

# *Tone-association in English*\*

MERCEDES CABRERA-ABREU

---

## **0 Introduction**

In the present article, I propose that tone is associated to a single type of Tone Bearing Unit (henceforth TBU) —boundaries of prosodic domains. This radical view has far reaching consequences for current phonological models of intonation which are mainly characterised by positing two types of TBUs —boundaries **and** accented syllables. I shall start by focusing on Grice's intonation model for English (which is mainly based on Pierrehumbert and Beckman's (1988) model for Japanese). This is the latest work on the phonology of intonation which not only counts as one of the major contributions to intonational studies, but also, gathers the fundamental tenets of previous proposals into a single piece of work. Then in §2, I shall investigate the possibility of adapting a single type of tone-association to Grice's model: first at the foot level in §2.1, and later at the levels above the foot in §2.2. Due to reasons related to the defectiveness of tree-format, such an attempt fails, and thus, in §3 I suggest that a grid-format should be used as an alternative. Its advantages over the other will be borne out in §3.1 where an effort is made to incorporate tone into the representation.

## **1 Grice's model of intonation**

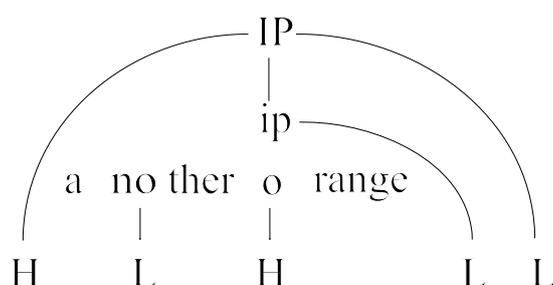
Since the publication of *The phonetics and phonology of English intonation* (Pierrehumbert 1980) and of subsequent work on English (Ladd 1983, Gussenhoven 1983, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, and Grice 1992 among others), a widely shared assumption has been that tones are associated to two types of TBUs: accented **syllables** and **boundaries** of some higher domains in the prosodic hierarchy, namely the Intonation Phrase (IP) and the intermediate phrase (ip). The former type of association is normally referred to as *central* and the latter as *peripheral*. I illustrate this in (1) below with a phrase taken

---

\*I would like to thank Phillip Backley, Jill House, John Maidment, Kuniya Nasukawa and Toyomi Takahashi for their valuable comments and discussion of parts of the material contained in an earlier draft of this paper.

from Pierrehumbert (1980) and with a highly simplified version of her analysis included in her later work with Beckman (1988)<sup>1</sup>.

(1)



In figure 1 (see appendix) I illustrate the F<sub>0</sub> trace expected for the utterance *Another orange*. At the beginning of *A-* pitch is relatively high, and then it drops to a relatively low level in order to interpret the initial H(igh) boundary tone and the L(ow)<sup>2</sup> tone associated to *-no-*. After that, the F<sub>0</sub> trace rises gradually to reach relatively high pitch, and thus the H tone associated to *o-* in *orange* is interpreted. Finally, pitch falls sharply to relatively low pitch and the last two tones (L- and L%) are interpreted.

In addition to central and peripheral association, Pierrehumbert and Beckman make use of what they call a *secondary attachment*: in their description of Japanese, tones which are associated with a mora, can also be associated with nodes above the word. Grice also suggests a similar type of association for Palermo Italian. I shall not go into any details about secondary attachment here, since they claim that this type of hybrid association does not take place in English.

The major concern of this paper is to argue that, with association to two different categories of TBUs, there is no way of making any fundamental generalisations about tone-association, because at least two independent phonological operations need to be stipulated for each type of association. It is preferable to claim that tone uniformly associates to a single type of TBU through the whole phonological structure, either to a central unit or to a peripheral unit. Bearing in mind this assumption, I shall

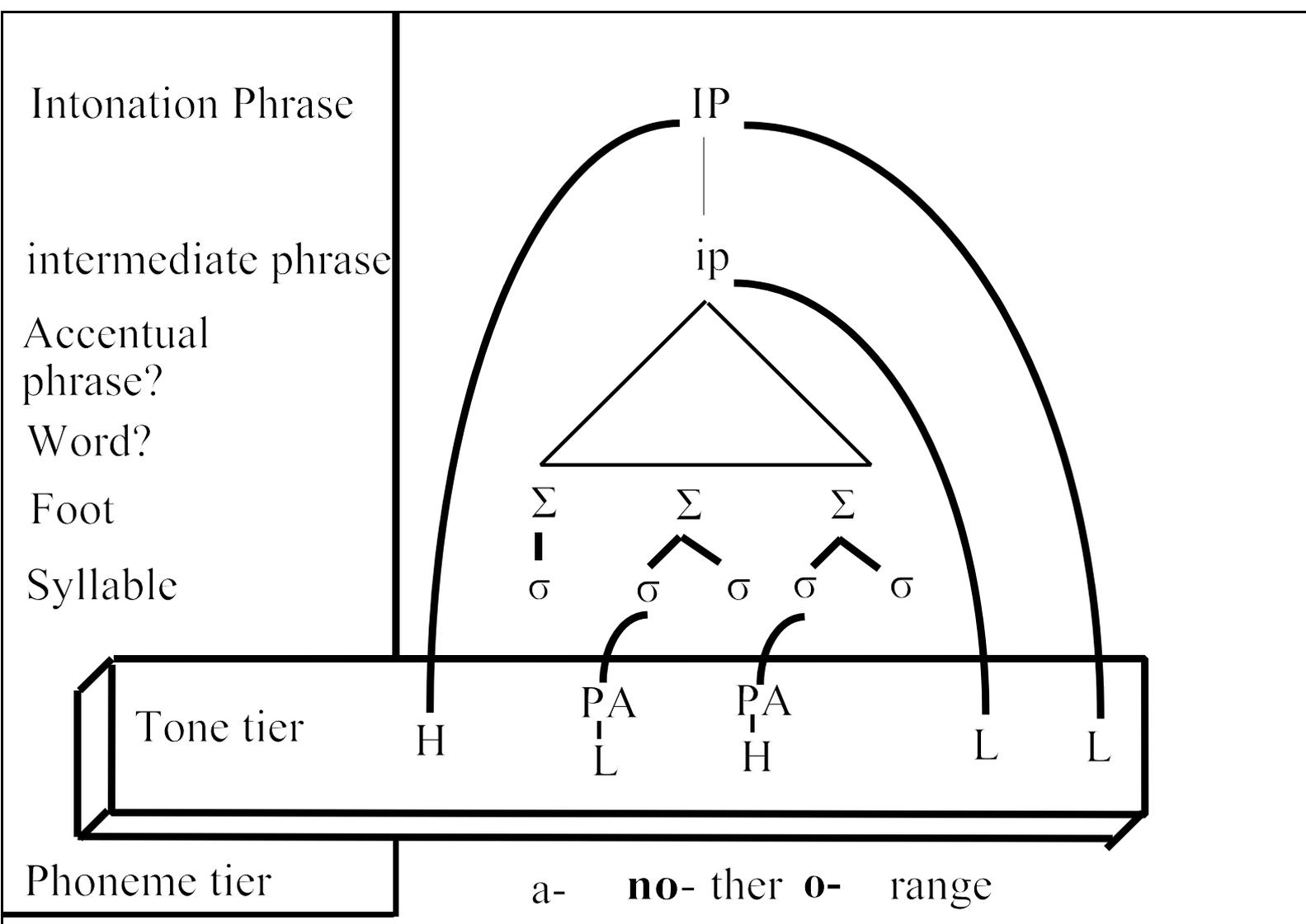
<sup>1</sup>Pierrehumbert and Beckman (1988) associate tone to morae in their account of Japanese. Hirst (1983) takes the stressed foot as a TBU. These types of association can be grouped together under the heading of central association.

<sup>2</sup>Pierrehumbert (1980) uses diacritics to signal boundary tones (%), tones mapped onto accented syllables (\*), and phrase-accent tones (-), which are mapped onto unaccented syllables.

investigate how to arrive at a way of associating tone in a single fashion, in order to establish the framework of a **restrictive** theory of phonological representation. By *restrictive* I mean a context in which overprediction is avoided and a certain degree of generalisation is achieved. It will turn out that the best way to satisfy my theoretical assumption is to associate tone to the boundaries of prosodic domains exclusively.

Let us now focus on some of the details of Grice's model. There are various structural components in the phonological representation of the phrase: the prosodic hierarchy, the tonal tier, the association lines and the phoneme tier. I shall discuss each of these in turn, with the exception of the phoneme tier, since this lies beyond the study of intonation. As can be seen in (2), the prosodic hierarchy is made up of the following domains: *IP* (Intonation Phrase), *ip* (intermediate phrase), Accentual phrase, Word, Foot and Syllable; however, it is not clear exactly how many domains need to be recognised between the *ip* and the foot. This is the reason why the accentual phrase and the word are followed by a question mark.

(2)



In her description of English and Palermo Italian, Grice presents evidence for the word as a prosodic domain in the latter language, but the status of the accentual phrase remains unclear in both languages. All this seems to indicate that, although languages can be described in terms of the prosodic hierarchy<sup>3</sup>, they differ as to the actual levels or domains in the hierarchy, and the use they make of them. The issue related to which constituents are necessary for an account of intonation patterns falls outside the scope of this paper, since my principal aim is to put forward a **unified**

<sup>3</sup>See Nespor and Vogel (1986) for a detailed account of the prosodic hierarchy of Greek, Italian, Latin and various other languages.

**association of tone** to TBUs. Nevertheless, I shall make an informal use of labels such as *foot* and *intonation phrase* as it seems that two prosodic domains suffice in order to achieve an efficient account of a wide variety of pitch patterns at the present stage of investigation. The reader is referred to the main sources (Selkirk 1984, Nespor and Vogel 1986, Pierrehumbert and Beckman 1988, and Grice 1992) for their own accounts of the prosodic hierarchy.

Another structural component in (2) is the tonal tier, which stands on a plane independent of the prosodic hierarchy, and adds to the autosegmental flavour of the model. The example illustrated above contains the simplest intra-tonal structure available, which is a single tone, either H or L. There can be other complex intra-tonal structures which are associated centrally through P(itch) A(cent)S and which may have, for example, two different tones (L+H\*<sup>4</sup>), or even three (HL+H\*). I would like to emphasize that this characteristic is exclusive to tones which associate centrally; tones which associate directly to the edges of the *IP* and the *ip* lack this enriched structure.

One of the shortcomings of this system is that it is impossible to make any kind of generalisation as to the occurrence of phonological phenomena which involve a specific tone structure. For instance, Pierrehumbert and Beckman (1986) contend that the stepping effect (the lowering of one tone with respect to the preceding one) is triggered by a bitonal PA. According to this claim it would be impossible to find a structure which triggers stepping between two *IP*s, since bitonal patterns are excluded from this environment; they are only found phrase-internally. But, in fact, it is perfectly possible to find such phenomena among *IP*s as well as phrase internally, and yet the present state of the model prevents us from capturing this generalisation. Consider the following sequence of three *IP*s: *Another apple, another orange and another banana* which could very well be uttered in a stepping fashion as illustrated by the F<sub>0</sub> trace in figure 2 in the appendix. Now consider a single *IP* like the one in *A big shaggy caterpillar*, but made up of a sequence of stepping pitch accents (see figure 3 in the appendix). According to Pierrehumbert (1980), the sequence of steps in the *IP*-set is not to be accounted by phonological rules, whereas the steps in the *ip*-set are attributed to the presence of a L tone in phonological representation.

---

<sup>4</sup>The asterisk adjacent to H or L is a shorthand device to show prominence relations between the tones. As I said in footnote 2, the starred tone is mapped onto the accented syllable and the unstarred tone leads it. The alternative option whereby the starred tone is trailed by the unstarred tone is also possible.

One of the fundamental tenets of Grice's model is that she formalizes this relationship between tones by labelling them as strong or weak, in the same fashion as constituents in prosodic hierarchy. In other words, she makes the prediction that the same organisation that is available in prosodic structure is also present in tone structure.

But surely, the same phenomenon **is** taking place in both examples, and therefore this should be captured by means of the same phonological representation. The only difference between them is that the stepping effect manifests itself in two different domains; at the *IP* level in *Another apple, another orange, another banana*, and at the foot level in *A big shaggy caterpillar*. Quite clearly, Grice's model is, as it stands, far from capable of capturing this generalisation.

In addition to this flaw, nothing in the theory can prevent us from assuming that there may be other languages in which the reverse type of association is found; that is, systems in which rich structures of tone are associated peripherally and conversely, where impoverished structures are associated centrally; until now, there has been no evidence which supports this prediction. Thus, Grice's model (and also Pierrehumbert and Beckman's) can also be criticised on the grounds that it allows for the prediction of unattested tonal structures.

Given that I intend to adhere to the most restrictive representation, in this paper I shall assume that phonological pitch patterns can be accounted for exclusively by means of single tones. Under this approach, the two problems relating to the specific case of stepping and to the model's potential for overprediction are avoided. However, there still remains the question of how to represent stepping, now that the model lacks bitonal structures. In addition, whatever solution is proposed, ideally, it must have the capacity to generalise about the occurrence of such a phenomenon. The reader is referred to Ladd (1990) for a proposal of how this could be done in terms of a metrical relationship between nodes of trees. Nevertheless, even if this problem is solved, it may be the case that this impoverished tonal organization is still insufficient to provide an explanation for an exhaustive range of pitch patterns at this stage. However, those which cannot be accounted for now, will be the subject of future work. Finally and most crucially, let us concentrate on the actual association of tone to TBUs.

As illustrated in (2), tone is associated to two categories of TBUs (which define peripheral and central association): edges and accented syllables. There are a H tone and a L tone associated to the left and right **edges** of the *IP* respectively, and also, there is a L tone associated to the **right edge** of the *ip*. This is shown by means of curved lines connecting tone to the sides of the phrase nodes; on the other hand, the remaining tones are associated to **accented syllables** (*no-* and *o-*); this is shown by short curved lines connecting tone (through PAs) to the syllable nodes. In order to capture the difference between accented and non-accented syllables, Grice has recourse to the labelling of syllables as strong or weak respectively, together with the representation of branching and non-branching structures (I have excluded these details from the illustration in (2) for the purposes of clarity in that diagram. A simplified version and an extensive discussion is included in §2.1).

In the higher levels of the hierarchy, there is already a factor which, once again, prevents us from making any kind of generalisation as to the association of tones: a given tone has to be associated to the rightmost edge of the *ip* only, and this type of constraint does not apply to the association with the *IP*<sup>5</sup>; at this level, tone is obligatorily associated to the right edge, but optionally associated to the left edge; (the reverse argument also applies: the constraints of tone-association to the *IP* do not apply to the *ip*). Thus, any statements that could be formulated about associating tone in English will require a different set of conditions depending on whether such a tone is associated to the *IP* or the *ip* (and it seems that such conditions will be far from straightforward). In addition, an extra set of conditions will be needed depending on whether tone is associated to boundaries or to syllables, (and even within syllables, an extra mechanism is still necessary in order to single out those which will be accented from those which will remain unaccented). In short, in this model, at least four independent sets of conditions are necessary for tone-association.

All this leads back to the main question in this paper, whether it is really necessary to have two completely different categories of TBUs in the phonological representation of intonation contours. In the context of a restrictive methodology, it is reasonable to investigate the possibility of applying the same procedures in the association of tones throughout the whole of the prosodic hierarchy, and thus, to bypass having to posit an array of conditions operating at each prosodic level. In the framework of Grice's model, this could be accomplished in two different ways: either by suggesting that the conditions in central association are also applicable to peripheral association; or alternatively, by proposing that the conditions in peripheral association operate everywhere in the hierarchy. Notice that the latter option is more restrictive in the following ways: first, tonal structure is highly constrained to a single tone; and secondly (as I shall suggest in §3), peripheral association does not require two independent mechanisms (*s(trong)/w(eak)* labelling of syllables and branching non-branching tree-structures) for the association of tones. In other words, peripheral association can be said to be more economical. In view of these advantages of peripheral association over central association, the former may well be pursued.

In brief, all these issues related to peripheral association entail the following assumptions:

- (3) (a) Tone is no longer associated to syllables **in** prosodic domains, but instead all tones are associated **to** the boundaries of prosodic domains.

---

<sup>5</sup>Another difference is that, unlike peripheral association, central association requires syllables to be labelled *s(trong)* or *w(eak)*, so that tone 'knows' which segments it has to associate to. This issue is discussed in §2.

- (b) Tonal structures are highly constrained.
- (c) Prominence relations are signalled by means other than strong/weak labelling of syllables.

As we shall soon discover in the following sections, by strictly adhering to (3a) the model's potential to generalise is boosted: the principles of association are the same everywhere in phonological representation and those phenomena which formerly appeared to behave differently (for instance stepping) can now be said to be the same, only that they manifest themselves at different levels in the prosodic hierarchy. Furthermore, by (3b) the model's generative power is kept to a minimum and therefore, under control. Finally, as I shall explain in the following sections, by (3c) the model is highly constrained in the sense that peripheral association is free from a battery of stipulative rules which are necessary in central association. In fact, peripheral association is subject to a set of principles which are much more general than rules, and which do not allow for overprediction.

Another point in favour of peripheral association is that its interpretation at the phonetic level closely reflects phonological structure. Because of this, there is a better chance of mapping phonological representation directly onto phonetic interpretation, without running the risk of having to suggest the alleged existence of an intermediate level in which interpretation rules would operate. There is ample evidence in the literature (cf. Hirst 1988) which suggests that F0 characteristics are not simply mapped onto the vocalic segment of the accented syllable, but rather that they are a property of a series of segments. It seems to me that these segments may very well correspond to those enclosed in a metrical domain in phonological representation. In the specific case of a smaller domain such as the syllable, it might be reasonable to map tone to the entire syllable domain; similarly, in larger domains like the foot, pitch characteristics are manifested throughout the whole domain. If this is how pitch characteristically behaves, there might be adequate motivation for the association of tone to the boundaries of the prosodic domain, as an indication that those pitch features belong to the entire domain. In this way, phonetic details are captured in a more realistic way.

So far, I have explained the components of Grice's model and, at the same time, I have argued for an extensive revision of some of them; I have rejected central association and I have cast some doubt on the use of the metrical tree for the representation of metrical relations, albeit without any full discussion.

## **2 Towards a single-type of association applied to Grice's model**

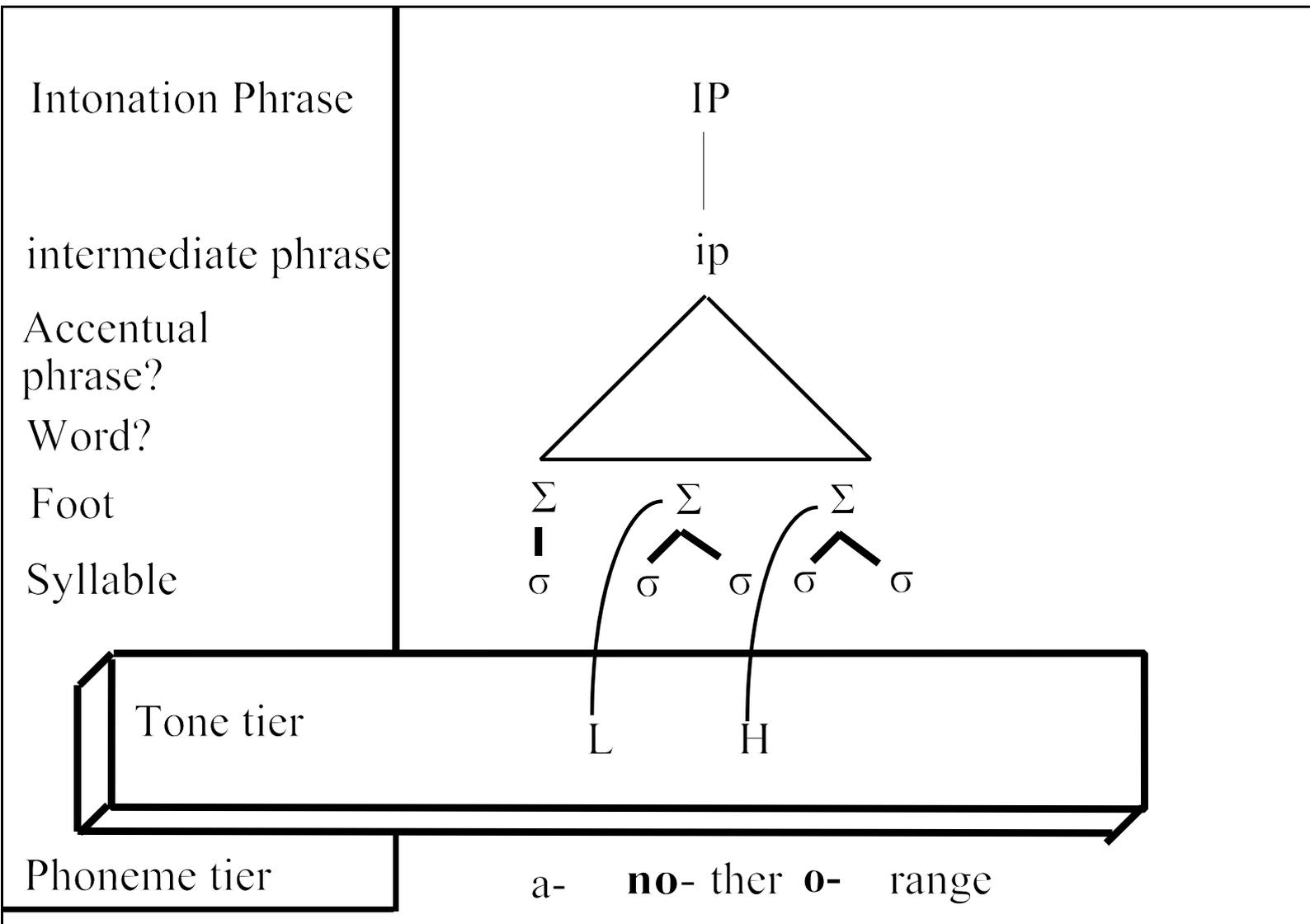
In the proposal I shall put forward below, peripheral association remains unchanged from Grice's model, although a uniform association to the *IP* and the *ip* is still needed. In addition, a principled account of tone-association at these levels remains to be made explicit. This is the topic addressed in §2.2, where I attempt to produce such an account by means of tree-structure, although, as I shall show, without success. As for tone-association to lower levels in the hierarchy (§2.1), tone is no longer associated to accented syllables (as in (2)), but instead, can only be associated directly **to** the edges of a prosodic domain—in this case, the foot<sup>6</sup>. In order to achieve this, it will turn out that prominence relations between constituents need to be expressed. The widespread mechanism used in previous works, and indeed the one suggested by Grice herself, has been the labelling of constituents as strong or weak; therefore, later I shall make an attempt at combining strong/weak labelling with peripheral association to the foot. Let us now look at a preliminary indication of association to feet, which is illustrated in (4) below.

---

<sup>6</sup>In the major work of which this paper forms a part, I suggest that it is possible to present a unified account of the so-called *Tone*, *Pitch Accent* and *Intonation* languages by means of associating pitch characteristics (tone) to different domains in the prosodic hierarchy, which roughly correspond to the syllable, the foot and the intonation phrase, respectively.

2.1 Association to feet

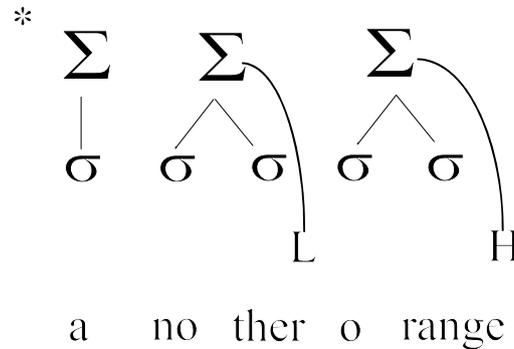
(4)



A question immediately arises as to whether the L and H tones which are associated to the left side of the foot in (4) can also be associated to its right side. Given that in both feet (*no-ther* and *o-range*) the leftmost syllable is the prominent one, then it seems intuitively correct to associate tone to the left side. In addition, the former option (to associate tone to the right side of the foot as in the simplified version in (5)) seems to be counterintuitive since it looks as if tone were to be understood as

belonging to the rightmost syllable, which is clearly a non-prominent syllable, and thus does not count as a legal landing site for tone:

(5)

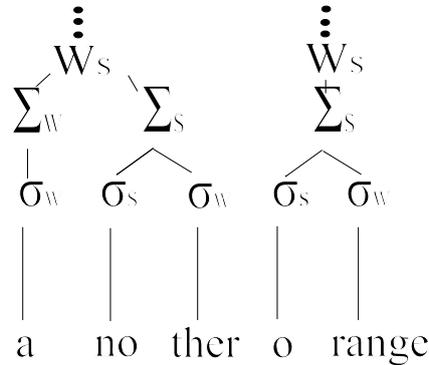


On the other hand, if the right syllable of a foot were the prominent one, then the right association represented in (5) would be perfectly acceptable. Hence, we are led to conclude that association is highly dependent on the definition of prominence relations among syllables within a foot<sup>7</sup>. If this is the case, then some way of formally representing prominence relations in the model is required. This has been done in various ways, the predominant one (Lieberman 1975, Lieberman and Prince 1977) being by means of treating syllables as *s*(trong) or *w*(eak) depending on whether they are prominent or non-prominent, respectively. For instance, *Another orange* is analysed as follows:

---

<sup>7</sup>This assumption begs the question of how to treat a case in which tone must be associated to an accented monosyllabic foot, as in for instance, *Yes*, where there is no prominence relation between two syllables which might help to decide to what edge tone should be associated. At this stage, it seems that tone could be associated to any side, without this having any effect on the overall pitch pattern. However, as will become clear later when the grid-format is used, association to the right or left edge of a domain will have immediate consequences for the resulting intonation contour.

(6)



Those syllables which are prominent (*-no-* and *o-*) are labelled *s*, and non-prominent syllables are labelled *w* (*a-* *-ther* and *-range*). At the same time, syllables are grouped into feet, which are also labelled *s* or *w*. Normally, nonbranching feet are labelled weak, and those feet which branch are labelled strong, although nonbranching feet can also be labelled strong. In order for a constituent to branch or not to branch, a series of conditions need to be met. For a comprehensive outline of such conditions see Hogg and Mc Cully (1987). At the foot level there are now two competing feet which show the same amount of prominence: *-nother* and *orange*. With the purpose of choosing the more prominent of the two, a further level is constructed which, according to the representation in (2), is the Word. As stated by tree-building-rules, more levels would have to be introduced in the hierarchy (this is shown in (6) by means of dots over the Word nodes), until only one single constituent is promoted onto the highest level, and is labelled *s*; in other words, there has to be an ultimate node. The constituent which is labelled *s* all through the hierarchy is the most prominent one in the entire structure; in (6) this is *o-* in *orange*. The whole of this procedure is normally performed on the basis of prominence-relations and rhythm rules (Selkirk 1984) which I have excluded from the above explanation (see also Liberman and Prince 1977, Hayes 1995 for other ways of building up the tree).

Now that prominence relations have been made explicit in the representation, it is understood that if tone is associated to a strong foot<sup>8</sup> which contains a weak and a strong syllable, then the aforementioned tone will be interpreted on the segments belonging to the strong syllable. At the phonological level, whether there is right or left association of tone to a strong foot is irrelevant, as long as the syllables in the foot

---

<sup>8</sup>It is understood that no tone is associated to the weak foot *a-* because it contains a weak syllable which cannot bear an accent.

are properly labelled. In other words, the decision as to which edge of a foot a tone is associated to depends on foot-internal syllable-prominences; the *s/w* labelling of a syllable conditions tone-association in the following way<sup>9</sup>:

(7) In a bisyllabic foot, tone is associated to that syllable which is labelled strong.

Given this, the structure in (4) rather than the one in (5) shows the proper association of tones, and therefore (5) is immediately ruled out.

As I said above, underlying the illustration of tonal association to the edges of feet, there is a battery of powerful rules in charge of defining *s/w* relations as well as branching versus non-branching constituents. I have informally referred to these rules, without elaborating any further on them, since I regard them as highly stipulative. In addition, from the point of view of an economical theory of intonation, having recourse to two independent mechanisms, *s/w* labelling and branching/non-branching trees, is too costly as a vehicle for the representation of pitch patterns. Ideally then, at the level of the foot, it would be preferable to reject stipulative rules and to develop general principles to determine tone-association, and also to simplify the dual task of *s/w* labelling and tree construction into a unified representation. For these reasons, I reject tree-format at this level.

Before completely rejecting the combination of tree-format with peripheral association from the model, it must be firmly established that this option is also defective at the *ip* and *IP* levels. This is the topic of the following section. Then, after having gathered sufficient evidence against this type of model, such a format will be discarded.

## 2.2 Association to levels above the foot: *ip* and *IP*

We ought to be suspicious of the fact that *s/w* metrical relations and tree configuration are an option which is unavailable at the level of the *IP* and the *ip* in Pierrehumbert's or Grice's model. As it will soon be discovered, the main reason for excluding this alternative is that it leads to further undesirable stipulations.

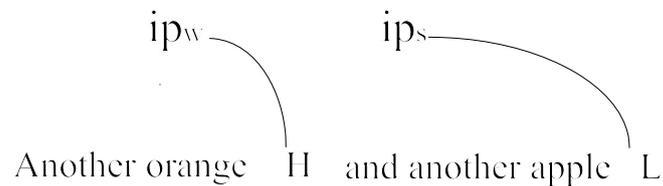
Having the peripheral association shown in (2) in mind, let us begin our task by making an attempt to apply the tree structure to a couple of *ips*, like the ones in (8) below: *Another orange* followed by a H tone, and *and another apple* followed by a

---

<sup>9</sup>But see Selkirk (1984) for a completely different view of the relationship between stress and pitch accents, which is mainly characterised by the idea that pitch accents are defined prior to stress relations.

L tone. I have chosen to show two of them so that the *s/w* relationship can be captured in a better way. The rightmost *ip* is labelled *s* in the same fashion as the rightmost foot was labelled *s* in §2.1 above, and consequently, the leftmost is labelled *w*. Once they are properly labelled, I can now proceed to associate tone as stipulated first by Pierrehumbert for Japanese and then by Grice for English: since there are two tones and each one has to be associated to the right edge of the *ip*, then each *ip* (whether it is strong or weak) must have a tone associated to it<sup>10</sup>.

(8)



But the association of L to the weak *ip* violates some of the conditions on association to feet which were mentioned above. Recall that only strong constituents count as landing sites for tone. Thus, I am confronted with having to make a choice between two alternatives: either to re-formulate tone-association in such a way that association to weak constituents is sanctioned at the *ip* level, or to suggest that a metrical representation based on *s/w* labelling is inadequate as far as it overrules tone association, and therefore, must be discarded. Before arriving at a decision, however, it is worth pointing out that in Grice's model, the fact that tone cannot associate to the *ip*'s left edge, but has to associate obligatorily to the right edge, is based on mere stipulation and is not theoretically motivated.

Let us now investigate what the situation is like for the *IP*, since all this might shed some light on which of the above options should be chosen. In Grice's model, there is no motivation for *s/w* labelling, since there are no prominence relations to be captured. Another aspect to take into consideration is that tone has to associate to the right edge obligatorily, but optionally to the left edge.

The above description of tonal behaviour indicates that the restriction imposed on the association of tones to edges of the *IP* is completely different from that imposed

---

<sup>10</sup>The possibility of bitonal association to *ips* is not considered by Pierrehumbert since there is no evidence in the data suggesting that this might be so. However, as I said in §1, there is nothing in the theory which would prevent us from doing so.

on the edges of the *ip*: (optional) left and (obligatory) right association for the former, and only (obligatory) right association for the latter. Unfortunately, these different restrictions (which seem to be universal) are not theoretically motivated, and yet they seem to constitute a crucial constraint in the system.

From an endeavour to combine a metrical tree-format with peripheral association in the analysis of intonation contours then, the following points can be concluded:

- (9)
- (a) the postulation of *s/w* labelling, and branching and non-branching structures, entails having recourse to stipulative rules;
  - (b) the design of two independent mechanisms (labelling and tree structure) for an account of intonational patterns is too costly for a model which belongs in a restrictive theory;
  - (c) at the *ip* level, *s/w* labelling seems to overrule tone association;
  - (d) *s/w* relations vanish at the *IP* level
  - (e) the alleged rules underlying the association of tone to some levels, namely to the *ip* and to the *IP*, are arbitrary and not theoretically motivated;
  - (f) due to (e) the model's ability to generalise is greatly diminished.

In view of all this evidence against this type of model, I suggest that the mechanisms for showing metrical relations and the arbitrary tone-association to *IPs* and *ips* can no longer be seriously maintained. Now, this opens the way to the proposal of an alternative framework<sup>11</sup> which offers solutions to all these points. This constitutes the topic of the following section.

### 3 Metrical grid

The principal aim of this section is to suggest that the alternative to the tree-format model is made viable by the use of the metrical grid and its inherent flat-format (against an arboreal structure), and that many of the inconveniences created by the former format can be overcome by the latter. Due to space limitations, I shall not develop a complete account of grid construction here. The reader is referred to

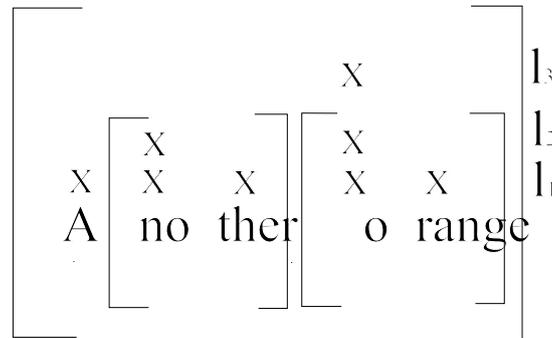
---

<sup>11</sup>The reader is referred to Hirst (1988:51-165) for a very attractive model which is mainly characterised by peripheral association. Unfortunately, his mechanism for generating prosodic boundaries is obscure, in the sense that he does not follow any kind of theoretical principle.

Selkirk (1984) and to Halle and Vergnaud (1987) as the main sources of this topic,<sup>12</sup> and to Cabrera-Abreu (1994) for a preliminary attempt at its application in the description of pitch patterns. What I shall do instead, is to illustrate the grid representation of *Another orange*, and I shall also discuss some issues connected to the representation of prominence relations.

In (10) each head (*head* as understood in the framework of Government Phonology, Kaye et al. 1990, Charette 1991, Harris 1994) of the syllable (that is, nucleus) is indicated by a position *x* on the lowest row<sup>13</sup>. Then, the positions which act as the heads of feet are projected onto the row one level higher.

(10)



This results in the projection of *no* from *another* and *o* from *orange* onto level 2. The domains corresponding to these feet are graphically represented by the inner boundaries, and they result in left-headed, bounded feet (Halle and Vergnaud 1987). From this it follows that each new foot begins with a position on level 2. Unlike Grice's model, *A* does not count as a foot since neither does this position itself act as a head nor it is integrated into a foot. At this stage of investigation, I suggest that positions before the head of feet are incorporated into the *IP* domain. Lastly, the head

---

<sup>12</sup>At first glance, it seems that Halle and Vergnaud also have recourse to a series of rules for grid construction. In my major work, I shall endeavour to improve the theory by proposing a more restrictive version of principles and parameters approach, mainly based on licensing principles.

<sup>13</sup>It is possible to assign boundaries to these heads, but for the purpose of illustrating intonational patterns in English, they are irrelevant. Nevertheless, this possibility opens the way to the representation of tonal characteristics in tone languages: it can be assumed that tonal characteristics are a property of syllable heads in Tone languages, and that they are a property of larger domains in Pitch Accent languages and of even larger domains in Intonation languages.

of one foot is projected onto the third level, and thus, the most prominent nuclear head of the entire structure is singled out (as I shall briefly discuss below, factors which determine which foot is projected are complex in English). This is the head of the intonation phrase, which constitutes the largest domain in the prosodic hierarchy; it is represented by the outer boundaries. The domain at level 3 is bounded and, unlike level 2 domains, is right-headed.

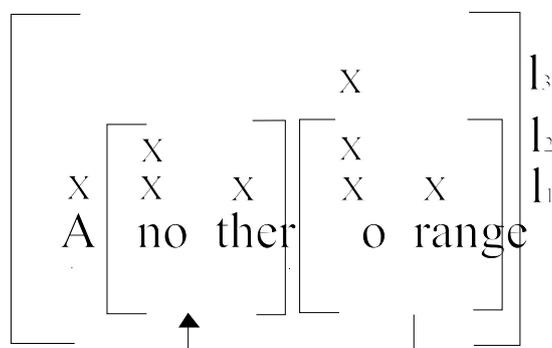
Given the above metrical description of *Another orange*, I claim that constituent boundaries arise as an automatic consequence of grid construction, unlike tree structures and *s/w* labelling, which are generated by two independent mechanisms. This counts as one of the advantages of the grid format over the tree format.

With the prosodic structure illustrated above, in which there are only two constituents (the foot and the *IP*), the problem related to tone-association to the *ip* being overruled by *s/w* labelling is circumvented (and indeed, any other issues related to the *ip*, as I shall discuss shortly).

At this juncture, it may be argued that the projection of *o-* in *orange* to the highest level looks like a rather arbitrary choice; for example, *-o-* in *another* may well have been chosen instead. A preliminary response to this question can be found in the framework of Metrical Phonology, where it is a general and straightforward assumption that sentence-stress is determined by repeatedly projecting the strongest position in the last syntactic domain until there are no other positions competing with it at a given level. Yet, with this type of assumption, we are once again driven by bare stipulations: why should the right syntactic domain be projected onto the next level, and not the left one; and furthermore, why should a syntactic domain be involved in metrical relations (see Selkirk 1984, chapter 4). In looking for a reasonable explanation for the first issue, I would like to draw a parallel between this type of situation and that found between the onset and the nucleus constituents in the context of Dependency Phonology (Anderson and Ewen 1987) and Government Phonology. In the framework of this type of phonology, the fact that the nucleus is always obligatory is symptomatic of an asymmetric relationship between the onset (which is optional) and the nucleus, whereby the nucleus is the licensor (or head) and the onset is the licensee (or complement). Hence, it is said that the nucleus and the onset are in a licensing relationship. Similarly, it can be said that a licensing relationship of this type is found between the two feet under discussion; the fact that the rightmost is obligatory is symptomatic of it being the licensor (or the *nuclear* foot) and the leftmost (which is optional) being the licensee (or the *onset* foot). *O-* in *orange* is projected onto level 3 because it belongs to the *nuclear* constituent. Licensing at the level of the foot is from right to left, or head final; this is shown by means of the arrow in (11):



(11)



It might be the case that in other languages licensing relationships at the level of the foot are reversed, that is, from left to right. If this line of investigation is pursued, it may well be possible to establish some general principles of licensing which are universal and therefore, part of Universal Grammar, and also to set parameters to which languages conform. In this way, languages which have been previously described as completely different in terms of their metrical patterns, would then be described as rather similar in that they obey the principles of licensing. Any differences would arise as a consequence of the selection of a different setting of the parameter.

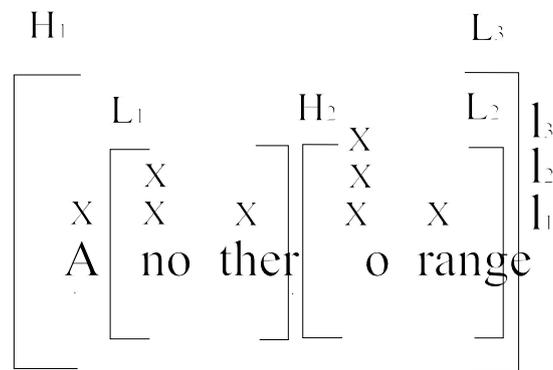
Another advantage of using the principles of licensing over the rules of *s/w* labelling is that the former conform to a set of principles and parameters (left/right headship) which operate throughout phonological representation, from the highest constituent to the lowest, whereas the latter seem to work satisfactorily only for levels below the foot, but unsatisfactorily for levels above the foot (recall that in §2.2 I pointed out that *s/w* labelling may overrule tone association to the *ip*. It appears that licensing has no immediate effect on tone association).

In the previous paragraph, I have suggested that one of the fundamental tenets of the phonological representation of intonation which I propose is the principle of licensing and the parametric variations that come with it. However, I have put forward a sketchy presentation of how this can be achieved, since this will be further formalised in the major work of which this article forms a part.

### 3.1 Tonal association and the metrical grid

Let us now proceed to associate the tones in (2) to the boundaries of the structure in (10); the sequence  $H_1 L_1 H_2 L_2 L_3$  (I have numbered the tones only for the purposes of clarity) needs to be incorporated into the representation. This is tentatively done in (12)<sup>14</sup>.

(12)



$H_1$  and  $L_3$  are associated to the outermost boundaries since they are a property of the *IP*.  $L_1$  and  $H_2$  are associated to the remaining left boundaries which enclose prominent feet. I have decided to associate  $L_2$  to the right boundary of the right foot provisionally as no other place is available. This is due to the following reasons: first, grid-construction does not allow for an alleged *ip* constituent and second, I assume that multiple association of tones to boundaries is not favoured in the restrictive framework which I have adopted. Hence, all the problems involving the *ip* which were encountered in Grice's and previous models, are now avoided. In addition, following claims made in the literature (Ladd 1983, Lindsey 1985) which cast some serious doubt about the phonological status of this tone, in subsequent work I shall also suggest that it should be excluded from phonological representation altogether.

I admit that, at this stage, the mapping of tones to boundaries is only descriptively adequate, and that it lacks any explanatory power; for instance it is still not clear why

---

<sup>14</sup>In the model I propose, there seems to be no need for association lines. As long as tone is placed over a boundary, its interpretation is quite straightforward: its effect will be manifested within the boundaries of that domain. Nevertheless, for purposes of clarity, in the appendix, I have included association lines. The issue of whether association lines are indeed a phonological entity is controversial (see Takahashi 1994) and is outside the scope of this paper.

only one boundary is left without association (instead of following Universal Association Conventions (Goldsmith 1990) whereby no TBUs are left unassociated in phonological representation).

A preliminary phonetic interpretation of (12) is now straightforward: the default value for pitch is to be relatively low (see Cabrera-Abreu 1994 and references therein for the reasons why relatively low pitch is taken as the default value), unless a H tone associated to one of the boundaries is found, which indicates that pitch is relatively high<sup>15</sup>.

Finally, notice that in (12), boundaries can be associated to H, to L, or they can remain unassociated (let us informally refer to this alternative as  $\emptyset$ ). This may lead to the assumption that in the model I propose there is a three-way contrast in phonological representation: H, L and  $\emptyset$ . However, it must be made clear that this assumption is totally misleading. As can be seen in the phonological structure proposed for *Another orange*, whether a boundary is associated to L or  $\emptyset$ , its interpretation remains the same: as relatively low pitch. Unless motivation is found in support of the idea that L is a phonological prime<sup>16</sup>, I shall adhere to the proposal put forward by Cabrera-Abreu and Takahashi (1993) and also by Cabrera-Abreu (1994) that phonological pitch patterns can be accounted for with a two-way contrast represented by the presence/absence of H. Under this view, the aforementioned examples can be re-analysed in terms of their boundaries being associated to H or remaining unassociated.

Bearing in mind the latter proposal, in the appendix I illustrate how this model can account for other pitch patterns which are associated to the same grid configuration. In addition, I include some examples of similar intonation contours mapped onto different grid configurations (the F0 traces of such structures, as well as Pierrehumbert's own account can be found in Pierrehumbert (1980).

---

<sup>15</sup>For a similar phonetic interpretation of a structure without L tone, see Cabrera-Abreu and Takahashi (1993).

<sup>16</sup>Pierrehumbert and Beckman (1988) argue that L is needed for the representation of stepping; in fact, they claim that a bitonal pitch accent made of the sequence H+L triggers downstep. On the other hand, some doubt is cast upon this argument by Ladd (1993), who suggests that stepping arises as a consequence of a shift in register, which he represents by means of tree structures. In any case, the crucial issue here is that L is no longer necessary for the representation of stepping; hence, the stepping effect no longer counts as evidence for the need of L in phonological representation.

## 4 Conclusion

The advantages of using the grid format over the tree format in the phonological representation of intonation are the following:

- (13) (a) the grid is subject to a set of general principles, whereas *s/w* labelling and branching/non-branching structures have recourse to stipulative rules;
- (b) grid-structure and its automatic generation of boundaries avoids having recourse to two independent mechanisms (labelling and tree structure) for the analysis of intonational patterns;
- (c) the principles of grid-construction are not overruled by tone association;
- (d) in the model I propose, there is no motivation for the *ip*; in this way, the problems relating to the model's inability to generalise about tone-association are circumvented;
- (e) overall, my model is constrained in all its structural components: the prosodic hierarchy is minimised to two levels; and the tonal tier is restricted to a single tone.

## References

- Anderson, J. and Ewen, C. (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.
- Beckman, M. and Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook 3*.
- Cabrera-Abreu, M. and Takahashi, T. (1993). A single tone representation of intonation. *Speech, Hearing and Language: work in progress UCL*. Vol 7.
- Cabrera-Abreu, M. (1994). An alternative model for intonation. *Speech, Hearing and Language; work in progress: UCL* Vol 8.
- Charette, M. (1991). *Conditions on phonological government*. Cambridge: Cambridge University Press.
- Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.
- Grice, M. (1992). *The intonation of interrogation in Palermo Italian; implications for intonation theory*. PhD dissertation. London: University College London.
- Gussenhoven, C. (1983). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
- Halle, M. and Vergnaud, J. (1987). *An essay on stress*. Cambridge, Massachusetts: The MIT press.
- Harris, J. (1994). *English sound structure*. Oxford: Blackwell.
- Hayes, B. (1995). *Metrical stress theory*. Chicago: Chicago University Press.
- Hirst, D. (1983). Structures and categories in prosodic representations. In A. Cutler and R. Ladd (eds.), *Prosody: models and measurements*. Berlin: Springer.

- Hirst, D. (1988). Tonal units as constituents of prosodic structure: the evidence from English and French intonation. In H. van der Hulst and N. Smith (eds.), *Autosegmental studies on pitch accent*. Dordrecht: Foris.
- Hogg, R. and McCully, C. (1987). *Metrical phonology: a coursebook*. Cambridge: CUP.
- Kaye, J., Lowwinstamm, J. and Vergnaud, J. (1990). Constituent structure and government in phonology. *Phonology* 7.
- Ladd, R. (1983). Phonological features of intonational peaks. *Language* 59.
- Ladd, R. (1990). Metrical representation of pitch register. In J. Kingston and M. Beckman (eds.), *Papers in laboratory phonology I*, Cambridge: CUP.
- Lieberman, M. (1975). *The intonational system of English*. PhD dissertation, MIT, Massachusetts.
- Lieberman, M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8-2.
- Lindsey, G. (1985). *Intonation and interrogation: tonal structure and the expression of a pragmatic function in English and other languages*. PhD dissertation, University of California, Los Angeles.
- Nespor, M. and Vogel, I. (1986). *Prosodic Phonology* Dordrecht: Foris.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. PhD dissertation: Bloomington, Indiana University Press.
- Pierrehumbert, J. and Beckman, M. (1988). *Japanese tone structure*. Cambridge, Massachusetts: The MIT Press.
- Selkirk, E. (1984.) *Phonology and syntax: the relation between sound and structure*. Cambridge, Massachusetts: The MIT Press.

Appendix

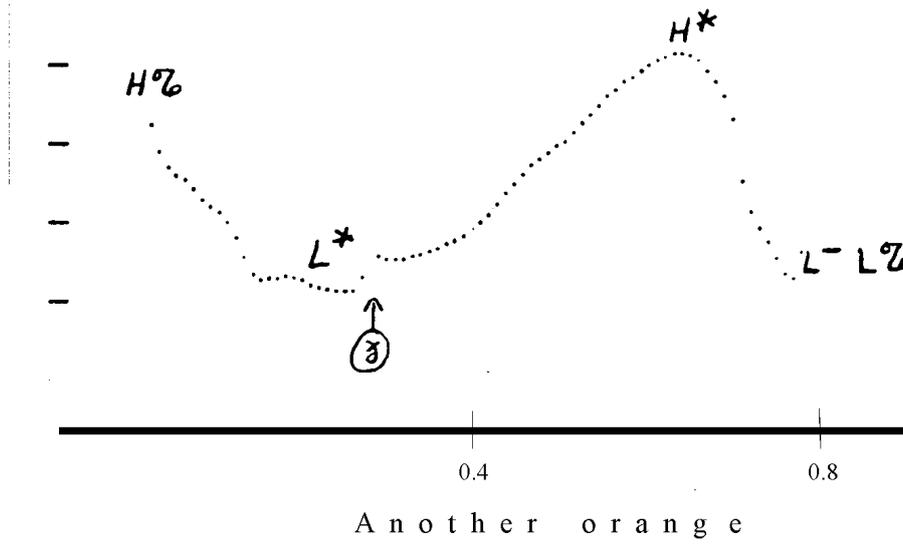
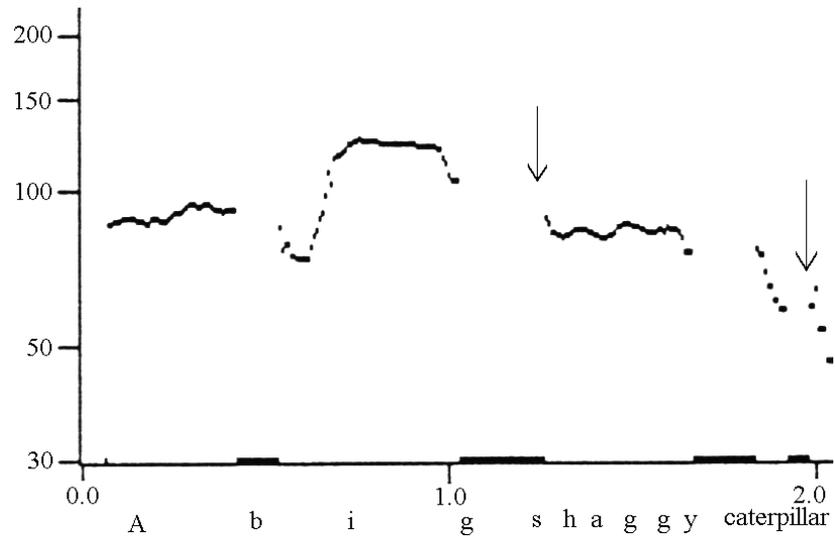


Figure 1



**Figure 2**

(The arrows indicate where the steps take place.)



**Figure 3**

(The arrows indicate where the steps take place.)

