Leading tones and downstep in English

Martine Grice

Leading tones and downstep in English*

Martine Grice
University of the Saarland

1 Introduction

1.1 Association of tones

In early autosegmental studies certain segments were specially marked with a 'star' diacritic. These starred segments had a priority association; only in a later stage of the derivation were other segments associated. Starred associations were resistant to modification by subsequent rules: 'the point is to preserve the prominence of the star' (Goldsmith 1979: 129). An associated tone had as its phonetic exponent a high or low pitch occurring approximately at the same time as the associated item(s) on the phoneme tier, usually a vowel.

Most current autosegmental studies of English intonation (Pierrehumbert 1980; Ladd 1983, 1990; Gussenhoven 1984; Lindsey 1985; Beckman & Pierrehumbert 1986; Pierrehumbert & Hirschberg 1990) allow for the grouping of tones into Pitch Accents. It is generally agreed that only one tone in each Pitch Accent may be properly associated with an item on another tier, that item usually being a syllable rather than an individual vowel phoneme. However, there is no consensus as to which tone in the Pitch Accent may be associated or as to how associated tones and syllables are synchronised.

Pierrehumbert (1980) follows Goldsmith in her use of the star to denote association. Her Pitch Accents contain one starred tone and optionally one unstarred tone, either to the left, referred to as a leading tone, or to the right, referred to as a trailing tone. In the Pitch Accent L*+H, L* is associated and H is trailing, and in L+H*, L is leading and H* is associated. Other Pitch Accents are H+L*, H*+L, H* and L* (H*+H also appeared in the original inventory but has been dropped in subsequent work). Starred tones are synchronised – or aligned – with their associated syllable, occurring within the same time slot, and leading and trailing tones occur at a stipulated distance in time before or after the starred tone. Whether the first or second of two Pitch Accent tones is starred, and therefore associated, is a phonological choice.

Other models (Ladd, Gussenhoven, Lindsey) differ from Pierrehumbert's in that they automatically associate the leftmost tone of a

Pitch Accent. This means that a Pitch Accent comprising two tones in a given order (e.g. LH) always has the first as the associated and the second as the trailing tone (in this case, L*+H), thus explicitly precluding leading tones.[1] But additionally, unlike in the model developed by Pierrehumbert and colleagues, the association of a tone to a syllable in the above models does not necessarily mean its alignment with that syllable. The possibility of specifying a different alignment from that simply determined by association is examined in more detail in the following section.

1.2 Alignment of tones

Ladd (1983), in which the first explicit differentiation between association and alignment was made, incorporates a mechanism for shifting peaks along the time axis. This includes negative pitch peaks – that is, valleys – as well as positive ones, although he, in fact, only gives features operating on valleys a brief mention. The peak shifting mechanism involves a feature [±delayed peak]. When the negative value of this feature is operative, an associated syllable is aligned with its associated tone, just as in the work of Goldsmith and Pierrehumbert. When the positive value is operative, however, the associated tone is aligned later, often with the following syllable, if there is one. The following examples of HL (H*+L) illustrate this:

(1) a. [−delayed peak]   b. [+delayed peak]

   OH$_H$ O^H$_H$

   WON$_{derful}$ WON^{der}$_{ful}$

These contours were originally used by Ladd to illustrate 'scoop', where (b) contours are scooped versions of their (a) counterparts. This corresponds to a meaning relation in which (b) contours are intensified versions of (a) contours, where the intensification adds 'a degree of emphasis, insistence etc.' (1980: 112). Although Ladd (1983) abandoned this gradient analysis in favour of a categorical one, the possibility of capturing the similarity in meaning between (a) and (b) contours was maintained, in that they were both H*+L. Furthermore, [±delayed peak] is distinctive in a pair of contours ending in a rise rather than a fall, such as the following, both H*+L H%:

(2) a. [−delayed peak]   b. [+delayed peak]

   JON$_{athon\ did}$ ^{it} JON^a$_{thon\ did}$ ^{it}

Both utterances could be used to contradict the proposition that a particular action cannot be done. (a) might be a simple counterassertion, whereas (b) can be said to have added insistence in that it is more cajoling,

e.g. in the context 'Try again and you'll manage it too!'. The following is the resulting cross-classification, where each column and each row represents a natural class of contours:[2]

(3) [−delayed peak] [+delayed peak]

H*+L L% WON$_{\text{derful}}$ WON$^{\text{der}}_{\text{ful}}$

H*+L H% JON$_{\text{athon did}}$ $^{\text{it}}$ JON$^{\text{a}}_{\text{thon did}}$ $^{\text{it}}$

Although Ladd's approach is, in principle, flexible enough to describe peaks before the accented syllable – it would involve incorporating a feature [±anticipated peak], with (4) represented as H*+L [+anticipated peak] – this opportunity is not taken.

(4) It's $_{\text{WONderful}}$ $\searrow$ OH

Neither is it taken by Gussenhoven or Lindsey, although they both make use of a delay feature similar to Ladd's. In these models, only tonal material on or following the accented syllable is treated as part of the accentual domain; the pitch preceding it is excluded from the analysis of that accent. It is in principle possible to have a feature-based analysis using [delayed peak] and [anticipated peak], where the meaning associated with the underlying tonal configuration (H*+L) is modulated, possibly along a continuum of given to new information (see Kohler 1991, although he suggests a closer link between medial and late than between early and medial). This possibility is briefly examined in §4.

An analogous situation obtains in British-style analyses (Crystal 1969; O'Connor & Arnold 1973; Couper-Kuhlen 1986) which, too, have no equivalent to leading tones. The nuclear component of a British tone unit begins on the nuclear syllable, equivalent to the final accented syllable. This is explicitly stated; Couper-Kuhlen, for instance, maintains that 'what the pitch of the voice does on the syllable before the prominent syllable or in order to get to the starting point for pitch movement on the prominent syllable is of no relevance in determining whether the nucleus is falling or rising' (1986: 80). British school intonation analysts are trained to ignore pitch movements which are considered to be onglides – as interpolation between essentially independent components of the tone unit (between head, or prehead, and nucleus). For instance, in a 'falling' head + 'high fall' nucleus sequence, the skip up between the end of the falling head and the start of the falling nucleus is treated simply as an automatic transition between the two components, with, consequently, no linguistic status.[3]

The division of the tone unit into component parts referred to here relies on a combination of rhythmic and pitch prominence. That is, unlike in autosegmental accounts, tonal and rhythmic structure are not treated on separate levels. As a result, the nuclear tone domain coincides with that of

the nuclear stress foot (plus any following feet). Just as the foot structure is left-headed, so is the part of the tone unit carrying the nuclear tone, consisting of the nucleus and the following tail. And despite the opportunity for formally separating tone and rhythm in the autosegmental models, the majority of studies already cited also maintain a *de facto* connection between them in making the Pitch Accent begin on a foot-initial syllable and precluding the shifting of tonal peaks to the left of this boundary. Pierrehumbert's model is the only one implicitly to break this connection.

1.3 Overview

This paper looks at the way different intonation models define the domain of the Pitch Accent and the ramifications a particular such definition has for an analysis of downstep and the completeness of a description of English intonation. The leading tone as introduced by Pierrehumbert (1980) will be shown to capture some, though not all, possible preaccentual pitch distinctions. The omission of leading tones in the other models is investigated, and their corresponding accounts of a number of intonation contours are compared. The use of leading tones in an analysis affects the way contexts which trigger downstep are described, and the scaling of L tones in Pierrehumbert (1980) and Beckman & Pierrehumbert (1986). These issues are discussed in §2.

   In §3 it will be shown that there is a particular type of intonation contour which, within a model which equates association with alignment, demands the use of a leading H tone to describe it, but whose resultant description conflicts with current accounts of downstep in such a framework. In §§4–6, three different resolutions of this problem are discussed and compared. The first is an extension of a feature-based alignment model along the lines of Ladd (1983). The second follows the analysis in Grice (1992), which presents a solution in terms of a hierarchical structure on the tone tier. The third is a new analysis, drawing partly on Yip (1989) in terms of tonal structure, which provides a consistent analysis of downstep. In §7, the merits and demerits of the different analyses are discussed.

2 The representation of preaccentual pitch

In this section we shall examine how perceptually salient pitch (i.e. pitch on a detectable speech segment potentially identifiable by ear as different from neighbouring segments) prior to an accented syllable can be accounted for in different models. The pitch in question occurs on a syllable just before the accented syllable or, if the accented syllable is the first in a phrase, at its very beginning, and is referred to here as PREACCENTUAL PITCH. Here the perceptually salient pitch is taken to be the pitch at a local maximum or minimum. That includes cases where the

pitch on the preaccentual syllable is at the end of a train of syllables uttered at the same pitch (which does not preclude a degree of declination) and the accented syllable is at a higher or lower pitch. Preaccentual pitch here is therefore understood to be perceptually salient, in the preliminary sense defined above.

2.1 The determinants of preaccentual pitch

There are two possible determinants of preaccentual pitch in the treatments discussed in this paper: (a) the preaccentual pitch results from a tonal value, and (b) the preaccentual pitch results from interpolation from a tone earlier in the utterance. In the case of (b), what preaccentual pitch value actually results from interpolation depends on the interpolation type. In Pierrehumbert (1980) and Ladd (1983) there are essentially two types of interpolation: linear interpolation between L and H tones and between two L tones, and sagging interpolation between two Pitch Accent H tones, shown schematically in (5):

(5)   H   H   H   H
     /   \ / \   \   /
   /   \   \   \ /
  L   L   L - - - - L

In Gussenhoven's work – e.g. Gussenhoven (1984) – only linear interpolation is used to connect tonal values.

Although the sagging between two H tones does not usually reach a very low level, the distance between the tones tends to increase the amount of sag. Conversely, should the H tones lie close together, there is no sag at all. These rules are purely phonetic; however, they must be sensitive not only to tone type (H or L) but also to whether the tone is part of a Pitch Accent, is a phrase accent or an intonation phrase boundary tone. (Phrase accents involve a plateau stretching from roughly the end of the accented word to just before the intonation phrase boundary.)

For pitch accents, the pitch on or around the preaccentual syllable can only be perceptually salient, according to our above specification, if it is a local maximum or minimum. In this situation, if the preaccentual pitch is *higher* than that on the accented syllable, only possibility (a) above could be the cause of it; that is, the higher pitch results from a tonal value. If the preaccentual pitch is *lower* than that on the accented syllable, this could be a result of either of possibilities (a) and (b) above, but in the case of (b), only as a result of 'sagging' interpolation.

However, the local minimum on a syllable just before an accented syllable which results from sagging interpolation presupposes accent peaks which are only a few syllables apart. This is because 'just before an accented syllable' typically means on the syllable before the accented syllable or, in certain circumstances, e.g. if that syllable is short, on the syllable before that.

(6)   H*   H*   H*   H*
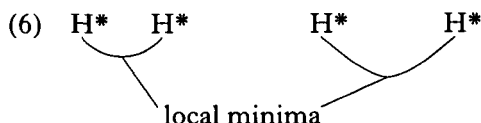
local minima

If that is the site of the local minimum, then only a reasonably short stretch can occur between the two accented syllables. This means that the depth of the local minimum cannot be very great. So if the distance between a local minimum and the following tone is as large as is commonly found in such configurations, then the local minimum cannot be considered to lie on a sagging interpolation line.[4] Therefore, generally speaking, the only cases relevant to the discussion in the rest of this paper are the cases where preaccentual pitch results from a tonal value.

A model which includes leading tones can use one to account for the perceptually salient pitch on a preaccentual syllable. For a model without leading tones, such preaccentual pitch can only be accounted for either by a tone belonging to a previous accent or an initial boundary tone. In Ladd's model, a pitch value which appears in this preaccentual position may be determined by a tone in a preceding Pitch Accent, the height of which is spread rightwards using a feature [sustained pitch]. The positive value of this feature also has the effect of preventing sagging interpolation between two H tones which, with the negative value, is assumed to take place. In Gussenhoven's model (e.g. 1984), this tone can belong to a previous Pitch Accent through the process of partial linking, which shifts a trailing tone from its original postaccentual position to the preaccentual position of the next Pitch Accent. Gussenhoven employs a rule which actually reassociates this tone to the preaccentual syllable. (This insight will be taken up again in §6.)

As already indicated, one important question for the phonological analysis of intonation at and around the left boundary of the accented tone is whether a local maximum or minimum in the preaccentual syllable slot should be considered a tonal category separate from that in the accented syllable slot, or whether it could be considered a shifted version of a tone in the accentual slot. In either case, the question arises as to what point in time a peak or trough can be considered early enough to be given its own tonal (preaccentual) denotation. A similar question arises on the frequency axis; that is, given different tones in preaccentual and accentual position, at which point in the frequency range does the accented tone acquire different polarity from its immediate predecessor?

The question of tonal categorisation on the time axis will be considered in §4 in a consideration of a feature-based analysis of temporal tonal variation. That of tonal categorisation on the frequency axis is considered below.

2.2 Vertical scaling of tones in Pitch Accents

In modern autosegmental intonation models tones are, in a paradigmatic sense, scaled relatively; in any given context, H tones are higher than L tones. The higher the prominence value, the higher the H tone and the lower the L tone, where prominence is a gradient parameter. But it isn't only prominence which determines the height of a H or L tone. A number of phonological rules may also affect the scaling of tones. They may be raised above the level at which they would otherwise be, or lowered below it; the cases in which the tonal value is lowered are those which exemplify the most commonly described phonological process in intonation and African tone languages, i.e. downstep. But tones are also scaled relatively in a syntagmatic sense, and it is this sort of scaling that the question at the start of the subsection is concerned with.

2.2.1 *Syntagmatic relative scaling of tones.* The following tonal permutations at the leading edge of a Pitch Accent are implied by the previous section:

(7)   Preaccentual   Accentual
           tone   tone

H_____H*
           ╳
L_____L*
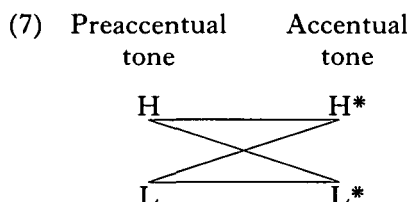
Although equal-polarity Pitch Accents are absent from current autosegmental models, they must be considered in order to clarify questions regarding leading-edge tonal structure. Our analysis begins with such potential Pitch Accents:

(8)   L——T*

In this configuration, how high does the starred tone have to become to be considered H*? The answer is that if a starred tone is even marginally higher than a preaccentual L, it can be considered a H*. The reason for this is clear. The prominence of a L tone is greater the lower it is, that of a H tone greater the higher it is. Since the leading L tone of a bitonal Pitch Accent has the same prominence as the starred tone of the same Pitch Accent, then increased pitch on the accented syllable would represent a mismatch of prominence if the starred tone were a L tone; therefore it has to be a (fairly low) H tone. Consideration of the shape of the resulting contour is also necessary in order to verify this judgement: the starred tone has to be a local maximum to be analysed as H*.

   In the configuration below, how low does a starred tone have to become to be considered a L* tone?

(9)   H——T*

Here the situation is more uncertain. There are two well-attested processes which prevent the difference in pitch between the starred and unstarred tone being of the same order as in the previous case for the tones to be considered of opposite polarity. The first of these is declination. This is currently most commonly thought of as a global phenomenon with only a slight downscaling effect on consecutive tones of equal polarity (Ladd 1986; Pierrehumbert & Beckman 1988). The second is downstep, and it is this process which constitutes the second theme of this paper.

2.2.2 *Downstep*. In this section we shall first examine the most simple case of downstep stemming from accounts of African languages. We shall then look at contours where the height of each H* tone is sustained until just before the following accent. Finally, we shall deal with calling contours, which are central to discussions of downstep in the work of Beckman & Pierrehumbert, but which are treated as raised L tones by Ladd.

The classic downstep shape, prevalent in much of the Africanist literature and discussed by Goldsmith for English, is roughly of the following shape:

(10)

Accent 1   Accent 2
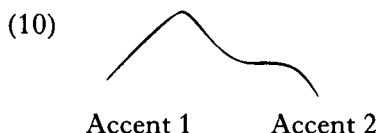
In English it consists of two Pitch Accents, both containing H tones, where the second is downstepped. The two H tones appear at the local peaks in the contour. Here it is apparent that two H tones, as elements of Pitch Accents, may have equal prominence, and yet appear at quite different pitch levels within the operative range. Unfortunately, the question of how wide a range such downstepped H tones can appear in is muddied by a requirement that certain contours which on the surface appear to require a downstep analysis need to be analysed by current analysts as having a L* tone in the position where a H* tone might be expected. This can be seen from the following.

Pierrehumbert claims that downstep is triggered by a particular tonal configuration: in earlier work, a HLH sequence incorporating at least one bitonal pitch accent downsteps the second H, and in later work (Beckman & Pierrehumbert 1986), any bitonal pitch accent downsteps the following tone. The above schematised shape can be represented as two Pitch Accents, H*+L H*. Here the second H* is downstepped, according to Pierrehumbert (1980) by virtue of the HLH sequence, and according to Beckman & Pierrehumbert (1986) owing to the fact that the first Pitch Accent is bitonal, this being their downstep trigger (though the fact that the bitonal Pitch Accents always have tones of opposite polarity means that it could be the alternation of their polarity which is the downstep trigger as much as their duality).[5] A case in which Pierrehumbert's model would predict no downstep, whereas Beckman & Pierrehumbert's model

would, is on the second Pitch Accent of the sequence L+H* H*. We shall return to this case in §6.4.

Ladd (1983), on the other hand, represents downstep as a feature, stipulating that it cannot occur on peaks which are initial in a phrase. In later work he represents it as a consequence of a metrical relationship between two nodes in a tree structure of the following type (1990: 44):

(11)

h   l   l   h

H ⌐__   H ————
L ⌐__   L ————
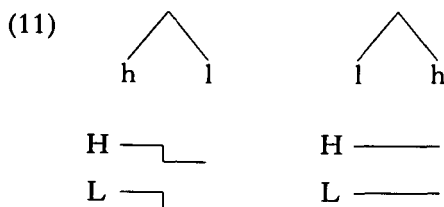
representing the register specification for two successive Pitch Accents. The horizontal lines represent the register lines on which H and L tones are scaled. Note that where h is to the left and l to the right, the register is lowered, i.e. the l triggers a lowering of the register. This means that the first Pitch Accent would have a higher H or L tone, *ceteris paribus*, than the second. By contrast, when l precedes h, the register lines are unaffected. The metrical tree thus constrains the occurrence of downstep to Pitch Accents which are on the right of a branching structure, thus ruling out downstep on initial Pitch Accents.

Thus, in all approaches examined, downstep occurs on non-initial pitch accents. It is also generally agreed to be a transitive process, in such a way that all peaks after a downstepped peak undergo the effect of lowering until a phrase break is encountered.

So far we have illustrated downstep on peaks which are quite close together. Peaks which are further apart have the following shape (where [d.s.] = [downstep]) (to avoid the inclusion of pitch movements due to intonation phrase peripheral tones, the contour is incomplete; the down-stepped H could, in fact, be followed by another pitch accent as well as by boundary tones):

(12)

| H*+L | H* | Pierrehumbert |
| H* | H* | Ladd |
| | [+d.s.] | |
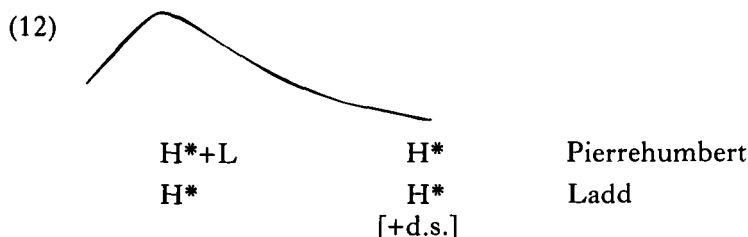
However, if the high pitch from the first Pitch Accent is sustained, necessitating an abrupt step down onto the following one, Pierrehumbert's

(1980) analysis is different, as is that of Beckman & Pierrehumbert (1986):

(13)

| H*+H | H+L* | (1980) |
| H* | H+L* | (1986) |
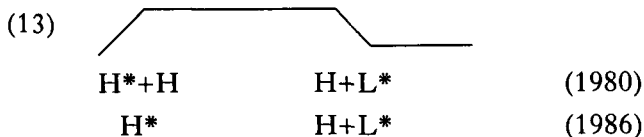
There is no downstep trigger, neither the sequence HLH needed in Pierrehumbert (1980), nor, in the later model, a bitonal pitch accent on the first peak, the necessary trigger according to Beckman & Pierrehumbert (1986). The drop to a mid-like level is a result of the rule used to calculate the height of L* in a H+L* accent. It involves multiplying the height of the previous H by a constant $k\,(0 < k < 1)$; this is the same constant as is used for calculating the height of a downstepped tone. In order to see the effect of actual downstep, a longer utterance is necessary, as in the following schematisation of a sentence such as *There are millions of intermediate levels*. Here we use the 1986 model for the Pitch Accent analysis.

(14)   There are MILLions of inter MEDiate   LEVels

| H* | H+L* | H+L* | L L% |
| Accent 1 | Accent 2 | Accent 3 | |
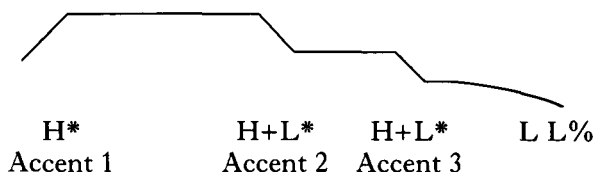
In this utterance the height above the baseline of the downstepped H is the same as that of the preceding L*, because both are calculated by the same formula. In the sequence $H_1+L^*{}_1\,H_2+H^*{}_2$, the pitch of H_2 is calculated as k(value of H_1); likewise, $L^*{}_1$ has a pitch of k(value of H_1).

   In this situation, Ladd's model involves a sequence of H* accents and the features [+sustained pitch][6] and [+downstep] as follows ([+s.p.] = [+sustained pitch]):

(15)   There are MILLions of inter MEDiate   LEVels

H*	H*	H*	L%
[+s.p.]	[+d.s.]	[+d.s.]	
	[+s.p.]		
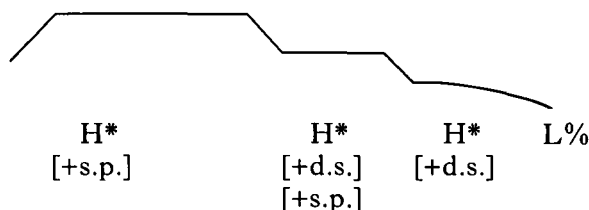
Here we can see that what Pierrehumbert analyses as L* in H+L* is analysed by Ladd as a downstepped H tone.

   However, it is unlikely that the H*[+downstep, +sustained pitch] would additionally be used to account for mid accentual pitch in another class of contours involving mid pitch, i.e. stylised contours. This can be

seen from the following. Stylised contours are used in situations involving a degree of routine or stereotype. A subset is treated by Beckman & Pierrehumbert (1986): three 'calling contours', so named because they are used for attracting attention. They all involve a step down in pitch from high to mid. The first simply involves a step down from high on the stressed syllable *to a mid pitch (usually on the next strong syllable, if there is one)*, as in (16a):

(16) a. An_{na} $\text{Aber}_{\text{nathy}}$

b. $\text{Ma}^{\text{ri}}\text{anna}$

c. $\text{Mari}^{\text{an}}\text{na}$ $\text{A}^{\text{n}}\text{na}$

The second, (16b), nowhere previously discussed, has a 'wheedling' quality (1986: 278), while the third, (16c), involves a low pitch before the high–mid step down. They treat these as a 'general class of contours ending at a mid pitch level' (1986: 276) and propose the following analysis:

(17) a. H*+L H L %
   b. L* H+L* H L %
   c. L+H* H L %

Downstep occurs on all of the H phrase accents. This is triggered in each case by the previous bitonal Pitch Accent. In (a) and (c) the stepped down (mid) part of the contour is determined by the downstepped H phrase accent tone alone. This is not the case in (b), which contains the H+L* Pitch Accent already described above in the stepping contour. In contour (b) it is the L* of H+L* which is responsible for (at least the first part of) the mid pitch.

As suggested, Ladd's (1983) model would not represent the mid pitch as a downstepped H* tone. His analysis of (a) and a projected analysis of (b) and (c) is as follows:

(18) a.   H*   +   L
            [+s.p.]   [+s.p.]
                      [−floor]

b.   L*+H   L*
                [+s.p.]
                [−floor]

c. L%   H*   +   L
            [+s.p.]   [+s.p.]
                      [−floor]

where [−floor] is a feature used to account for the fact that the L pitch is mid rather than low; it is above the floor level. He omits the boundary tone in these contours; the feature value [+sustained pitch] controls the pitch right up to the end of the phrase and maintains it at the same level as the [−floor] L tone. In the only case where mid pitch is represented by a

starred tone, (b), Pierrehumbert & Beckman and Ladd would have a similar analysis involving L*.

As in (14), the height of adjacent L* and H tones may be identical if the second is downstepped:

(19)   H + $\widehat{(L^*)}$ $\widehat{(H)}$

Thus, there is evidence that for both Ladd and Pierrehumbert & Beckman, the answer to the question of how low a tone after a preaccentual H has to be before it is identified as L is: 'not very low'. Although the rules which generate tonal sequences determine when a particular salient pitch in these contexts should be H or L, the resultant analysis is not as elegant as it could be. In the next section it will be shown that not only is the aforementioned analysis inelegant, it also does not account for a tonal distinction, after a preaccentual H, which is apparently categorical.

3 Early peak contours

In this section, a class of contours will be examined which throw further light on the question of the category boundary between H and L tones following a preaccentual tone. In the following examples, the difference in pitch between the preaccentual and accentual syllable is, in principle, maximal. That is, the high pitch could be as high in the speaker's range as desirable and the low as low as is possible for a L* tone. The examples are typical of Southern Standard British English, the dialect spoken by the author. Although American and British English may have similar inventories of intonation contours, it is unclear whether they have the same semantic interpretation. This section should therefore be taken to refer specifically to Southern Standard British English, with a possible extension to other varieties of English.

(20)   *Context*: 'We go all the way to Sicily, having planned months in advance to get a decent warm break, and when we get there, what's the weather like?'

       ^a
a. BYSmal   b. ABsol^u^{tely} ^aBYSmal   c. ^{it's} ^aBYSmal

(21)   *Context*: 'The practice is going fine until some of the brass start getting it wrong. Two guesses who it is...'

     the
a.   TUbas   b. the TRUMP^{ets} ^{and} ^{the}TUbas

    ^{it's} ^{the}
c.   TUbas again

F0 contours of these examples are given in Figs. 1 and 2.

A British-style analysis would involve the following descriptions: (a) high prehead, low falling nucleus, and (b) rising head, low falling or level
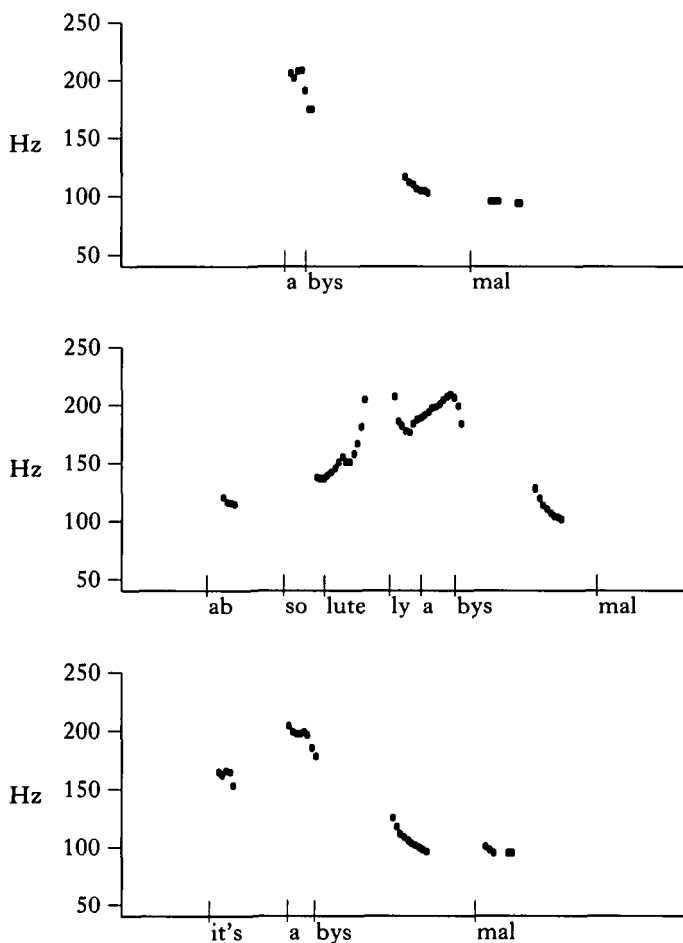
Figure 1
F0 contours for (20a–c): *abysmal, absolutely abysmal, it's abysmal*

nucleus. Example (c) is problematic in that it would require the prehead to be described as rising, a category which does not form part of the consensus British model where preheads are generally analysed as simple, containing no distinctive pitch movement.

(20) and (21) signal that the listener should already share the knowledge expressed in the proposition or should at least be in a position to share it by virtue of its predictability. It can follow a rhetorical question by the same speaker. The low pitch on the accented syllable can be sustained, this signalling suppressed emotion of some kind (pique, impatience, etc.) and can be accompanied by a rather tense laryngeal posture (see Johnson & Grice 1990 for further discussion of level contours).
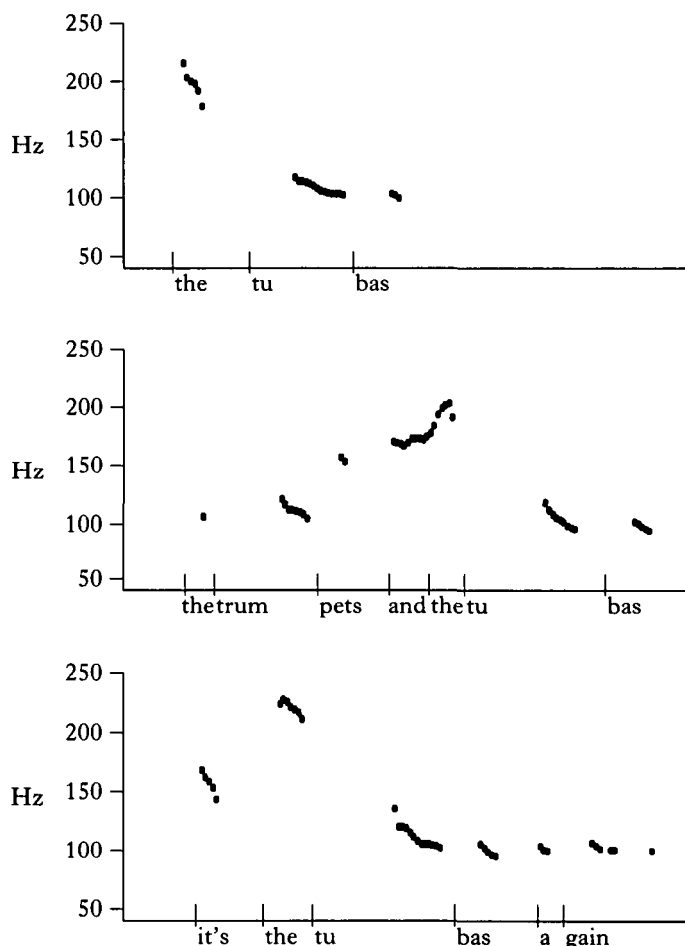
250 –
200 –
Hz   150 –
100 –
50 –

the   tu   bas

250 –
200 –
Hz   150 –
100 –
50 –

the trum   pets   and the tu   bas

250 –
200 –
Hz   150 –
100 –
50 –

it's   the tu   bas   a gain

Figure 2
FO contours for (21a–c): *the tubas, the trumpets and the tubas,*
it's the tubas again

On semantic grounds, one might expect the above examples to be
distinguished from the following set, where the high pitch is on the
nuclear syllable rather than before it:

(22) a. $a^BY S_{mal}$ b. AB$solute^{ly}$ $a^BY S_{mal}$ c. it's $a^BY S_{mal}$

Here the listener is not assumed to share the knowledge expressed in the
proposition. Within the British approach, the two sets are distinguished:
the first has low falls (or level nuclei) and the second high falls. The high
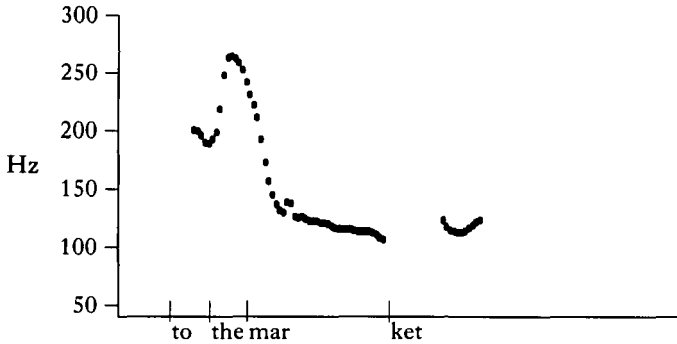*vs.* low distinction can refer to position within a speaker's range (O'Connor

Figure 3
F0 contour for (23): *to the market*

& Arnold 1973) or can relate to a previous pitch (Crystal 1969; Couper-Kuhlen 1986; Cruttenden 1986).

Another contour which emphasises the problems of the British school in analysing the preaccentual contour of examples such as (20c) and (21c) is the following, the F0 of which is given in Fig. 3:

(23)   *Context*: A neighbour keeps asking you where you are going. You try to ignore the question but they insist. You reply:

To ^{the} MARket

Here neither *to* nor *the*, if pronounced with reduced vowels, can be accented. This is then a clear case of a rising prehead rather than head.

But it is the problems in the treatment of the pitch obtrusion on the unaccented syllable *the* within autosegmental accounts which are the main issue here. That pitch peak can be neither part of a preceding accent nor attributable to an initial boundary tone; it is thus not analysable in Ladd's or Gussenhoven's model.

The same problem extends to the examples in (20) and (21). Since each instance within the set of contours in those examples has a similar meaning, it is expected that each should be analysed with the same Pitch Accent. Thus, although (20a) and (21a) might superficially look like a %H L* combination, the same analysis cannot be extended to examples (b) and (c). Nor, in any case, is an analysis of (20b) and (21b) as L*+H L* possible, because the Pitch Accent on H is too far from the initial accent for that to be feasible (except in Gussenhoven's model, where trailing tones may be shifted to just before the following accent through the process of 'partial linking'). For (20c) and (21c), similar considerations apply as to (23). Thus, it seems that the only consistent and correct autosegmental analysis of these contours is for the pitch peak before the

trough on the accented syllable to be identified as a leading H tone; that is, they are all examples containing nuclear H+L* accents.

That being the case, and if Pierrehumbert's H+L* accent is retained for description of certain downstep configurations (as detailed in §2.2.2), it is clear that the range of values for L* in such contexts is very much wider than is normally considered to be the case for L tones. This seems odd, but at first sight might be made acceptable; the threshold at which a starred tone following a leading H tone is considered to be a L tone is relatively high in the range because its height is calculated using the constant k, which is also used for calculating the value of downstepped tones. The cases of H+L* where L is low in the range, as in the above examples, would have to be analysed using a much lower value for k. This variation might not matter if it could be shown that the value for k could freely vary, both in downstep and outside downstep contexts. However, the following examples show that this cannot be the case. These contour types have a final 'continuation rise' rather than a level or slight falling terminal.

(24) (*Setting*: business meeting)

	Contour	*Context*
a.	The COST of the enterpris^e,	the difficulties in finding a good location, and the current economic climate, are making us seriously reconsider.
b.	The ASTronomical COST of the enterpris^e,	(context as in (a))
c.	It's the COST of the enterpris^e,	the difficulties in finding a good location, and the current economic climate: they've forced us to reconsider.

The contours in (24) are used as the first part of a coordinated constituent, as the context illustrates. (24b) can be contrasted with (25), which is not part of a coordinated subject but, instead, has a predicate immediately following it.

(25) (*Setting*: business meeting)

	Contour	*Context*
	The ASTronomical COST of the enterpris^e	is forcing us to reconsider.

F0 contours of these two examples are in Fig. 4.

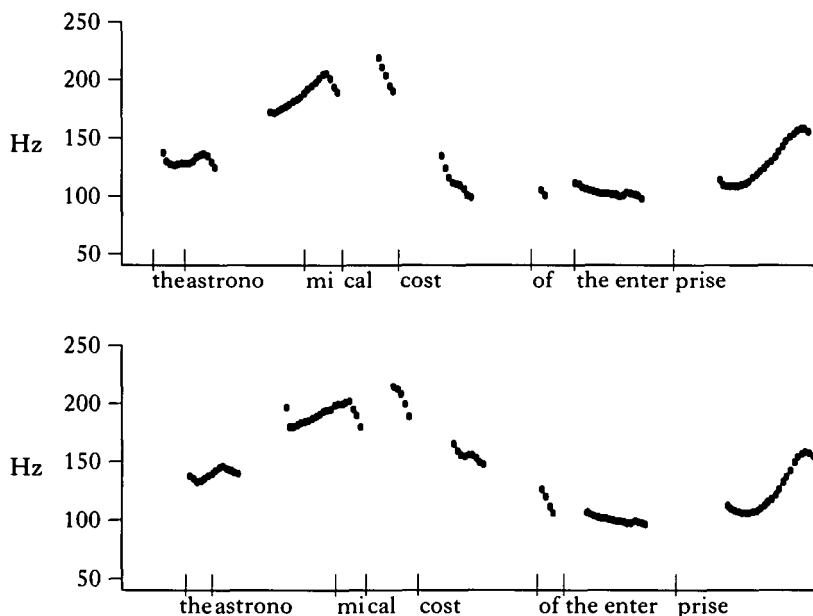A pair of utterances containing the same intonational contrast as

Figure 4

F0 contours for (24b) and (25): *the astronomical cost of the enterprise*

between (24b) and (25) but where the two pitch accents are on one word is as follows[7] (see Fig. 5):

(26) a. *Context*: 'What foods are we waiting for?'

The TAra^m^a^s^a LAta hasn't arrived ye^t,   nor has the Tsatsiki nor have the olives.

   b. *Context*: 'All the Greek foods are in now, aren't they?'

W^e_ll – the TAra^m^a^s^a^LA_ta hasn't arrived ye^t.

(26b) gives a counterexample to a previous proposition; (26a) signals, amongst other things, that the accented item is one of a set, other members of which are to follow in the same phrase. It is important to distinguish (26a) from a contour which it may superficially resemble – the 'contradiction contour', as introduced by Liberman & Sag (1974) to account for whole phrase contours which are used to express contradiction. Two such examples are given in (27):

(27) a. B_a^NAnas aren't POIson^o^u^s

   b. EL^e^ph_anTIasis isn't inCUR_ab^le
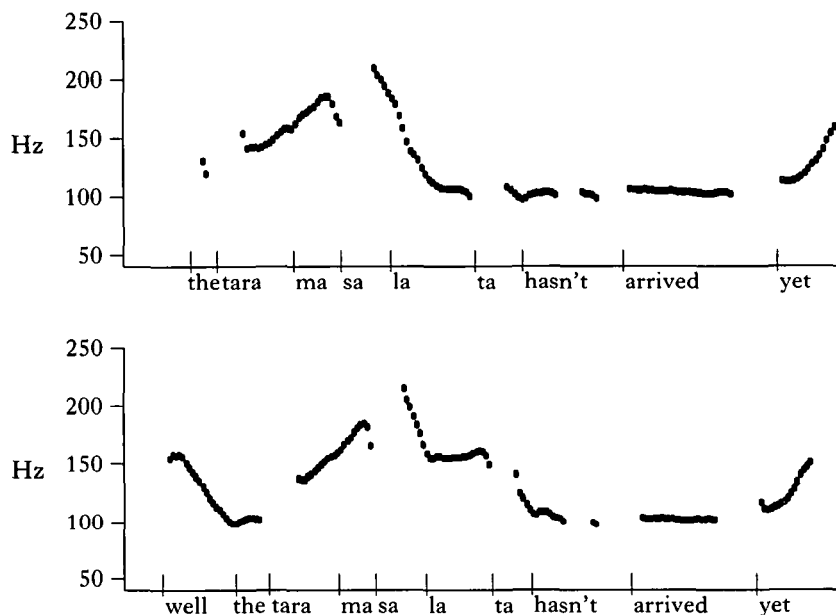
Figure 5
F0 contours for (26a, b): *the taramasalata hasn't arrived yet,*
well – the taramasalata hasn't arrived yet

F0 contours, which have been taken from the training materials for the ToBI transcription system (Beckman & Ayers 1994), are given in Fig. 6.

The contour in (26a) differs from those in (27) in that (26a) has high pitch immediately before the low accented syllable, whereas the examples of the contradiction contour have high pitch earlier in the utterance (either at a boundary, in the case of (27a), or on an accent, in the case of (27b)).

An equivalent distinction in meaning to that between (26a) and (b) can be made where the Pitch Accent being compared is the first in a phrase, as in the following examples:

(28) a. *Context*: 'What are we still waiting for?'

The $^{\text{to}}$MAtoes haven't arrived ye$^{\text{t}}$, the radishes haven't come, nor has the celery...

b. *Context*: 'We've got all the salad stuff now, haven't we?'

The $^{\text{to}}$MA$_{\text{toes}}$ haven't arrived ye$^{\text{t}}$,

F0 traces of an instance of both of these examples are given in Fig. 7.

The contours described above with early peaks have been mentioned by a number of authors, often in relation to British English. It is clear that,
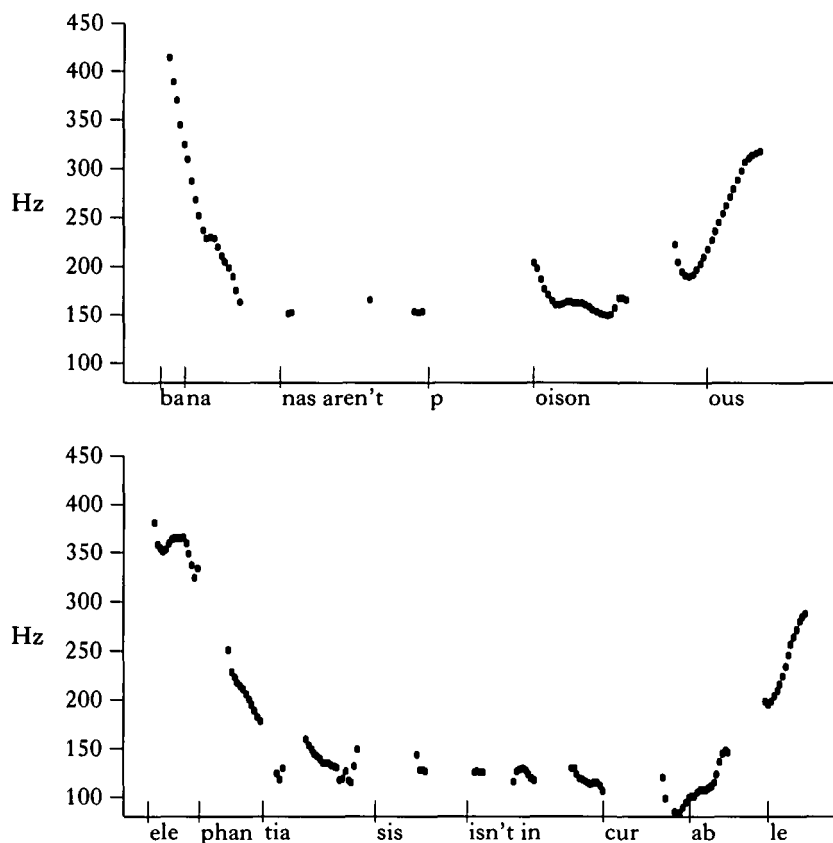
Figure 6
F0 contours for (27a, b): *bananas aren't poisonous, elephantiasis
isn't incurable*

because of the differences in meaning between these pairs of contours, the same Pitch Accent should not be used to mark their nuclei. That is, they cannot both be H+L*, even though the pitch value on the accented syllable in the second of the pairs is about the same height as a L* tone scaled by Pierrehumbert's downstep constant k. For this reason, we shall refer to the nuclear Pitch Accent on the second of each pair as H+!H* (high preaccentual pitch followed by a mid accented syllable), along the lines of the ToBI transcription system,[8] and H+L* for the first of each pair (high preaccentual pitch followed by a low accented syllable).

There is additional evidence in the literature on English for H+L* with a low accented syllable.[9] Bolinger (1989) cites Coustenoble & Armstrong's (1934) study of RP British English and matches their British contours with Network Standard American counterparts. One striking example of
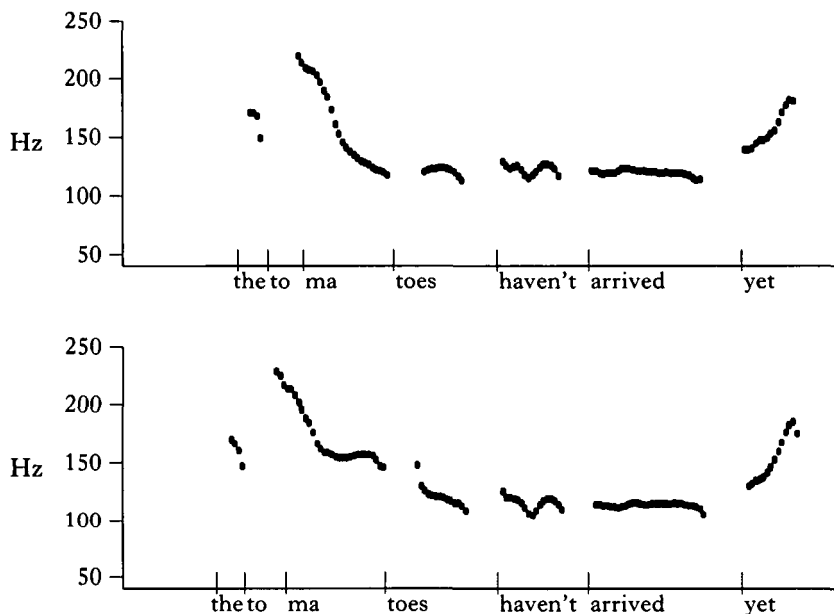
Figure 7
F0 contours for (28a, b): *the tomatoes haven't arrived yet*

a pronounced early peak in a British example is as follows (Bolinger 1989: 31):

(29)   It   won't
              BITE you

       H+L* L H%

Bolinger also cites a similar contour described by Shakbagova (1982). F0 traces of (29) and (30) are given in Fig. 8:

(30)   Can you come in De
                      CEMber

       H+L* L H%

where the high pitch on *De* is a leading H tone and the low pitch on the accented syllable CEM is a L* tone, thus H+L* L−H%. He notes that an 'abrupt drop is common in British English before the accent, where American English is more apt to use a gradual fall' (1990: 32). He points out that in the American English version of (30), the syllables *come in De* would be falling. The impression to an American ear of the wider
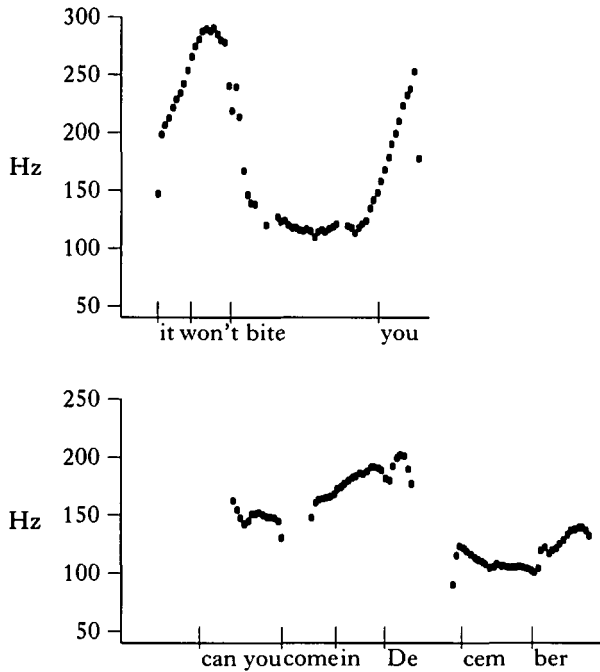
Figure 8

F0 contours for (29) and (30): *it won't bite you, can you come in December?*

intervals is of greater involvement. This means that an utterance with H+L* L H% sounds more involved than one with simply L* L H%.

It is important to point out that not all of the examples of early peak contours given by Bolinger refer to British English. The following example, which, as a reply to the question *How on earth are you able to do it?*, indicates that there is 'no reason to be concerned', refers to American English:

(31)   It's
          EA^{sy}

          H+L* L H%

It has been analysed here as having a H+L* Pitch Accent, although, since the early peak is phrase-initial, the high pitch could in principle be captured by means of an initial %H boundary tone.

Ladd (1980: 183) gives an example of the low rise preceded by a high pitch which is not phrase-initial:

(32)   Is   she
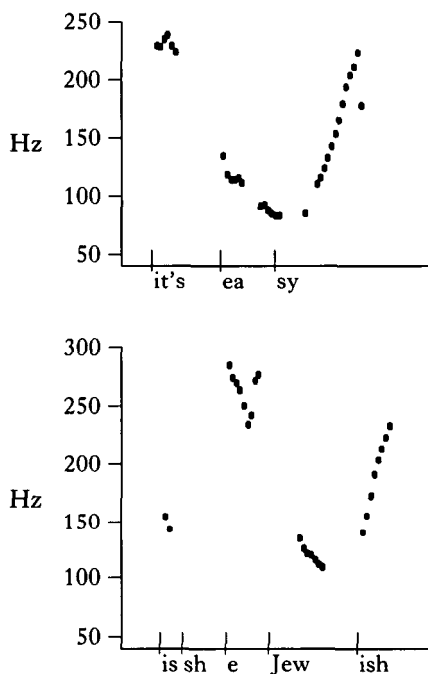   JEW^{ish}

Figure 9
F0 contours for (31) and (32): *it's easy, is she Jewish?*

(32) is uttered as an interruption in a conversation about a mutual acquaintance with the meaning: 'I'm surprised to infer from the drift of this conversation that she is Jewish; is that true?'. It is assumed here that Ladd is describing American English. This is the same Pitch Accent as in (26a) and (28a). The author's judgement of this contour is that it should be transcribed H+L*LH%, the element of surprise increasing the overall range but the final rise never reaching the height of H in H+L*, thus precluding the analysis H+L*HH%, which would involve an upstepped final boundary tone. F0 traces corresponding to (31) and (32) are in Fig. 9.

It is clear from the above examples that the H+L* where L* is scaled at or near the bottom of the speaker's range is common in British and American English when followed by L H%. This is not the same Pitch Accent as H+!H* in H+!H* L H%, as shown in examples (26b) and (28b). These data indicate that Pierrehumbert's claim that H+L* is realised as H+mid pitch is inaccurate.

In the next sections, different approaches to accounting for the two Pitch Accents referred to as H+L* and H+!H* are examined. The first will be feature-based. The second uses more than one level on the tone

tier: the supertone and the tone. The third will build on work by Yip, and will propose a distinction between Pitch Accents with a single branching tonal root node and those which consist of adjacent tonal root nodes.

4 A feature-based analysis

It was pointed out in §1.2 that Ladd's feature-based approach accounts for both downstep and peak alignment, and that the formal analysis does not, in principle, rule out an extra alignment feature [anticipated peak], although he does not propose it. This option is explored below as a way of accounting for the distinction discussed in the previous section between what we have called H+L* and H+!H*. The advantages and disadvantages of such an approach are discussed, both for these two contours and in respect of formal phonological analysis in general.

Using an alignment feature, the final Pitch Accent of the two contours, as given in (26) (Fig. 5), could be analysed as follows (the star is omitted here to emphasise the point that the alignment is determined by features rather than by the star; [a.p.] = [anticipated peak]):

(33) a. The TAra^{m}^{as}^{a} LAta hasn't arrived ye^{t},

$$\begin{array}{cc} H & + & L \\ [+a.p.] & \end{array}$$

b. ^{W}e$_{ll}$ – the TAra^{m}^{as}^{a}^{L}A$_{ta}$ hasn't arrived ye^{t}.

$$\begin{array}{cc} H & + & H \\ [+a.p.] & [+d.s.] \end{array}$$

This involves assigning features to single tones in a Pitch Accent. If we specify the negative as well as positive values, then the following representation would follow:

(34) a.　　H　　+　　L　　　b.　　H　　+　　H
　　　　[+a.p.]　[−a.p.]　　　　[+a.p.]　[−a.p.]
　　　　[−d.s.]　[−d.s.]　　　　[−d.s.]　[+d.s.]

If we assume that the absence of a specification for a particular feature is equivalent to its negative value, the value of [delayed peak] is always negative in (a) and (b). It might, in fact, be argued that [delayed peak] and [anticipated peak] are mutually exclusive since they may not occur with the same polarity on the same tone, rather as [+high] and [+low] are mutually exclusive in the specification of tongue height for vowels. This would rule out:

(35)　　H
　　　[+d.p.]
　　　[+a.p.]
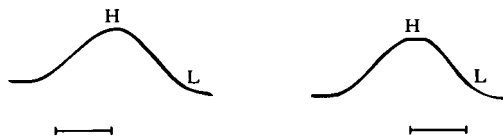
Figure 10
Ladd's Fig. E, showing 'a constant contour HL and different peak
alignment' (1983: 730). The horizontal bar indicates the accented syllable.

(where [d.p.] = [delayed peak]). Furthermore, it is logical to rule out
such feature combinations on *different* tones within the same Pitch Accent
so as to avoid the possibility of a [+delayed peak] tone preceding a
[+anticipated peak] tone. The need for this type of constraint could be
obviated if features were applied simply to the Pitch Accent as a whole. A
closer look at Ladd's exposition will provide additional motivation for this
option.

Ladd follows Bruce & Gårding (1978) in claiming that it is structurally
significant *points*, rather than whole shapes, which undergo alignment
modification. To this end, he applies a [delayed peak] feature to the H tone
of the Swedish HL Pitch Accent. However, he does represent this
graphically as a *shape*, involving a rising element, in addition to the peak
and falling elements; and the delay feature is shown as applying to the
whole shape, rather than simply to the target representing the H tone.
Figure 10, taken from Ladd (1983: 730), illustrates this, where the shape
on the left represents the [+delayed peak] contour. Thus, if the peak is
shifted to the right, say to the postaccentual syllable, the whole shape is
shifted with it. This not only implies a previous L target, as mentioned by
Bruce (1983: 230) in respect of a tonal 'copy rule' which inserts a L tone
before the configuration in cases of extreme delay, but also implies that the
delay feature, although supposedly a property of the H, has an effect on
the alignment of the whole Pitch Accent. Thus, the delay feature can be
considered to apply to the whole Pitch Accent.

Unfortunately, this mode of application does not work as well for
features relating to vertical scaling, where it must be specified which
features may apply to which types of tone. For instance, in the case of the
feature [floor], as in the calling contour already given in (16a) in §2.2.2,
and its unstylised counterpart:

(36) a. H+L   b. H+L
          [−floor]   [+floor]

      e.g. An_{na} e.g. An_{n_a}

a general stipulation would be required to the effect that this feature only
has a phonetic effect on L tones, otherwise [+floor] would make not only
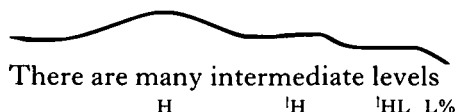the L tone but also the H tone reach floor level.

There are many intermediate levels

H   ¹H   ¹HL L%

Figure 11
Ladd's example (2b) (1983: 733), corresponding to Fig. 13a of this paper

Applying features to whole shapes or Pitch Accents would also pose a problem for the analysis of (26b), reproduced as (33b) above, where the [downstep] feature must be made to apply to only the second of the two H tones (the [anticipated peak] feature simply making the whole shape earlier than it would otherwise be):

(37)   H+H
       [+a.p.]
       [+d.s.]

A general rule to the effect that only the second tone of a [+downstep] Pitch Accent is lowered is impossible because there are cases where the first tone of a Pitch Accent does undergo lowering, i.e. Ladd's sequence for the sentence *There are MANy interMEDiate LEVels* from Ladd's example (2b), reproduced in Fig. 11 (1983: 733), which is analysed as:

(38)   H   H   H+L
              [+d.s.] [+d.s.]

The stipulation would, instead, have to be more specific: given the value [+downstep], only the last H tone in a Pitch Accent undergoes lowering (if it is the only H tone, it satisfies the condition, as it is at once the first and last). This would then be added to the stipulation mentioned in §1.2 that only non-initial Pitch Accents can be [+downstep], with an exception to account for cases such as (28b), where downstep occurs on part of the first Pitch Accent in a phrase.

Ladd's later metrical approach to downstep is unable to account for contours with downstep on the second part of an initial Pitch Accent, as (a) the register specifications reported on in §2.2.2 relate to the scaling of whole Pitch accents, not parts of them, and (b) for a lowering of the register to take place, it must be to the right of a branching node, thus requiring another Pitch Accent to precede it.

It is clear from the above discussion that more research needs to be done on the feature-based generation mechanism to ascertain whether principled ways of constraining the output to rule out illegal contours can be found. However, the greatest problem with this approach is not the generation of illegal contours, but the ambiguity in analysing run-of-the-mill legal ones. This problem arises whether features apply to single tones or to whole Pitch Accents. One such example is considered next.

Ladd's analysis of Swedish is feasible because Swedish has only HL Pitch Accents. English, by contrast, has a much larger inventory. This means that the analogous contours he proposes for English, given in (1) and reproduced below, cannot each be analysed unambiguously:

(39) a.   WON
               derful   b.   WON^der ful

(39b) could be analysed as either (40a), as discussed in §1.2, or alternatively as (40b), using the Pitch Accent which Ladd employs for the description of finally rising contours in questions of the type in his (14) (1984: 741) *Does Manitowoc have a bowling alley?*:

(40) a. H+L   L%   b. L+H   L%
        [+d.p.]

From a formal point of view, both analyses are possible. It is only a long-established meaning relation between (a) and (b) which leads to the choice of the [±delayed peak] with H+L.

If, additionally, the feature [anticipated peak] is an option (which it must be if the early peak contours are to be accounted for), then (a) could also be alternatively analysed using the same basic Pitch Accent, as long as a preaccentual syllable is present, such as *it's* in *it's wonderful*, which would usually be low-pitched:

(41) a. It's   WON
                   derful   b.   It's WON^der ful
             L+H   L%   L+H   L%
             [+a.p.]

The 'peak' alignment feature can be employed here on a L tone since Ladd explicitly states that the term 'peak' should apply to L as well as H tones (1983: 728). This use of [±anticipated peak] would thus also allow a meaning relation between the two contours to be captured. (This relation has much in common with that proposed by Pierrehumbert & Hirschberg (1990), mentioned in note 2, where both contours are L+H and the position of the star distinguishes them.)

Such ambiguity in the analysis of common English contours is a problem for a feature-based account, whether features apply to single tones or to whole Pitch Accents.

Given these difficulties with a feature-based account, we shall, for the rest of this paper, treat association as a rough indicator of alignment. That is to say, it is assumed that phonetic realisation rules determine the alignment of some part of the tone-bearing unit (in Germanic languages, the syllable) with a pitch event corresponding to a tone. Where exactly this event occurs relative to the syllable can depend on a number of factors. It can be a language-specific parameter: late in American English and early in Swedish (Pierrehumbert & Beckman 1988), late in Standard German (Kohler 1991) and even later than that in certain Southern German

dialects, such as Swabian.[10] It may depend on the internal structure of the Pitch Accent (as in Palermo Italian: see Grice 1992). It can also be affected by processes such as final lengthening (Silverman & Pierrehumbert 1990).

To say that an item on the tone tier is associated with a syllable need not mean that the pitch or F0 event occurs exactly within the time slot occupied by that syllable (or its rhyme or nucleus). It is expected, however, that they would not occur too far apart, so that the prominence of associated segments is preserved, as advocated by Goldsmith (1979) and discussed in §1 of this paper. A syllable may of course be perceived as accented (starred) through means other than pitch (e.g. through implicit knowledge of lexical stress, or because other correlates of stress are present) but it is reasonable to assume that the phonetic exponents of the starred tonal entity and the syllable occur in each other's vicinity. This assumption is in fact made in the following sections.

5 A two-level Pitch Accent analysis

In Grice (1992) an analysis was proposed which assigned a more elaborate structure to Pitch Accents. The main ideas are summarised in §5.1 below. In the subsequent sections, the advantages and disadvantages for the analysis of early peak contours, downstep and calling contours are discussed.

5.1 Structure of the Pitch Accent

In this account, the Pitch Accent has a structure analogous to the prosodic word as proposed by Nespor & Vogel (1986). Within the prosodic phonology framework, a three-syllable word such as *abysmal* would have the following analysis:
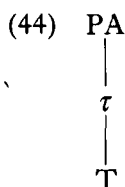
(42) *Prosodic tier*
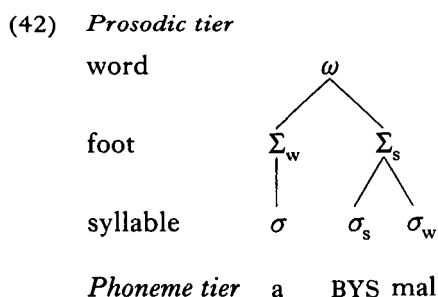
The main part of the prosodic word is the part under Σ_s; this dominates two syllables, a strong one to the left (*BYS*) and a weaker one to the right (*mal*). The strong syllable is the head of the strong foot. It is thus the designated terminal element of the prosodic word. The initial weak foot dominates only one syllable: *a*.

As suggested, the Pitch Accent could be seen to have an analogous structure. In the analogy, the PITCH ACCENT node is equivalent to the prosodic word node. This dominates two nodes: a strong one and a weak one. These are called SUPERTONE nodes, and are represented by the symbol τ. The strong τ node (τ_s) dominates the main part of the Pitch Accent (PA), which must have a head (a strong node) and can optionally have another weak node. These are called TONE nodes, which we represent as T_s and/or T_w. The weak τ node dominates a T which has no strength value because it has no sister notes. This structure can be drawn as follows:

(43)   PA

τ_w τ_s

T T_s T_w
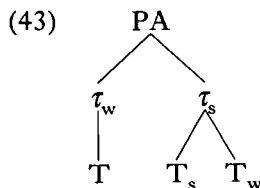
It is organised in a way which is comparable to Pierrehumbert & Beckman's (1988) prosodic structure, in that it obeys the tenets of Selkirk's Strict Layer Hypothesis (Selkirk 1984). There are constraints on the above canonical structure as to what form actual Pitch Accents may take. They are summarised in the following two points: (i) If branching occurs, it occurs *either* at the PA node *or* the τ_s node; that is, it is never the case that all three tone slots in the canonical PA are simultaneously filled.[11] The minimal representation is one supertone (τ) node and one tone (T) node; that is, a PA must contain at least one τ and one T node:

(44)   PA

τ

T

(ii) Headedness is a parameter setting for each level in the tone tier hierarchy, just as it is (at least for feet and prosodic words) in the prosodic tier. PAs are right-headed (analogous to the word) and supertones are left-headed (analogous to the foot).

It should be emphasised that the analogy between the PA and Σ structures does not imply a direct mapping between the two along the lines of Liberman (1975); the PA and Σ nodes do not have to have the same type of subtrees. Leading and trailing tones are realised in the same way as specified in Pierrehumbert (1980), at a given time respectively before and after the starred tone.

This introduction of a hierarchical structure with two levels beneath the PA node on the tone tier constitutes an elaboration of the tonal structure as it is commonly viewed. This elaboration of tonal structure allows for a clearer differentiation to be made between leading tones and other types of tone, where leading tones are dominated by an optional part of the PA

(τ_w), whereas other tones are within the core, i.e. the part dominated by τ_s. This, amongst other things, allows for a rather more principled analysis of downstep than is afforded by the feature-based approach.

As in Beckman & Pierrehumbert (1986), it is a bitonal Pitch Accent which triggers downstep on the following Pitch Accent. This automatically means that only non-initial PAs can be downstepped. It is proposed that when a Pitch Accent is downstepped, lowering only occurs on tones within its core. This means, in effect, that leading tones are not lowered, whereas starred and trailing tones are.

The inventory of Pitch Accents which falls out from the above specifications has eight members, rather than the six proposed by Beckman & Pierrehumbert (1986):

(45)

PA		PA		PA
τ_w	τ_s		τ	τ
T	T	T_s	T_w	T
L	– H*	L* + H		L*
H	– L*	H* + L		H*
H	– H*			
L	– L*			
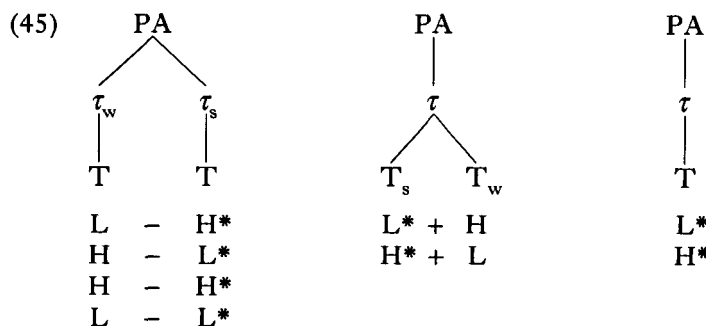
A dash is used to indicate that the tones are not dominated by the same τ, and a plus sign to indicate that they are.

The extra two Pitch Accents are bitonal, with member tones of like polarity. They do not, however, violate the Obligatory Contour Principle, as long as it is restricted to apply only to sister tones (i.e. those dominated by the same node one level up the tree). This means that H*+H and (L*+L) are ruled out because they are sister nodes under τ_s, but H–H* and L–L* are not. These two Pitch Accents are used in the account of early peak contours and are examined in more detail in §5.2 below.

5.2 Analysis of early peak contours

The two types of early peak contour referred to as H+L* and H+!H* in §3 are examined separately, the former in §5.2.1, and the latter along with downstep in general in §5.2.2.

5.2.1 *Analysis of early peak contours with low L* tone.* This account proposes a similar scaling for L* tones, whether they are in a monotonal pitch accent (L*) or part of a bitonal accent (H–L*, L*+H). Early peak contours such as the examples in (20) and (21) can be described as H–L*, as shown below:

(46) a.   ᵃBYSmal   b. ABsoluᵗᵉˡʸ ᵃBYSmal   c. ɪt's ᵃBYSmal

       H–L*   L*   H–L*   H–L*

(47) a.   ^{the} TUbas   b. the TRUMP^{ets and the} TUbas

         H–L*   L*   H–L*

    c. i^{t's the} TUbas again

       H–L*

The nuclear pitch accent is always H–L*; the similarity within each set of contours is thus captured. (26a) can also be analysed with H–L*:

(48)   The TAra^{mas}a LAta hasn't arrived ye^t,

          H–L* L   H%

5.2.2 Analysis of Downstep. As already mentioned, any bitonal pitch accent can be a downstep trigger; it triggers downstep in the following Pitch Accent or phrase accent. Only tones of a downstepped Pitch Accent which are within its core can undergo lowering.[12] Since the rule incorporating downstep is transitive, all tones after the lowered tone are affected in their scaling. An analysis of the downstepping sequence given in (14) and (15) illustrates how this works:

(49)   There are MILLions of inter MEDiate   LEVels

     L%   H–H*   H–H*   H–H*   L L%
     PA   1   2   3
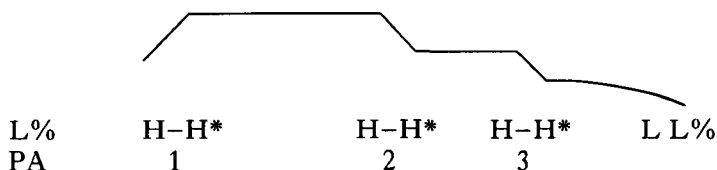
In PA 1, the value of H is the same as that of H*. In PA 2, which is downstepped because it follows a bitonal Pitch Accent, H does not undergo lowering but H* does, as it is in the T_s domain. In PA 3, the scaling of H is affected by the downstep which the previous PA has undergone – an example of the effect of downstep on the scaling of succeeding tones. H* in this last pitch accent undergoes lowering itself, because this third Pitch Accent is downstepped as a result of the bitonal pitch accent in second position.

    The early peak contour in (26b) can also be analysed using the H–H* Pitch Accent:

(50) ^We^{ll} – the TAra^{mas}a^{LA}ta hasn't arrived ye^t.

             L–L*   H–H* L   H%

The Pitch Accent L–L* on *TA* triggers downstep on H* of H–H* (L–L* is otherwise not distinct from L*). The contour in (26b), however, is not accounted for, as it entails a single Pitch Accent which cannot be

downstepped because it is not preceded by another PA which could act as a trigger.

5.3 Analysis of calling contours

In the proposed analysis, all three calling contours given in (16) can be analysed as having a downstepped tone – either a phrase accent or pitch accent tone. The upstep rule proposed by Pierrehumbert (1980), which raises a L% after a H phrase accent, is retained. For clarity of exposition, downstepped tones are underscored.

(51) a. ^{AN}na   b. MA^{ri}ANna

    H –H* H̲ L%   L* H–H̲* H L%

    H*+L   H̲ L%

    c. Mari^{AN}na

    L–H* H̲ L%

H−H* and H*+L are neutralised in initial position, just as H* and H*+L are in final position before a L phrase accent.

5.4 Other contours

A reanalysis using the above inventory of Beckman & Pierrehumbert's contours, referred to in (1986), is reported in Grice (1992). There are, in fact, only four cases where the analysis differs. These all involve early peak contours which are analysed here as having a H−H* nuclear Pitch Accent instead of H+L*. The H* is lowered as a result of downstep triggered by L−L*. One such example is given below:

(52) I REALly don't allow Mari^{AN}$_{na}$

    L−L*   H–H̲* L L%

5.5 Evaluation of the two-level Pitch Accent model

This model is able adequately to account for both types of early peak contour and has a consistent scaling of L tones, whether starred or unstarred, and whether part of a monotonal or bitonal Pitch Accent.

The predominant role of accentual and postaccentual tonal phenomena in a number of autosegmental as well as traditional accounts is reflected in the structure proposed here, in that it is the tones within the core of the Pitch Accent which represent that part of the contour. Leading tones are incorporated into the analysis as part of the PA structure, though not in the core. The extra level allows for headedness to be a parameter setting for each level on the tonal tier, rather like prosodic words and feet in the

prosodic hierarchy. Rhythmic and tonal structure are thus related: left-headed ω and PA, right-headed Σ and τ.

Constraining the output of a generator of PAs in this analysis is less problematic than for the feature-based analysis. Given the constraints advanced in this section, all and only the Pitch Accents used in the above analyses are generated.

In certain positions, a number of Pitch Accents are neutralised, but the ambiguities occur in restricted contexts. In initial position and after a monotonal accent, the distinction between H*+L and H−H* is neutralised; in this context, both have the effect of triggering downstep on the following accent and both involve a single phonetically realised peak. As in Beckman & Pierrehumbert's model, H* and H*+L are virtually indistinguishable in final position before a L phrase accent. This is because downstep is difficult to detect on the L tone, since 'higher values are more affected than lower ones' (Pierrehumbert & Beckman 1988: 90).

The way L−L* and H−H* are realised, in relation to L* and H* respectively, can best be explained with reference to Pierrehumbert & Beckman's type of phonetic implementation (see Grice 1992 for more discussion on this point). In short, leading tones, having no direct association to a syllable, are initially represented as points, whereas starred tones are represented as horizontal lines (i.e. with a given duration). These points and the ends of the lines are interpolated. After a few minor adjustments, the whole contour is low-pass filtered. The L or H points (if they are from a L−L* or a H−H* Pitch Accent) are so close to the L* or H* lines that they are effectively filtered out.

Unstructured strings of tones from studies of African tone languages have long been seen as inadequate for the analysis of languages with contour tones and those with complex Pitch Accents. The need to analyse contour Pitch Accents (rising or falling as opposed to high or low) led to an initial simple packaging of pairs of tones into Pitch Accents, later analysed as branching nodes on the tone tier. This appears to be the extent to which analyses of intonation and tone allow for structure on the tone tier, making the introduction of a hierarchical structure below the PA node an unprecedented elaboration which might require further justification. It might be argued that the account of downstep is not as straightforward as it might be, since a stipulation is required to limit its application to tones within the core. Apart from the fact that the core appears to have represented what was regarded as the whole Pitch Accent for the majority of autosegmental analysts, there is little motivation for such a constraint.

As in the feature-based approach, H−H* can only have downstep on its second tone if there is a preceding trigger Pitch Accent. This means that examples such as (26b), where there is no trigger, cannot be analysed. This is perhaps the greatest weakness of the two-level analysis.

However, it is possible to produce an analysis which does not necessitate an extra level on the tone tier, accounts consistently for downstep, distinguishes between the different early peak contours in initial as well as

non-initial position in a phrase, and provides a treatment which accounts for the difference in behaviour of leading tones compared with other tones. This last analysis of the paper is outlined in the next section.

# 6 A flat-structured Pitch Accent analysis

In this section, an analysis is proposed which has branching at the root level of the tone tier, the equivalent of the PA node in the two-level analysis, but has no intermediate level between the root level and the individual tones; there is no supertone equivalent.[13] The return to a flatter structure does not, however, mean a straightforward return to Beckman & Pierrehumbert's model.

The first section will outline Yip's (1989) analysis of contour tones in Chinese and African languages. This will provide the basis for the account of English Pitch Accents given in the subsequent sections.

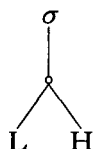## 6.1 Theoretical background

Yip (1989) has convincingly argued for two types of contour tone, which she calls 'clusters' and 'melodic units'. She schematises them as in (53):

(53) a. *Cluster*   b. *Unit*
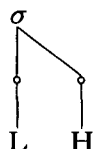
Tonal root level

Although in both cases one syllable bears two tones, there is a crucial difference in the way the tones are associated to the syllable. Whereas in clusters two tonal root nodes are attached to the syllable (here the tone-bearing unit), in melodic units there is only one tonal root node, and it is branching.

Clusters occur domain-finally or are derived and are typical of African languages. They are structurally similar to consonant clusters. The schema below shows how a cluster is typically obtained:

(54)
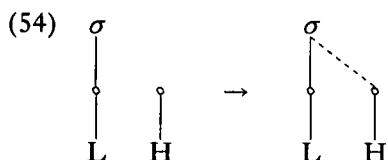
In this example, the H tone does not have a free syllable with which to associate, so it docks onto a syllable which has already been assigned a tone earlier in the derivation.

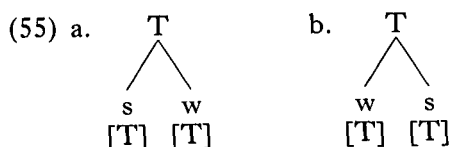Melodic units are freely distributed and are typical of Chinese

languages. They are not derived, and are structurally similar to affricates and diphthongs.

In the next section we shall see how these structures are employed in an analysis of English.

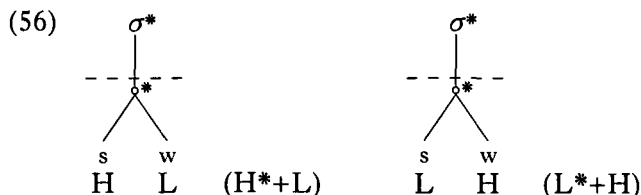## 6.2 Melodic units and tonal sequences in English

The analysis of English Pitch Accents offered in this section draws on the work of Yip in distinguishing between clusters and melodic units and, as in the two-level Pitch Accent analysis, on that of Pierrehumbert & Beckman in allowing for strong/weak labelling of branches on the tone tier. It extends Yip's discussion of canonical tone languages to allow for an analysis of a Pitch Accent language. It constrains the strong/weak labelling of branches, allowing only strong to be on the left, but balances this reduction in permutations by letting two types of bitonal Pitch Accent coexist in the same language, as well as by introducing the possibility of tritonal Pitch Accents. We shall examine each type of Pitch Accent in turn.

6.2.1 *Melodic units.* In Pierrehumbert & Beckman's account of English, the strong tone may be either on the left or on the right, as in (55) (they use T instead of Yip's tonal root node):

(55) a.
```
        T                    b.     T
       / \                         / \
      /   \                       /   \
     s     w                     w     s
    [T]   [T]                   [T]   [T]
```

These are, as they point out, essentially melodic units. The tone on the strong branch is considered to be the equivalent of a starred tone: T*, making (a) T*+T, representing L*+H and H*+L, and (b) T+T*, representing H+L* and L+H*.

The analysis currently being proposed also has melodic unit Pitch Accents. However, where there is branching in these Pitch Accents, the strong branch is always on the left, rather like the τ node in the two-level analysis. This means that the only Pitch Accents that can be accounted for in this way are left-headed ones. They can be represented thus:

(56)
```
      σ*                         σ*
      |                          |
 - - -τ*- - -              - - - τ*- - -
      /\                         /\
     /  \                       /  \
    s    w                     s    w
    H    L    (H*+L)           L    H    (L*+H)
```

Enclosing tones in parentheses indicates that they are dominated by the same tonal root node. The dashed horizontal line serves as a reminder that the line between the σ and tonal root nodes is an association line linking items on different tiers, whereas the lines from the tonal root node to the

tones (T) constitute a hierarchical structure where the tonal root node dominates T. This means that the structure proposed in this section is flatter than that discussed in § 5, having essentially only one level below the tonal root node instead of the two below the PA node. In all autosegmental studies surveyed in this paper, the trailing tone is reported to occur at a given distance in normalised time after the starred tone, rather than being aligned to a particular syllable. That is also the case in the current analysis. Additionally, as in Pierrehumbert (1980) and subsequent work, and in the two-level analysis, L in (H*+L) does not have a phonetic target value; it simply serves as part of the downstep trigger. All of the above-mentioned models have an equivalent of the 'floating' tone which has been widely reported and amply motivated in African tone languages, although none of the models have independent motivation for such a tone in English. As we have seen in § 2 of this paper, the only models which can do without this type of tone are those which achieve downstep by other means, using features or a metrical approach to register. Features have been shown in § 4 to have other problems and the metrical approach will be shown below to be inadequate. In the current analysis, a rule of the type: 'Delink L tones on weak branches after all other phonological rules (principally the downstep rule) have applied' would account for the non-realisation of L in (H*+L) and in other cases, discussed in §6.2.3.

The star on each tier marks items which are to be associated (the syllables with the tonal root nodes). Here the starred tone is indicated by a marker at the tonal root level. The use of the star is akin to Goldsmith's original use, as described in § 1.1 of this paper. The strong/weak labels are not substitutes for the star, but simply give information as to which part of an associated branching structure must be aligned. However, since it has become accepted practice to use the star notation to mark the strong tone of a Pitch Accent, it will be used in the 'surface' shorthand versions presented in this paper. We shall use it to mark specifically the strong branch of a tonal root node which itself has a star in the underlying structural representation.

6.2.2 *Tonal sequences*. Pitch Accents with leading tones are represented in a different way. A Pitch Accent such as L+H* or H+L* has two tonal root nodes, as schematised below:

(57)   σ   σ*   σ   σ*
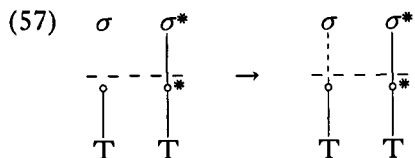
Here the two tones of a bitonal Pitch Accent are associated and aligned with different syllables: the starred tone with the starred syllable and the unstarred tone with the syllable before it. The association of the unstarred tone takes place at a second stage in the derivation, after all starred items from the two tiers have been associated. The fact that a leading tone is

separately associated to a preaccentual syllable is rather controversial. However, the association of tones to the preaccentual syllable is not new. This is achieved in Gussenhoven's (1983) model by 'partial linking': the association of a trailing tone from a Pitch Accent to the syllable before the next starred syllable. This association which arises from partial linking takes place after an initial association between Pitch Accents and starred syllables. It too is said to take place at a later stage in the derivation.

Since only starred items are associated at the first stage of a derivation, as in African tone languages, the only prerequisite for a well-formed structure is an equal count of starred items on each tier (here the tone and prosodic tiers). Either tier may or may not contain unstarred items. The existence of leading tones does not depend on the existence of an unstarred preaccentual syllable, as the tiers are autonomous.

Where the starred syllable is the first in the phrase and the Pitch Accent is L+H* or H+L*, although the starred tone will have already been associated with the starred syllable, the leading tone must also associate with it. This produces a tonal cluster, as represented below, which might occur on, for instance, *ANna*:

(58)
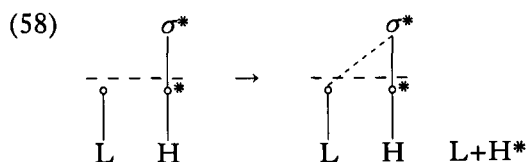
There is another context in which clusters occur: if two successive syllables are starred and therefore associated to different Pitch Accents, the second of which has a leading tone, there would be no free syllable to carry that leading tone. In this case, it is associated to one of the syllables which already has an associated tone. In (59), it is shown to associate with the second accented syllable. The text might be, for instance, *WELL CHOsen*:

(59)
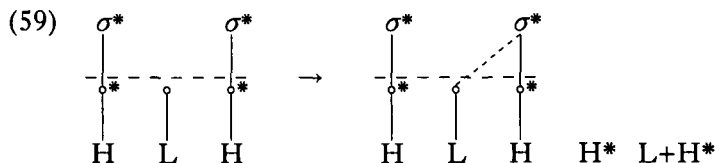
The example can be compared with Pierrehumbert's (1980) Fig. 4.8, in which the same sequence of tones occurs on *CAREfully seLECted*.

Not only is there a difference in the phonological representation of leading and trailing tones, one associating with an unstarred syllable, the other simply specified to occur after the starred tone, but there is also a difference in their phonetic realisation; there is a considerable degree of reduction in the pitch excursion in cases where the leading tone does not have a corresponding free syllable. This reduction process is examined at some length in §6.6, where a phonetic reduction rule is proposed to

account for the behaviour of such tones in situations where the tones are crowded together with respect to the corresponding tone-bearing units.

The two bitonal Pitch Accents which are analysed as tonal sequences are the one represented in (58), L+H*, and the following Pitch Accent, H+L*, which may occur with a preaccentual syllable, in which case the H is associated with it, or without one, in which case the H is associated with the starred syllable and may undergo reduction.

(60)

$$\sigma^*$$

H   L   H+L*
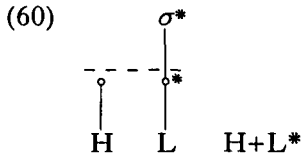
Beckman & Pierrehumbert's (1988) downstep trigger is a bitonal Pitch Accent. This is equivalent to positing a branching node as trigger, as they represent bitonal Pitch Accents as branching. This formulation is adopted in the model currently being proposed – which means that L+H* and H+L* do not trigger downstep, whereas (L*+H) and (H*+L) do.

6.2.3 *Tritonal Pitch Accents.* Tritonal Pitch Accents constitute a horizontal elaboration on the tone tier which may at first appear as problematic as the vertical elaboration proposed in the two-level analysis. For, in the first place, tritonal Pitch Accents are unusual in the analysis of English intonation because of their implicit greater complexity,[14] although they have been proposed by Yip to account for contours in a number of Chinese dialects. And secondly, the introduction of tritonal Pitch Accents threatens to increase substantially the number of possible permutations of tones, and to lead to ambiguities in the analysis. However, both of these problems can be attenuated by the constraint on the surface form of the Pitch Accent that it must maximally be binary, limiting tritonal Pitch Accents to ones containing a floating tone. This rules out Pitch Accents with an underlying structure of the type H+(L+H), as the surface form would comprise three tones, but allows Pitch Accents containing (H+L), which surfaces as a single tone.

An example of an admissible tritonal Pitch Accent is as follows:

(61)

$$\sigma^*$$

s   w
H   L   H   (H+L)+H*
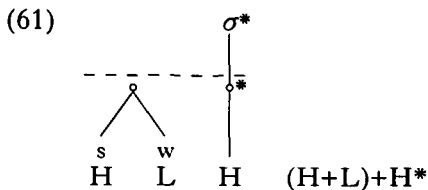
It contains a downstep trigger (the branching tonal root node) and undergoes downstep as a result of this trigger. Downstep thus occurs on the second H, the tone dominated by the node immediately following the branching one. The L tone of (H+L) in the above Pitch Accent has no

phonetic target. It is, like the L tone of (H*+L), on a weak branch and is therefore subject to the delinking rule.

The s/w ordering within tonal root nodes and the fact that starred nodes are always to the right reduces the potential inventory further. Application of the Obligatory Contour Principle could be said to operate at every level within the Pitch Accent, disallowing within a Pitch Accent (a) like sequences of tonal root nodes, ruling out contours such as (H+L)+(H*+L) or H+H*; and (b) like sequences of tones, ruling out (H*+H) or (L*+L) within a node, and H+(H+L) or (H+L)+L across nodes. The only other admissible Pitch Accent is then the following:

(62)

L   H   L   L+(H*+L)
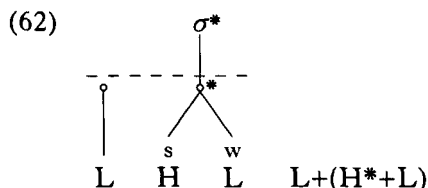
The surface form of this Pitch Accent is not distinct from that of L+H*, because the final L is delinked. The difference lies in the effect this Pitch Accent has on the following node: in the case of L+(H*+L) the following node is downstepped, and in the case of L+H* it is not.

6.2.4 *Inventory of flat-structured Pitch Accents.* We are now in a position to draw up a list of the Pitch Accents generable using the constraints introduced in the preceding sections. The building blocks are binary branching tonal root nodes with HL or LH, and non-branching nodes. A Pitch Accent comprises a starred node and, optionally, a preceding unstarred node. The inventory of Pitch Accents used in the above analysis is given below:

(63) a. *two monotonal Pitch Accents*

H   L
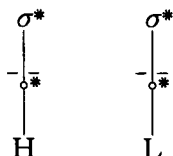
b. *two bitonal Pitch Accents with a branching tonal root node (melodic units)*

s   w   s   w
H   L   L   H
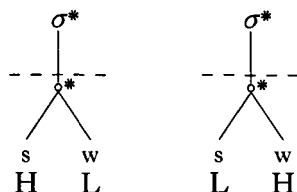
c. *two bitonal Pitch Accents with two tonal root nodes*
   (sequences of tones, potentially clusters)
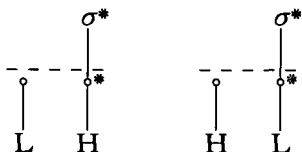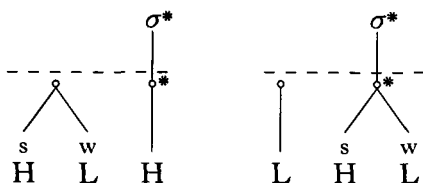


```
        σ*              σ*
   ┌────┤          ┌────┤
 - - - -│σ̅*      - - - -│σ̅*
   │    │          │    │
   L    H          H    L
```

d. *two underlyingly tritonal (but superficially bitonal)*
   Pitch Accents with two tonal root nodes


```
           σ*                 σ*
       ┌───┤             ┌───┤
- - - - - -│σ̅*     - - - -│- -
   /\      │          │   /\
  s  w     │          │  s  w
  H  L     H          L  H  L
```

In the following sections, we shall examine how this inventory accounts for early peak contours, calling contours and the other intonation patterns discussed in this paper.

6.3 Early peak contours and downstep

In this analysis, as in the feature-based and two-level analyses, L tones are always scaled in the lower part of the pitch range, and H tones in the upper part. This means that the early peak contour with low pitch on the accented syllable is H+L*. Thus, (20), (21) and (23) all contain H+L* and are followed by LL%. For instance, (20) would be analysed as (64):

(64) a.   ᵃBYSmal   b. ABsolutely ᵃBYSmal   c. it's ᵃBYSmal
       H+L* LL%   L*   H+L* LL%   H+L* LL%

(24a–c), (26a) and (28a) have the same Pitch Accent as the above set but a following LH% boundary combination, e.g.:

(65)   The TAramasa LAta hasn't arrived yet
        L*   H+L*   L   H%

        to
    The   MAtoes haven't arrived yet
        H+L*   L   H%

For early peak examples with a mid accented syllable such as (24d) and (25b), the tritonal Pitch Accent structure (H+L)+H* is proposed, where the leading unstarred part is branching. This analysis is the only one so far proposed which allows for (28b) to be accounted for, as the trigger is

within the Pitch Accent rather than on a previous one. For instance, (26b) and (28b) would be analysed as follows:

(66) ^{W}e$_{ll}$ – the TAr^{a}m$^{as^{a}}$L^{A}ta hasn't arrived y^{et}
L*   (H+L)+$\underline{H}$*

The ^{to}M^{A}toes haven't arrived y^{et}
(H+L)+$\underline{H}$*

It can also be used for the downstep sequence given in (14):

(67)   There are MILlions of inter$_{MEDiate}$ LEVe$_{ls}$

H*(H+L)+$\underline{H}$*(H+L)+$\underline{H}$*LL%
1$$2$$3

As already outlined, the reduction and lowering of register is triggered by the branching tonal root node and takes effect from the following node (here underlined). The first downstepped tone is therefore H* of the second Pitch Accent. The fact that downstep is transitive means that the initial H of the third pitch accent is at the same level as the downstepped H* of the previous one, although it is not itself downstepped. The level stretch between H* and H of (H+L) is achieved by linear interpolation, sagging interpolation taking place only between two *starred* H tones, all other interpolation being linear.

6.4 The L+H* Pitch Accent and downstep

As mentioned, L+H* does not involve a branching structure and therefore cannot be a trigger for downstep. There is, in fact, evidence in the established intonation literature for a L+H* Pitch Accent which does not trigger downstep. For instance, Ladd (1980: 184) gives examples such as the following:

(68)   'What a ridiculous day I've had. I spent the whole morning running around downtown.'
'What for?'

had ^{to} $_{get}$ CAR registered –STAMPS –
Imy$$and ^{buy} some$$and…
L+H*H–$$L+H*H–

The use of this Pitch Accent in the above context indicates, according to Ladd, the intention to suggest a loose grouping which the hearer can fill out for himself. In both cases there is level pitch from the accented syllable right up to the end of the phrase. The H phrase accent cannot be downstepped, otherwise the contour would involve a chanting tune, as in (69):
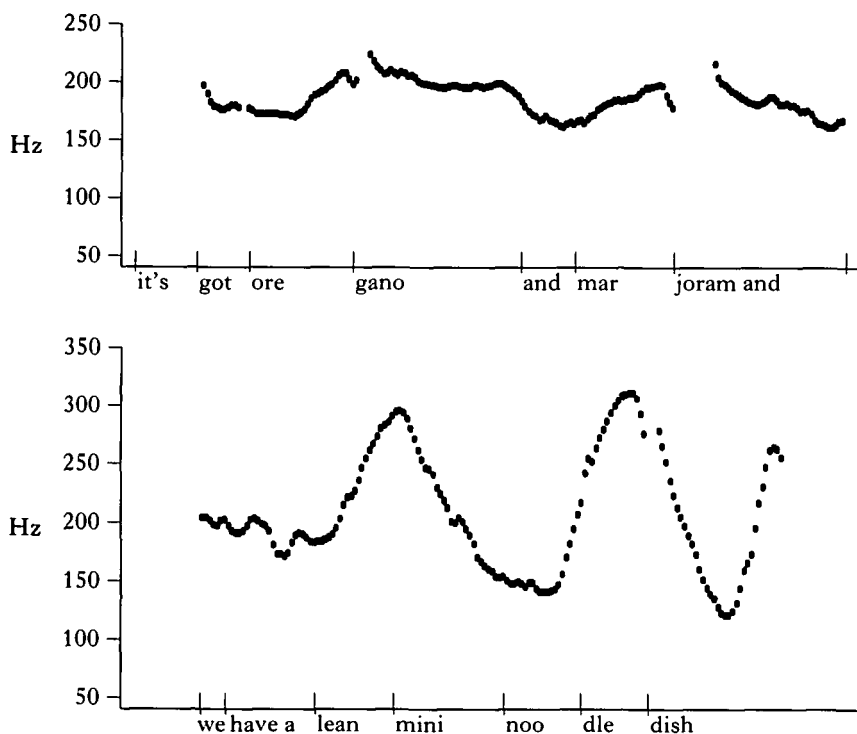
Figure 12

Display of original F0 contours provided within files in the ToBI training materials. (71) corresponds to part of the file «oregano», the part shown here has the text *It's got oregano and marjoram and...* (72) corresponds to part of «noodle2», with the text *We have a lean mini noodle dish.* Annotations have been provided for orientation by the author.

(69)   I ^{had} to $_{get}$ $_{my}$ ^{CAR} registered –

A similar contour to that in (68) appears in the ToBI training materials in the file «oregano», reproduced in Fig. 12.

(70)   It's got oREGano 'n MARjoram   ('n some fresh basil)
        ToBI: H* H–   H*   H–

The ToBI tones are below the text (H— is the phrase accent, equivalent to H in the foregoing transcriptions). Level stretches are transcribed with a H* Pitch Accent, followed by a H— phrase accent. The use of monotonal H* despite a considerable dip in F0 before the accented syllables of *oREGano* and *MARjoram* is difficult to justify. Instead, we propose a L+H* Pitch Accent, as in:

(71)   It's got oREGano 'n MARjoram   ...
         L+H* H   L+H* H   ...

If L+H* were to trigger downstep, as it does in Pierrehumbert & Beckman's analysis, then the following H phrase accent would be downstepped, which is clearly not the case. Another example appearing in the ToBI training materials as «noodle2», also in Fig. 12, contains a transcribed L+H* followed by a non-downstepped Pitch Accent, the text and transcription of which are reproduced below:

(72)   We have a LEAN mini NOOdle dish
         ToBI: L+H*   L*+H L H%

That is, L*+H is not downstepped despite the fact that the bitonal L + H* Pitch Accent precedes it.[15]

However, there are circumstances in which it is necessary for a Pitch Accent with a low leading tone and a high starred tone to act as a trigger for downstep. This is evident in the third calling contour of the set given in (16), which is analysed as having L+(H*+L). This gives the following analysis for (16c):

(73)   Mari^{an}na
         L+(H*+L) H̲ L%

6.5 Calling and other contours

The other two calling contours can be analysed as follows:

(74) a. AN_{na}   b. Ma ^{ri} ANna
       (H*+L) H̲ L%   L* (H+L)+H̲* H L%

In (74a) the branching tonal root node triggers downstep on the following phrase accent H. (74b) uses the Pitch Accent with internal downstep, which means that the mid pitch is represented as a downstepped H tone, as in the other calling contours. The advantage of this analysis of (b) is that the presence in the phrase of two Pitch Accents, a downstepper and a downstepped, is not necessary. For instance, such a contour could be produced on a name with only one accented syllable, such as *MaRIna*, which would be analysed thus:

(75)   Ma _{RIna}
         (H+L) H̲* H L%

The example from Beckman & Pierrehumbert (1986), discussed in §5.4 as (52), can be accounted for in the current analysis, using the same nuclear Pitch Accent as in (75). The first Pitch Accent does not have to constitute a trigger for downstep, as the trigger is the branching unstarred node (H+L) of the second Pitch Accent:

(76) I REALly don't allow Mari^{A}N$_{na}$

      L*   (H+L)+<u>H*</u> L L%

This analysis of the second calling contour and (52) discussed above is preferred to that of the two-level analysis, where a special trigger Pitch Accent, L−L*, had to be invoked, because we have shown in (75) a calling contour similar to (16b) without the possibility of a prenuclear Pitch Accent as trigger.

Another contour which Beckman & Pierrehumbert discuss involves the same nuclear Pitch Accent but a bitonal first Pitch Accent:

(77) I REAL^{ly} d$^{on't}$ believe Mari^{A}N$_{na}$

   (L*+H)   (<u>H+L</u>)+<u>H*</u> L L%

Here there are two downsteps in the same Pitch Accent. The first is triggered by the branching Pitch Accent (L*+H) and has an effect on the following H tone, the first tone in (H+L)+H*. The second is triggered by the (H+L) branching structure and has an effect on the following H* tone.

In §6.3 above, we looked at a contour involving the sentence *There are millions of intermediate levels*. This is a modification of a contour discussed at length by both Pierrehumbert and Ladd and considered to be part of a 'family' of contours. They all have the text *There are MANy interMEDiate LEvels*. Figure 13, taken from Ladd (1983), illustrates these contours. Ladd argues that what the contours have in common is downstep on the second and third accent, which he marks with a [downstep] feature. Pierrehumbert's analysis does not have downstep until the third accent on one of the contours. This is Fig. 13b, where the second accent is her H+L*, which is not downstepped. The four contours are analysed in the current analysis as follows:

(78)   There are MANy inter   MEDiate   LEVels

     a.   (H*+L)   (<u>H*+L</u>)   (<u>H*+L</u>) L L%

     b.   H*   (H+L)+<u>H*</u>   (H+L)+<u>H*</u>   L L%

     c.   L+(H*+L)   <u>L</u>+(H*+L)   <u>L</u>+(H*+L) L L%

     d.   (L*+H)   (<u>L*+H</u>)   (<u>L*+H</u>) L L%

We can see from the underscored Pitch Accents that the second and third Pitch Accents all undergo downstep, although in the underlying structure this is only implicit, as in Pierrehumbert's analysis. What can be observed at the underlying level is that the second and third Pitch Accents always contain a node preceded by a branching trigger node. Whether this trigger is in the previous Pitch Accent or not is irrelevant.

As already pointed out in §5.5, downstep causes L tones to undergo
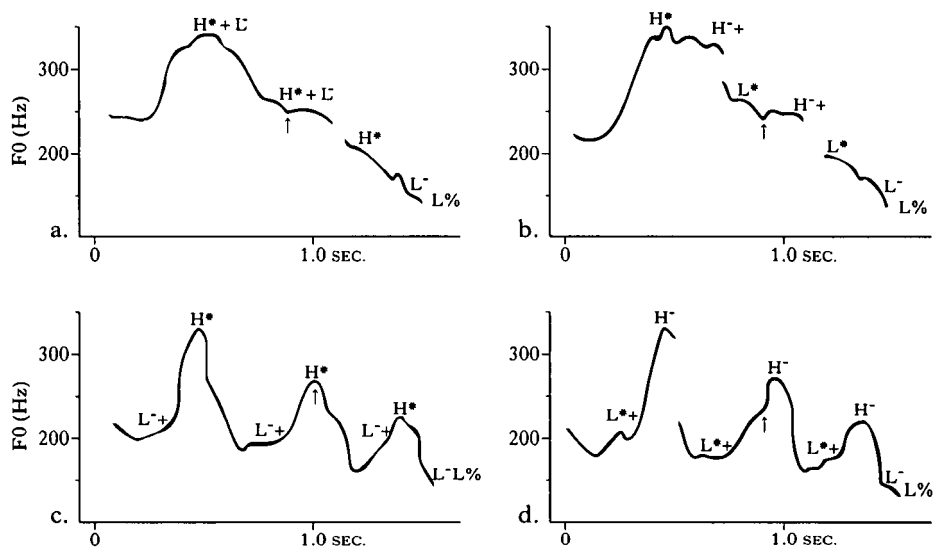
Figure 13
Ladd's Fig. B (1983: 725), illustrating four downstepped contours on the
utterance *There are many intermediate levels*, from Pierrehumbert (1980),
with her analyses

minor F0 reductions but causes H tones to undergo substantial ones. Once
the range parameter is set, it will remain reduced until it is reset at an
intermediate phrase boundary. This means that (c), which Pierrehumbert
refers to as a sequence of L+H* Pitch Accents, can here be analysed as
a sequence of L+(H*+L) Pitch Accents; the fact that the leading L of the
second and third Pitch Accents is downstepped means that the following
H is lower than the one preceding it, although it is not itself downstepped.
In (a) and (d), the whole of the second and third Pitch Accents are
underlined as, in these examples, each Pitch Accent constitutes one tonal
root node.[16]

6.6 Reduction in leading tones

It has already been observed that L+H* is not as easy to distinguish
from H* when the accented syllable with which it is associated is first in
a phrase. This is discussed in the ToBI transcription labelling guide
(Beckman & Ayers 1994):

> In theory, this contrast between H* and L+H* can occur anywhere
> within a phrase. However, the distinction is difficult to make when the
> accented syllable is the first in the utterance ... because the word 'Anna'
> has no unstressed syllables before the main stressed one, it is difficult to
> realise the low tone.

This is accounted for by the hypothesis that not all tones are associated at the same stage in the derivation.[17]

Beckman & Ayers advise labellers of performance data to opt for H* rather than L+H* in cases where the Pitch Accent is phrase-initial. It is generally agreed, however, that, in the phonological representation, the L tone is not deleted altogether, rather it undergoes a degree of reduction. In fact, one of Beckman & Pierrehumbert's (1986) examples of L+H* is an utterance comprising simply the word *ANna*, which has initial stress. This was alongside *MariANna* and other texts where there was a free syllable before the starred syllable to carry the L.

An early proponent of the British school, Kingdon, had already described preaccentual phenomena and how they can appear in cases where the accent is initial. This is no longer discussed in more recent studies within this framework. He would have analysed the non-stylised utterance in (79) (he did not give an analysis of stylised or calling contours):

(79)   Mari ^AN^ na   ⌐^AN^ na

in the following way: In *MariANna*, *Mari* would constitute a low prehead and *ANna* a high falling nuclear tone. In *ANna*, there would be a low 'homosyllabic prehead' followed by a high falling nuclear tone. That is, he allows for distinctive 'prenuclear' pitch movement on the accented syllable as long as it is at its very beginning. He points out that sonorant syllable onsets are better carriers of this pitch movement than, say, obstruents, and that voiced obstruents are better than voiceless ones. In the following three examples from Kingdon (1958: 55):

(80)

REAL ly   THERE   SHE   does

the leading low homosyllabic prehead in *really*, equivalent to a leading L, is likely to be more salient than in *there*, and even more so than in *she does*, where there is no vibration but where 'the pitch change is suggested by a change in the tension of the vocal cords' (Kingdon 1958: 54), indicated by the dashed line in Kingdon's representation. Jones (1940: 281–282) had said something similar: 'This rise is thus clearly audible in the words 'reasonable' and 'you'...; but in such words as 'think', 'ask'...it is too short to be easily heard, and may be objectively absent. The speaker has, however, a subjective feeling of its presence in all such cases'.

Abduction of the vocal folds for producing the voiceless consonant does not imply absence of cricothyroid activity, reported to have a high correlation with pitch variation (e.g. Löfqvist *et al.* 1989). Although this activity can have no effect in these circumstances on the pitch of the

consonant, there being no vocal fold vibration, it could account for the perceived change in tension referred to by Kingdon and contribute to the subjective feeling described by Jones.

Since it is the production constraints of the individual segments which can prevent the changes in pitch from taking place, and since it is argued that the syllable rather than the segment is the tone-bearing unit, these processes are best treated at a phonetic level. That is, the phonological analysis should remain the same, regardless of whether voiced segments exist to bear the leading tone.

Speech style and rate can also affect the realisation of a leading tone. The two renditions of *Anna* in (81) could be said to represent two points along a stylistic or rate continuum.

(81) ⌐AN na   ⌐AN na

Such gradient differences should also be accounted for in the phonetics; they should result from a phonetic reduction process affecting the realisation of the onglide rather than a phonological all-or-none delinking or deletion rule. A rule accounting for such a reduction process is proposed below, inspired by a phonetic reduction rule at the level of articulatory gestures proposed by Hayes (1992) to account for data observed in an electropalatographical study by Nolan (1992). In this study, considerable variance was found in the degree of alveolar closure at the end of the first word of sequences such as *late calls*, where partial assimilation can be said to have taken place. He refers to this as a rule of alveolar weakening and informally states the rule as follows: 'Depending on the rate and casualness of speech, lessen the degree of closure of a coronal autosegment, if it is [−cont] and syllable-final.' A similar gradient phonetic rule is proposed here for reducing the phonetic value of leading tones, where reduction implies making L tones less low and H tones less high than they would otherwise have been: 'Depending on the rate and casualness of speech, adjust appropriately the value of a tone target if that tone's root node is not starred and is associated to a syllable with more than one association.'

The above accounts for leading L in L+H* being variably realised according to factors such as rate and syllable structure.

Phrase-initial position for leading H is not discussed in the autosegmental literature. However, Kingdon gives examples of high homosyllabic preheads with differing segmental structures at the syllable onset. This indicates that leading H tone should behave in a similar way to leading L.

6.7 Evaluation of flat-structured Pitch Accent analysis

We have seen that bitonal Pitch Accents can be described as two tones packaged together, either tightly, as if they were affricates or diphthongs,

e.g. (L*+H), or more loosely, as if they were consonant clusters or vowel sequences, e.g. L+H*. The potential complexity of a system which allows for both branching nodes, for the former type, and sequences of nodes, for the latter type of Pitch Accent, is reduced by a number of constraints on the underlying and surface forms. These constraints lead to the generation of all and only the Pitch Accents necessary for the analysis of the contours discussed in this paper.

Rather than bitonal Pitch Accents triggering downstep on the next tone, as in Beckman & Pierrehumbert (1986), a branching tonal root node triggers downstep on the following node. The analysis of 'high plus mid accented syllable' as (H+L)+H* frees the H+L* tone for analysing 'high plus low accented syllable' which has been shown to be qualitatively different from the former accent. This analysis has an advantage over the two-level analysis in that it can account for cases where this contrast is on utterances with a single Pitch Accent. It also has a straightforward way of accounting for downstep, both from one Pitch Accent to another or simply within one Pitch Accent, whereas the feature-based and two-level analyses have to resort to stipulations as to its domain of application.

An added advantage of leading tones having a separate node is that it provides a way of capturing variability in their phonetic realisation. In cases where there is no preaccentual syllable, leading tones are associated to the starred syllable, along with the starred tone. This double association leads to a reduction in the size of the pitch excursion for the tone in question.

7 Conclusion

From the discussion in this paper, it can be argued that there are four principal measures of success of an analysis of English Pitch Accents: (a) to account for the difference between mid and low accentual pitch when preceded by a high preaccentual pitch (referred to as H+!H* and H+L*), (b) to provide a satisfactory account of downstep, (c) to minimise complexity and the concomitant potential for ambiguity in the analysis and (d) to capture the difference between leading and other tones, given the inconsistencies in the literature regarding preaccentual pitch. We shall examine each of the analyses in turn with regard to how they fare on the above points. Except in cases of floating tones, we assume a straightforward scaling of Pitch Accent L tones, such that a L tone is always low, regardless of its position within the Pitch Accent.

The feature-based approach offered in §4 is able to capture the H+L*/H+!H* distinction, as long as each tone can be specified as [±anticipated peak] and [±downstep]. The account of downstep is hampered by constraints on its domain of operation, particularly with regard to the contour given in (28b) of a downstepped tone within the first Pitch Accent of a phrase. The main problems with this approach are the generation of a great number of illegal and, in some cases, logically

impossible forms, and the ambiguity in the analysis of commonly occurring contours. Leading tones are not given special treatment.

The two-level analysis presented in §5 captures the distinction between the two early peak Pitch Accents, except in cases where the H+!H* Pitch Accent is first in the phrase. The domain of operation of downstep is restricted to tones within the core part of the Pitch Accent. Parameter settings for headedness at each level in the structure and a constraint on branching mean that there are no problems of overgeneration, although there are a few cases of neutralisation. The inventory contains eight Pitch Accents, two more than proposed by Beckman & Pierrehumbert (1986). However, since their analysis could not capture the established distinctions, the increased number is justifiable. Leading tones are outside the core of the Pitch Accent and are thus set apart from other tones which are inside the core.

An analysis involving flatter structures was proposed in §6. Pitch Accents with a leading tone are represented as having two separate nodes on the tone tier, one for the leading and one for the starred tone; whereas those with a trailing tone have one node which branches. The fact that Pitch Accents may have more than one node and that these nodes may branch leads to the possibility of Pitch Accents with more than two tones, and might therefore appear to be over-elaborate. However, a number of constraints on the underlying form (such as that tonal root nodes are left-headed, starred nodes are to the right), and on the surface form (such as that Pitch Accents have maximally two surface tones, adjacent like tones are disallowed) leads to the generation of eight Pitch Accents, all of which are shown to be necessary for the analysis of English contours. The downstep rule is simple: a branching node on the tone tier triggers downstep on the following one. In the analysis of high plus *mid* accentual pitch, downstepper and downstepped nodes are in the same Pitch Accent. This is the only analysis which can analyse this Pitch Accent successfully in initial position in a phrase. Leading tones are treated as structurally different from trailing tones, and their different phonetic behaviour is captured by a reduction rule.

Both the two-level and the flat-structured analyses generate a modest number of Pitch Accents with minimal ambiguity. They both distinguish structurally between leading and other tones. But the account of downstep in the two-level analysis requires a special stipulation to limit its operation to core tones, and its failure to account for H+!H* in initial position is a notably weak point. By contrast, the account of downstep in the flat-structured analysis is simple and accounts for H+!H* in any position in a phrase. For these reasons, the latter analysis is preferred.

NOTES

Johnson, Bob Ladd and Stefanie Shattuck-Hufnagel. I would also like to thank two anonymous reviewers for their generous comments and critiques.

[1] Although the star diacritic is unnecessary, it will be used in the present paper for ease of comparison between models, as will the + sign to show when tones are part of the same Pitch Accent. Ladd's inventory is H, L, HL and LH (represented here as H*, L*, H*+L, L*+H); Gussenhoven's H*L, L*H, H*LH (H*+L, L*+H, H*+L+H); Lindsey's H, L, L+H, H+L (H*, L*, L*+H, H*+L).

[2] Pierrehumbert & Steele (1989) would not agree with this cross-classification, arguing that there is an initial low pitch in Ladd's [−delayed peak] contours represented by their Pitch Accent L+H* which then becomes L*+H in the [+delayed peak] contours. The similarity lies in the L+H tone sequence and the difference lies in the position of the star. Such an option is not available to Ladd, as he has no leading tones.

[3] The British approach is, it must be emphasised, auditory and introspective. However well-trained listeners might be, they are not in a position to record exactly the relative cooccurrence of acoustic episodes corresponding to oral and glottal gestures. A given pitch movement may be perceived as occurring on the nuclear syllable when instrumental analysis shows an F0 movement before or after that syllable. This means that a 'high fall' may involve a rising pitch movement within the nuclear syllable, as can often be found to be the case. Similarly, a 'rise–fall' may involve a rise not only on but also before the nuclear syllable.

[4] In later work by Pierrehumbert, as reported in Anderson *et al.* (1984), H* targets are represented not as high points, but rather as shapes entailing a small rise up to a high level stretch. A sequence of two H* tones is realised as follows: there is linear interpolation between the end of the high level stretch of the first H* tone and the beginning of the short rise up to the high level of the second. Once this interpolated shape is smoothed, there is a sag nearer to the second than to the first peak. However, the sag is fairly shallow, as the rise up in the unfiltered shape is short.

[5] The scaling of L in H*+L is such that it is generally at about the same level as a following H* (which is downstepped by virtue of it following the H*+L Pitch Accent). This means that there is typically no local minimum corresponding to L of H*+L.

[6] Beckman & Pierrehumbert (1986) refer to a reduced range parameter in order to account for the sustained pitch between H* and H of H+L*. Another simple way of accounting for it is to restrict sagging interpolation to between two *starred* H tones. Interpolation would be linear in all other cases (i.e. those involving at least one L or unstarred H tone). This is suggested for the analysis in §6.

[7] This is similar to a pair of examples given by Bruce Hayes (personal communication to Carlos Gussenhoven), involving the distinction in (i):

(i) Win$^{\mathrm{nepe}}$SAUkee Street Clu$^{\mathrm{b}}$ Win$^{\mathrm{nepeSAU}}$kee

This is the first explicit mention of this distinction that the author is aware of.

[8] The tonal part of the ToBI transcription system is a consensus system based largely on Pierrehumbert (1980), developed for the labelling of large corpora of various varieties of English. One aim has been to make the system less abstract and consequently easier to learn. This has meant marking downstep explicitly as '!' before H tones and marking the starred tone in Pierrehumbert's H+L* as !H* because it resembles a downstepped H tone in other environments.

[9] This distinction has also been illustrated by Prieto (forthcoming) for Central Catalan, where, on the same text, H+!H* L−L% implies a pedagogical declarative and H+L* L−L% an interrogative.

[10] I would like to thank Bob Ladd for discussion on this point.

[11] Use can be made of licensing principles from a theory such as Government Phonology (Kaye *et al.* 1990) to obtain this constraint. Applied to the PA node,

the designated terminal element can license adjacent slots from left to right or from right to left, but not both. This means that T* can be preceded *or* followed by a T, but it cannot be preceded *and* followed by a T.

[12] It has been suggested that downstep can involve language-specific stipulations. For instance, Japanese requires the stipulation that the second tone of the *triggering* pitch accent undergo lowering (Pierrehumbert & Beckman 1988).

[13] This analysis was in part motivated by the comments of one of the anonymous reviewers, to whom I am grateful.

[14] An apparent exception to this is the analysis of Gussenhoven (1983). However, he only uses tritonal Pitch Accents where other analysts would use a combination of a bitonal Pitch Accent and a boundary tone.

[15] It is not transcribed with a '!' diacritic, which is used in ToBI to explicitly mark downstep.

[16] They are all sequences of like Pitch Accents except for (b). It could be argued that a contour beginning high would be a better candidate for 'family' membership so that it too is a sequence of Pitch Accents of the same type, as in (i):

(i) (H+L)+H̲*   (H+L)+H̲*   (H+L)+H̲*   L L%

[17] The idea that tones are not all associated at once was first proposed for Japanese by Pierrehumbert & Beckman (1988). They propose that first Pitch Accents then boundary tones are associated to individual morae, and thereby account for boundary tones which are not associated to a mora, because the Pitch Accent tones have priority.

REFERENCES

Anderson, M. D., J. B. Pierrehumbert & M. Y. Liberman (1984). Synthesis by rule of English intonation patterns. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. New York: IEEE. 2.8.1–2.8.4.

Beckman, M. E. & G. M. Ayers (1994). *Guidelines for ToBI labeling*. Columbus: Ohio State University.

Beckman, M. E. & J. B. Pierrehumbert (1986). Intonational structure in English and Japanese. *Phonology Yearbook* 3. 255–309.

Bolinger, D. (1986). *Intonation and its parts*. London: Edward Arnold.

Bolinger, D. (1989). *Intonation and its uses*. London: Edward Arnold.

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

Bruce, G. (1983). Accentuation and timing in Swedish. *Folia Linguistica* 8. 221–238.

Bruce, G. & E. Gårding (1978). A prosodic typology for Swedish dialects. In E. Gårding, G. Bruce & R. Bannert (eds.) *Nordic prosody*. Lund: Gleerup. 219–228.

Couper-Kuhlen, E. (1986). *An introduction to English prosody*. London: Edward Arnold.

Coustenoble, H. & L. Armstrong (1934). *Studies in French intonation*. Cambridge: Heffer.

Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.

Docherty, G. J. & D. R. Ladd (eds.) (1992). *Papers in laboratory phonology II : gesture, segment, prosody*. Cambridge: Cambridge University Press.

Goldsmith, J. (1979). *Autosegmental phonology*. New York: Garland.

Grice, M. (1992). *The intonation of interrogation in Palermo Italian : implications for intonation theory*. PhD thesis, University College London. Published 1995, Tübingen: Niemeyer.

Gussenhoven, C. (1983). Focus, mode and the nucleus. *JL* 19. 377–417.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents.* Dordrecht: Foris.

Hayes, B. (1992). Comments on chapter 10. In Docherty & Ladd (1992). 280–286.

Johnson, M. & M. Grice (1990). The phonological status of stylised intonation contours. *Work in Progress, Speech, Hearing and Language, UCL* 4. 227–256.

Jones, D. (1940). *An outline of English phonetics.* 6th edn. Cambridge: Heffer.

Kaye, J., J. Lowenstamm & J.-R. Vergnaud (1990). Constituent structure and government in phonology. *Phonology* 7. 193–231.

Kingdon, R. (1958). *The groundwork of English intonation.* London: Longmans, Green & Co.

Kingston, J. & M. Beckman (eds.) (1990). *Papers in laboratory phonology I : between the grammar and physics of speech.* Cambridge: Cambridge University Press.

Kohler, K. J. (ed.) (1991). *Studies in German intonation.* Kiel: AIPUK.

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. *Lg* 59. 721–759.

Ladd, D. R. (1990). Metrical representation of pitch register. In Kingston & Beckman (1990). 35–57.

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

Liberman, M. (1975). *The intonational system of English.* PhD dissertation, MIT. Published 1980, New York: Garland.

Liberman, M. & I. Sag (1974). Prosodic form and discourse function. *CLS* 10. 416–427.

Lindsey, G. A. (1985). *Tonal structure and the expression of a pragmatic function in English and other languages.* PhD dissertation, UCLA.

Löfqvist, A., T. Baer, N. S. McGarr & R. S. Story (1989). The cricothyroid muscle in voicing control. *JASA* 85. 1314–1321.

Nespor, M. & I. Vogel (1986). *Prosodic phonology.* Dordrecht: Foris.

Nolan, F. (1992). The descriptive role of segments: evidence from assimilation. In Docherty & Ladd (1992). 261–280.

O'Connor, J. D. & G. F. Arnold (1961). *Intonation of colloquial English.* London: Longman.

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

Pierrehumbert, J. B. & M. E. Beckman (1988). *Japanese tone structure.* Cambridge, Mass.: MIT Press.

Pierrehumbert, J. B. & J. Hirschberg (1990). The meaning of intonation contours in the interpretation of discourse. In P. R. Cohen, J. Morgan & M. E. Pollack (eds.) *Intentions in communication.* Cambridge, Mass.: MIT Press. 271–311.

Pierrehumbert, J. B. & S. A. Steele (1989). Categories of tonal alignment in English. *Phonetica* 46. 181–196.

Prieto, P. (forthcoming). Aproximació als contorns entonatius del Català Central. *Caplletra* 18.

Selkirk, E. O. (1984). *Phonology and syntax.* Cambridge, Mass.: MIT Press.

Shakbagova, D. A. (1982). *Varieties of English pronunciation.* Moscow: Vishcha Skola.

Silverman, K. A. & J. B. Pierrehumbert (1990). The timing of prenuclear High accents in English. In Kingston & Beckman (1990). 72–106.

Yip, M. (1989). Contour tones. *Phonology* 6. 149–174.