pitch accents beginning with H*. The second was the generation of a set of contours with final truncated falls, 'slumps', which have not been attested in MAE and would probably be considered alien to that dialect if presented to its speakers. It is to be noted, however, that in both cases the overgeneration concerns identifiably different contours from other contours generated by the grammar, which was not true for the overgeneration of representations in MAE_ToBI. In the first case, more empirical evidence is required to validate the distinction predicted by the off-ramp analysis between L*H and H* prefixed by L*. To cover the second class of contours, we have appealed to a wider coverage of the grammar than that for any specific variety of English, such that varieties may fail to use contours that are legitimate products of the grammar. Varieties are known in any event to differ in the frequency of use of contours (Section 3.5), and a stochastic structure as envisaged by Dainora (2006) may be a goal of future research.

The above suggestion of a grammar which serves a group of closely related varieties of a language is not intended to blur the fact that we exclusively evaluated a phonological analysis of English, MAE_ToBI, and compared it with an alternative analysis. That is, there is no direct implication that analyses of other languages should be revised in similar ways. Phonological diversity is likely to apply to intonational structure as much as it does to segmental structure. A two-level intonational phrasing structure of the type that was introduced by Beckman and Pierrehumbert (1986) appears to be well-motivated in the case of varieties of Bengali (Hayes & Lahiri, 1991; Kahn, 2014), to give just one example. On-ramp and off-ramp analyses appear to apply to similar rising-falling contours in different Romance languages (Frota, this issue). Empty space between a nuclear pitch accent and an IP-final boundary tone is pronounced with left-aligned targets of the boundary tone in the tonal dialect of Roermond Dutch, but with right-aligned tones of the pitch accent in non-tonal Dutch (Gussenhoven, 2000, 2004), and so on. More empirical research into issues of the phonological representation of intonation is a desideratum. Pierrehumbert's (1980) conceptualization of the difference between phonological structure and phonetic implementation will provide an important background here, given that many communicative effects of pitch variation are non-structural, i.e., paralinguistic (Ladd, 2008, p. 34).

Supplementary Files

For accompanying TextGrid, Pitch, and wav files, go to http://dx.doi.org/10.5334/labphon.30.smo

Acknowledgements

Competing Interests

References

Arvaniti, A. 2016. Analytical Decisions in Intonation Research and the Role of Representations: Lessons from Romani. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 7*(1): 6, pp. 1–43, DOI: http://dx.doi.org/10.5334/labphon.14

Barnes, J., Veilleux, N., Brugos, A., & Shattuck-Hufnagel, S. 2010. Turning points, tonal targets, and the English L- phrase accent. *Language and Cognitive Processes, 25,* 982–1023. DOI: http://dx.doi.org/10.1080/01690961003599954

Barnes, J., Veilleux, N., Brugos, A., & Shattuck-Hufnagel, S. 2012. Tonal center of gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology, 3,* 337–383. DOI: http://dx.doi.org/10.1515/lp-2012-0017

Beckman, M. E., & Ayers, G. M. 1994. *Guidelines for ToBI labeling.* Retrieved from http://www.speech.cs.cmu.edu/tobi/ToBI.0.html.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. 2005. The original ToBI system and the evolution of the ToBI framework. In Jun, S.-A. (Ed.), *Prosodic typology: The phonology of intonation and phrasing.* Oxford: Oxford University Press, pp. 9–45. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199249633.003.0002

Beckman, M. E., & Pierrehumbert, J. B. 1986. Intonational structure of English and Japanese. *Phonology Yearbook, 3,* 255–309. DOI: http://dx.doi.org/10.1017/S09526757 0000066X

Bing, J. M. 1979. *Aspects of English prosody* (Unpublished doctoral dissertation). University of Massachusetts.

Boersma, P., & Weenink, D. 1992–2009. *Praat: doing phonetics by computer.* Retrieved from http://www.praat.org.

Bolinger, D. 1951. Intonation: Levels vs. configurations. *Word, 14,* 109–149. DOI: http://dx.doi.org/10.1080/00437956.1958.11659660

Borrás-Comes, J., Sichel-Bazin, R., & Prieto, P. 2015. Vocative intonation patterns are sensitive to politeness factors. *Language and Speech, 58,* 68–83. DOI: http://dx.doi.org/10.1177/0023830914565441

Brazil, D. 1985. *The communicative value of intonation in English.* Birmingham: Bleakhouse Press and English Language Research.

Bruce, G. 1977. *Swedish word accents in sentence perspective.* Lund: Gleerup.

Chen, A. 2011. What's in a rise: Evidence for an off-ramp analysis of Dutch intonation. In Lee, W.-S., & Zee, E. (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* [ICPhS XVII]. Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong, pp. 448–451.

Chen, A., den Os, E., & de Ruiter, J.-P. 2007. Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *The Linguistic Review, 24*, 317–344. DOI: http://dx.doi.org/10.1515/TLR.2007.012

Chen, Y., & Xu, Y. 2006. Production of weak elements in speech: Evidence from F0 patterns of neutral tone in Standard Chinese. *Phonetica, 63*, 47–75. DOI: http://dx.doi.org/10.1159/000091406

Cole, J., & Shattuck-Hufnagel, S. 2016. New Methods for Prosodic Transcription: Capturing Variability as a Source of Information. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1): 8, pp. 1–29, DOI: http://dx.doi.org/10.5334/labphon.29

Collier, R., & 't Hart, H. 1980. *Cursus Nederlandse intonatie*. Leuven: Acco.

Cruttenden, A. 1997. *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Crystal, D. 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9781139166973

Dainora, A. 2006. Modeling intonation in English: A probabilistic approach to phonological competence. In Goldstein, L. M., Whalen, D., & Best, C. T. (Eds.), *Papers in Laboratory Phonology 8: Varieties of Phonological Competence*. Berlin/New York: Mouton de Gruyter, pp. 107–132.

Dilley, L. C., Ladd, D. R.., & Schepman, A. 2005. Alignment of L and H in bitonal pitch accents: testing two hypotheses. *Journal of Phonetics, 33*, 115–119. DOI: http://dx.doi.org/10.1016/j.wocn.2004.02.003

Frota, S. 2016. Surface and Structure: Transcribing Intonation within and across Languages. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1): 7, pp. 1–19. DOI: http://dx.doi.org/10.5334/labphon.10

Gibbon, D. 1976. *Perspectives on intonation analysis*. Bern: Lang.

Goldsmith, J. A. 1980. English as a tone language. In Goyvaerts, D. (Ed.), *Phonology in the 80s* (Unpublished mimeographed version, 1974). Ghent: Story-Scientia, pp. 287–308.

Grabe, E., & Post, B. 2004. Intonational variation in the British Isles. In Sampson, G., & McCarthy, D. (Eds.), *Corpus linguistics: Readings in a widening discipline*. London and New York: Continuum International, pp. 474–481.

Grice, M. 1995. Leading tones and downstep in English. *Phonology, 12*, 183–233. DOI: http://dx.doi.org/10.1017/S0952675700002475

Grice, M., Baumann, S., & Benzmüller, R. 2005. German intonation and autosegmental-metrical phonology. In Jun, S.-A. (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 55–83. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199249633.003.0003

Grice, M., Baumann, S., & Jagdfeld, N. 2009. Tonal association and derived nuclear accents: The case of downstepping contours in German. *Lingua, 119,* 881–905. DOI: http://dx.doi.org/10.1016/j.lingua.2007.11.013

Gussenhoven, C. 1983. A semantic analysis of the nuclear tones of English. Distributed by IULC. Included as ch. 5 in Gussenhoven, C. 1985. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

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

Gussenhoven, C. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511616983

Gussenhoven, C. 2008. Semantic judgments as evidence for the intonational structure of Dutch. *Proceedings of Speech Prosody 2008*, Campinas, Brazil, pp. 297–300.

Gussenhoven, C., & Rietveld, A. C. M. 1991. An experimental evaluation of two nuclear tone taxonomies. *Linguistics, 29*, 423–449. DOI: http://dx.doi.org/10.1515/ling.1991.29.3.423

Gussenhoven, C., Rietveld, T., Kerkhoff, J., & Terken, J. 1999–2003. Transcription of Dutch intonation. Retrieved from todi.science.ru.nl.

Hayes, B., & Lahiri, A. 1991. Durationally specified intonation in English and Bengali. In Sundberg, J., Nord, L., & Carlson, R. (Eds.), *Music, language, speech, and brain.* London: Macmillan, pp. 78–91. DOI: http://dx.doi.org/10.1007/978-1-349-12670-5_7

Kahn, S. D. 2014. The intonational phonology of Bangladeshi Standard Bengali. In Jun, S.-A. (Ed.), *Prosodic typology II: The phonology of intonation and phrasing.* Oxford: Oxford University Press, pp. 81–117. DOI: http://dx.doi.org/10.1093/acprof:oso/9780199567300.003.0004

Ladd, D. R. 1978. Stylized intonation. *Language, 54*, 517–540. DOI: http://dx.doi.org/10.1353/lan.1978.0056

Ladd, D. R. 1980. *The structure of intonational meaning: Evidence from English.* Bloomington: Indiana University Press.

Ladd, D. R. 1983. Phonological features of intonational peaks. *Language, 59*, 721–759. DOI: http://dx.doi.org/10.2307/413371

Ladd, D. R. 2008. *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511808814

Ladd, D. R., & Schepman, A. 2003. "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics, 31*, 81–112. DOI: http://dx.doi.org/10.1016/S0095-4470(02)00073-6

Leben, W. R. 1976. The tones of English intonation. *Linguistic Analysis, 2*, 69–107.

Liberman, M. Y. 1975. *The intonational system of English.* MIT dissertation. Published 1979 by Garland Publishing.

Lickley, R., Schepman, A., & Ladd, D. R. 2005. Alignment of 'phrase accent' low in Dutch falling rising questions: Theoretical and methodological implications. *Language and Speech, 48*, 157–183. DOI: http://dx.doi.org/10.1177/00238309050480020201

Mayo, C., Aylett, M., & Ladd, D. R. 1997. Prosodic transcription of Glasgow English: An evaluation study of GlaToBI. In Botinis, A., Kouroupetroglou, G., & Carayannis, G. (Eds.), *Intonation: Theory, models and applications. Proceedings of an ESCA Workshop.* Athens: ESCA and University of Athens, Department of Informatics, pp. 231–234.

McCarthy, J. J., & Prince, A. 1993. Generalized alignment. *Department of Linguistics Faculty Publication Series.* Paper 12. DOI: http://dx.doi.org/10.1007/978-94-017-3712-8_4

Nolan, F., & Grabe, E. 1997. Can ToBI transcribe intonational variation in English? In Botinis, A., Kouroupetroglou, G., & Carayannis, G. (Eds.), *Intonation: Theory, models and applications. Proceedings of an ESCA Workshop.* Athens: ESCA and University of Athens, Department of Informatics, pp. 259–262.

O'Connor, J. D., & Arnold, G. F. 1973. *Intonation of colloquial English* (2nd ed.). London: Longman.

Palmer, H. E. 1922. *English intonation with systematic exercises*. Cambridge: Heffer.

Peters, J., Hanssen, J., & Gussenhoven, C. 2015. The timing of nuclear falls: Evidence from Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and High German. *Laboratory Phonology, 6*, 1–52. DOI: http://dx.doi.org/10.1515/lp-2015-0004

Pierrehumbert, J. B. 1980. *The phonetics and phonology of English intonation* (Unpublished doctoral dissertation). MIT.

Pierrehumbert, J. B. 1993. Alignment and prosodic heads. In *Proceedings of the Eastern States Conference on Formal Linguistics (ESCOL)*. Linguistics Graduate Student Association. Cornell, pp. 268–286.

Pierrehumbert, J. B. 2000. Tonal elements and their alignment. In Horne, M. (Ed.), *Intonation: Theory and experiment*. Dordrecht: Kluwer, pp. 11–36. DOI: http://dx.doi.org/10.1007/978-94-015-9413-4_2

Pitrelli, J. F., Beckman, M. E., & Hirschberg, J. 1994. Evaluation of prosodic transcription labeling reliability in the ToBI framework. *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, Yokohama, Japan, pp. 123–126.

Riad, T. 2014. *Swedish phonology*. Oxford: Oxford University Press.

Ritchart, A., & Arvaniti, A. 2014. The form and use of uptalk in Southern California English. *Proceedings of Speech Prosody 7*, Dublin. Retrieved from http://www.speechprosody2014.org.

Shattuck-Hufnagel, S., Dilley, L., Veilleux, N., Brugos, A., & Speer, R. 2004. F0 peaks and valleys aligned with non-prominent syllables can influence perceived prominence in adjacent syllables. *Proceedings of Speech Prosody, 2*, 705–708.

Silverman, K., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C., Price, P. J., Pierrehumbert, J. B., & Hirschberg, J. 1992. TOBI: A standard for labelling English prosody. In *Proceedings of ICSLP*, Banff, Alberta, Canada.

Steedman, M. 1991. Structure and intonation. *Language, 67*, 260–296. DOI: http://dx.doi.org/10.1353/lan.1991.0098

Steedman, M. 2014. The surface-compositional semantics of English intonation. *Language, 90*, 2–57. DOI: http://dx.doi.org/10.1353/lan.2014.0010

't Hart, H., & Collier, R. 1980. *Cursus Nederlandse Intonatie*. Louvain: Acco.

't Hart, H., Collier, R., & Cohen, A. 1990. *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511627743

Tench, P. 1996. *The intonation systems of English*. London/New York: Cassell.

Trim, J. L. M. 1959. Major and minor tone groups in English. *Le Maitre Phonétique, 112*, 26–29.

Uldall, E. 1961. Review of R. Kingdon. 1985. *The groundwork of English intonation* (Oxford University Press). *Archivum Linguisticum, 13*, 214–218.

Vanderslice, R. 1972. The binary suprasegmental features of English. In Rigault, A., & Charboneau, R. (Eds.), *Proceedings of the Seventh International Congress of Phonetic Sciences* (Montreal 1971). The Hague/Paris: Mouton, pp. 1052–1057.

Vanderslice, R., & Pierson, L. S. 1967. Prosodic features of Hawaiian. *English.Quarterly Journal of Speech, 53*, 156–166. DOI: http://dx.doi.org/10.1080/00335636709382828

van de Ven, M., & Gussenhoven, C. 2011.The timing of the final rise in falling-rising intonation contours in Dutch. *Journal of Phonetics, 39*, 225–236. DOI: http://dx.doi.org/10.1016/j.wocn.2011.01.006

Wells, J. C. 2006. *English intonation: An introduction*. Cambridge: Cambridge University Press.