A PHONOLOGICAL MODEL FOR INTONATION WITHOUT LOW TONE

Mercedes Cabrera-Abreu

Submitted for the degree of PhD in Linguistics

University College London

1996

ProQuest Number: 10045826

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10045826

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

To my parents and Kenny

Abstract

Using the evidence of intonational phenomena drawn mainly from English, I propose a phonological model for intonation characterised by a single tone (T), and boundaries of prosodic domains as Tone Bearing Units (TBUs). Such domains are defined in terms of licensing principles, which manifest themselves throughout the whole prosodic hierarchy.

In the framework of a restrictive theory of phonological representation, I argue that intonational pitch patterns can be accounted for with just T, and that L can be discarded. I defend this claim against evidence from earlier models which have recourse to L in the account of an extensive range of intonation contours, and of tonal phenomena such as contour-tones, tone-spreading, and downstep.

Contour-tones are defined by assigning different pitch specifications (presence/absence of T) to TBUs. Tone-spreading is treated as a matter of interpolation between boundaries with equal tonal association. Downstep is accounted for by the presence of toneless boundaries which correspond to empty categories. The phonetic interpretation of such categories (and toneless boundaries in general) is controlled by Gravitation Effect, a pulling down force which has a similar effect on domains to its right.

I support this view by subjecting earlier analyses to an examination based on maximally constrained principles in phonology. Under this approach, such analyses to some extent lose their validity, and no longer constitute a challenge for my model.

The extensive revision of previous models results in a parsimonious and homogeneous model, and at the same time, permits the formulation of phonological generalisations.

Acknowledgements

I owe a considerable debt to John Maidment and Toyomi Takahashi for their support, guidance, dedication, encouragement, and patience. It is hard to imagine how this piece of research would have looked without them.

Special recognition and gratitude go to Phillip Backley, whose painstaking, but highly critical review of the manuscript of this thesis led to numerous improvements both in form and content. I am especially grateful to John Harris and Kuniya Nasukawa on the one hand, and Martine Grice on the other, for lively discussions and suggestions on the phonological and intonational aspects, respectively, of the thesis. I also wish to thank Jill House for suggestions and criticisms of the first draft of this thesis. These turned out to be highly stimulating.

Thanks are also due to the members of the Department of Phonetics and Linguistics at UCL, especially to Judith Crompton and Steve Nevard, who helped me with technical and computer matters.

My final acknowledgements go to my family, especially to José Cabrera for his constant support and concern for my non-academic life and health; and to all my friends, especially María Luisa García Lecumberri for her moral support.

The work presented in this thesis has been partially supported by grants from the Gobierno Autónomo de Canarias, and the Fundación Universitaria de la Universidad de Las Palmas de Gran Canaria.

Table of contents

Chapter 1. Introduction

1 Scope of this thesis

1.1 From linear to non-linear phonology

The present work constitutes the pursuance of current trends in phonological theory to account for tonal phenomena in the intonation domain. It is fully embedded in the tradition of *Generative Phonology.*

The publication of the *Sound Pattern of English* (1968) (henceforth, *SPE*) by Chomsky and Halle, constitutes the foundation of a wide range of current phonological theories. The *SPE*-type of phonology is mainly characterised by a set of rules which operate on a **linear** arrangement of sound-segments (described in terms of bundles of binary features), and which describe a range of phonological phenomena which manifest themselves in the derivation of a surface from an underlying form.

A few years after the publication of Chomsky and Halle's influential work, the formalisation of an idea from the Firthian school, which offers an entirely different perspective of the nature of phonological theory, was developed as **non-linear** phonology (or *autosegmental*[1] phonology), covering a variety of frameworks including Government Phonology, Metrical Phonology, etc. One of the main tenets of the Firthian school is that sound properties should be allowed to span domains of varying sizes, rather than being restricted to segment-sized units. The motivation for this claim arises from the analysis of the properties of languages like Chinese in which some features should not be associated with a particular segment, but rather to a set of segments in a syllable. For example, *palatalisation* is better understood as the property of a syllable, rather than of the individual segments which comprise such a syllable (Firth (1937); quoted in Anderson (1985)).

Firth's views have had far-reaching consequences in the phonological study of languages, as they suggest a completely new approach to the organisation of phonological

[1]See example of spreading and contour-tone below, which describes the characteristics of an *autosegmental* model.

representations. Due to such influence, modern theories of phonology distinguish the following components in phonological representations: on the one hand, structural information defines the organisation of sounds into domains, and also demarcates the limits of such domains like, for example, syllables, feet, intonational phrases, etc. In an analogy with the human body, this can be understood as the skeleton, which organises flesh into what constitutes the limbs, trunk and head. On the other hand, the structure is independent of other properties like *palatality* or *roundness* — that is, those characteristics which define the quality of a sound, that is customarily referred to as *melody*[2]. In my analogy with the human body, this equates with its flesh, and all the properties that define it, like for example, hair, fat, skin, etc. Note that although structure and substance (and for that matter, skeleton and flesh) are intimately connected (as I shall explain shortly), they constitute independent components.

Another idea which stems from the Firthian school is related to the nature of the factor which establishes a link between the structure and the melody. In some current theories of phonology, specifically in Government Phonology, this factor is captured under *the principles of licensing* which control the organisation of both structural and interpretive units, and thus, define the connection between both components. Returning to the parallel with the human body, the principles of licensing correspond to the principles of chemistry, which govern the organisation of elements present both in our bones and flesh.

The departure of phonological studies from the *SPE* tradition, which is characterised by describing segments in terms of an unordered set of features grouped in a matrix, leads on to other studies which are characterised by assigning **independent** status to former features. Studies which follow the latter approach are said to be embedded in the framework of *Autosegmental Phonology* (Goldsmith 1976). The term *autosegmental* refers to the organisation of phonological features, which, under this view, enjoy a higher degree of autonomy from each other.

As an illustration of the difference between a linear and a non-linear representation, and of the motivation for the development of an autosegmental approach, I shall outline an example included in Carr (1993). For this purpose, he examines briefly

[2]See Goldsmith (1976).

the phenomenon of nasal[3] assimilation (or spreading) in Terena. In this language, forms like those in the left-hand column which express 'third person singular' correspond to the forms in the right-hand column which express 'first person singular'.

(1)	(a)	[emoʔu]	*his word*	[ẽmõʔ̃ũ]	*my word*
	(b)	[owoku]	*his house*	[õwõŋgu]	*my house*
	(c)	[piho]	*he went*	[mbiho]	*I went*

According to Carr, the phonological phenomenon which signals 'first person singular' involves nasality, and this could be informally stated as follows: nasalise the segments in the left-hand column, from left to right, up to the first obstruent, which itself will become nasalised, such that p → mb, and k → ŋg. The question to be addressed is how a linear notation can capture this generalisation. The case in (1)(a), which contains a [+nas] segment, suggests that we could spread this value over all the segments in the domain (evidence from the same type of phenomenon in other languages favours this type of analysis). This would be expressed in terms of an *SPE* rule as follows:

(2)
$$[+\text{syll}] \rightarrow [+\text{nas}] \;/\; \left\{ \begin{array}{l} [+\text{nas}]\ \_ \\ \_\ [+\text{nas}] \end{array} \right\}$$

However, notice that the iterative application of this rule accounts for (1)(a) only. Nevertheless, due to the failure to express the above generalisation, it cannot account for the other forms in (1). The main reason for this is that such forms do not contain the [+nas] value, which seems to be the source from which this phenomenon originates. Clearly, by adhering to an *SPE*-based analysis, important generalisations are missed, and therefore, this type of model proves to be inappropriate.

Carr then suggests that the phonological form of the 'first person singular' morpheme ([+nas]) is 'overlaid' from left to right on the morphemes in the left-hand

---

[3]Nasal spreading is just a token of other varieties of feature agreement among segments like the ones I have already mentioned, palatality, and roundness.

column in order to yield those in the right-hand column. For this, it is necessary to represent [+nas] on a separate tier (or on a different dimension) from the segmental tier, as follows:

(3) [+nas]

/emoʔu/

In order to show the overlaying of the [+nas] feature over the morpheme, and therefore, to capture the actual phonological process, some means of connecting both tiers is now necessary. This is achieved by means of association lines, as illustrated below:

(4) [+nas]

/emoʔu/

This type of *non-linear* representation is at the core of studies which have departed from the *SPE linear* representation. The new arrangement of features (now referred to as *autosegments*), which are independent from one another (and therefore, organised in different tiers), gives the name to the *autosegmental* framework in phonology.

The application of autosegmental ideas to the account of what were formerly treated as **segmental** phenomena (assimilation of consonants, vowel harmony, etc.) has flourished greatly, and, as I claimed earlier, today offers a different perspective on phonological events.

However, of greater importance here is the **formalisation** of all these ideas in phonological analyses of **tonal** phenomena (tone spreading, tonal stability, etc.). Work on tone languages included in Leben (1971, 1973), Goldsmith (1976), and subsequent research by Clements and Goldsmith (1984), and many others, marks another turning point in the study of the phonological behaviour of tone. As an illustration of the flaws present in *SPE*-type phonology in the description of tonal phenomena, (for instance, contour-tones), and their revision in the framework of autosegmental phonology, let us turn to an example included in Durand (1990), chapter 7.

A possible way of representing a contour-tone in traditional *SPE* terms, say, a falling tone, is by means of the combination of two primitive level tones, (+H, +L); so, for example /â/[4], is specified in the fashion of a matrix as follows:

(5)

$$\begin{bmatrix} \text{â} \\ \text{+syll} \\ \text{-cons} \\ \text{+low} \\ \text{-round} \\ \text{+H+L} \end{bmatrix}$$

However, Durand points out that this feature matrix violates the fundamental convention of the internal structure of segments within the *SPE* tradition, since, within a column of features only one feature is allowed in each row. Here, the offending factor is the juxtaposition of [+H+L].

There are two possible alternatives to this representation, which adhere to the formalism: either to specify each feature on a separate line, as I illustrate below in (6)(a), or to show two feature matrices as in (6)(b).

(6)

(a)

$$\begin{array}{c} /\text{â}/ \\ \begin{bmatrix} \text{+syll} \\ \text{-cons} \\ \text{+low} \\ \text{-round} \\ \text{+H} \\ \text{+L} \end{bmatrix} \end{array}$$

(b)

$$\begin{array}{c} /\text{â}/ \\ \begin{bmatrix} \text{+syll} \\ \text{-cons} \\ \text{+low} \\ \text{-round} \\ \text{+H} \end{bmatrix} \begin{bmatrix} \text{+syll} \\ \text{-cons} \\ \text{+low} \\ \text{-round} \\ \text{+L} \end{bmatrix} \end{array}$$

The problem with the representation in (a) is that it is impossible to distinguish a falling from a rising contour-tone. In relation to (b), this structure constitutes the representation of not one segment, but two. Therefore, these representations fail to achieve descriptive adequacy. Durand presents further evidence which emphasizes the inability of the *SPE*

[4] The circumflex is used here to represent a falling tone. Later in this chapter, a grave accent represents a low tone, and an a cute accent represents a high tone.

formalism to account for contour-tones. However, the example I have presented above ought to provide sufficient evidence to highlight the problems encountered by the *SPE*-type framework.

In view of the case of nasality in the example from Carr (1993), we might expect the first step to be taken by an alternative autosegmental model to be one of separating tone from the rest of the segmental representation. This procedure is followed for the representation of contour-tones. In this way, the representation of a falling contour-tone is straightforward, and looks as follows:

(7) +H +L

/â/

A rising movement would be represented by reversing the ordering of features. This type of configuration allows for the sequencing of features, the impossibility of which proved to be a fundamental shortcoming within *SPE.*

The most dramatic development in the understanding of phonological systems in the framework of non-linear phonology, is found in the study of **intonational** phenomena. Although some of the ideas included in *SPE* are applied to research in intonation (see for example, the use of distinctive features in Vanderslice and Ladefoged (1972), and Ladd (1983) to account for the intonation of English), the majority of work in this field during the 70's amounts to the construction of models which produce detailed and excellent descriptions of the languages for which they were created. However, due to this restricted scope of application, those models are rarely applied to a wider range of languages. On the other hand, the non-linear approach constitutes an advantage over former descriptions, and nowadays, we encounter work like that of Liberman (1975), Gussenhoven (1983), Beckman and Pierrehumbert (1986), Hayes and Lahiri (1991), Grice (1992), (and many others), which offer an insightful view of intonational phenomena in English, Dutch, Japanese, Bengali, and Palermo Italian, respectively. Thus, non-linear phonology has broadened the scope of its application to phenomena formerly recognised as segmental and non-segmental.

However, the majority of the intonational work mentioned so far suggests that the phonological organisation of pitch has little in common with the phonological organisation of other (spectral) qualities of sounds. Take, for example, the organisation of tones (represented by means of H(igh) and L(ow)) at the intonation level as included in Pierrehumbert (1980) (henceforth P'80), and the organisation of roundness and palatality (represented by [U] and [I]) as included in Kaye, Lowenstamm and Vergnaud (1990). H and L are somehow combined to form pitch accents, intonational phrases and the like. On the other hand, the combination of [I] and [U] to convey a particular sound quality (for example [y]), and then, its presence in either the nucleus or the onset of a syllable, is controlled by a specific set of well-established principles, which fall under the principles of licensing. The issue at stake here, is that the development of the autosegmental framework has reached a point at which there has been a theoretical bifurcation into those models which specialise in the account of, on the one hand, tonal phenomena, and on the other, non-tonal phenomena.

The present thesis constitutes an attempt to reconcile the two types of model. It shares a few ideas already scattered among intonational studies (like, for example, the intonation domain as a well-defined structural domain), but most importantly, it develops those ideas much further, following well-established principles included in models which are centred around the study of non-tonal phenomena, like for example, Government Phonology (henceforth GP). Let me justify the motivation for the reconciliation of the two approaches.

It must be made clear that the object of study in phonology is linguistic sounds which are produced by human beings. If this is so, I can conceive of no motivation whatsoever to examine tonal and non-tonal phenomena under different principles. I would like to clarify this point by turning back to my analogy between phonology and the human body. To suggest that tonal and non-tonal phenomena are governed by the same principles parallels the widely accepted principle of biology, that the same laws of chemistry control the development of both flesh and bone. No matter which part of our body we may want to analyse, the combination of its chemical components will always follow the same principles. For example, hydrogen and oxygen are allowed to merge (wherever they may be in the body), and their combination results in H_2O. The stability of the development of the body would be disrupted if we suggested that the principles of chemistry which operate in bones were different from those which operate in blood. By the same token, tonal and non-tonal properties — as components of sounds — ought to be governed by the principles which operate in phonology.

The motivation for adopting the principles assumed in GP — which has been developed as a highly impoverished theory, derives, first, from the need to impose a strict control on the generative capacity of the intonational model, and second, from the fact that those principles have been applied successfully in the account of an extensive range of phonological phenomena over many of the world languages.

1.2 Principles and parameters

Implicit in the adoption of the fundamental tenets of GP is the idea that the present model of intonation is fully embedded in the tradition of Generative Phonology[5], and consequently, of Generative Grammar (Chomsky 1986). In order to set this model in such a context, it is necessary to present a brief overview of other notions developed in Generative Phonology, and then, to explain how all these relate to the model of intonation.

One of the goals in Generative Phonology is to capture only those phonological processes that actually occur; that is, it should have the capacity to account for well-formed structures, and to exclude ill-formed structures. The procedure by which this goal is achieved in later generative theories defines a set of **general principles**, and then refers to a relatively small number of **parameters** which allow for cross-linguistic variation (Chomsky 1981).

General constraints operate on phonological representations, and offer principled accounts of phonological processes. For instance, rather than stipulating a rule of assimilation for consonants which is independent of another rule of harmony for vowels, GP has the capacity to account for a general process of *spreading* of resonance elements, regardless of whether it applies to vowels or to consonants. In *SPE*-type phonology, information relating to both the structure and the substance of a sound are thrown together into a feature matrix. On the other hand, in GP, structural and melodic information are treated as wholly independent properties. For example, in *SPE*, vowels and consonants were defined in terms of features; vowels as [+syllabic], [-consonantal],

[5]See Harris (1994), chapters 1 and 4 and references therein.

[±round], [±high], etc; and consonants as [+consonantal], [±coronal], [±labial], etc. By contrast, in GP, vowels and consonants share the same building blocks, which they refer to as *elements* ([I], [U], etc). However, the exact interpretation is captured after the association of these blocks to phonological structure becomes explicit. That is, whether they are associated to a nuclear position or to an onset position. The former would be understood in traditional terms as a vowel, and the latter as a consonant. I illustrate this with the word *pop* below:

(8)[6]

O	N	O	N
x	x	x	x
h		h	
U	U	U	
?		?	
H		H	
	A		

For the purposes of clarity, I shall explain in very crude terms a few points relating to the above structure. [h], [U], [?], [H], [A] are some of the elements used in GP to describe the acoustic properties of sounds. The salient property of [h] is noise; [U] is labiality; [?] is occlusion; [H] is stiff vocal folds; and [A] is openness (Harris 1990). These elements stand on independent tiers. They are associated to the skeletal tier (represented by means of *x*s). Above the skeletal tier, there is the prosodic structure composed at the lowest level of a sequence of onsets and nuclei[7]. The crucial point to notice here is that, in the phonological representation of the word *pop*, the element [U] is present in the composition of both 'p' and 'o'.

---

[6]The phonological structure for words like *pop* which show a final consonant are characterised in GP by showing a final empty nucleus. I shall address this topic in chapter 6.

[7]For further details about this representation, see Harris (1994).

By way of an example, let us now turn to an account of labial assimilation of a nasal when it occurs before a labial consonant, and rounding harmony in two adjacent vowels. In the *SPE* phonology, the following rules would be needed:

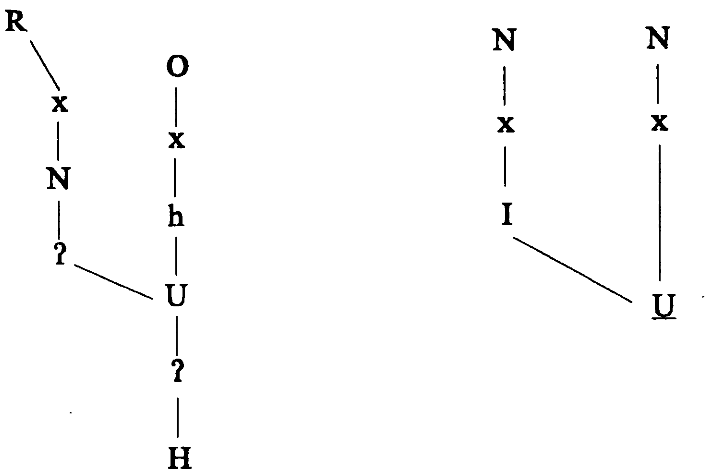
(9)[8]

(a) $\overset{C}{[+\text{nasal}]} \rightarrow [+\text{labial}] \; / \; \_ \; \overset{C}{[+\text{labial}]}$

(b) $[+\text{syllabic}] \rightarrow [\alpha\text{round}] \; / \; \_ \; [\alpha\text{round}]$

(a) would account for /np/ becoming /mp/, for example, and (b) would account for /i/+/u/ becoming /y/+/u/. Notice that here, we need two independent rules to account for exactly the same type of process. In terms of GP these examples would look as follows:

(10) (a) /np/ → /mp/ (b) /i/+/u/ → /y/ +/u/



By contrast, here, assimilation and/or vowel harmony is accounted for by the *spreading* of a particular element, more precisely, by the linking of one element to two positions. This is represented by means of the slanted line which joins the relevant element to those positions to which it spreads. In short, by separating structural from melodic information GP presents a unified account of former consonant assimilation and vowel harmony in

[8]Notice also that, in *SPE* phonology, two different features, [±labial] and [±round] are necessary to describe what in GP terms is described by means of a single element, [U]. Thus, GP has a capacity for generalisation which is greater than that of *SPE* phonology.

terms of linking an element to the relevant position[9].

The variability of language systems is defined upon a small number of parameters, each of which defines a choice (usually binary) between particular typological characteristics. One well-established example in Government Phonology controls whether or not the final nucleus in a word can be left empty: in languages such as Japanese or Maori, which show a -CV]typology, it cannot; on the other hand, in languages such as English or Spanish, which show a -CVC]typology, it can. Another example has to do with whether or not the initial nucleus in a word can be left empty: in languages such as Spanish, it cannot, in others such as Italian and English it can[10] (cf. Spanish *estadio*, English *stadium*, and Italian *stadio*).

1.3 Competence/performance

Another idea embedded in Government Phonology, and shared by other theories of phonology (for example, Dependency Phonology) is concerned with the respective domains of inquiry of phonology and phonetics. I shall offer here an outline of the view presented in Lindsey and Harris (1990), since I shall take these assumptions for granted in the present work.

Lindsey and Harris adopt the *competence/performance-based* view which, in broad terms, assigns phonology to competence (these are the constraints on speech signals of type (11)(a) below), and phonetics to performance (these are the constraints of type (11)(b)). They define this view in the following way:

(11) *Competence/performance-based view*
There are two levels of constraints on speech signals:
(a) Cognitive, grammatical constraints, common to all members of a speech community and independent of any specific utterance.
(b) Non-grammatical constraints of various subtypes:
(i) physical constraints common to all normal humans; and
(ii) physical constraints peculiar to each individual.
(Lindsey and Harris 1990:357)

[9]Obviously, linking and delinking does not occur freely. It is constrained by other principles in the theory, the details of which are not relevant here.

[10]Note that English *stadium* and Italian *stadio* are both analysed in GP terms as showing an empty nucleus before the consonant, which is treated as belonging to a coda (Kaye 1991/92).

In the context of Lindsey and Harris's view, distinctions formerly made between underlying and surface phonological representations are abandoned, and instead, the input and the output representations in phonological derivation are isotypic. In their own terms, expressed in a later article,

(12) processes map phonological objects onto other phonological objects rather than onto phonetic ones [...]. If phonological processes map like onto like, it follows that initial representations should be no less phonetically interpretable than final representations. In fact, initial representations can in principle be envisaged as being wholly indistinguishable in kind from final representations (Harris and Lindsey 1995:49).

In the model I develop, the above views are understood as follows: the principles of licensing — which are responsible for the construction of prosodic structure, and for the well-formed association of tone to boundaries (see chapter 4) — are defined under the constraints in (11)(a). Gravitation Effect (henceforth GE)[11] — which is the acoustic manifestation of the physical constraints imposed on the production of pitch — is defined under the constraints in (11)(b)(i). Phonological representations are directly interpretable by means of GE. In the next paragraph I compare this view with that followed by P'80, since her work constitutes the foundation of current models of intonation.

In the model presented in P'80 for English, Pierrehumbert assigns only a peripheral role to phonology. Embedded in the tradition of the distinction between underlying and derived representations, she claims that in the field of intonation, no such distinction is relevant. That is, in English, there are no rules which alter tonal values or delete tones in the derivation from underlying to surface representation. Given this view, it follows that, as long as the model incorporates a battery of phonetic rules to interpret phonological representations, which we may assume is indeed the case, such representations are directly mappable onto a tune, and hence, there is no longer any need for the underlying/surface dichotomy. She expresses this view in Pierrehumbert and Beckman (1988:5) in the following terms: 'They [phonetic rules] take as input phonological representations, but their output consists of quantitative functions'. In short, they suggest that the phonetic component bears a greater burden than the phonological component in the study of pitch

[11]See chapter 4.

patterns. Let me illustrate this point with the treatment of the phenomenon of downstep (informally defined as the down-shifting of the scaling of the second H in a sequence such as HH[12]). P'80 accounts for downstep by means of a phonetic realisation rule. In this rule, the input is a phonological representation which contains the sequence HLH, and the output consists in scaling the second H lower than the first. The actual rule causes the second H to have the F0 value of the first, multiplied by a constant factor, whose value is stipulated to be between 0 and 1 exclusive. This being so, the value of the second H will always be lower than that of the first.

The approach adopted here differs from that of P'80 on two counts: first, with respect to the notion of phonological representation, and second, regarding the understanding of the **interpretability** of such representation. Notice that, under P'80's view, phonological representations are meaningless and uninterpretable until the phonetic rules are established. Thus, this assigns an insignificant role to the phonological component, and a major role to the phonetic component.

On the other hand, in the model I propose, the linguistic study of pitch contours is primarily concerned with the identity of the units it is composed of, and how these can be combined to build the structure which underlies such pitch patterns. Thus, the focus is primarily set on the cognitive, grammatical constraints on the pitch aspects of the speech signal.

Let us now turn to the interpretability of phonological representations. As I have already stated above, according to Pierrehumbert, in order for a phonological representation to become interpretable, a battery of interpolation rules is necessary. These rules compute the F0 values of tones, which are treated as phonological targets. However, the formalisation within which these rules are conceived is extremely powerful and grossly arbitrary. By contrast, in the model I propose, phonological representations are interpretable at any stage in phonological derivation by means of a single universal principle, GE, and not from a series of interpolation rules. The advantage of the principle over the rules is that the former has a universal scope of application, whereas the latter turn out to be language specific.

Finally, I unite the fundamental tenets of non-linear phonology, and the notion of

[12]See chapters 5 and 6.

principles and parameters, in an attempt to construct a coherent and restricted phonological model which can account for intonational phenomena. I claim that the phonological unit responsible for the description of pitch patterns is a single autosegmental prime, which I shall refer to as *T*. This unit can be associated to a structure in a formal manner — that is, by following a set of principles already mentioned, the principles of licensing. The variability of intonational patterns in the world languages is characterised in terms of parameters, like, for example, whether or not a final nucleus in the intonation domain can remain empty. As far as the investigation in this thesis is concerned, in English it can. Although not proven, it may be the case that, in other languages, it cannot.

In short, this thesis is primarily concerned with the **form** (as opposed to the function) of pitch patterns at the level of **intonation**. Nevertheless, it should be borne in mind that the model is general enough to be applied to the behaviour of pitch at other (pitch accent or tonal) levels.

2 What this work is not about

The **function** of intonation falls outside the scope of this thesis. The motivation for leaving this out results from the following argument: I suggest that the purpose of an object (in this case, intonation) can be analysed only after such an object has been identified, and consequently, defined. This entails that, prior to trying to define the object, there is only a **general** (and not detailed) idea of its function.

P'80 follows a similar line when she suggests that the approach whereby intonational patterns are identified, by relation to the similarity or difference of the meanings conveyed, is to be avoided. This is due to the fact that it may result in false predictions about the contrastiveness of F0 contours. For example, even if the meaning of two contours is the same, we should not allow ourselves to conclude that they correspond to the same phonological category. Furthermore, intonational meanings are extremely context dependent, and this results in the fact that a given contour may have one particular meaning in one instance, but a completely different one in another situation.

Due to the high degree of inaccuracy inherent in an approach based on meaning, Pierrehumbert rejects it immediately, and instead, adopts another approach, which she defines in the following way:

(13) One approach attacks the problem by attempting to deduce a system of phonological representation for intonation from observed features of F0 contours. After constructing such a system, the next step is to compare the usage of F0 patterns which are phonologically distinct (Pierrehumbert 1980:31).

Pierrehumbert defends the study of phonological units prior to the study of meaning. In her thesis, she makes absolutely clear that her work is solely concerned with the first step stated in the quotation, leaving the field of meaning for future research.

As I have already stated above, in this thesis I adopt an approach similar to Pierrehumbert's to the extent that I also stop short at the comparatively indeterminate field of meaning. However, my approach completely differs from hers in the fact that I follow a deductive approach which starts from a well-established theory, and ends in the analysis of intonational data, rather than vice versa, as she states in the above quotation by 'deduce a system [...] from observed features of F0 contours'. If Pierrehumbert's approach were to be adopted, we would run the risk of mistreating what constitute idiosyncratic characteristics of pitch contours as fundamental properties. For instance, the manifestation of relative high pitch at the beginning of the F0 contour leads Pierrehumbert to analyse such protrusion as the presence of H in phonological structure[13]. The pairing of the manifestation of high pitch in phonetic interpretation with the presence of H in phonological representation, should not be taken for granted, as it may be the case that a particular instance of high pitch is predictable, and therefore, need not be encoded in phonological structure[14]. On the other hand, the approach I defend saves us from falling into this trap, since the theory does not predict the existence of a tone in that position in the first place.

I would like to conclude this section with a brief discussion of the fact that some intonationalists have assigned specific types of meaning to H and L, and of the way this

[13]In the rest of the present work, I shall refer to 'relative high pitch' in phonetic interpretation as simply 'high pitch'. The same applies to 'relative low pitch' and 'low pitch'.

[14]See specifically chapter 4 in which I claim that the left edge of the intonation domain is incapable of bearing tone due to the ban imposed by the directionality of licensing relations between boundaries.

claim relates to my model, which is characterised by a single tone, T (which corresponds to the former H). The issue at stake here is that, if the above idea is to be accepted, then it may impose a threat to the model I develop here, since there is only one tone. The argument against the model I defend would follow this line of reasoning: only one set of meanings can be captured under T, and therefore, the other set (formerly captured by L) would not be represented by any unit in phonological representation, and consequently, would be lost. In order to find a solution to this problem, I shall present the definition that Pierrehumbert and Hirschberg (1990) assign to H* and L*[15], and after that I prove that their findings impose no such threat to the model.

According to Pierrehumbert and Hirschberg,

(14) The H* accents [...] in general convey that the items made salient by the H* are to be treated as "new" in the discourse (1990:288).

S [the speaker] may also employ L* accents when the instantiated expression is believed already part of H's [the hearer] mutual beliefs (1990:292).

The ideas presented by Pierrehumbert and Hirschberg do not undermine the model I present here, since the meaning of H* ('new' information) is the **opposite** of the meaning of L* ('old' information). Moreover, even if they were not opposite, notice that there is only **one** contrast represented by these two units. Perhaps the most logical approach would be to capture a single contrast by the presence versus absence of **a single** unit. Let me illustrate this idea with a simple example. Let us assume that 'new' information is represented by *x*, and that 'old' information is represented by *y*, as I show below:

(15) H means *x*

L means *y*

This single contrast is captured perfectly in terms of the presence vs absence (informally represented as Ø) of T, as I illustrate below:

[15]The same argument that I present here for H* and L* can be extended to H-, H%, L- and L%. Also, see Hobbs (1990).

(16)	T	means	*x*
	Ø	means	*y*

In view of the fact that the ideas included in Pierrehumbert and Hirschberg are still captured straightforwardly in the present model, they do not constitute a challenge.

3 Outline

In this chapter, I have introduced general notions which are fundamental to the establishment of the framework in which the present model of intonation stands.

Chapter 2 presents other basic concepts pertaining to the nature of phonological units. With the intention of developing a restricted phonological model, I adopt a privative approach to phonological contrasts. This view is characterised by assigning a **unary** value to phonological units, so that contrasts are manifested by means of the presence or absence of those units. The decision to pursue this metatheoretical condition raises the question of the identity of such units in intonational analyses.

Due to the fact that the model I propose departs from earlier models of intonation which resort to a binary feature, a decision must be made as to which value of the binary feature is to be adopted; whether it is H or L. Evidence gathered from the behaviour of tone in a variety of contexts, such as two-tonal systems in both tonal and intonational languages, leads us to conclude that H constitutes the unit which must be encoded in phonological structure.

The above conclusion prompts an investigation into the representation of contour-tones, which result from the combination of H and L. In order to present a proposal which accounts for them, I take a step back, and I investigate the strategy of associating former H and L to the boundaries of prosodic domains, exclusively. The results of such investigation strongly suggest that this procedure is viable. Moreover, the subsequent application of such a procedure to the account of contour-tones confirms the hypothesis that L is redundant in phonological structure.

Chapter 3 is concerned with yet another obstacle which questions the validity of this hypothesis: claims made in the literature on tone maintain that H and L participate in the phenomenon of tone-spreading. In view of this observation, it would be unwise to reject L from phonological representation. With the aim of overcoming this stumbling block, I examine and revise the procedure followed by former analyses in an array of languages which count as reference points in the tonal spectrum. Evidence provided by the results of this investigation indicates that, in the majority of those analyses, spreading can be treated as a phonetic phenomenon, rather than a phonological one. This finding suggests that spreading occurs independently of the number of tones present in phonological structure and therefore, that L-spread no longer constitutes a challenge to my hypothesis. The way is now cleared to introduce the other components which, altogether, constitute the grammar of the present model.

Chapter 4 stands at the core of the present thesis. It introduces the basic grammar of intonation, the fundamental component of which comprises the principles of licensing. Phonological structure is characterised by the prosodic hierarchy, and the melodic content is provided by a single tone which can associate to boundaries of prosodic domains. In addition, I present the general principle of Gravitation Effect, which enables us to achieve a straightforward interpretation of phonological representations. Thus, with all this at our disposal, we may begin to generate phonological representations which account for intonation contours.

The last obstacle to overcome in order to succeed in the project of a single-tone model is introduced in chapter 5. There, I present and evaluate the latest accounts of *stepping*. In general, such accounts are characterised by having recourse to L in phonological representations, which accounts for an observed shift in the F0 trace. Obviously, this constitutes yet another challenge for the single-tone model. However, a detailed revision of these models discloses the fact that, in order to account for stepping, they all show the tendency to enrich phonological structure. Naturally, this characteristic shared by earlier models cannot be adopted by one designed upon the conditions which lead us in the opposite direction. The single-tone model aims at a maximally restricted phonological component.

Given this conflict, and in pursuit of the metatheoretical conditions presented in chapters 1 and 2, chapter 6 develops some of the ideas of phonological structure presented in chapter 4. Specifically, it presents an alternative account of stepping which is characterised by the presence of extra toneless boundaries, belonging to *empty* *categories.*

Finally, Chapter 7 presents the conclusions drawn from the present work.

Chapter 2. A valediction for L

1 Introduction

In this chapter I take for granted one of the fundamental ideas developed in GP, namely, that phonological contrasts are captured by the presence or absence of monovalent melodic units. I shall follow Trubetzkoy (1969) in referring to this idea as the privative approach. This entails that, in the model I propose, tonal contrasts are understood as the **presence or absence** of a **monovalent** tonal property, which I shall refer to as **T(one)**[1].

In order to define the nature of such a property later in the chapter, it is necessary to remind the reader that the present model has developed from earlier systems of intonation which, nevertheless, adopt an alternative view of phonological contrasts. Such a view is characterised by the representation of phonological contrasts in terms of opposite, and in principle, equally weighted values of a bivalent tonal feature, like for example, [+high] versus [-high], or as it is also represented, simply H(igh) versus L(ow). Following Trubetzkoy's terms, I shall refer to this approach as the equipollent view.

The motivation for adopting the privative view rather than the equipollent view is that the former approach has the potential for tightly **constraining** the generative capacity of the model I develop here. In order to clarify this claim, let us consider for a moment the predictions inherent in an equipollence approach. For instance, a model in which a bivalent feature is used, predicts that both values stand an equal chance of being specified in the lexical form of phonological representations, and of participating in phonological processes. This issue has long been recognised, and addressed within various forms of phonological underspecification. On the other hand, a model with a monovalent feature reduces the amount of lexical forms and phonological processes by half. In the context of a restrictive phonology, then obviously, this approach is preferred.

This leads us on to the question of the nature of this monovalent feature. In

[1]This view is well-established in other theories of phonology, like Dependency Phonology (Anderson and Ewen 1987), etc. For a detailed discussion on the topic related to the preference for privativeness rather than equipollence, see specially, van der Hulst (1989), den Dikken and van der Hulst (1988). For an outline, see Harris and Lindsey (1995).

principle, it could potentially be either H or L. As I shall show shortly, in some analyses of languages employing two tones, L is either lost in their historical evolution (for example, from Proto-Bantu to Bantu), or it is not present at all in lexical representations. If this is the case, that is, that L is assigned to toneless units **after** all postlexical rules have applied, then it is hard to conceive of a situation in which L is involved in any phonological processes in the first place. So, there is already some evidence which casts doubt on the fact that L may be a monovalent feature. If we are faithful to our intention to design a constrained model, we are forced to reject L, and consequently, rely only on the presence or absence of H; in this way we achieve a significantly higher level of restrictiveness.

Evidence in favour of rejecting L is provided by the examination of the representation of tone in two-tone language systems, including, for example, some African languages (in §2), and also, English (in §3). Having done so, there is then no longer sufficient motivation to retain H as the label of the phonological property which represents pitch, as this could lead to the misconception that L may still be referred to. Thus, after discarding L, I shall represent the phonological property of pitch as T.

The rejection of L from the model raises a fundamental question of how to account for contour tones (a falling or rising movement on a single syllable), since earlier models propose the combination of H and L[2]. Notice that, with the presence or absence of T as a means of showing phonological contrasts, it appears to be impossible to capture any sort of pitch movement. This proposal allows the model to go as far as capturing high (presence of T) and/or low (absence of T) manifestations of pitch.

With the goal of finding a satisfactory solution to this shortcoming without having to reject our hypothesis of a single-tone model, I analyse previous models of intonation which propose that some tones can be associated to the edges of intonation domains, and extend this idea to all tones. That is, I shall demonstrate that it is possible to treat

[2]Another phenomenon which may pose a problem for a single tone model is downstep, since it is generally agreed that L is the phonological factor which is responsible for this. In chapter 6, I propose that the presence of toneless boundaries belonging to empty categories is the main factor which triggers downstep in intonation. The same can be said to apply in tone languages, although the details of such an account fall outside the scope of the present work.

boundaries of prosodic domains exclusively as Tone Bearing Units (henceforth TBUs)[3]. This alternative will allow me to capture elegantly a four-way distinction as follows:

```
(1)  (a)  high            (b)  falling
          T   T                T
          |   |                |
          [   ]                [    ]

     (c)  low             (d)  rising
                                    T
                                    |
          [   ]                [    ]
```

2 H and L in tone languages

My motivation for briefly looking at tone languages before turning to intonation languages in the following section is that the system proposed for their description by tonologists like Goldsmith (1976), Hyman (1982), Pulleyblank (1986), and others stands as the foundation for current models of intonation. Thus, tone languages may well provide an insight into the nature of tone, which later may prove crucial in the account of pitch patterns in intonation languages. For the purposes of supporting the argument that L can be dispensed with in phonological representation, I shall briefly examine evidence from the historical evolution of Bantu languages in the following section. Such evidence suggests that L is unstable to the extent that, in particular instances, it can disappear from lexical representations. In addition, Carter (1973) suggests that H is regularly active in phenomena like 'tone copying'. After that, I shall present an overview of some synchronic analyses of two-tone languages which indicate that L, but not H, may be omitted from lexical representations.

The fact that I shall turn to tone languages in search for evidence to support the exclusion of L from phonological representation, should not be taken as an indication that I shall perform a detailed analysis of them, although the framework to be presented in this thesis can potentially be applied in this area.

[3] A similar idea has already been proposed by Hirst (1988). See §4.1 for an outline of this proposal.

2.1 Historical evidence

A widespread observation made in relation to Bantu languages is that they seem to be, or have been, in the process of transition from purely tonal languages, to accentual languages[4], and that many exhibit properties of both tone and *pitch accent* languages in varying degrees. Before I present an example which illustrates this tendency, let me clarify what is generally understood by *tone*, *pitch accent*, and *stress* languages.

Informally, if *tone* languages and *stress* languages are said to stand at opposite ends of a continuum, *pitch accent* languages are located at the same point in between, in the sense that they share characteristics of both of them. In very crude terms, **tone** languages (like Kikuyu) are characterised by the location of tone being predictable; that is, the association of tone to vowels follows from a set of fixed principles. Additionally, tonal contrasts are unpredictable, and the tonal specification of vowels has to be registered in the lexicon. For example, according to Goldsmith (1990), and his Universal Association Conventions, in Kikuyu, lexical tones are associated to vowels in a one-to-one fashion, and from left to right.

By contrast, **pitch accent** languages show an unpredictable location of tone, and therefore, tone is assigned to particular vowels in the lexicon. In addition, tonal contrasts are fixed. Take for example Japanese, where the tonal melody is fixed as HL, but the location of this accent is registered in the lexicon.

Finally, **stress** languages do not make essential use of pitch to show prominence relations; rather, they choose other cues as being primary, like for example, duration and loudness[5].

The transition from tone languages to pitch accent language is clearly manifested diachronically in Proto Bantu and Sukuma. According to Goldsmith (1988), tonal contrasts are captured in Proto Bantu by H or L being associated to vowels. He adds that those vowels which were associated to H in Proto Bantu, show a melody which could be H, or LH, or HL in Sukuma. On the other hand, other vowels which were associated to

[4]See for example, Clements and Goldsmith (1984), and articles in van der Hulst and Smith (eds) (1988).

[5]I remind the reader that this is a very crude classification of *tone, pitch accent*, and *stress* languages, but that it suffices for the purposes of an introductory idea to the differences among them. See Beckman (1986) for a detailed discussion of this topic.

L in Proto Bantu, show no tone at all in Sukuma. This is illustrated below:

(2) (a) Proto Bantu: V
|
L

Sukuma: V

(b) Proto Bantu: V
|
H

Sukuma: V
|
H or LH or HL

In general (except for H), vowels have the tendency to show a fixed tonal melody, rather than simply a level melody. However, the crucial factor to notice here is that, in this evolution, L no longer appears on its own, but instead is integrated into a melody. On the other hand, the relatively autonomous behaviour of H affords it the choice either to stand on its own or to become part of a melody. I understand these observations as symptomatic of the fact that H constitutes a robust phonological unit. The same cannot be said of L.

Let us now focus a while longer on Goldsmith's claim that this shift in tonal behaviour from Proto Bantu to Sukuma can be understood as a transition from equipollence — either H or L is associated to vowels — to privativeness — either H, HL, LH *or nothing* is associated to vowels. Notice that privativeness in Sukuma is inconsistent with observed tone patterns. Let me present Goldsmith's motivation for his claim, and after that, my statement about the inconsistency of privativeness should become clearer.

The fact that, in Sukuma, vowels may show HL or LH leads Goldsmith to assume that in Proto Bantu, both components (H and L) were present in the lexicon. Given that Proto Bantu is assumed to be purely tonal, and therefore, that all vowels must have been associated to either H or L, he is led to conclude that tonal contrasts must be analysed under an equipollent treatment. The fact that, in Sukuma, vowels may show HL, LH, H, or nothing at all, leads him to assume that each one of these melodies participates as a privative unit, so that lexical contrasts are defined by the presence or absence of each one of them. Note that here, Goldsmith makes the assumption that L and *nothing* refer to exactly the same property: absence of any tonal specification. However, this conflicts with the fact that L can still be present in some melodies, like (HL) or (LH). It is precisely here where the inappropriateness of the privative view manifests itself; notice that L cannot be fully rejected from the model. Obviously, this type of model, which relies on two tones to capture contour tones, cannot afford to exclude L from its structure. Instead it is doomed to resort to some kind of arbitrary stipulation in order to exclude L whenever it is considered inconvenient. Ideally, though, a principled account should be pursued; one in which there is only H, and in which contour tones are represented without having recourse to L.

As I have already stated in chapter 1, and illustrated towards the end of the previous section, a central goal in this thesis is to develop such a model. However, due to the diversity of tonal systems in the language spectrum, and also due to the fact that I am concerned here with intonation, and not directly with tone languages, I set aside testing the validity of this model against these languages for future research.

In relation to the tendency toward a re-analysis of tonal systems into pitch accent systems, and to the fact that H is the favoured tone in such a trend, Clements and Goldsmith (1984) report that Carter (1973) proposes the following rule-based typology of Bantu tone systems:

(3)
- i. "etymological" or "clear tone" systems, which show a point-for-point correspondence with Proto-Bantu.
- ii. "reversive" systems in which H replaces etymological L and L replaces etymological H.
- iii. "tone doubling" systems in which H tone is copied onto a following syllable.
- iv. "tone anticipation" systems in which a H tone is copied onto a preceding syllable.
- v. "displacement" or delayed realization systems in which basic H tone is realized on a tone-bearing element later than that which it is lexically associated.

With the exception of (ii), these points can be grouped into a single statement, that H, and not L, spreads.

2.2 Non-historical evidence

Pulleyblank (1986)[6] claims that H and L show an unequal behaviour in phonological systems, and that L, rather than H, is treated as **default** by insertion rules. I assume that the manifestation of this behaviour counts as evidence that H is somehow more robust than L.

Pulleyblank argues that in the General Past and Recent Past forms from the verb paradigm of Tiv, either H or L is specified for particular vowels in the lexicon. In addition, some vowels can remain toneless, and that it is precisely these vowels which surface as L. Take, for instance, the following underlying form of the General Past meaning *to flee*, that is, *fled*:

(4)[7]

$$\left[\begin{array}{llllll} \text{y} & \text{e} & \text{v} & \text{e} & \text{s} & \text{e} \\ & \vdots & & & & \\ & \text{H} & & & & \end{array}\right]$$

The surface form is interpreted as *'yévèsè*[8]:

(5)

$$\left[\begin{array}{ccc} \text{ye} & \text{ve} & \text{se} \\ | & | & | \\ \text{H} & \text{L} & \text{L} \end{array}\right]$$

Pulleyblank's motivation for this analysis stems from the rejection of a previous proposal made by Goldsmith (1976). In short, there has been a conceptual tension between Pulleyblank's *non-automatic* spreading and Goldsmith's *automatic* spreading. Let me briefly illustrate Goldsmith's analysis in the following paragraph.

Goldsmith's (1976) Well Formedness Condition (WFC) associates lexical tones

---

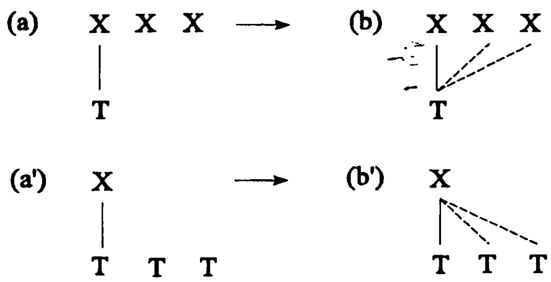
[6]Pulleyblank's analysis is embedded in the framework of *Lexical Phonology* (Mohanan 1982, and Kiparsky 1982).

[7]Dashed lines indicate that lexical tones associate by convention. See Pulleyblank (1986:11) for his understanding of Association Conventions of tone.

[8]I present a simplified version of the example included in Pulleyblank. The first tone has undergone downstep (indicated by '), but the reasons for why this has happened are not relevant to the matter under discussion here.

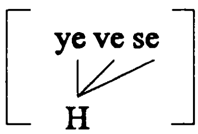
(here represented by means of *T*, which covers both H and L) to positions in a one-to-one and left-to-right fashion, spreading the melody-final tone onto toneless positions, as in (6)(a) and (b), or linking extra tones onto the final TBU, as shown in (6)(a') and (b'), so that all tones and TBUs are associated. This is illustrated below:

(6)



Note that, had Goldsmith's idea been pursued in the analysis of Tiv, then H in (4) would have been automatically associated to the remaining vowels, consequently yielding an incorrect result, as I illustrate below:

(7)*



It is clear that Goldsmith's proposal is unsatisfactory here, since the surface form is *ˈyévèsè*, and not *ˈyévésé*. In view of this, Pulleyblank's account constitutes an improvement on Goldsmith's in that it achieves observational adequacy, and also in that it offers an elegant and maximally simple account of tonal behaviour in the verbal system of Tiv[9].

Another example of this type of tonal behaviour is reported by Pulleyblank (1986) and Goldsmith (1988) to occur in Tonga. Although their approaches are different — the former is totally against automatic and universal associations whilst the latter is in favour of them — their analyses suggest that H is more robust than L. Even though a detailed examination of their proposals is outside the scope of this thesis, let us, nevertheless, turn briefly to the basic points of their analyses, which I summarise in the following paragraphs.

[9]Nevertheless, the universal scope of application of Goldsmith's proposal is also highly favoured in current trends of phonological theories.

According to Pulleyblank, there is only H-tone in the lexicon, and 'after the application of phonological rules in each cycle of the derivation, vowels that have no associated tonal autosegment will be assigned L-tones by default'. (1986:92). As in his previous example, the motivation for such a proposal is to support the argument that tone does not spread automatically.

On the other hand, Goldsmith (1988:90) claims that in languages like Tonga 'a trend is found towards replacement of an original Low tone by no tone at all'. Specifically, vowels in a verb-morpheme can be toneless. In the circumstances in which a toneless morpheme is attached to a H-tone morpheme, H automatically spreads to the vowels of the toneless morpheme.

Finally, in an analysis outside the context of Lexical Phonology, but still in the framework of Autosegmental Phonology, Hyman and Ngunga (1994) propose that

(8) only H tones occur underlyingly and in the lexical tonology. At a late stage in the postlexical tonology, toneless moras receive a default L. Default L tone assignment takes place after all the postlexical tone rules (Hyman and Ngunga 1994:29).

The motivation for this claim is that they defend the view that all H associations are accomplished by language-specific rules. If L were to be present in the lexicon, it would interact with the application of such rules, making it highly complicated. On the other hand, by confining the association of L to the latest stage of derivation, they secure a clear path for H association rules.

In the following paragraph, I present an outline of how contour tones are represented in current models of tone languages. Then, I give a preliminary indication of how the model I propose would account for them.

Yip (1989), Clements (1983) and others, represent contour tones as follows:

(9) (a) a (b) a phoneme tier

| |

x x skeletal tier

/ \ / \

H L L H tone tier

There are several independent tiers in this representation: the phoneme tier, where non-tonal segments stand; the skeletal tier, which is responsible for timing relations between segments; and the tonal tier, where Hs and Ls stand. Tones are attached to the skeletal tier by means of association lines, which also synchronise the phoneme tier and the skeletal tier. Depending on the ordering of tones, as illustrated in (9)(a) or (b), a falling or rising contour is accounted for[10].

In a single-tone model, an optimal account of contour tone is apparently unachievable, since, as I have already stated, the presence or absence of H only accounts for high level or low level pitch (if the absence of tone is interpreted as low level pitch). In order to overcome this, and following P'80's suggestion to associate some tones to the boundaries of prosodic domains in the field of intonation, I also propose to associate tone to the **boundaries** of a given domain, rather than to the syllable (or nucleus) **in** the domain. It is important to observe that such boundaries are already implicit in the arboreal structure shown below, and that they are not generated by an independent mechanism. For this reason, (9), which I reproduce below, can also be represented as follows:

(10) (a) a (b) a phoneme tier

| |

x x skeletal tier

/ \ / \

H L L H tone tier

(a')[11] $[\, \hat{a}\,]_x$ (b') $[\, \check{a}\,]_x$

The representations illustrated in (10) express exactly the same idea, that in the *x* domain,

---

[10]See Lombardi (1990) for a criticism of this proposal.

[11]Following the IPA marks, in the description of tone languages, I use (^) as a shorthand device to represent a falling movement, and (ˇ) to represent a rising movement.

there is a segment characterised by a falling pitch movement (as in (a) and (a')) or a rising pitch movement (as in (b) and (b')). Thus, the presence of boundaries follows from the assumption of a hierarchical structure[12].

Having adopted boundaries as potential TBUs, it is now possible to represent contour tones without referring to L, as I illustrate in (11) below:

```
(11)  (a)    H                (b)      H
             |                         |
             [ a ]x                  [ a ]x
```

The presence or absence of H is interpreted as high or low pitch respectively, so that (11)(a) is interpreted as a falling pitch movement, and (b) as a rising movement.

3 Intonation languages

3.1 The leftmost boundary tone

In this section, I introduce and discuss the nature of tone in the model proposed by P'80 for English. I argue that particular points in her own proposal can be understood as further evidence in support of the rejection of L from phonological representation. Specifically, I examine tonal associations to the initial intonation-phrase boundary, which include H and L or *zero*; that is, L and absence of tone are treated as equivalent. After that, I present a proposal made by Lindsey (1983), where he suggests that L% (both initial and final) should be discarded in the analysis of standard English dialects. The advantage of Lindsey's proposal over Pierrehumbert's is that his analysis restricts the generative capacity of the model, and also, shows a higher degree of generalisation.

In P'80 tone is associated to metrically accented syllables (accented syllables are identified as the highest positions in a metrical grid, which is constructed according to the principles included in Liberman and Prince (1977)), and to the edges of the intonation phrase, as illustrated in the following example[13]:

---

[12]This alternative representation will be formalised in chapter 4.

[13]See appendix to this chapter for the F0 contours included here (*That's a remarkably clever suggestion*, *Try occasional moderate agitation*, and *Another orange*, all from (Pierrehumbert 1980)).

(12) That's a re MARkably clever sugGEStion
H% L* H* L- L%[14]

Notice that, at the level of intonation, it is not the case that every single syllable is associated to tone, contrary to what we saw in former models of tone languages, where all syllables were associated to tone before they could be interpreted (in fact, this is one of the aspects which differentiates tone languages from intonation languages). Those syllables which remain unspecified in phonological representation receive their pitch specifications by means of a battery of interpolation rules in the phonetic component. For example, *-kably clever sug-* receive their pitch specifications by means of a rule which interpolates a gradient line between the low pitch on *MAR-* and the high pitch on *-GES-*.

In relation to the phrase-accent, this is not associated to any specific syllable, but is just in control of pitch specifications between the last accented syllable and the rightmost boundary tone[15]. In short, in Pierrehumbert's model the phonological characterisation of intonation has three components: a grammar which generates sequences of Hs and Ls, a metrical grid which captures prominence relations, and finally, rules for lining up the tune with the text.

The crucial point to observe in this model, is that, under certain conditions, the initial boundary tone can remain unspecified in phonological representation, and that it is under these conditions specifically, that L and *zero* convey exactly the same information. However, this approach to tonal distribution does not apply to accented syllables, which must always be specified as either H or L. Similarly, the phrase accent must always be present. As an illustration of this, consider the example below:

(13) Try oc CAsional MOderate agi TAtion
H* H* H* L- L%

[14]P'80 uses diacritic symbols to signal boundary tones (%), tones mapped onto accented syllables (*), and phrase-accent tones (-), which are mapped onto some unaccented syllables.

[15]Note that, in this particular structure, specifically in the sequence H* L- L%, L- is redundant, since its presence or absence in phonological representation has no difference in effect in phonetic interpretation. Hence, this is yet another argument in favour of excluding L from the representation.

In cases like this, the absence of tone at the left edge of the intonation domain is interpreted as mid or low pitch, or in Pierrehumbert's own words, 'the F0 onset may be some kind of neutral value' (1980:43). The remaining relevant tonal locations, accented syllables and phrase accents, are all associated to either H or L. It is on this type of evidence — the behaviour of the initial boundary tone — which I focus, in order to support the argument that it is possible to exclude L from phonological representation. But before we turn to examine yet another piece of evidence, let us probe a little deeper into Pierrehumbert's model.

Pierrehumbert's treatment of the initial boundary tone is based on an extension of an idea included in Liberman (1975:120-125), where, based on observation, he stipulates that an initial boundary tone cannot be aligned with an accented syllable. So, for example, an utterance like *Peter has a dog* — with tones associated as in (14) — is ill-formed, due to the fact that *Pe-* from *Peter* is already accented (L*), and hence, cannot also be associated to H%:

```
(14)*          Peter      has a  dog
       |         |                |    |
       H%        L*               H* L- L%
```

Liberman's proposal is restricted to this particular metrical structure, in which the first syllable is accented. He claims that in this case, "the meter would be wrong, since it is normally the beginning of a note that is felt as the 'beat'". If this is the case, then, *Pe-* would be interpreted on low pitch, and this would leave no space available for the interpretation of H%.

The main shortcoming of Liberman's idea is that the elimination of the above structure is not captured formally, and for this reason, the model can still generate ill-formed structures. As I shall show below, Pierrehumbert develops Liberman's idea further, by excluding L% from the representation, regardless of the metrical structure, but a principled account of the tonal specifications of this boundary is still missing.

As I stated earlier, Liberman's condition, in which an initial boundary tone can remain unspecified only if the first syllable of the phrase is accented, is reinterpreted by

Pierrehumbert in slightly different terms[16]. For instance, notice that in the structure in (13), in which the first accented syllable occurs well away from the left edge — and therefore, the presence of a left boundary tone is perfectly possible — the left edge itself remains unassociated. In fact, the absence of tone here is clearly not conditioned by the metrical structure. Instead, Pierrehumbert stipulates that if the F0 onset shows this type of neutral value, then the boundary tone can remain phonologically unspecified. On the other hand, if the F0 shows a rather high value, then the boundary tone should be specified as H. Thus, the choice of the presence or absence of a boundary tone at the left edge of the intonation phrase is primarily conditioned by the shape of the F0 contour in that particular location. In addition, she claims that the occurrence of H% in phonological representation is further justified by its correlation with a particular meaning; that is, H% adds a note of vivacity to an intonation pattern.

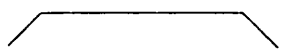
An important point to notice here is that the presence or absence of L% does not necessarily imply the existence of two different phonological representations. In addition, as I have already illustrated above, the presence or absence of L% results in identical phonetic interpretations. In view of these circumstances, in which L% has no particular task to perform (in fact, it is clearly redundant), and in which I pursue a minimalist model of intonation, I argue that the leftmost L% ought to be excluded from phonological representation. Thus, either H can be associated to the leftmost boundary, or otherwise, such a boundary is to remain unspecified. This is formalised by adopting a purely privative approach. This line of argument is also developed in Lindsey (1983), where he claims that if L% does not convey any semantic meaning (contrary to H%, which clearly does), then he can find no reason for maintaining L% in phonological representation any longer.

3.2 The rightmost boundary tone

In the context of a privative approach to tonal association, let me present yet another proposal, this time by Lindsey (1983), which suggests that this approach should also be adopted for the rightmost boundary tone (in relation to this metrical position, Pierrehumbert's model allows for either H or L to be associated, and the option for L to

[16]Note that Liberman's original condition still applies for H% in Pierrehumbert's model.

remain unspecified is not available). The motivation for this proposal stems from two facts. First, some patterns predicted by Pierrehumbert's model, for example, a H phrase accent followed by a L boundary tone, do not occur phonetically. Such a pattern, which would show the F0 shape illustrated below, is interpretable only with the final fall as nuclear (that is, H*L-L%)[17]:

(15) 

John went to the park

Secondly, her model misses an important semantic generalisation. According to Lindsey,

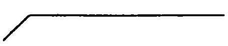
(16) All contours which terminate away from a neutral low pitch (i.e. terminal rises and mid or high levels) have a characteristic of 'incompleteness'. It would be nice to relate this to a common phonological characteristic, and the most obvious candidate would be the H% boundary tone (1983:31).

From this point of view, the contour included in footnote17 is accounted for by a H phrase accent followed by another H boundary tone.

With this account, Lindsey treats both edges of the intonation phrase **uniformly**: H can be either present or absent. In addition, this analysis nicely correlates with the events in the semantic component of intonation. Thus, these two factors, which allow for a greater degree of generalisation, offer a range of benefits over Pierrehumbert's proposal.

As the reader has doubtless noted, both Pierrehumbert's and Lindsey's model, and for that matter, any model of intonation which combines privativeness with equipollence, is destined to bring about some degree of indeterminacy to phonological components, especially to the nature of tonal units. Under an equipollent view, the predictions are that tonal units have two equally weighted values. Therefore, we understand that they stand

[17] Pierrehumbert is aware of this, and in order to save this phonological structure, she invokes a rule of up-step, whereby the phonetic value of L% is computed at the same level as that of H-. The interpretation of this structure shows the following shape:



John went to the park

the same chance of being present in phonological representation[18]. As far as pitch accents and phrase accents are concerned, this prediction is confirmed. However, soon after gaining some kind of understanding of the nature of tonal units, we encounter cases such as boundary tones, which show a privative behaviour. This is certainly discouraging, since this finding throws doubt onto our initial assumptions. The equally weighted values of tonal units are no longer balanced; one of the values is now more robust than the other, in that it enjoys a wider distribution.

Another point of indeterminacy is how the model decides which TBUs should bear a unary feature and which should bear a binary feature. It is obvious that a particular location of tone in the structure is not the factor. Note that pitch accents and phrase accents show the same type of equipollent tonal contrasts, and yet their distribution in phonological structure is completely unrelated. The former is associated to accented syllables and the latter is, in fact, floating[19].

In view of the questions raised above, and of the tendency in current generative studies to develop maximally restricted theories, one wonders why previous models of intonation have not rejected an equipollence approach to tonal contrasts and adopted a privative approach instead. The answer to this question is straightforward: with a single tone it is impossible to show a falling or rising pitch movement on a single syllable. In addition, in the event of rejecting L (as I suggest), those former models which had recourse to L as the trigger for **downstep** would be unable to account for this phenomenon. Moreover, models which allow L to **spread** would leave many contours unaccounted for. Hence, the expulsion of L from phonological models creates problems which are even more fundamental than the original ones. Ideally then, what we need is a coherent model which explicitly defines the nature of tone, and can account for falling and rising pitch movements, and for the phenomena of spreading and stepping.

Having presented evidence which points to the possibility of rejecting L from phonological representation in tone languages, and from specific locations in models

---

[18]See discussion of this issue in chapter 1.

[19]Recall that in subsection 3.1 I pointed out that in P'80, the phrase accent is in charge of pitch specifications between the last accented syllable and the rightmost boundary tone. Note that it is not actually associated to any syllable in particular. Therefore, it can be treated as *floating*.

which describe intonation, in the following section I begin the pursuit of an alternative model without L tone, which can account, at least, for contour tones. In essence, I claim that the key factor to such an alternative is to be found in the nature of TBUs. I shall propose that, rather than associating tone to accented syllables, one of the alternative options, which is to associate tone to the boundaries of prosodic domains, should be pursued. Since, by definition, each domain shows two edges (or boundaries), then it is possible to account for contour tones by associating different values to each boundary, as in the fashion already illustrated in §1.

4 Towards the representation of contour tones

In this section, I propose that tone is associated to a single type of TBU, which are the boundaries of prosodic domains. The views presented here constitute an outline of the ideas I developed in Cabrera-Abreu (1995). There, I claimed that the association of tone to two different types of TBUs (boundaries and accented syllables) forces us to posit independent operations for tone association, and hence, this prevents us from producing a unified account of tone association. On the other hand, by pursuing a single type of association, a certain degree of generalisation is achieved. In the interests of restrictiveness, I choose the option of associating tone to boundaries exclusively; (this idea was developed, initially, in the context of a model with two tones, H and L).

Here, however, the motivation for such a choice becomes much clearer, when we acknowledge a model in which only a single tone is present. In order to clarify this point, let us briefly remind ourselves of some of the implications of choosing the option to associate tone to accented syllables exclusively. First, if there is only one tone which can be associated to a syllable, then only level pitch specifications can be accounted for. There is no way in which a contour tone can be represented. This is already highly undesirable since it would render the model incapable of accounting for very common features of intonation. Secondly, in a situation in which the accented syllables are located away from the edges of the intonation contour, it is impossible to account for any of the F0 shapes in these particular locations. This would entail the loss of a great deal of information, which is crucial in the encoding of intonation patterns. In view of all these complications, a single-tone model in which accented syllables are the only TBUs is impossible to maintain.

Having clarified this point, let us turn to an investigation of the viability, and at the same time, the benefits of models in which tone is associated to boundaries only. In order to focus on this topic, momentarily, I shall take a step back in the argument concerning the rejection of L from the model, and allow for both H and L to be present in the description of pitch contours. I shall devote the following section to an overview of a proposal included in Hirst (1988), and after that (in §4.2), I shall focus on another model, the ideas of which are adopted by and developed in Grice (1992). By the end of the section, however, it will become clear that, by adopting the 'boundary only' proposal, the representation of contour tones as an argument against the exclusion of L from phonological representation, loses its validity.

4.1 Tonal association in Hirst (1988)

Hirst's (1988) approach to tonal association is mainly characterised by assigning tonal templates such as H-L or L-H[20] to **phonological domains**, which he refers to as the Tonal Unit (TU) and the Intonation Unit (IU). The former domain corresponds roughly to Bolinger's (1958) pitch accent, and the latter corresponds to the intonation or tone group. Let me illustrate this with an example from English[21]. First, in (17) below I show the prosodic structure of a phrase such as *a cup of coffee*:

(17) $\left[\text{a} \left[\text{cup of} \right]_{\text{TU}} \left[\text{coffee} \right]_{\text{TU}} \right]_{\text{IU}}$

The tonal template for TU is (H-L), and for IU can be either (L-H) or (H-L), depending on whether we are required to account for a rising or a falling intonation respectively. Let us now assign these templates to the structure in (17) in order to generate such representations:

[20] See discussion below for the make-up of tones.

[21] See Hirst, Di Cristo and Nishinuma (1994) for an analysis of French.

(18) (a) $[\text{a} \; [\text{cup of} \;\; \text{H} \;\; \text{L}]_{\text{TU}} \; [\text{coffee} \;\; \text{H} \;\; \text{L}]_{\text{TU}} \; \text{H}]_{\text{IU}}$ (b) $[\text{a} \; [\text{cup of} \;\; \text{H} \;\; \text{L}]_{\text{TU}} \; [\text{coffee} \;\; \text{H} \;\; \text{L}]_{\text{TU}} \; \text{L}]_{\text{IU}}$

In relation to the identity of TBUs, Hirst claims that

(19) There are [...] no *tone-bearing segments.* Rather, tones and phones are held to be produced by separate mechanisms with the only temporal constraint being that the given sequence be realised within the limits of the Tonal Unit (1988:161-162).

In other words, tones are not associated to segments, but instead, tonal patterns specify the pitch characteristics of the segments of a whole domain, by virtue of being associated to prosodic boundaries.

Hirst's motivation for such association is that it captures the timing relationship between tones and segments which does 'not generally coincide in any simple one-to-one way with the stressed vowels, nor even with the stressed syllables'. He then adds that if a phonological/acoustic model were designed to interpolate between F0 steady states, these states should be defined at the edges of something like the TU, and not by accented syllables. Given this situation, Hirst no longer sees any motivation for maintaining the metrical grid in his model.

However, this idea is dropped in a later paper. In Hirst, Di Cristo and Nishinuma (1994), prominence is informally determined by following a battery of rules, and then, this is formalised in terms of the construction of prosodic structure. For instance, after describing rules relating to the division of an utterance into intonation units, and to the assignment of accent to various syllables of such domains, they suggest that TUs in French are right-headed, rather than left-headed (as in the case of English). Although they do not illustrate prosodic structure, they acknowledge that this analysis runs along the same lines as that of Halle and Vergnaud (1987), which shows constituent structure by means of grid construction.

The model proposed by Hirst et al. (1994) is certainly attractive for the following reasons:

(20) (a) it has a universal scope of application;

(b) it has a fairly simple tonal structure;

(c) tones are uniformly associated to boundaries of prosodic domains;

(d) it shows prominence relations by means of prosodic structure.

In relation to (20)(a), the model has already been applied to languages such as English, Spanish, Arabic, Italian, German and Dutch. The universality of the model is achieved by the formulation of a set of parameters, each of which can assume different values; languages will show different typologies depending on the value of the parameter they choose. For example, Hirst et al. claim that right/left headedness of accent groups establishes a typological difference between Romance and Germanic languages, respectively. It follows from (20)(b)) that there are no composite tones. This is in sharp contrast with models *a là* Pierrehumbert which allow bitonal pitch accents. The fact that the tonal structure is so simple, together with a unified association of tone to boundaries, renders this model a highly restrictive, and also, attractive proposal for an account of intonation.

As we shall soon see in chapter 4, the model I propose inherits those aspects from Hirst and his followers, which I summarised in (20). Therefore, to this extent they are both on a par. However, as I discuss below, there are a few aspects in which my model constitutes an advantage over his.

First of all, tones in Hirst (1988) are defined by means of binary features. [±high] and [±low] account for 4 different tones:

(21)

Features	*Tones*
[+hi -lo]	H(igh)
[+hi +lo]	D(ownstep)
[-hi -lo]	M(id)
[-hi +lo]	L(ow)

P'80 has already argued in favour of a model with a single binary feature to account for two-tone systems like English, for which it has been proved that D and M are no longer necessary. In addition, as widely argued in the literature (see van der Hulst and Snider (eds) (1993) and references therein), not only does this type of model fail to capture natural classes, but also, it allows for two-tone systems to be described in eight different ways. Furthermore, as I shall prove conclusively in chapters 4 and 6, English can even be described in terms of a single unary feature. Thus, Hirst's model is not maximally restricted, and consequently, allows for overgeneration.

Another aspect to take into consideration is that it is not immediately clear how the model under discussion captures the observation that *coffee* is more prominent than *cup*, and also that, as a consequence of this, the resulting phonetic interpretations of each of these two prosodic constituents are rather different from one another: in *cup of* the F0 trace falls gradually, whereas in *coffee*, it falls sharply. By contrast, in chapter 4, I shall suggest that these differences constitute a reflection of licensing relations between units in phonological structure, and furthermore, that prominence relations between **all** units are formally captured under the principles of licensing.

Having presented some evidence which supports the viability of a model characterised by the association of tone exclusively to boundaries, let us now turn to the analysis of another model which points in the same direction. Although this one is characterised by a dual-type of association (to accented syllables and boundaries), I shall prove that, after performing the necessary adjustments, this too can dispense with association to accented syllables. Thus, the path is clear for a unified association of tone only to boundaries.

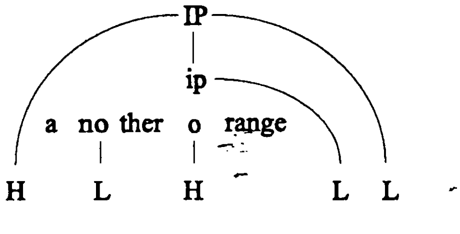
4.2 Tonal association in Grice 1992

Earlier works on intonation (P'80, Ladd 1983, Gussenhoven 1983, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, and Grice 1992 among others) share the assumption 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)[22] and the intermediate phrase (ip). The former type of association is normally

[22]The IP is equivalent to Hirst's Intonation Unit.

referred to as *central* and the latter as *peripheral.* This is illustrated below[23]:

(22)



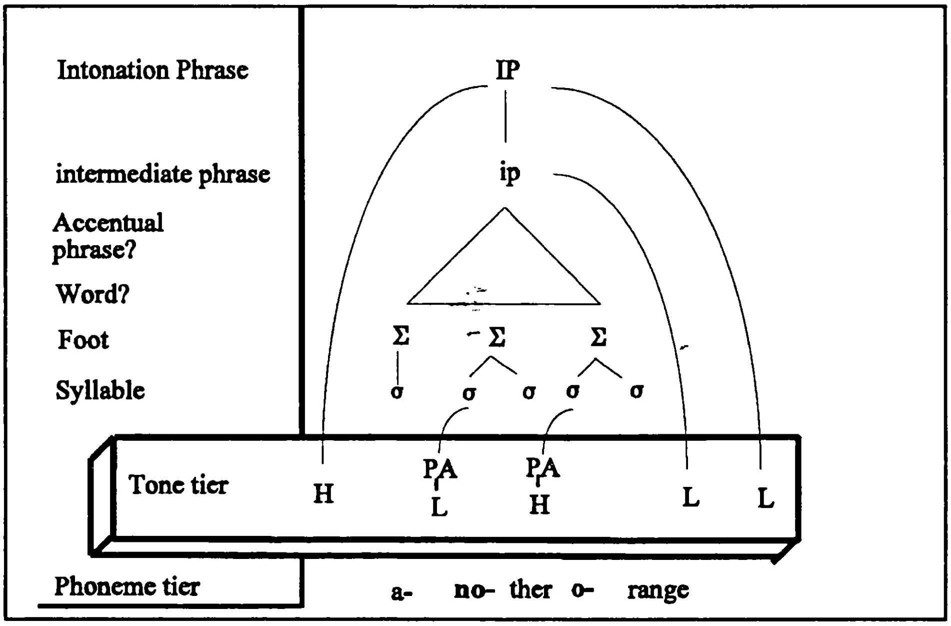
Let us focus only on those details which are relevant to the present discussion. For that, I shall concentrate on the model put forward in Grice (1992), since this constitutes some of the most recent work on the phonology of intonation which not only counts as one of the major contributions to intonational studies, but also, gathers the fundamental tenets of previous proposals into a single piece of work.

As illustrated in (23), there are various structural components in the typical autosegmental representation of a phrase: the prosodic hierarchy, the tonal tier, the association lines and the phoneme tier[24].

[23]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.

[24]For the purposes of the present discussion, I shall concentrate on the actual tonal association of tone to TBUs. The reader is referred to the original sources for a detailed account of the other components of the structure.

(23)



Also illustrated in (23), tone is associated to two categories of TBUs: 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 in addition, 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 (23) for the purposes of clarity in that diagram. (A simplified version and an extensive discussion is included in §4.3.1).

In the higher levels of the hierarchy, there is already a factor which 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*[25]; at this level, tone is obligatorily associated to the right edge, but optionally associated to the left edge; (the reverse argument also applies:

[25]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.

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 the formulation of 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 issue addressed in this section, 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 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 constrained in the following ways: first, tonal structure is restricted to a single tone; and second (as I shall suggest later in §4.4, 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 presents itself as the favoured option.

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

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

(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 (24)(a) the model's potential to generalise is boosted: the principles of association are the same everywhere in 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 (24)(b) the model's generative power is kept to a minimum, and therefore, under control. Finally, as I shall explain in the following sections, by (24)(c) 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. This assists in the mapping of phonological representations directly onto phonetic intepretation, without the risk of having to postulate the existence of an intermediate level in which interpretation rules would operate. There is evidence in the literature (Hirst 1988) which, as we saw in the previous section, 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. I shall tentatively suggest that these segments may 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 sufficient 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 presented an outline of some of the components of Grice's model, and I have also argued for an extensive revision; I have rejected central association and

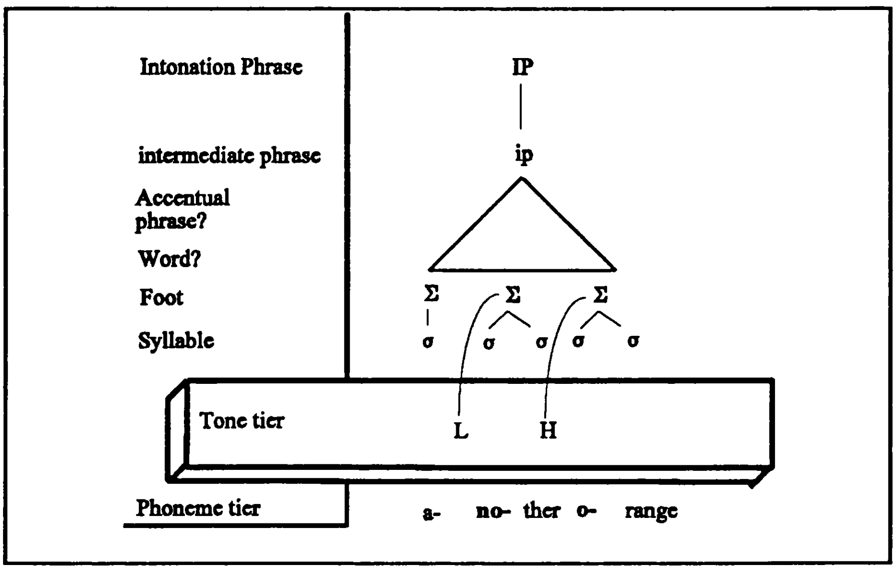
I have cast some doubt on the use of the metrical tree for the representation of metrical relations, albeit without any detailed discussion.

4.3 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 §4.3.2, 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 (§4.3.1), tone is no longer associated to accented syllables (as in (23)), but instead, can only be associated directly to the edges of a prosodic domain — in this case, the foot. 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 (25) below.

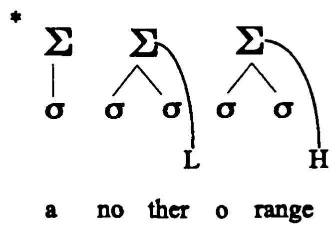
4.3.1 Association to feet

(25)



A question immediately arises as to whether the L and H tones which are associated to the left side of the foot in (25) 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 (26)) 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:

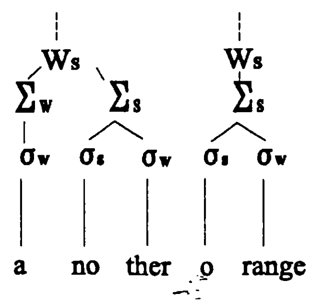
(26)



On the other hand, if the right syllable of a foot were the prominent one, then the right association represented in (26) 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[26]. 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 (Liberman 1975, Liberman 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:

[26]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 either 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.

(27)



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 degree 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 (27), is the Word. As stated by tree-building-rules, more levels would have to be introduced in the hierarchy (this is shown in (27) by means of a dashed line 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 deemed the most prominent one in the entire structure; in (27) this is *o-* in *orange*. This entire 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[27] 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

---

[27]It is understood that no tone is associated to the weak foot *a-* because it contains a weak syllable which cannot bear an accent.

association of tone to a strong foot is irrelevant, as long as the syllables in the foot 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-prominence relations; the *s/w* labelling of a syllable conditions tone-association in the following way[28]:

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

Given this, the structure in (25), rather than the one in (26), shows the proper association of tones; therefore, (26) is immediately ruled out.

As I mentioned 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 suggest that tree-format at this level be rejected.

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.

4.3.2 Association to levels above the foot: *ip* and *IP*

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

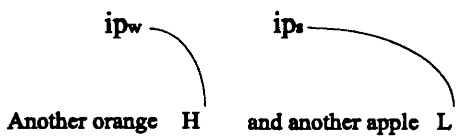
---

[28]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.

it leads to further undesirable stipulations.

Having the peripheral association shown in (23) in mind, let us begin our task by making an attempt to apply the tree structure to a couple of *ip*s, like the ones in (29) below: *Another orange* followed by a H tone, and *and another apple* followed by a 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 §4.3.1 above, and consequently, the leftmost is labelled *w*. Once they are properly labelled, I can 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[29].

(29)



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

[29]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.

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 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:

(30)
- (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 the 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[30] which offers solutions to all these points. This constitutes the topic of the

[30]The reader is referred to Hirst (1988:151-165) 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.

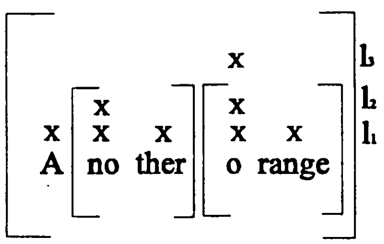
following section.

4.4 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. For this, I shall present a crude construction of the grid[31], since I shall devote the following chapter to a formal construction of the prosodic hierarchy based on the principles of licensing.

For the purpose of illustrating grid-structure, consider the structure in (31).

(31)



In (31) 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[32]. Then, the positions which act as the heads of feet are projected onto the row one level higher.

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).

[31]The reader is referred to Selkirk (1984) and to Halle and Vergnaud (1987) as the main sources of this topic, and to Cabrera-Abreu (1994) for a preliminary attempt at its application in the description of pitch patterns.

[32]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.

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 this position neither acts as a head nor is it integrated into a foot. I suggest that positions before the head of feet are incorporated into the *IP* domain. Lastly, the head 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[33], 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 relating 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 point, I would like to draw a parallel between this type of situation and that found between the onset and the nucleus

[33] In the following chapter, we shall see that this idea is formalised by adopting the principles of licensing.

constituents in the context of Dependency Phonology (Anderson and Ewen 1987) and Government Phonology. In these frameworks, 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 enter into 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 (32):

(32)

			x		L_3
	x		x		L_2
x	x	x	x	x	L_1
A	no	ther	o	range	

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 had been previously described as quite different in terms of their metrical patterns, would then be united by the claim that they obey the principles of licensing. Any differences would arise as a consequence of the selection of a different setting of various parameters.

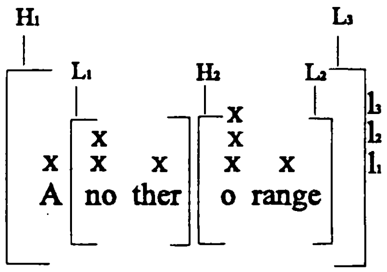
Another advantage of using the principles of licensing over the rules of *s/w* labelling is that the former belong to a set of principles and parameters (left/right headship) which operate throughout phonological representation, from the highest constituent to the lowest, whereas the latter seem to work satisfactorily only for levels below the foot.

In the previous paragraph, I suggested that one of the fundamental tenets of the phonological representation of intonation which I propose are the principles of licensing and the parametric variation that they entail. However, I have put forward merely an outline of how this can be achieved, since this will be further formalised in chapter 4.

4.4.1 Tonal association and the metrical grid

Let us now proceed to associate tones in (23) to the boundaries of the structure in (31); 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 (33)[34].

(33)



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 *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 other models, are now avoided.

At this stage of the present work, the mapping of tones to boundaries is only descriptively adequate, and it lacks any explanatory power; for instance it is still not clear why only one boundary is left without association (instead of following Universal Association Conventions (Goldsmith 1990), which require that no TBUs are left unassociated in phonological representation). This will become clearer in chapter 4,

[34]In the following chapter, association lines constitute the graphic representation of Licensing Inheritance.

where I shall invoke the principles of licensing to control tone association. Specifically, I shall claim that due to licensing relations, and to the fact that this boundary counts as a licensee, it does not inherit enought licensing potential to sustain tone.

A preliminary phonetic interpretation of (33) is now straightforward: H is interpreted as rather high pitch, and L (or *zero*) as rather low pitch. (See appendix to this chapter for the F0 contour of this structure included in Pierrehumbert (1980)).

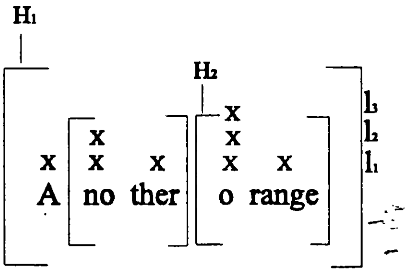
Thus far, the advantages of using the grid format (or a similar type of format), over the tree format in the representation of intonation are the following:

(34)
- (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 can be restricted to a single tone (as I shall show in the following section).

5 Contour tones

In this section, I bring together the claims made at the beginning of this chapter, and those presented in the immediately preceding sections: respectively, that L can be excluded from phonological representation, and that tone can be associated to boundaries of prosodic domains throughout phonological structure. Thus, the final task in presenting a satisfactory account of contour tones is to revise (33) according to the assumptions mentioned above. I do this in (35) where L is no longer associated to any boundaries, and therefore, is completely absent from the representation.

(35)[35]



6 Summary

In this chapter I have presented part of the foundations of a phonological model of intonation without L tone. I have also overcome a major obstacle towards the achievement of such a model, namely the representation of contour-tones. In order to account for such shapes, I have argued in favour of associating tone to the boundaries of prosodic domains.

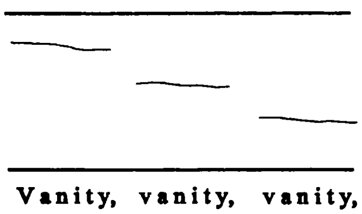
[35]This structure should be treated as a preliminary attempt to account for this F0 contour, since in chapter 4 the manifestation of licensing relations between boundaries dictates that the leftmost boundary of the intonation phrase must remain toneless. See §6.2 for a proposal of an alternative account.

Chapter 3. Spreading

1 Introduction

Another task to be carried out before I introduce the reader to the grammar of the single-tone model in the following chapter, is to present a preliminary description of how such a model can account for *spreading*, a phenomenon which is phonetically manifested as level pitch in the F0 contour. For instance, the utterance *Vanity, vanity, vanity* can be uttered in a sequence of **levelled** F0 terraces, as illustrated in the following diagram, where the outer horizontal lines correspond to the speaker's F0 range, and the inner lines correspond to F0 traces:

(1)



In the spirit of autosegmental phonology (Goldsmith (1976) and followers), a possible representation of the pitch pattern of these words would be as follows[1]:

(2)[2]

```
Vanity  vanity  vanity
  |       |       |
  H      !H      !H
```

Let us assume (by Goldsmith's WFC) that H can propagate to the right so that all TBUs are specified for tone. This is illustrated below:

[1]See §2.2 in the preceding chapter for an illustration of spreading in tone languages.

[2]The exclamation mark to the left of H indicates that H to its right has undergone downstep; that is, that the F0 level is shifted downwards. However, the representation of downstep is not relevant for this example. See chapter 5 for a detailed discussion related to this topic.

(3) Vanity vanity vanity

H !H !H

Thus, level stretches of pitch may arise as a consequence of the multiple association of a given tone to a number of TBUs. Another case which can result in level pitch is the association of a set of equally valued tones to a single position. An example of this can be seen in §2.2 in chapter 2.

Claims in the literature (Goldsmith (1976), P'80, Lindsey (1985), Pulleyblank (1986), Johnson and Grice (1990) (henceforth J&G), and others) suggest that both H and L tones can spread. In this context, the model I attempt to develop may face some difficulties, since the fact that it is characterised by a single tone raises the question of how to account for those cases which were formerly described in terms of L-spread. As I shall explain in §2 and §3, such instances can be treated as the result of phonetic interpolation between toneless units, and not as a phenomenon which is triggered by the presence of L in phonological representation. In this chapter, I shall present an overview of how previous models for tone and intonation languages deal with spreading. Specifically in relation to tone languages, I shall turn briefly to the analysis proposed by Pulleyblank (1986) for Tiv and Tonga, which I introduced in §2 of chapter 2, and I shall suggest that in the light of the revision of his account in terms of the phonological framework I have adopted here, spreading can be treated as a matter of phonetic interpolation.

As far as intonation languages are concerned (in §3 below), earlier models have fuelled some degree of conceptual antagonism between phonological and phonetic spreading. By means of phonetic interpolation rules, P'80 and Pierrehumbert and Beckman (1988) ensure that toneless vowels in English and Japanese, respectively, receive tonal specifications. Arguments against phonetic spreading are present also in the literature. By means of distinctive features, Ladd (1983) and Lindsey (1985) supply enough information to tones in phonological representation to be interpreted as level pitch in phonetic representation. In addition, in a pilot study, J&G also propose that spreading is accounted for by means of a feature [±spread], but they add that *stylised intonation* (F0 patterns showing long stretches of level pitch with a semantic correlate of routine and lack of emotion) is accounted for by means of an extra feature ([±laryngeal precision]) in phonological representation.

Due to the shortcomings in the use of features — which can be summarised as the potential for overgeneration — these views are not pursued any further in later models of intonation (for example, Grice (1992), and the one I develop here), but instead, the alternative approach — to treat spreading as a matter of phonetic interpolation — is widely accepted.

2 Tone languages

In chapter 2, §2, I mentioned briefly the analyses of Tiv and Tonga proposed by Pulleyblank (1986) on the one hand, and by Goldsmith (1976, 1988) on the other. Recall that Pulleyblank defends the view that in Tiv, for example, those tones which are specified in the lexicon, do not spread automatically. Instead, they are associated to vowels in the first cycle. Later in the derivation, those vowels which remain toneless are assigned L-tone by default. By contrast, Goldsmith has recourse to universal spreading of lexical tones as a device to achieve the well-formedness constraint that all TBUs must be specified for tone. Even though these analyses show different views of the characteristics of phonological systems, the results presented at the end of this discussion can be applied to both systems.

Pulleyblank favours the view that L-insertion rules operate in the last stage of phonological derivation, and that the output of these rules are phonetic shapes characterized by stretches of low pitch. Therefore, phonological representations are uninterpretable during derivation, as the value of toneless vowels is unspecified. Vowels become interpretable only after a value is specified for them.

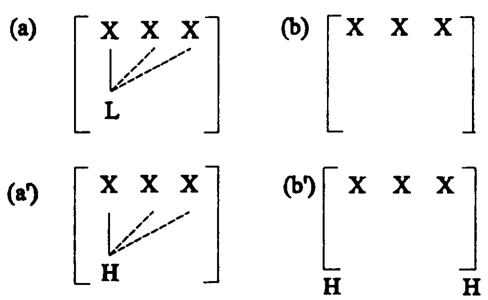
Let us now revise this type of analysis in terms of the *competence/performance* view which I introduced in chapter 1. Recall that, according to this approach, phonological structures are interpretable at all stages of derivation, and that the input and output of phonological processes are of the same type. Clearly, if such a framework is followed to its ultimate consequences, then there is no place for the type of derivation adopted by Pulleyblank. That is, insertion rules which operate in the last stage of the phonological component in order to assign a value to unspecified units, do not constitute a conceivable option, since such units would remain uninterpretable throughout derivation, until this stage is reached[3].

Also, notice that if L is treated as the default tone which is inserted by rule at the very end of derivation, as Pulleyblank claims, it will never stand a chance of participating in any phonological process involving such a value. Thus, there is no reason for assuming that L-spread may occur in the phonological component in the first place. Toneless units are automatically interpretable as relative low pitch, and do not have to wait until the last stage of phonological derivation in order to be interpreted.

The argument against Pulleyblank's type of analysis is based on the assumptions of those who favour a radical privative approach, that is, presence/absence of a unary phonological value (Anderson and Ewen 1987). This contrasts with an equipollent approach based on phonological units with binary values, which is typically adopted in, for example, Underspecification Theory (Archangeli 1984) and Lexical Phonology (Mohanan 1982).

Thus, in the model I propose, the manifestation of stretches of low level pitch in the F0 contour constitutes the interpretation of phonological structures in which H is absent. According to this, previous representations of L-spread in tone languages, which look like that illustrated in (4)(a), are now captured as in (4)(b); that is, TBUs remain toneless (assume for the time being, that the square brackets in (4) correspond to the limits of a prosodic domain, say, the word):

(4)



In the case of (a'), I illustrate H-spread as is customarily found in previous autosegmental analyses, while (b') is couched within the proposed single-tone model[4]. As

[3]See §1.1 in chapter 1, and §1 in chapter 2 for a discussion of some of the benefits of this approach.

[4]I shall adopt this structure later in chapter 6, §5 to account for stretches of level pitch in the so-called *calling* contours.

can be seen, TBUs are no longer the positions inside a domain, but rather, they are the **boundaries** of domains[5]. Since both boundaries show H associated to them, the interpretation of this structure results as high level pitch. In other words, I suggest that H-spread arises as a consequence of **interpolation** between two Hs rather than as a consequence of a phonological item being added to the phonological representation. This item could be an association line, as in autosegmental analyses, or a distinctive feature [+spread], as I shall show in the following section about intonation analyses.

The view that H-spreading arises as a consequence of phonetic interpolation constitutes an advantage over the view which treats spreading as a phonological phenomenon. The reason for this is that the phonological component is released from what turns out to be an unnecessary burden. I have shown that Goldsmith's use of extra association lines, and also Pulleyblank's L-insertion rules, may be dispensed with. This manoeuvre results in a gross simplification of phonological representation. Moreover, by reducing the tone inventory to a single tone, the number of possible phonological representations and processes is dramatically reduced, thus yielding a maximally constrained model. In addition to this, by associating tones to boundaries in a straightforward fashion, it would be possible to dispense with Goldsmith's Association Conventions, thus rendering a simplified representation of tonal association. However, the details of the application of such a model to the analysis of tone languages falls outside the scope of this thesis, and is left for future research.

We may conclude this section by claiming that what was formerly treated as L-spread, is analysed in the present work as a matter of interpolation between toneless TBUs in phonological representation. In relation to H-spread, preliminary analyses indicate that this may be treated as a matter of interpolation, too. With these assumptions in mind, let us turn to the field of intonation in the following section, in order to test their validity against some data from English.

---

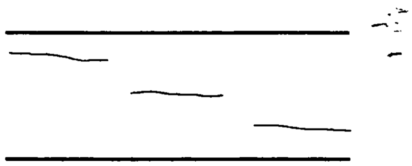
[5]Recall that this has already been anticipated in chapter 2.

3 Intonation languages

3.1 Spreading results from interpolation

I begin this section by illustrating the analysis proposed by P'80 to account for the contour included in (1):

(5)



Vanity, vanity, vanity,

H* H-+L* H-+L* H-L%

P'80's model allows for tones to be grouped into pitch accents, which can be either monotonal (as H*), or bitonal (as H-+L*). The starred tone of a bitonal pitch accent associates to the accented syllable, in this case, *va-*. The other tone is set afloat, that is, it is not associated to any particular syllable. Its only function is to act as a trigger of downstep[6]. The rightmost boundary tone is interpreted towards the edge of the utterance (and it could equally be associated to the boundary of the domain, as illustrated in the preceding chapter, §3.1). The phrase accent, which is the tone in between the last pitch accent and the boundary tone, is responsible for pitch specifications in this area.

The first pitch accent H*, together with the following H-, account for the first terrace of pitch. Moving on to the right, L* is responsible for pulling down the interpretation of the following H-, so that it is manifested at a level lower than that of the preceding H-. Note that the stretch between the first L* and the third H is characterised by level pitch, and that this pattern is repeated again to the right. In relation to the interpretation of L%, its value is computed at a level higher than in other circumstances, for example, when preceded by L-. This particular height is achieved by means of a rule of upstep, which ensures that the value of L% is computed at the same height as that of H-. Note that, so far, H- and L% are interpreted as two targets at the same height, but this is insufficient to achieve the type of level interpolation we aim at. Hence, P'80

[6] I shall ignore stepping in this chapter, since I devote chapters 5 and 6 to a detailed discussion of this phenomenon.

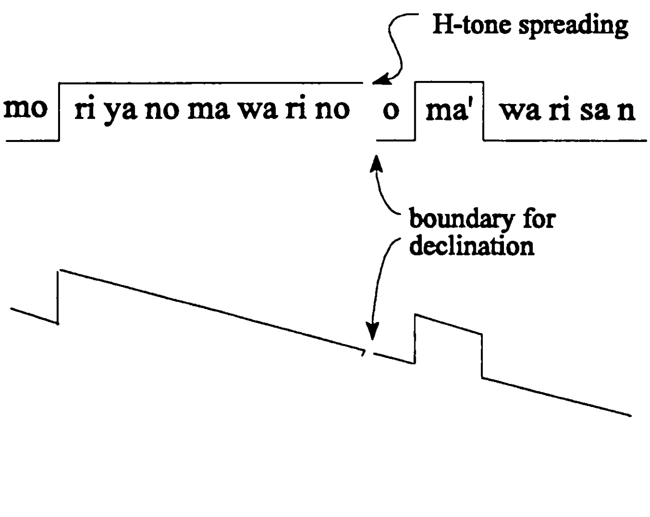
suggests a rule of spreading, whereby the value of H- extends its association to the right, and thus, the F0 results in the desired terrace between H- and L%. This is formulated in terms of the following rule:

(6) $T\text{-}_i$ spreads towards T_{i+1} if $/T_{i+1}/ \geq /T\text{-}_i/$

This type of rule, together with the upstep rule and others, which are in control of the implementation of phonological representations as F0 contours, lead us to conclude that, in Pierrehumbert's terms, spreading is a matter of phonetic interpolation.

Additional evidence in favour of a phonetic treatment of spreading is put forward by Pierrehumbert and Beckman (1988) in their analysis of Tokyo Japanese, which is a pitch-accent language. They claim that the F0 shape which results from an unaccented phrase, like, for example, *moriya-no mawarin-no* in an utterance such as *moriya-no mawari-no oma'warisan* 'the policeman of the Forrests' neighbourhood', is a far cry from an interpretation which could correspond to H-spread. On the contrary, the unaccented phrase shows traces of having undergone declination[7]. In the following diagram I reproduce their figure 2.7 (p.37) which illustrates, first, the expected stylised shape of H-spread, and second, the actual stylised shape which is found not as a consequence of H-spread, but instead, as a consequence of declination.

(7)



[7]According to Cruttenden (1986:167) 'declination is the phenomenon whereby pitch [...] is on average lower at the end of an intonation-group than at the beginning.'

The break in the F0 shape between the two phrases corresponds to the manifestation of L%, which signals the limits of the phrases. This type of evidence leads Pierrehumbert and Beckman to reject H-spread as a phonological phenomenon.

Although I shall follow Pierrehumbert in the treatment of spreading as a matter of interpolation, I object to her proposal in that, in order to achieve a phonetic interpretation for a given phonological structure, she has recourse to a set of interpolation rules. One of the reasons for rejecting Pierrehumbert's battery of rules is that the input and the output of these rules are representations which are dissimilar in nature: the input is a phonological structure, and the output is a phonetic interpretation[8]. Another, and more important reason is that, under the competence/performance view, there is no place for interpolation rules, since these do not qualify either as cognitive, grammatical constraints, common to all members of a community, or as physical constraints common to all humans or peculiar to each individual. In the model I shall present in the next chapter, phonetic interpretation follows from a general principle (*Gravitation Effect*) which stems from the manifestation of physical constraints, rather than from rules.

3.2 Stylised versus non-stylised contours[9]

3.2.1 Ladd (1983)

Ladd (1983) suggests that spreading arises as a consequence of encoding distinctive features in phonological representation[10]. The motivation for making a distinction between spreading and non-spreading (or *stylised* and *plain* in Ladd's own terms) contours is that this correlates with events in the field of meaning. Basically, all spreading versions share the same type of meaning, which, he claims, is added onto the non-spreading version. For example, the difference between the contours illustrated below in (8) would rest on the idea that the first version signals routine, stereotype, 'lack

---

[8]Recall that this constitutes one of the reasons for also rejecting Pulleyblank's model in the preceding section.

[9]See House (1995) for a re-analysis of this type of contour associated with the specific case of telephone enquiry service dialogues.

[10]See also Lindsey (1983), (1985) for a similar proposal.

of information transfer', even ritual (Ladd 1983), whereas the second version is free from such an array of meanings, being understood just as a question.

(8) (a)  (b) 

Can I go now   Can I go now

H- - -   L* H H%

The stylised contour is analysed by Ladd as H followed by (- - -), which is a shorthand device to refer to a feature [stylised]. The plain contour lacks this feature.

Ladd's proposal achieves a high degree of phonological generalisation by having recourse to features in the classification of intonation contours. By means of the feature [stylised] he classifies contours into natural classes: stylised and plain. Another benefit gained from his proposal is that the degree of abstractness inherited in earlier models, especially P'80, is drastically reduced. For instance, although in P'80's model the natural class of stylised contours is characterised by having L% at the right edge, the actual interpretation of this tone is far from being low pitch, which is what would have been normally expected; however, contrary to our expectations, L% is interpreted rather high.

Let us now examine the shortcomings present in Ladd's proposal. For this, I shall concentrate on the criteria he adopts for the classification of contours as distinctive.

First of all, Ladd distinguishes between plain and stylised contours by means of the meaning they convey. Recall that, in chapter 1 I have already pointed out the potential problems in the adoption of such an approach based on meaning, namely that it can be extremely vague in intonation. Furthermore, very frequently the nature of meaning which the intonologist is referring to is not made explicit as to whether it is semantic or pragmatic. Thus, meaning as a criterion for defining contrastive F0 contours may indeed lead to very different — and sometimes contradictory — taxonomies.

Another problem in Ladd's model is related to an inconsistent use of features, namely, whether they display a privative or an equipollent distribution. For example, in his account of a sequence of stepping levelled terraces, not only does he use a unary feature [stylised] to account for the flat terraces, but also, he falls back on a binary feature,

[±floor][11], to account for the fact that the pitch does not fall as low as the baseline. As I stated in §3.2 in chapter 2, the combination of unary and binary features in a single model raises fundamental questions as to the criterion that defines which approach to follow in particular circumstances. In addition, it appears that the creation of features to describe intonational patterns is totally unconstrained, since there seems to be no limit to the number that can be postulated. For instance, we could similarly posit a feature [±declination] (or for that matter [declination]) to account for the general tendency of intonation contours to move towards the baseline. However, this proposal is highly disfavoured, and thus far, has not been adopted. This may be due to the fact that this phenomenon results as a consequence of a universal physical constraint, say, breathing, and that it does not arise from a cognitive constraint. But notice that, there are no means by which the formalisms in the theory could prevent us from suggesting this feature.

In the following section, I shall present an outline of an alternative view to spreading. This is also characterised by the idea that features in phonological representations are responsible for spreading in the F0 contour.

3.2.2 Johnson and Grice (1990)

Inconsistencies in the categorisation of patterns showing spreading in earlier analyses, (P'80, Beckman and Pierrehumbert 1986, Ladd 1983, and Gussenhoven 1984) lead J&G to examine this type of contour in more detail. Their findings suggest that pitch contours characterised by flat stretches of pitch should be categorised as showing an exponent of *stylisation.* They add that such an exponent arises as a consequence of 'increased laryngeal control'. Contours which lack such control may still be uttered with monotone intonation. This distinction is captured by the use of a categorial feature, [laryngeal precision], which is attached to the tone undergoing stylisation. For instance, *Daddy fell downstairs*, as illustrated below, is a stylised contour. By contrast, *I wouldn't be so sure of that!* is not stylised, although it also shows monotone terraces:

[11]See also Johnson and Grice (1990:237) where the authors point out Ladd's inconsistent manipulation of these features.

(9)[12] Daddy fell downsta:
irs

I
that!
sure of
wouldn't be so

Thus, the former would be analysed as [+laryngeal precision], [+spread] (as the feature which accounts for monotone), and the latter as [-laryngeal precision], [+spread].

A third type of contour, characterised by pitch movements (peaks and troughs), is accounted for by [-laryngeal precision], [-spread].

According to J&G, their analysis has certain advantages over former analyses in that the categorisation of stylised contours is more straightforward, since the feature [laryngeal precision] can accompany tones in both nuclear and prenuclear position. Nevertheless, they acknowledge that a **formal** analysis of this type of contour still remains undefined.

J&G's study of contours involving spreading is inconclusive, and therefore, their suggestions should be considered with caution. For instance, their use of features allows for a four-way contrast — the three mentioned above, plus [+laryngeal precision], [-spread] — and yet, only three are exploited. J&G are aware of this fact when they claim that, 'the feature [+laryngeal precision] occurs when [+spread] occurs' (p. 245). But notice that this claim is driven by what is observed in the articulation of our speech organs, and not by any phonological motivation. If this type of articulation-driven procedure is pursued further, it may lead us to the erroneous prediction of pitch phenomena which are idiosyncratic, and therefore, not relevant for phonological studies. For example, the observation that the use of creaky voice is a regularly observed phenomenon utterance finally, may lead us to postulate the presence of a feature [+creaky voice] next to L%. If creaky voice is not observed, then the negative value is chosen. The crucial point here is that, by following this procedure, there are no limits as to what may be treated as potential phonological phenomena. This would yield an extremely rich phonological model, which is precisely what we wish to avoid. In addition to this, there

[12] The underscore is an informal marker of monotone.

is no theoretical motivation for the rejection of the fourth combination of feature values I presented above, and yet, intuitively, we would wish to do so. This unsolved matter leads to the common problem inherent in models based on distinctive features, which is the potential for overgeneration.

Another point to take into consideration is J&G's definition of *stylised* contours in terms of a semantic correlate with 'routine' and 'lack of emotion'. It is not my intention to move into the field of meaning here, and it is left for future examination whether 'routine' and 'lack of emotion' are associated with semantic or pragmatic meaning. The results of such investigation may shed light on the nature of these contours; whether they should be accounted for by the phonological component, or left as a matter for phonetic interpolation.

4 Conclusion

In view of the fact that there is inconclusive evidence to support the claim that spreading arises as a consequence of a special phonological factor, such as association lines or distinctive features, I shall follow the idea that spreading is a matter of phonetic interpolation, and that stretches of levelled pitch arise as a consequence of TBUs showing the same type of pitch specifications, either presence or absence of T, without having to posit any other factor alien to the present model. Thus, L-spread is no longer an obstacle for the development of such a model.

In order to produce an account of contours which show spreading, the reader is referred to §5, chapter 6, since it is necessary for me to present other fundamental components of the model in the following chapter.

Chapter 4. The grammar of intonation

1 Introduction

In an attempt to develop a **restrictive** phonological model of intonation, I have proposed that phonological oppositions must be captured in terms of monovalent primes. In the case of English, an intonational language in which it can be argued that there is a two-way contrast, phonological oppositions are represented by the presence/absence of Tone. This led us to a discussion of which tone is to be discarded, and I have argued that this should be L, based on evidence drawn from analyses of two-tone languages. After that, I put forward an alternative representation of the contour-tone by means of a single tone and the association of T exclusively to the boundaries of prosodic domains. Thus far, however, my presentation of the prosodic hierarchy to which these domains belong, has been rather informal.

In this chapter, I present the basic grammar of intonation, the fundamental components of which are the *Principles of Licensing*.

2 The prosodic hierarchy

In chapter 2 I informally referred to boundaries of prosodic domains as TBUs. Of immediate relevance is the question of the origin of these boundaries in the grammar of intonation. In this chapter, I argue that they correspond to the edges of prosodic domains which arise as a consequence of *head-dependent* relations between the items within that domain[1]. Let me clarify this point by drawing an example from the field of stress. Given a pair of positions, ('x_1 and x_2) in which the first one is stressed and the second one is unstressed, we would like the model to capture this asymmetry. This can be done by assuming that there is a relation between them, which I represent by means of the arrow thus[2]:

[1]See Halle and Vergnaud (1987) and Hayes (1995) for a full account of stress patterns in a variety of languages.

[2]Lines and arrows in this chapter represent licensing paths.

(1) $'x_1 \quad x_2$

The fact that $'x_1$ is more prominent than x_2 is captured by means of an arrow which originates from $'x_1$. It can also be said that in the relationship between $'x_1$ and x_2, $'x_1$ is the head and x_2 counts as its complement.

In addition to this, let us assume that each prosodic domain must have a head. If $'x_1$ is a head, as I have already stated, then, it is possible to claim that $'x_1$ and x_2 constitute a domain, which I represent by means of boundaries in (2):

(2) $\left[\, 'x_1 \quad x_2 \,\right]$

All this has been formalised in terms of the **principles of licensing** by Kaye, Lowenstamm and Vergnaud (1990), Charette (1991), Harris (1994). It must be made clear that in the present work — unless stated otherwise — I adopt the view of GP developed by Harris (1994). In the following section, I illustrate and discuss those licensing principles which underlie the phonological representation of intonation patterns.

3 Some fundamental principles of licensing (I)

(3) **Constituent licensing**

Within constituents, licensing relations are head-initial.

According to (3), the asymmetrical relationship between xs can be formalised in terms of licensing. Hence, in (2) x_1 *licenses* x_2.

(4) **Phonological licensing**

(a) Within a domain, all phonological units must be licensed save one, the head of that domain.

(b) Licensing relations are local and directional.

From (4)(a), only one head is recognised in each domain[3]. Unlicensed heads must be licensed by a head which is itself licensed. From the notion of **directionality** in (4)(b), it follows that all constituents must contain one head, and that licensing can go from right to left at one level, and from left to right at another level, but not in both directions at the same level. As an illustration of this, consider the following analogy with stress relations at the foot level and at the phrase level in English:

(5) (a) ['darkroom]

(b) [dark] ['room]

[*'darkroom*] is a compound which is stressed at the foot level. Licensing at this level goes from left to right, and therefore, *dark* bears stress. On the other hand, [*dark*][*'room*] is a noun phrase which receives stress at the phrase level, where the directionality of licensing is from right to left. Hence, *room* is stressed.

The constraint of **locality**, also in (4)(b), requires that the licensor and licensee be adjacent[4]. In a process such as vowel harmony, for example, locality operates as the manifestation of this licensing relation. Consider vowel harmony in some Bantu languages (see Harris and Lindsey (1995) and Mtenje (1985) on vowel harmony in Chichewa) in which a high vowel in a suffix is lowered due to the presence of a mid vowel in the root.

(6)		Causative	
(a)	*put-a*	*put-its-a*	'provoke'
(b)	*konz-a*	*konz-ets-a*	'correct'

In an analysis based on the principles of GP, it is proposed that the trigger vowel[5] in the root is adjacent to the target vowel in the suffix, and not, as it seems, separated from it

[3]In chapter 6 this constraint is loosened at the level of the intonation domain.

[4]This constraint is loosened in the analysis of structures which account for contours showing downstep in chapter 6.

[5]It must be made clear that in GP the terms *vowel* and *consonant* are informal labels to refer to melodic expressions in a nuclear and onset constituent respectively.

by intervening consonants. The adjacency of the vowels is captured by means of the projection of the relevant constituents: in (7) below, the trigger nucleus is adjacent to the target nucleus at the word level. The characteristics of the root vowel are manifested in the suffix due to the licensing relation between them, which can only involve adjacent units.

(7)

```
      N"--------------N"            Word
      |               |
      N'              |             Rhyme
      | \             |
   O  N  \         O  N...
   |  |   \        |  |
   x  x   x  [     x  x...    ]
   |  |   |        |  |
   k  o   n        z  e...
```

Let me now show how licensing relations can work for a phrase such as *Mary has a lamb.*

(8)

```
 X→x    X→x    X    L0
 |  |   |  |   |
 Ma ry  has a  lamb
```

I have projected all the syllable nuclei — represented by means of *x* — to L0 which, for the time being, can be understood as the sentence stress level. As can be seen, some of the xs are larger than others. This is an informal device to show that the larger ones are stressed, while the smaller ones are unstressed. This relationship is formally captured by the arrows, which represent the directionality of licensing from left to right (as in the case of [*'darkroom*]) between positions which are adjacent at a given level of projection.

In the model I propose, licensing relations are universal, and they are parametrically defined for each language. For instance, in the specific case of English, the directionality of licensing at the stress level is from left to right. On the other hand, in the case of French — a language in which words are generally stressed on the last vowel (cf. *originalité*) — the directionality of licensing at this particular level is reversed.

Let us now focus on L1. By (4)(a), unlicensed heads are projected onto the next level (L1), which can be understood as the sentence accent level. Here again, licensing is from left to right.

(9)

```
X--------> x        X   L1
|          |        |
X-> x    X -> x     X   L0
|   |    |    |     |
Ma  ry   has  a    lamb
```

There are still two constituents at level (L1) which themselves are not licensed. In order to see the licensing relationship between them, I shall project them onto a further level (L2):

(10)

```
x <------------------X   L2
|                    |
X--------> x         X   L1
|          |         |
X-> x    X -> x      X   L0
|   |    |    |      |
Ma  ry   has  a     lamb
```

At L2 licensing is from right to left, as in the case of [*dark*][*'room*]. There is now only one head which is the ultimate licensor of the entire prosodic structure.

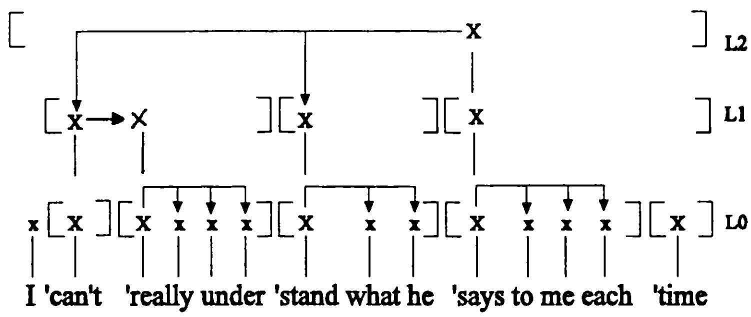
By phonological licensing, each head and the positions it licenses constitute a domain, so that (10) shows the following domains:

```
(11) [x <------------------X]   L2
      |                    |
     [X--------> x    ][   X]   L1
      |          |         |
     [X-> x][X -> x][      X]   L0
      |   |   |    |       |
      Ma  ry  has  a      lamb
```

Having all this in hand, in the following section I proceed to illustrate the phonological representations for some simple contours. For this, I shall follow the traditional distinction between *nucleus* and *head* which is customary in the British tradition of intonation, but without assigning any theoretical status to the items they refer to at this stage[6]. However, before that, let us examine another example which shows a more elaborate prosodic structure in order to illustrate the fact that constituents at L0 (or any other level) are not necessarily binary in intonation.

Consider the following sentence, *I 'can't 'really under'stand what he 'says to me each 'time*, whose prosodic structure I illustrate in (12):

(12)[7]



Notice that, at L0, the head of each constituent licenses all the positions in its domain. In addition, at L1 there are two positions (the projection of *can't* and of *stand*) which are licensed by the projection of *says* at L2. Thus, at all levels, a head can license more than one position.

This situation is well-established in the metrical theory of stress developed by Halle and Vergnaud (1987), in which *unbounded* constituents result as a consequence of

[6]It is not my intention to suggest that nuclei show a set of pitch accents which differ from those in heads, as O'Connor and Arnold seem to imply. Nevertheless, I believe that O'Connor and Arnold's insight that the nucleus is somehow different from the head, is valid. I shall capture this idea in §7 by means of constituent licensing, whereby I assign constituent labels to the nucleus and the head.

[7]In this example, it appears that *I* and *time* do not intervene in any licensing relationship. The case of *time* is similar to that of *really*: they are the heads of their domains at L0 and they are not projected to L1 because they do not carry any pitch information. The case of *I* is not so clear and requires further investigation. However, at this stage of the thesis, I claim that it is directly licensed by the ultimate licensor.

allowing a head position to license all dependent positions in its domain. As an illustration, they propose the case of French. For example, the word *originalité* is analysed as follows (line 0 and line 1 are levels for computations on stress):

(13)
```
 .  .  .  .  .  *     line 1
(*  *  *  *  *  *)    line 0
 o  ri gi na li té
```

Note that the last position is the head, which would act as licensor of the other positions in the domain. Such a domain is then said to be unbounded. We shall encounter a similar situation later in chapter 6.

4 Pitch contours in nuclear position

In order to illustrate pitch patterns in nuclear position, let us focus on *[lamb]* at L1 and L2[8] of the structure included in (11); this structure is repeated in (14) for convenience, where both levels are conflated into a single line. For the purposes of clarity, I have labelled the boundaries according to the level to which they correspond.

(14) [[lamb $]_{L1}$ $]_{L2}$

There is general consensus in the intonation literature that, at least, the following pitch contours should be accounted for in nuclear position[9]:

(15) (a) falling (c) fall-rise
(b) rising

[8]L0 is the level at which metrical computations of sentence stress level are performed. Therefore, no tonal specifications need to be defined at this level. This idea is borne out by the constraints imposed on the model in order to avoid the generation of unattested structures. See the discussion related to this issue at the end of this chapter.

[9]O'Connor and Arnold (1969) describe other pitch patterns in this position. For example, they refer to a rise-fall movement. I shall comment on this issue shortly.

The phonological representations I propose for the above pitch patterns are the following:

(16) (a) Fall

$$\begin{array}{ccccc} & T & & & \\ & | & & & \\ [& [\ \text{lamb} & &]_{L1} &]_{L2} \end{array}$$

(b) Rise

$$\begin{array}{ccccc} & & & & T \\ & & & & | \\ [& [\ \text{lamb} & &]_{L1} &]_{L2} \end{array}$$

(c) F-rise

$$\begin{array}{ccccc} & T & & & T \\ & | & & & | \\ [& [\ \text{lamb} & &]_{L1} &]_{L2} \end{array}$$

The representation for a falling pitch movement is characterised by associating T to the left boundary of L1; for a rising pitch movement, by associating T to the right boundary of L2; and for a falling-rising pitch pattern, by associating T to the left boundary of L1 and to the right boundary of L2.

Having illustrated the structures in (16), it is now reasonable to suggest that the structure in (17) below could account for the pattern which is described as a rise-fall,

(17) Rise-fall

$$\begin{array}{cccc} & & T & \\ & & | & \\ [\ \text{lamb} & &]_{L1} &]_{L2} \end{array}$$

in which T is associated to the right boundary of L1.

However, I agree with Ladd (1983) and Gussenhoven (1983) in their proposal that there is no categorial distinction between this nuclear tone and any of the others. Therefore, below I shall exclude this structure from my model. Ladd analyses the rise-fall and fall patterns with the same phonological category, both (HL). In addition to this, he has recourse to the feature [±delayed peak], which he uses as a descriptive device to account for the rising portion of the rise-fall pattern.

On a par with Ladd, Gussenhoven also treats the rise-fall as a modification of the fall (HL), which, for him, counts as one of the basic tones. He adds a phonetic variant to the basic HL, which he calls *delayed.* However, unlike Ladd's feature, which affects peaks, Gussenhoven's variant affects the mapping of tones onto segmental material.

Given that, in my model, there is no place for features, and that the rising portion is just a phonetic characteristic, I claim that the phonetic shape of a rising-falling pitch movement and of a falling pitch movement are accounted for in phonology by means of the same phonological representation: the one I have illustrated in (16)(a).

In relation to the rising portion of the rise-fall pattern, I adopt Gussenhoven's proposal that this is just a matter of **interpolation**. In fact, it is the same reasoning which has been applied to Crystal's (1969) proposal to include a rise-fall-rise pitch pattern as a phonological category. Recall that this option is not pursued by intonologists, on the grounds that the rising portion preceding the fall-rise movement is simply a phonetic effect. Therefore, neither the rise-fall nor the rise-fall-rise are treated as independent phonological units.

According to this line of thinking, the model should be constrained in order not to generate structures such as the one in (17). I shall return to this issue in the following section where I propose that tone association is subject to the constraining principles of licensing between boundaries.

Other nuclear tones reported in some of the literature (O'Connor and Arnold 1969), are the high-fall, the low-fall, the high-rise and the low-rise. However, the falling set is not treated as a distinctive phonological category by Ladd (1983), Gussenhoven (1983), Lindsey (1985) and others. With regards to the rising set, Gussenhoven is the only author who treats it as a single phonological category. In chapter 6, I shall defend the view that the falling set constitutes a single phonological category and that the phonetic difference between the so-called low-fall and high-fall nuclear tones arises as a consequence of the presence/absence of toneless boundaries to the left of the structure which represents a falling movement. I argue in the same terms for a single rising category.

The representations in (16) are descriptively adequate, but the question as to why some particular boundaries remain unassociated, may be raised. In other words, the issue of how to constrain the model in order to prevent it from generating structures such as the one in (17) remains unsolved.

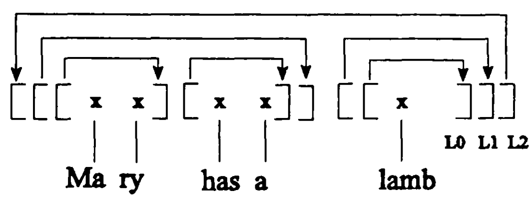
5 Some fundamental principles of licensing (II)

Let us focus on the above question of how to constrain tone association, and for this, once again I shall look to the theoretical principles of GP.

One of the claims of GP is that licensing principles run throughout the entire prosodic structure, allowing for a unified representation. Also, an item must participate in a licensing relation in order to be interpreted phonetically. According to this, it is fundamental that each unit (whatever its nature, a prosodic constituent, a skeletal position, a melodic prime) enters into a licensing relation with another item. Given this context, the prediction is borne out that boundaries of prosodic domains must also participate in a licensing relation.

I propose that licensing between boundaries behaves in the same way as licensing relations observed at other levels of structure. Thus, licensing at this level resembles that shown in (18) below, where I reproduce the structure in (11) but this time, I have conflated the levels, omitting the projection of heads.

(18)



At L0 and L1 the directionality of licensing goes from left to right, and at L2 from right to left. Note, however, that this is not the case for inter-constituent licensing (see p.101).

Another licensing principle to add to the present model of intonation is the *Licensing Inheritance Principle* (Harris 1992) as stated in (19):

(19) **Licensing Inheritance Principle** (henceforth LI):

A licensor has enough licensing potential to bear a wider set of contrasts (by virtue of being the licensor). A licensee receives a reduced amount of licensing potential, therefore its capacity to bear phonological contrasts is diminished.

Let me illustrate LI with an example taken from English, specifically the composition of its diphthongs illustrated in (20)(a) below.

(20) (a) /aɪ, eɪ, ɔɪ, ɪə, eə, ʊə, əʊ, aʊ/

(b) /ə, ɪ, ʊ/

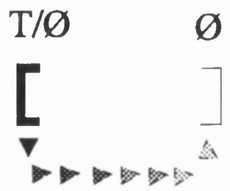
(c) /ɪ, ɔ, ə, ʊ, a, e/

The second position in each diphthong is restricted to one of the vowels in (b), whilst the first position may show up as having any one of the vowels in (c). In other words, assuming that the left position is the licensor, and that the right position is the licensee, it is possible to claim that the licensor shows a wider set of contrasts than the licensee.

5.1 Licensing Inheritance at L1

Let me now illustrate how LI is understood in intonation: given a pair of boundaries in a licensing relation,

(21)



the head position (shown by [) has enough licensing potential to bear a range of contrasts (like the first position in a diphthong); in this case, T versus absence of T (or as informally shown in the diagram, Ø). The dependent position, due to LI (shown by the fading arrows), receives less licensing potential and, therefore, is unable to support a similar degree of tonal contrast (like the second position in a diphthong). For, if T is allowed to appear in this position, then this would entail that we can still expect to find absence of T in some instances. In order to exclude this possibility, it is preferable to allow only for Ø, and in this way, we do not expect T to be a possible option in the first place.

Given what I have stated in the previous paragraph, the only boundaries to which T can associate are the ones in (22)(a) and (b). The structures represented in (c) and (d) cannot be generated by the model, owing to the restrictions imposed by LI.

(22)	(a)	(b)	(c)*	(d)*
	T | [ ]	 [ ]	T T | | [ ]	T | [ ]

(22)(a) represents a falling movement, and (b) represents non-high level pitch[10]. Due to the fact that the left boundary is the licensor, it can support either the presence or absence of T. Furthermore, the right boundary is the licensee, and therefore it cannot bear tone. Thus, these two structures illustrate the maximum number of contrasts that either boundary can bear. In relation to (22)(c) and (d), these structures are regarded as ill-formed in as much as they violate the principle of LI. In this way, the right boundary is predicted to be unable to bear tone.

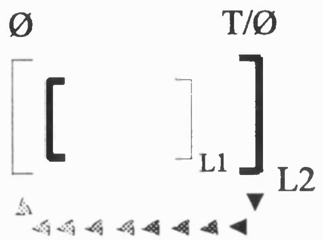
From (22) it follows then that rising and falling-rising nuclear contours cannot be generated by the model, and yet these are well-established patterns in intonational data. In order to propose a phonological account of these two contours, I shall illustrate the manifestation of boundary-licensing relations at the highest level of the hierarchy (L2), and show how they interact with tone association.

5.2 Licensing Inheritance at L2

Recall that in (11) the directionality of licensing between positions at the phrase level goes from right-to-left. As I have already stated in relation to (18), given that boundaries and positions show the same directionality in licensing, it is not surprising to find that the directionality between L2 boundaries goes from right to left, too. The distribution of T is then as follows:

[10] The crucial piece of information conveyed in this structure is that there is level pitch, regardless of its height. I comment on this structure below when I discuss (24)(b).

(23)



The right L2-boundary can bear a tonal contrast, whereas the left boundary cannot bear tone at all. This predicts the phonological representations illustrated in (24) below, which are interpreted as rising pitch and low level pitch respectively.

(24) (a)

$$\begin{array}{cccc} & & & T \\ & & & | \\ [& [&]_{L1} &]_{L2} \end{array}$$

(b)

$$\begin{array}{cccc} [& [&]_{L1} &]_{L2} \end{array}$$

In fact, (24)(a) is the only possible way in which the model can generate the representation for a rising movement[11].

Thus far, I have accounted for falling, rising and level contours. However, there still remains a representation for the fall-rise contour. In addition to this, formally, I must exclude the structure in (17) from the model.

Let us focus first on the fall-rise contour, which is accounted for by the following structure:

(25)

$$\begin{array}{cccc} & T & & T \\ & | & & | \\ [& [&]_{L1} &]_{L2} \end{array}$$

The left L1 and right L2 boundaries bear T. They are allowed to do so by virtue of being the licensors at their respective levels and also by the LI principle. It is possible to claim that the fall-rise structure is different from both the fall and the rise structures in the sense

[11] But see the discussion about rising contours in pre-nuclear patterns in §6.2.

that, in the first one, tones from two different prosodic levels are involved, whereas in the other two, only one tone from a single level is involved.

Finally, the rise-fall nuclear contour, whose structure I informally illustrated in (17), can now be formally excluded from phonological representation by LI. In the context of (21) and (23) there is no way in which the model can generate the structure in (17) because the right boundary of L1 corresponds to the licensee, and therefore, does not inherit enough licensing potential to bear T.

6 Pitch contours in PRE-nuclear position

6.1 Heads

O'Connor and Arnold include the following pitch contours in their inventory of simple heads:

(26)	(a)	falling	(↘)	(c)	high	(')
	(b)	low	(,)	(d)	rising	(↗)

Note that in their analysis of these contours in the head position, O'Connor and Arnold use a set of diacritic symbols which are different from those in the nucleus[12]. This could be understood as a means of categorising the set of contours in the head as being different from those occurring in the nucleus.

Another view of these smaller domains inside the intonation phrase is offered by Pierrehumbert and her followers. They do not distinguish between head and nucleus. Instead, they maintain that a set of pitch accents can occur anywhere in the intonation domain.

In the model I defend, I bring together both views and combine them in the following statement:

(27) (a) T in phonological representation is associated to boundaries of domains and this is subject to the constraining principles of licensing, more specifically, LI.

[12]High fall (`), Low fall (ˎ), High rise (´), Low rise (ˏ), Mid level (>).

(b) Prosodic domains at L1 are also subject to licensing principles, more specifically to *inter-constituent* licensing.

According to (27)(a), in general, the structures which are generated to account for contours in the nucleus, must also account for the same type of contours in the head. In this way, Pierrehumbert's idea is captured. From (27)(b), it is understood that the domains at L1 also participate in a licensing relation. In other words, I shall capture O'Connor and Arnold's insight in terms of establishing a licensing relationship between the nucleus and the head, in which the former is the licensor and the latter the licensee. I shall return to this issue towards the end of this chapter, where I argue for constituent structure and inter-constituent licensing in intonation.

Let us now concentrate on how to generate phonological structures which can account for the shapes in (26) over the domain *[Mary has a]* at L1 in (11). In order to do so, let us remind ourselves of the well-formed structures at L1 which can be generated by the model so far. These are the following:

```
(28)   (a)   T            (b)
             |
             [ ]               [   ]
```

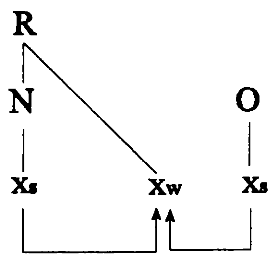
In a straightforward fashion, (28)(a) accounts for a falling movement and (b) for relatively low pitch. However, structures which can account for high and rising contours cannot be generated. This is due to the fact that LI prevents tone from associating to L1 right boundaries. In order to put forward a phonological representation for such contours, I must refer to another licensing relation which is proposed in GP: *Coda* licensing (Harris 1994).

(29) **Coda licensing:**

A dependent position can also be licensed by an immediately following head[13]. In this situation the dependent position receives extra licensing potential.

For example, a coda (x_w) is doubly licensed by a strong position to its right and another to its left. I illustrate this below.

(30)[14]



This is the reason why a coda can show a wider set of oppositions than those shown by any other weak position in the structure, in spite of being a licensee.

Let me now explain how (30) is understood in intonation. For this, I conflate L1 and L2 below and show two domains at L1: P-n and N, which correspond to the pre-nuclear and nuclear domains, respectively. I have labelled their boundaries as *s* and *w* so that I can draw a clear analogy between the events at the level of the Rhyme and at the level of the intonation phrase.

---

[13]In GP, an Onset, as well as a Nucleus, can show a branching structure with two positions. The position to the left is treated as the head, and the one to the right as the dependent. So far, branching constituent structures have not been attested in research on intonation. Therefore, I do not adopt this idea in the model for intonation.

[14]For the purposes of clarity, I shall use the terms *onset* and *nucleus* in lower case to refer to O'Connor and Arnold's *head* and *nucleus* respectively. I shall use *Onset* and *Nucleus* in upper case to refer to these constituents in GP. In this figure, *R* refers to the Rhyme.

(31)

$$[\ [_s\ \text{P-n}\ ]_w\ [_s\ \text{N}\ ]\ ]$$

(31) shows a licensing relationship between (P-n) boundaries (in the same way as there was licensing between x_s and x_w in the Rhyme of (30)). If we assume that $]_w$ corresponds to the coda in this analogy, it is possible to claim that it also participates in a coda-licensing relationship, whereby this position is coda-licensed by $[_s$ from (N). Consequently, $]_w$ is doubly licensed: once by $[_s$ in (P-n), and again by $[_s$ in (N). Hence, $]_w$ has gathered enough licensing potential to bear a wider set of contrasts: presence/absence of T.

By conforming to (29), the model can now generate representations which account for high and rising pre-nuclear patterns as follows:

(32) (a) high pre-nuclear contour (b) rising pre-nuclear contour

$$\begin{array}{cccccc} & T & T & & & \\ & | & | & & & \\ [& [_s &]_w & [_s &]_w &] \end{array} \qquad \begin{array}{cccccc} & & T & & & \\ & & | & & & \\ [& [_s &]_w & [_s &]_w &] \end{array}$$

High pre-nuclear contours are accounted for by showing both boundaries associated to T, and rising pre-nuclear contours are accounted for by showing the right boundary of their domain associated to T. All these associations are subject to the principles of licensing.

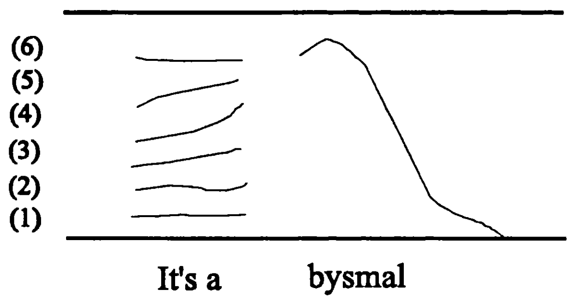
6.2 Pre-heads

O'Connor and Arnold include the pre-head as another distinctive category in their intonation group. They also distinguish between high and low pre-heads. Similarly, Pierrehumbert and her followers refer to H and L as boundary tones which are found at the edges of the intonation domain[15]. Moreover, as I shall illustrate and discuss in chapter 5, Grice (1995) elaborates further on the internal pitch accent structure in order to account for what could be called a rising pre-head.

The question to consider now is whether or not these differences should be encoded in phonological representation. As I shall argue below, the shape of the pre-head is irrelevant in phonological representation, and in any case, the model does not allow for such a domain in prosodic structure.

With the goal of clarifying this point, let us first examine the contours represented in (33)[16] in the framework of previous models of intonation. They illustrate a range of pre-heads going from low to high through rising, all followed by a falling nuclear tone.

(33)



In O'Connor and Arnold's terms, given that all the pre-heads are lower than the initial point of the following high fall nuclear tone, these would have to be treated as **low** pre-heads. However, it could also be argued that (4), (5) and (6) in (33) are rather high in the range, and therefore, should be categorised as **high** pre-heads.

From the point of view of Pierrehumbert and her followers, (1), (2), (3), and possibly (4) in (33), would have to be treated as L% or otherwise, they would have to

[15]But see discussion about L% in chapter 2.

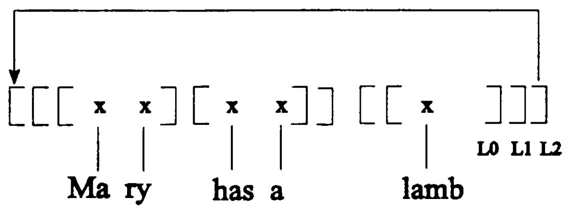
[16]The utterance *It's abysmal* comes from Grice (1992) and (1995), but the nuclear tone is different. Here, I use a high fall, while a low fall is used in the original sources. Also, this figure is a drawing based on the originals.

remain unspecified in phonological representation. (5) and (6) would have to be analysed as H%.

I conclude from the above discussion that it is highly unlikely that an agreement will ever be reached as to how to categorise pre-head movements of pitch (if they should be categorised at all). The fact that this is a controversial matter indicates that more research must be conducted. A possible way of doing so in the future, would be to perform a categorical perception test, in which listeners are presented with a pair of stimuli whose content are similar to the stylised diagram in (33). For example, the first pair might be (1) and (2), the second pair (3) and (4) and so forth. Subjects would then be asked whether they can distinguish the stimuli in each pair.

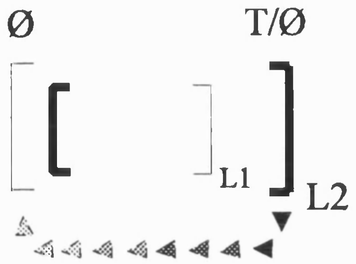
Having discussed the issue of the shape of pre-heads, I now claim that previous proposals, which put forward a set of different categories in this position, have adopted a highly descriptive approach to account for intonational data and that the shape of pre-heads is irrelevant in phonological representation. This idea is straightforwardly borne out in my model by means of licensing conditions operating between L2 boundaries. I have already illustrated these conditions in (18) above, which I now reproduce in (34) in a simplified version.

(34)



Recall that licensing at L2 goes from right to left. In order to demonstrate the association of tone at this level, the reader is reminded of LI, which I introduced in (19). According to this principle, L2's right boundary can bear presence/absence of T, while L2's left boundary can only show absence of T. I have already illustrated this in (23), but I repeat it here for the sake of clarity:

(35)



Given (35), there is no possible way in which T could be associated to L2-left boundary because the latter does not receive enough licensing potential to bear a tonal contrast. Hence, phonologically, there is only one pitch specification for this boundary. It follows then, that the phonological representation for (33) must be the following:

(36)
```
          T
          |
[It's a [ bysmal ]L1      ]L2
```

T associated to the left boundary of L1 accounts for the falling pitch movement over *bysmal*. This entails that the shapes illustrated in (36) arise as a consequence of phonetic interpolation.

In the context of the present discussion it appears that previous proposals have failed to achieve a coherent phonological account of intonation, due to the fact that they have treated the pre-head and head as two absolute categories. As an alternative to this proposal, I suggest that, instead, the shape of the contours at the left edge of the intonation domain are the result of phonetic interpolation between L2's left boundary and L1.

From a descriptive point of view, the height of the starting point of the nuclear tone **in relation to** the preceding pitch specification is the point to take into consideration, in order to put forward a descriptively adequate phonological representation. For example, let me briefly illustrate how a step-down versus 'non-step' from the initial pitch pattern to the falling nuclear tone is captured in phonological representation (the contours I refer to here can be seen in the appendix to this chapter):

(37) (a) step-down + fall

```
                 T
                 |
[ It's a  [@ ]L1   [ bysmal ]L1 ]
```

(b) no-step + fall

```
                 T
                 |
[ It's a           [ bysmal ]L1 ]
```

I shall discuss the representations in (37) in chapter 6 in more detail. Here, it suffices to say that in (a), the fall over [-bysmal] is preceded by a domain with toneless boundaries, ([ ]), whereas in (b), such domain is absent. The falling pattern in the former structure is observed to be interpreted as starting lower than the one in the latter structure. This difference arises as a consequence of toneless boundaries being encoded in the phonological representation of (37)(a), but not (37)(b).

I shall return to the issue of the exclusion of pre-heads from phonological representation later in this chapter where I discuss *inter*-constituent licensing. There, I shall present further motivation for excluding them from phonological representation altogether.

7 The intonation domain and its constituents

So far, I have proposed that licensing relations manifest themselves between **positions** in prosodic domains, and between the **boundaries** of prosodic domains. Also, I have made an informal distinction between nuclear and pre-nuclear contours in order to put forward a set of phonological structures which account for pitch patterns in these locations. In order to formalise the nuclear/pre-nuclear bifurcation (and thus to capture O'Connor and Arnold's insight which I mentioned in §6.1), in this section I argue that there is a licensing relation between these two domains whereby the nuclear domain is the licensor, and the pre-nuclear domain is the licensee. In this way, I abide by the principles of licensing in GP (which I briefly mentioned at the beginning of §5) which state that all units in phonological representation must participate in a licensing relation. Hence, in intonation as well, there is licensing between prosodic **constituents**. But before I can develop an argument along these lines, I need to present independent evidence that these constituents exist in phonology.

In GP, the manifestation of stress provides evidence for the existence of two independent constituents — Onset and Nucleus — at a lower level in the prosodic hierarchy. For instance, the internal structure of the Nucleus — whether it branches or not — is the factor which determines the placement of stress in a word, regardless of what appears in the Onset. Therefore, the different behaviour of these two units provides evidence for their status as different constituents.

Similarly, in the field of intonation, I propose that *focus* placement is the factor which provides evidence for the existence of two independent constituents, the onset and nucleus. Before I illustrate this, let me explain briefly the notion of focus[17]. In general, speakers use focus to refer to that part of the intonation domain which conveys new information, or to highlight some information which they want to draw attention to. For example, given the phrase *Mary has a lamb*, the speaker may decide to draw attention to certain parts, which I have underlined, as is customary in the field of meaning:

(38) (a) <u>Mary has a lamb</u>
(b) Mary has a <u>lamb</u>
(c) Mary <u>has</u> a lamb
(d) <u>Mary</u> has a lamb

In (38)(a) the speaker has placed the whole phrase in focus, whereas in (b) to (d), only *lamb*, *has*, and *Mary* are in focus, respectively.

Of immediate relevance here is the question of how the speaker shows all these different distributions of focus. I propose that this is accomplished by treating specific parts of a phrase as having a status different from others, and that this status is achieved by requiring that a specific item bears the most important pitch accent (that is, the nuclear tone). For instance, in the case of (38)(b), (c) and (d), the speaker has identified *lamb*, *has* and *Mary*, respectively, as the bearers of this special status by virtue of carrying the

[17]Here I assume a general idea about focus, without going into any discussion related to *broad* and *narrow* focus. I acknowledge that this topic is certainly extensive, but is also outside the scope of this thesis. For an up to date summary and discussion of focus, see García-Lecumberri (1995) and references therein.

nuclear tone. Thus, in order to signal focus, speakers select an item from the intonation phrase on which they place the main (or nuclear) pitch accent. In the cases of (b) to (d) in (38) above, the location of pitch accent is illustrated in (39) by means of uppercase.

(39) (b) Mary has a LAMB

(c) Mary HAS a lamb

(d) MARY has a lamb

Notice, however, that the scope of focus and the edges of the nuclear domain do not necessarily have to coincide. For example, in (38)(a), the nucleus is placed on the last item, *lamb*, and yet the complete phrase is in focus. I illustrate this below:

(40) (a) <u>Mary has a LAMB</u>

Nevertheless, this example does not undermine the claim I argue for here. Even if the entire phrase is in focus, there is only one item which has been selected to bear the main pitch accent, which is *lamb*.

There is a close analogy between the phenomena of stress and of focus distribution: at the prosodic level below the foot, stress defines constituent structure in terms of Onsets and Nuclei, and only the latter constituent counts as relevant for metrical purposes; similarly, at the prosodic level above the foot, the distribution of focus also defines constituent structure in terms of onsets and nuclei, and again only nuclei play a fundamental role for such purposes.

According to this we are now in a position to assign constituent structure to the phrases in (39) and (40):

(41) (a) and (b)

```
        O                        N
        |                        |
[[      x   x     x   x  ] [     x      ]]
        |   |     |   |          |
       Ma  ry    has  a         lamb
```

(c)

```
        O       N
        |       |
[[      x   x] [x    x    x  ] ]
        |   |   |    |    |
       Ma  ry  has   a   lamb
```

(d)

```
       N
       |
[[     x   x     x   x   x  ] ]
       |   |     |   |   |
      Ma  ry    has  a  lamb
```

Let us now return to the starting point of this argument: focus placement as evidence for the existence of **two** constituents in intonation. Recall that, in (40)(a) and (b), *lamb* has been selected as the unit to bear the main pitch accent (which is obligatory) and that no other unit has been selected to perform this task. Given this evidence, I propose that *lamb* corresponds to the nucleus or licensor and that *Mary has a* corresponds to the onset or licensee, which signals the information which is out of focus.

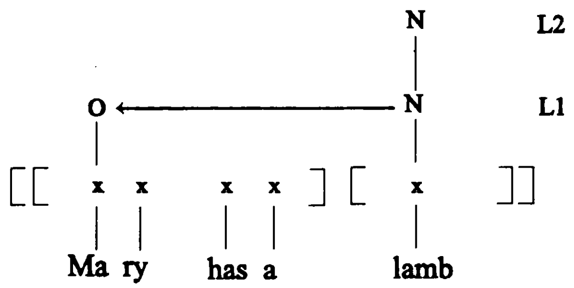
This proposal is perfectly valid in the framework of GP, where it is taken for granted that the obligatory presence of a constituent in prosodic structure is an indication of its headship. Thus, the fact that the nucleus is the only obligatory constituent in intonation, is reason enough for defining it as the licensor and, consequently, the onset as the licensee. This idea is developed in GP under the label of *onset licensing*, which is stated as follows:

(42) **Onset licensing**

An Onset head position must be licensed by a Nuclear position.

I exemplify (42) with *Mary has a lamb* below[18]:

(43)



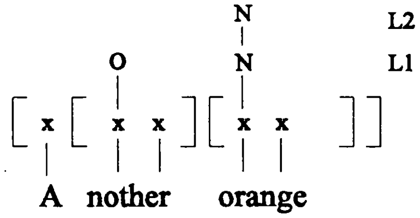
The nucleus (*lamb*) licenses the onset *(Mary has a)*. Note that the directionality of licensing between constituents at L1 (right to left) differs from that between boundaries (see ch.6).

Up to this point, I have presented evidence for admitting the existence of two constituents in intonation: the nucleus as a domain in which focus is defined, and the onset as a domain in which out-of-focus material is included. Given this restricted approach to constituency at L1 in intonation, there is no place for a 'pre-head' type of constituent in the phonology of intonation.

Bearing this in mind, I would like to turn now to the main example from chapter 2, *Another orange*, in order to introduce a structure which formally accounts for it. I illustrate such a structure below:

[18] The reader who is already acquainted with GP should not be misled by the association of an onset to the first vocalic element in *Mary*. *Mary* in intonation belongs in the onset constituent which is licensed by the following nucleus.

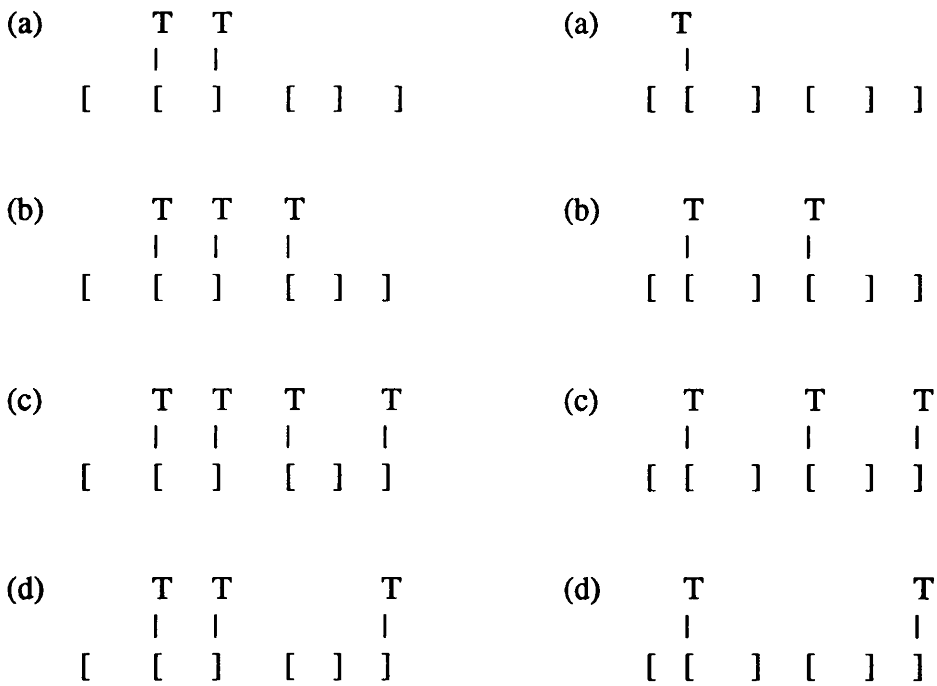
(44)



-nother and *orange* are two domains — the onset and nucleus, respectively — at L1, and *A-* is not assigned an independent constituent, but instead, is left as part of L2.

The number of possible tonal associations to nuclear and onset constituents is severely restricted. Thus far, only 16 (that is, 4x4) structures can be generated by the model. I illustrate them below, and include their phonetic interpretation in the appendix to this chapter.

(45) High onset + nucleus: Falling onset + nucleus:



Low onset + nucleus:

```
(a)

        [   [   ]   [   ]   ]

(b)                 T
                    |
        [   [   ]   [   ]   ]

(c)                 T       T
                    |       |
        [   [   ]   [   ]   ]

(d)                         T
                            |
        [   [   ]   [   ]   ]
```

Rising onset + nucleus:

```
(a)             T
                |
        [  [    ]   [    ]   ]

(b)             T   T
                |   |
        [  [    ]   [    ]   ]

(c)             T   T        T
                |   |        |
        [  [    ]   [    ]   ]

(d)             T            T
                |            |
        [  [    ]   [    ]   ]
```

Recall that, as I have already illustrated and discussed, one of the features of GP is that its principles apply throughout phonological representation, enabling the model to make generalisations about specific aspects of phonological structure. One such principle is, for example, that of phonological licensing. Given this view, and after having presented LI as another fundamental principle which constrains the association of T to boundaries, we now predict the manifestation of LI also at the level of constituent structure. In fact, this prediction is confirmed by the distinction already made between nucleus and onset. The former possesses a greater licensing potential by virtue of being the licensor, and the latter receives a reduced amount, by virtue of being the licensee. As we shall soon see, this is correlated with the slightly different interpretation of falling movements in nuclear or onset position.

In the following section, I illustrate how LI between boundaries is manifested in phonetic interpretation. In preparation for this, I first describe the phenomenon of *Gravitation Effect* (henceforth GE) (Cabrera-Abreu and Takahashi 1993) in phonetic interpretation, and then I examine how GE can be affected by the reflection of an array of factors present in phonological representation, namely LI and its correlation with constituent structure.

8 Interpretation of phonological structures: Gravitation Effect

According to Cabrera-Abreu and Takahashi (henceforth C&T), GE is a phonetic phenomenon which accounts for well-attested F0 patterns, and it counts as a fundamental tool for the understanding of the description of pitch contours.

More specifically, C&T contend that the dip which is normally found between two distant Ts (or Hs as was customary in previous models) (cf. P'80) is due to the down-pushing force of GE. In addition to this, they claim that the straight interpolation between two former L tones is also due to GE's continuous down-pushing force. Thus, they understand GE as the general tendency of pitch to drop towards the baseline.

They also claim that there are two factors which intervene in the manifestation of GE: the distance between target tones, and constraints which derive from physiological limitations. In relation to the former, they state that the greater the distance between two Hs, the greater the time for GE to exercise its force, and consequently, the interpolation between these two tones results in a clear dip. With regards to the latter, they suggest that the different gradient between a fall and a rise reflects the fact that the deactivating of vocal fold vibration needed to bring down F0 does not necessitate as much air pressure, muscular effort and time as the activating of vocal fold vibration to accomplish a high F0.

In short, in order to interpret the structures in (45), the following simple statements which result from GE must be taken into consideration:

(46) (a) a boundary associated to T is interpreted by GE as pitch going up or as relatively high pitch.

(b) an unassociated boundary is interpreted by GE as pitch going down or as relatively low pitch.

The interpretation of tonal configurations in the nucleus constituent donot arise as a direct reflection of the above statements only; as I shall show below, additional factors like, for example, LI in phonological structure, may participate in their interpretation.

In order to illustrate how LI controls the interpretation of melodic units, let us concentrate, first, on the falling onset plus falling nucleus sequence, and after that, compare their interpretation with that of the rising onset plus rising nucleus sequence.

In relation to the former, both onset and nucleus show identical phonological structures: their left boundaries are associated to T, while their right boundaries remain toneless. However, the falling movement in the onset and nucleus are manifested differently in phonetic interpretation: as a gradual fall in the onset, and as a sharp one in the nucleus. I argue that this difference is due to LI in phonological representation; specifically, to the fact that the right boundary of L2 counts as the licensor, which enhances its pitch specification. Since this boundary remains toneless its interpretation results in falling pitch and moreover, because it is enhanced, the fall is sharp. However, the right toneless boundary of an onset does not suffer this type of enhancement because this boundary inherits less licensing potential, by virtue of being the licensee.

As far as the rising sequence is concerned, note that the onset and nucleus show different phonological structures: in the onset, T is associated to the right boundary, whereas in the nucleus, the equivalent boundary remains toneless. Given this state of affairs, we may predict that different phonetic interpretations will result for each structure. This prediction appears to be confirmed in the light of the F0 contour (labelled as (ROND)) included on p.210 in the appendix to this chapter. However, we must take this observation cautiously, since there are durational differences between the material which is analysed as belonging to the nucleus and that belonging to the onset.

By analogy to the case of the falling sequence, for which I argued that LI enhances the effect of the right L2 boundary — by virtue of this being the licensor — here, I claim that, for the rising sequence, LI enhances the effect of T associated to the right L2 boundary. Thus, because the effect of this boundary is enhanced, we do not find the rising that we might have expected initially: low pitch (as the interpretation of two toneless boundaries would lead us to assume) and then a rising movement. On the contrary, the low portion is practically missing, and pitch rises earlier.

Given this evidence, I conclude that these two structures have the following features in common: the right L2 boundary possesses a greater licensing potential by virtue of being the **licensor**, and this is understood as tonal specifications being **enhanced** in phonetic interpretation. In addition, by LI, the tonal specification of boundaries which count as **licensees** is understood as being **attenuated** in phonetic interpretation.

The manifestation of LI in phonetic interpretation can be summarised as follows:

(47) the pitch specifications of a unit which is nearer to a licensor are enhanced, whereas, those of a unit which is farther away from a licensor, are attenuated.

As far as the interaction between LI and GE is concerned, I conclude that LI interacts with GE in the following way: the nearer to the licensor, the greater the gravitation effect on a boundary which remains toneless.

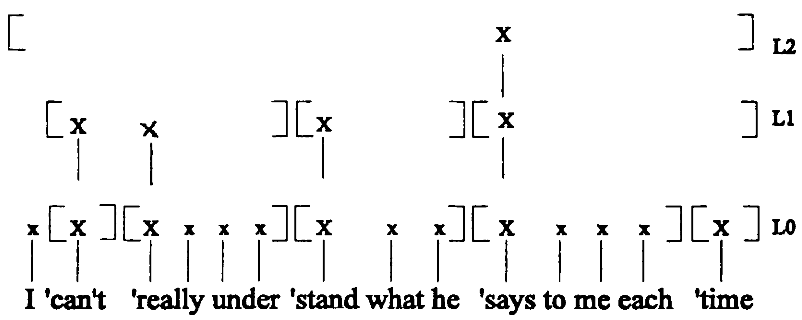
Notice that LI between boundaries also correlates with constituent structure. For example, as I have just illustrated, depending on whether the relevant melodic material is licensed by a nucleus (licensor) or an onset (licensee), a falling movement shows a sharp or a gradient slope, respectively. Let me draw an analogy between the interpretation of this pitch movement and the interpretation of the element [I] in GP: if [I] is licensed by a nucleus, then its phonetic interpretation is that of a so-called vowel, as in the word <*pit*>. If it is licensed by an onset, then it is interpreted as a glide, as in the word <*yes*>. So, in GP, the interpretation of a melody depends on which constituent licenses it; similarly, in intonation, the interpretation of a falling pitch movement depends on the constituent by which it is licensed.

9 Association of tone to prosodic domains

In §3 I took for granted the idea that L0 corresponds to the level at which stress computations are performed. For this reason, in §4, I associated tone to L1 and L2 boundaries only, to the exclusion of L0. This type of association has allowed me to present a maximally restrictive model, and yet to account for a large set of pitch contours. For, if I chose the option of associating tone to L0 as well, this would increase substantially the number of possible landing sites for tone. The consequences of this move would be as follows: this would immediately lead to the overgeneration of phonological structures; in other words, it would lead us towards the undesirable option of accounting for a given contour by means of various phonological structures. As a result, the model would show a high degree of indeterminacy. Surely, these potential shortcomings must be avoided.

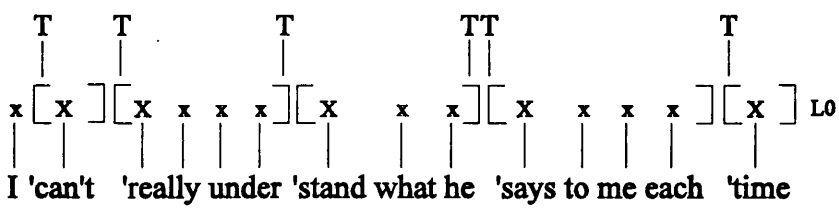
In order to illustrate the undesirability of overgeneration, consider the following example below (which I already mentioned in §3, in which I use the diacritic (') to indicate stress):

(48)



Let us assume that the following items are stressed, and that, therefore they constitute the domains which are shown at L0: *can't*, *real-* (from *really*), *-stand* (from *understand*), *says*, and *time*. Then, let us associate tone to a number of L0 boundaries. I select these TBUs by applying the principles of licensing which I have discussed above.

(49)



In (49), I have represented the following pitch contours: a falling movement in the first domain; a high level in the second; a rise in the third; a fall in the fourth, and another fall in the last domain. In view of this example, it is already highly unlikely that such an array of pitch movements will be found within an intonation domain, especially in pre-nuclear position. In fact, O'Connor and Arnold's description of pre-nuclear pitch patterns is highly restricted, and they do not describe this type of movement in this position. Moreover, the nuclear domain is not free of undesirable pitch movements either. Specifically, the association of T to the domain surrounding *time* allows for the possibility of generating extremely complex pitch movements within the nucleus[19], such as, for example, a sequence of two falling movements.

In addition to all this, bear in mind that the intonational structure as a whole remains incomplete, since L1 and L2 are also part of prosodic structure, and therefore count as legitimate candidates for conveying pitch information (I leave the reader to determine this himself/herself).

As an example of the danger of indeterminacy, let us focus on the nuclear constituent: as can be seen below, a falling movement can be represented by means of two phonological structures, as shown in (50):

(50) (a)
$$\begin{array}{llll} T & & & \\ | & & & \\ [& [_{L0} &] &]_{L1} \end{array}$$

(b)
$$\begin{array}{llll} T & T & & \\ | & | & & \\ [& [_{L0} &] &]_{L1} \end{array}$$

T can be associated either to the left boundary of L1 exclusively, or to both this boundary and the left boundary of L2.

Given the evidence which I have presented here, that the model runs the risk of overgenerating phonological structures and of becoming highly ambiguous in the hypothetical event of associating T to L0, I conclude that L0 must be reserved for the computation of metrical stress only, and not for matters relating to pitch. However, at the present stage of the model's development, this idea amounts only to a stipulation. Ideally, it should be borne out by theoretical principles which already exist in the model. A possible way to accomplish this would be by referring to LI in the following way: I suggest that L0 boundaries are too distant from the ultimate licensor, and therefore, they do not receive enough licensing potential to bear T. This would, at least, prevent T from appearing at this level.

[19]In fact, other models of intonation would treat these two domains as independent pitch accents, thus rendering the one on *time* the last one.

10 Summary

In this chapter, I have presented some of the fundamental principles and components of a model of intonation in the context of a restrictive phonological framework. With T as the only melodic component (which can be associated to the boundaries of prosodic domains), and the principles of licensing (which control the well-formedness of structures), I have designed a model which has the capacity to account for a subset of intonational data from English. Obviously, there is a wide range of intonational contours which the model in its present state cannot account for. In the following chapter, I shall illustrate and discuss those which are characterised by a step down in their F0 contour.

Another important issue developed in this chapter has been the notion of licensing inheritance, and its reflection in phonetic interpretation, where its manifestation interacts with the general principle of Gravitation Effect. LI in phonological representation and GE in phonetic interpretation, enable us to produce a direct mapping of phonological structure into phonetic interpretation, without resorting to a complicated and detailed body of interpolation rules.

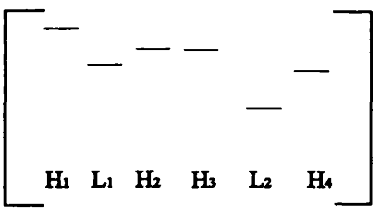
In the following chapter I examine and discuss earlier analyses of stepping, most of which have recourse to L (directly or indirectly, that is by means of complex pitch accents with L as a member). Naturally, this constitutes a challenge to the validity and appropriateness of the present model.

Chapter 5. The phenomenon of downstep

1 Introduction

In many of the world's languages (particularly some of those spoken in Africa), there is a phenomenon whereby the scaling of two H tones can be different; under some circumstances, the phonetic value of one of them can be shifted upwards or downwards within a particular pitch range[1]. For instance, consider the following example which is frequently mentioned in the literature (see Pulleyblank, 1986, van der Hulst and Snider (eds.) 1993, and references therein): in a sequence such as HLHHLH, tones can be interpreted as shown in (1).

(1)



In (1), H_2 and H_3 are interpreted on relatively lower pitch than H_1, even if phonologically they are the same H tone. The same effect is manifested in H_4 with respect to H_2 and H_3; that is, H_4 is interpreted on a level lower than H_2 and H_3. The general consensus among tonologists is that this type of lowering is conditioned by the existence of L[2] in phonological representation. For example, H_2 and H_4 are downstepped because there is an L tone to their left. On the other hand, the reason why H_3 is not lowered with respect to H_2 is that there is no intervening L.

[1]In the present work I shall analyse downward shifting only.

[2]L can be manifestly present or floating. Other accounts (see Pulleyblank 1986) suggest that stepping is not related to the presence of L in phonological representation specifically, but to whether or not two Hs belong to the same foot (no stepping) or to different feet (stepping).

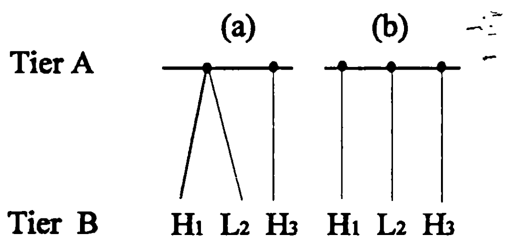
A characteristic frequently attributed to stepping is that the lowering effect can take place an unlimited number of times, so that, given a longer sequence of Hs and Ls under the appropriate conditions, an unlimited set of terraces will result. Thus, Snider (1988) claims that stepping manifests a *cumulative* trait.

Although the trigger for the stepping effect is generally formalised in terms of L in phonological representation, it is the formalisation of its cumulative nature which has consistently worried tone theorists. Early attempts to do so by means of rules and features have met with failure, owing to a number of clearly defined reasons. For example, one analysis was to posit a feature for each level; but this, of course, led to the stipulation of an unlimited number of features and consequently, to a model with excessive representational power. An alternative was to suggest that features have a multi-scalar nature; but this approach was rejected on the grounds that it did not capture any kind of generalisations.

In an effort to produce better analyses, researchers have become increasingly varied in their approach. Some of them treat stepping as a phonetic effect which is accounted for by phonetic interpolation rules (P'80, Pulleyblank 1986 Pierrehumbert and Beckman 1988). Basically, these researchers have recourse to a phonetic interpolation rule which applies iteratively to H tones which occur to the right of L tones. In order to calculate the value of these H tones, the value of a preceding tone is multiplied by a value of less than one. In doing so, the value which is being calculated will always be less than that of the preceding tone. By applying the rule iteratively, they capture the cumulative nature of stepping.

Other researchers treat stepping as a phonological phenomenon which is accounted for by a particular item in phonological structure (Ladd 1983 uses a feature) or by the phonological structure itself (Clements 1983, Hyman 1986, Ladd 1993, Grice 1992, Grice 1995 (henceforth Grice'92, and Grice'95, respectively)). The main argument for a phonological approach is that in tone and intonation languages, stepping alone can signal a contrast between two contours which are otherwise equal. The strategy adopted is generally to propose an enriched tonal structure such as the one I illustrate below in (2). It must be made clear, though, that the authors mentioned above differ with respect to the details of their approaches. Nevertheless, with the illustration in (2) I attempt to capture the spirit shared by all of them.

(2)



There are two tiers (A[3] and B). Tones stand on tier B. Depending on how tones are associated to tier A, a different phonological structure can be generated: in the case of (a), H_1 and L_2 share the same node; in (b) they do not. It is suggested that the branching structure in (a) triggers downstep on the following H, whereas the non-branching structure in (b) does not trigger downstep. In relation to how the cumulative nature of stepping is represented in phonological structure, authors put forward slightly different proposals, to which I shall refer as I examine each one in independent sections.

Another characteristic is that the item which undergoes downstep is located to the right of the trigger. For instance, in (1) H_2 is downstepped because it occurs to the right of L_1, while H_1 does not undergo stepping because it occurs to the left of L_1. In those models which account for stepping by means of interpolation rules, this property is captured by stating that it applies to H tones which occur to the **right** of L tones under certain conditions. Similarly, models which treat stepping as a phonological phenomenon rely on the same kind of stipulation: under certain conditions, a H tone which occurs to the right of the trigger is downstepped.

In the model I propose, it results as the interpretation of a particular phonological structure. In this sense, it proceeds along the same line as Grice'92 and Ladd (1993).

[3]A-tier corresponds to Grice's *root-node* tier and to Ladd's *register*-tier. *B*-tier corresponds to *tone*-tier in both models.

However, unlike these models, in which a richer tonal structure is put forward, I account for stepping by postulating toneless boundaries which correspond to empty nuclei. In this way, I maintain maximal restrictiveness within the model, since I do not introduce new substructures into the original structure. In other words, I avoid the risk of significantly enriching phonological structure.

2 Previous accounts of stepping in intonation

When scanning through the literature mentioned earlier in search of approaches to stepping in intonation, it becomes apparent that the focus of its analysis has shifted considerably: from being accounted for by a particular sequence of tones in the tonal tier (P'80, Pierrehumbert and Beckman1988), to being treated as a matter related to metrical structure (Ladd 1993). In the following section, I begin by discussing the earlier type of proposal. Later, I summarise the main problems associated with such an approach which have been pointed out by Grice'92. Then I outline Grice's (1992, 1995) alternative analysis of stepping and other related contours. Finally, I shall discuss and evaluate Grice's proposal in the framework of a restrictive phonological model for intonation.

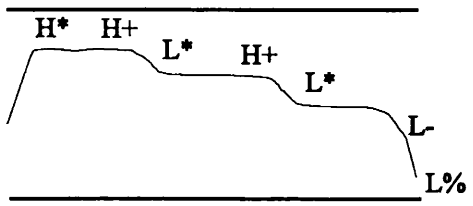
2.1 Stepping triggered by tonal structure

2.1.1 Pierrehumbert (1980), Beckman and Pierrehumbert (1986)

As an example of a particular sequence of tones being the trigger for stepping contours, consider P'80. Briefly, in those instances where L is flanked by Hs (that is, H-L-H), the second H is downstepped with respect to the first one.

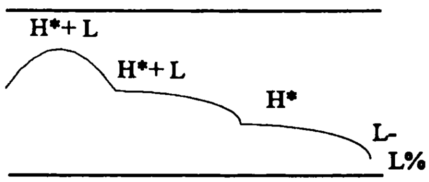
The four types of stepping contours described by Pierrehumbert are schematised in (3). The structure illustrated in (3)(a) is made up of a sequence of falling pitch accents (HL) in which L is associated to the accented syllable. Also, it can be seen that the first L, which belongs to a pitch accent and is flanked by Hs, triggers downstep on the following H.

(3) (a)



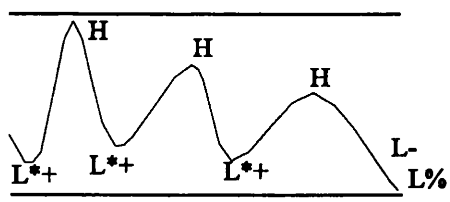
The pitch pattern in (3)(b) is essentially the same as in (3)(a). The only difference between them is that, in (b), H is associated to the accented syllable.

(b)

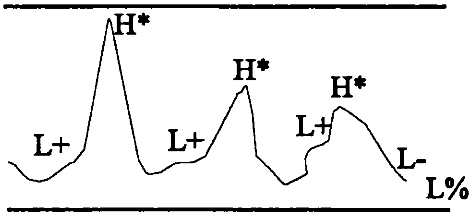


The diagrams in (3)(c) and (3)(d) below represent a sequence of rising pitch movements. This is captured in phonological representation by means of LH. Notice, also, that the relative height of the second and third Hs shifts downwards with respect to the previous H. Unlike (3)(c), in which L is associated to the accented syllable, in (3)(d) it is H which is associated to the accented syllable.

(c)

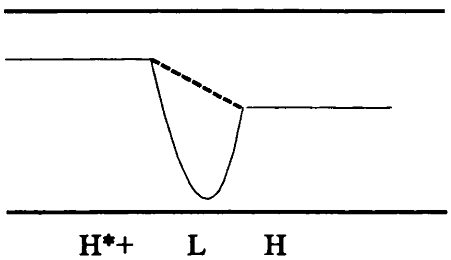


(d)



Pierrehumbert acknowledges that there is a slight irregularity in her analysis. This relates to the interpretation of the L tone which acts as the trigger for stepping. As can be seen in the figures above, the manifestation of L in phonetic interpretation varies a great deal from, for example, the representations in (3)(a) and in (3)(c). In the case of (a) the interpretation of L is extremely high in the frequency range, whereas in (c) there is a very clear dip in the F0 contour, an indication that the value for L is extremely low. Let me show a schematised illustration of what I am referring to here:

(4)



The interpretation of L should look like a deep valley, or at least L should be interpreted in the lower half of the frequency range (as in (3)(c) and (3)(d)). However, there are also cases like (3)(a) and (3)(b) whose schematic representation is illustrated in (4) by the dashed line; that is, a linear interpolation.

This finding in the data prompted Pierrehumbert to slightly modify her analysis in the following way: in those instances like (3)(a) and (3)(b), in which L is not interpreted phonetically, L is a floating tone of the type found in tone languages; it is present in phonological representation only as a trigger for stepping. It is for this reason that there is a rather shallow interpolation in (a) and (b), but a valley in (c) and (d).

After this work, a comparison of English with Japanese (together with strong criticisms from other intonologists, especially Ladd — see discussion below) caused B&P'86 (p. 277) to re-examine their analysis and to propose different conditions for stepping: rather than L being the trigger, it is the occurrence of a preceding bitonal pitch accent which brings about the stepping effect. Basically, they found out that in sequences of L+H H in which there are two consecutive Hs — and no L flanked by Hs — there is a clear step-down from the first H to the second, especially in calling contours. An example of this is shown in the appendix to this chapter.

B&P'86 analyse the pitch accent on *Anna* as either L+H* or L*+H. This bitonal accent accounts for the dip and rising movement before the peak. After the peak, F0 falls and stays on a plateau. They analyse this latter shape as the interpretation of a downstepped H. It is patterns like this which forced them to re-formulate their previous proposal (alternating tones) in terms of bitonal accents as being the only triggers of downstep.

Notice, however, that the proposal to treat bitonal pitch accents as the trigger for downstep is grossly arbitrary. It clearly inherits the spirit of the SPE-type of phonology in which many kinds of processes could be described in terms of features and rules, to the extent that even unattested ones could be predicted. For example, it was perfectly possible to design a rule whereby laterals became nasals in the context of laryngeals. Clearly this type of statement is highly undesirable, due to the fact that the three categories involved in this process have absolutely nothing in common, and yet they can be put together to predict a phonological process which is not even attested. The case of stepping in intonation is similar to the one from SPE, in the sense that bitonal pitch accents do not share any characteristics with the tone which undergoes downstep. Despite these factors against an account of stepping based on the existence of a bitonal pitch accent, B&P'86's followers adopt their view, as I shall explain in the following section.

A major concern of the following chapter is to present an alternative account with greater degree of coherence than the one I have just described. There I propose that stepping arises as a consequence of the presence of toneless boundaries in phonological representation. These boundaries are interpreted by GE as a strong down-pushing force which pulls down the interpretation of following boundaries. This effect is perceived as a step-down. As I shall discuss later, this account has some advantage over others, in the sense that it yields a natural and immediate relation between the trigger and the target involved in the down-step.

2.1.2 Grice (1992)

2.1.2.1 Preliminaries

Grice'92 puts forward an account of stepping which resembles that of B&P'86 in as much as she allows for bitonal pitch accents to act as the trigger for stepping. However, her proposal is more sophisticated than B&P'86's in the sense that tones are now organised into particular tonal structures, such as those illustrated below in (11). The motivation for this proposal is two-fold: to offer an optimal account of stepping contours (including *calling* contours), and to provide a unified analysis of these contours and the so-called *early peak* contours. In the following paragraph, I outline the problems found in P'80 and B&P'86 as discussed by Grice, in order to set the context of Grice's proposal.

The first problem is related to the phonological representation suggested for the contours in (3) above; namely (a) is different from the others in two respects. As can be seen, in (a) the second pitch accent (H+L*) is not downstepped, whereas in (b), (c) and (d), the second pitch accent (H*+L, L*+H, and L+H*, respectively) is downstepped. Furthermore, in order to account for the contour in (a), two types of pitch accents are used (H* and H+L*), whereas in the other examples, there is only one pitch-accent type. Ideally, given that the four contours are the interpretation of the same phenomenon, they should be captured by similar phonological structures. That is, in all representations, the second pitch accent should be downstepped, and bitonal pitch accents should be used.

Another issue discussed by Grice is the representation of downstep in calling contours as proposed in B&P'86. These involve sequences like those in (5). Following Grice's convention, those tones which undergo downstep are underlined for the purposes of clarity.

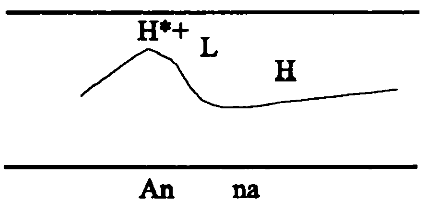
(5) (a) AN$_{\text{na}}$ (a') H*+L $\underline{\text{H}}$ L%

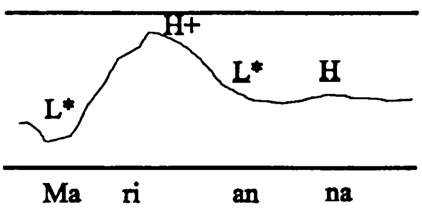
(b) MA ri ANna (b') L* H+L* H L%

(c) Mari AN na (c') L+H* H L%

All the examples in (5) involve a step down from relatively high pitch to mid pitch. The crucial point here concerns the phonological representation proposed to capture the stepped down and mid pitch portion of the contour. For example, in (5)(a') and (c') the underlined H is responsible for this part, whereas in (b'), L*, together with underlined-H, is also needed. I reproduce a schematised diagram of (5)(a, a') and (b, b') below in order to clarify this point.

(6) (a) (b)





The phonological representation responsible for sustained pitch in (6)(a) (see underlined tones) is different from that in (6)(b), and yet the same phonological phenomenon is being described. In other words, the analysis of calling contours does not display the degree of unity that we would wish.

In relation to the early peak contours, the main difficulty in previous models (O'Connor and Arnold 1973, P'80) has been how to account for a rising pitch movement over unaccented syllables before any accented material. For instance, consider (7):

(7) (a) It's a BYSmal (b) To the MARket

It's a in (7)(a) and *To the* in (7)(b) are chunks of unaccented speech which show a rising pitch movement which occurs just before the first accented syllable, *bys* in (a) and *mar* in (b). O'Connor and Arnold (1973) ignore this contour, as they only describe high or low preheads. P'80 suggests the sequence L% H+L*, in which H would account for the high peak.

This representation is rejected by Grice on the basis that it does not achieve descriptive adequacy. Notice that, according to Pierrehumbert, the interpretation of L* in H+L* is around mid pitch (as in (3)(a)). On the other hand, as observed in (7), pitch actually falls to a rather low level. This representation is left to account for the contour in (9)(b) below.

A possible solution to (7) would be to treat the accented syllable simply as L*, although this offers only a partial solution to the problem, since it still does not account for the early high peak. I shall add more about the contours in (7) later in this section.

Another issue pointed out by Grice'92 (and also by references therein), is that **after** the peak, the following accented syllable can be interpreted on low or mid pitch (that is, it can also be downstepped). This is illustrated in (8)(a) and (b) respectively.

(8)

(a) The ASTronomical COST of the...

(b) The ASTronomical COST of the...

In (8)(a), *COST* is rather low, whereas in (b) it turns out to be on mid level. According to Grice, B&P'86 propose the following structures for each contour:

(9)

(a) The AS Tr o no m i ca l COST of the...

L*(+H?) L* L

(b) The AS Tr o no m i ca l COS T of the...

L* H+L* L

In (9)(a) the first pitch accent must be L* +H so that H accounts for the rising pitch movement. The drop to low pitch is analysed as L*. In (9)(b), the first accent is just L*. The high peak is accounted for by the leading H-tone of the bitonal pitch accent H+L*. The mid pitch is analysed as the starred tone of the bitonal pitch accent.

As I have just illustrated above, B&P'86 can account for these patterns. Nevertheless, in §2.1.2.2 I shall demonstrate how Grice'92 accounts for these contours once she has rejected bitonal pitch accents with the starred member to the right, as in (H+L*), from her inventory of pitch accents.

2.1.2.2 Grice's proposal

In order to put forward a unified account of the stepping, calling, early peak, and contours in general, Grice proposes the structure in (10), in which tones are no longer ordered in a string, but instead show a hierarchical organisation:

(10)

```
        P A
       /    \
      /      \
    τ_w        τ_s
     |        /  \
     |       /    \
     T     T_s    T_w
```

The internal structure of the P(itch) A(ccent) consists of a branching or non-branching Supertone (τ), which itself consists of T(ones). The well-formedness of intonational structures is constrained by conditions which state, for example, that nodes must branch in a maximally binary fashion. For instance, only the PA node and the strong Supertone node may branch. A condition that they may not branch simultaneously is necessary in order to obey the binary constraint. Without such a condition, a sequence of three tones is predicted. In order to formalise this further, she refers to the principle of directionality in licensing (Kaye et al. 1990)[4], whereby a licensing position can license another adjacent position either to its right or to its left, but not both. In terms of the PA structure, this is understood in the following way: T_s can license T_w or it can license T, but not both at the same time since this would entail that licensing operates to the right and to the left simultaneously[5].

Unfortunately, Grice does not elaborate on the principles of licensing any further, and these are not mentioned elsewhere in her work. Had she done so, she would have been able to explain why binarity is a fundamental principle in current phonological theories: binarity is a reflection of licensing relations between a licensor and a licensee.

Other constraints stipulate that the strong Supertone node must be to the right, and that the strong Tone node must be to the left. This means that those pitch accents in previous models which showed the starred tone on the right are rejected and that the contours they account for are re-analysed in these new terms.

Taking all this into consideration, the structures which can be generated by her model are the following:

[4]See chapter 2.

[5]However, in §2.1.3 we shall see that she allows for pitch accents which are tritonal underlyingly, but which surface as bitonal.

(11) (a)

```
     PA
    /  \
   /    \
 τw      τs
  |       |
  |       |
  T       T

 L —  H*
 H —  L*
 H —  H*
 L —  L*
```

(b)

```
    PA
     |
     |
     τ
    / \
   /   \
  Ts    Tw

 L* + H
 H* + L
```

(c)

```
  PA
   |
   |
   τ
   |
   |
   T

  L*
  H*
```

Those tones which are dominated by (τ_s) belong to the core part of the pitch accent. Tones which precede the core are referred to as leading tones.

Let us see how these structures can account for some of the examples in (3), (5), (7) and (9)(b) (in (3'), (5'), (7') and (9')(b'), respectively). In (3') I reproduce the phrase *There are millions of intermediate levels* uttered in the fashion of stepped-down terraces. Each terrace is represented as H—H* as they are on level pitches. Additionally, they constitute bitonal pitch accents, so they can trigger stepping on the core of the following pitch accents. These are the second and third starred tones which are underlined.

(3') There are MILLions of inter
MEDiate
LEVels

L% H — H* H — <u>H*</u> H — <u>H*</u> L L%

The structure proposed by Grice to account for the contour in (3') is an improvement over Pierrehumbert's to the extent that, in Grice's proposal, the sequence of pitch accents are all alike, whereas in Pierrehumbert's, the first pitch accent (H*) differs from the rest (H+L*). Thus, Grice's proposal elegantly captures the generalisation that exact copies of the first pitch-accent pattern are involved in stepping structures.

However, this analysis is not without its shortcomings either. As claimed by Grice'95, there is not sufficient motivation for the stipulation that downstepping is limited to the core of pitch accents. In the next section, I shall examine how she solves this problem.

With reference to calling contours, I shall focus on the structure proposed by Grice for *Marianna*, with the accented syllable *-AN-* on a pitch lower than the preceding syllable, and compare it with the other structures in (5).

(5') (a') AN na

H*+L $\underline{\text{H}}$ L%

(b') MA ri ANna

L* H — $\underline{\text{H}}$* H L%

(c') Mari AN na

L—H* $\underline{\text{H}}$ L%

L* and the first H in (5')(b') account for the rising pitch movement over *Mari*. H—H* is a bitonal pitch accent whose second member is downstepped. After that, H* and H are interpreted on mid level pitch. Similarly, in (5')(a') and (c') the mid level pitch is accounted for by H.

Recall that Pierrehumbert's analysis of (5)(b) was undesirable since it prevented her from proposing a unified account of calling contours — specifically, she had L in this particular structure, whereas in the others, she had H. Grice's account on the other hand, overcomes Pierrehumbert's problem by proposing H for the entire set of calling patterns.

Another point which affects the viability of Pierrehumbert's proposal is that the interpretation of L* is not straightforward. L* is normally interpreted as rather low pitch, and here it manifests itself as mid pitch. On the other hand, by using H* (rather than L*), Grice's account is more consistent.

Taking all this into consideration, Grice's proposal is superior to that of Pierrehumbert's in that it accounts elegantly for the early peak in the contour, and also, that it puts forward a unified and coherent analysis of calling contours.

Notice, however, that the structure put forward in (5')(b') relies on yet another stipulation for the account of stepping: that is, tones **within** the pitch-accent trigger can also undergo downstep. This entails that the first stipulation — that pitch accents

following bitonal pitch accents undergo downstep — has to be modified in order to account for data like that presented in (5')(b'). Clearly, this needs to be revised, and Grice does so in her 1995 paper which I shall discuss below.

The last structure to comment on from Grice's early work is the one put forward for early peak contours. For example, the contour illustrated in (7) is accounted for as follows:

(7') It's a BYS mal

H — L*

H accounts for rising pitch over *It's a* and L* for low pitch[6] on *BYS*. Recall that, as I noted earlier, this pattern is not adequately accounted for by any of the previous models.

In relation to (9)(b), in which B&P'86 make use of a bitonal pitch accent with the starred tone in second position, Grice is forced to choose a different structure. This is due to the constraint that starred tones can only occur on the left branch of a Supertone node. Hence, (9)(b) is now analysed as follows:

(9') (b') The AS Tr o no m i ca l COS T of the...

L—L* H—H* L

The initial bitonal pitch accent L—L* accounts for the low pitch over *the* and the accented syllable *AST*. The first H accounts for the early high peak on *-stronomical*. After that, the stepped down syllable of *COST* is analysed as H*. The L-phrase accent accounts for the rest of the contour. The reason why H* in the second pitch accent is downstepped is that it is preceded by another bitonal pitch accent.

Let me now compare this structure with the one which Grice puts forward for (9)(a), in which there is no step down:

[6]Grice's interpretation of L* is on a rather low pitch, unlike L* in Pierrehumbert's H+L*, which is mid level.

(9') (a') The AS Tr o no m ical COST of the...
L* H—L* L[7]

In this case, the first pitch accent is not bitonal, and that is why there is no step down on *COST*; instead *COST* is an accented syllable on low pitch, which is represented by the starred tone of H—L*. (I shall return to this contour in §2.1.3.)

Summing up, Grice's model has the advantage over Pierrehumbert's, in that her account of early peak, calling and stepping contours is more consistent. However, as discussed in Grice'95, there are still some problems with the PA structure which remain unsolved. I shall summarise the main ones in (12).

(12) (a) Bitonal pitch accents trigger downstep only on the core component of the following pitch accent or on a following phrase accent, and not on the peripheral component.

(b) In order for a pitch accent to be downstepped, it must be preceded by a bitonal pitch accent.

(c) In initial and final position some of the structures proposed for the pitch accents are neutralised[8].

Grice acknowledges that (12)(a) is an unprecedented stipulation which requires further motivation. The statement in (b) does not allow her model to capture some contours which she would like to group with other downstepped ones. For example, there are cases in which a pitch accent can be downstepped, but in which this pitch accent is not preceded by a bitonal pitch accent which can act as a downstep trigger. In relation to (12)(c), Grice minimizes the seriousness of this issue when she states that 'the ambiguities occur in restricted contexts' (p. 214). But, in fact, (c) arises as a consequence of the model's enriched tonal structure, and therefore its excessive generative power.

[7]Grice claims that L*—L and L* can neutralise in initial position (1995:214). However, if she adheres to the constraint that only bitonal pitch accents can trigger downstep, then L*—L and L* do have different functions in phonological representation, and therefore, should not be neutralised.

[8]See Grice'95 for a full discussion of this issue.

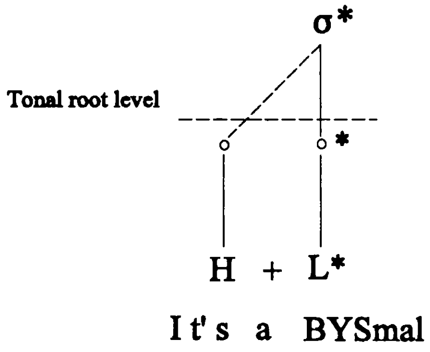
In acknowledgement of these problems, Grice revises her model by making it more constrained. For instance, she dispenses with the Supertone node, so that the resulting flatter structure reduces the possible patterns that can be generated. Additionally, she introduces some constraints which condition the occurrence of strong positions in the tonal structure. These issues are all addressed in the following section.

2.1.3 Grice (1995)

Grice borrows from Yip (1989) the idea of grouping tones differently to form *clusters* or *melodic* units; specifically, she claims that tones can be grouped in this fashion to create pitch accents in English. In addition to this, Grice adopts Yip's tonal structure which is simpler than the one put forward in Grice'92. Let me illustrate the revised model and present the structure suggested for the contour in (7), *It's abysmal* which, as I explain below, counts as a cluster of tones.

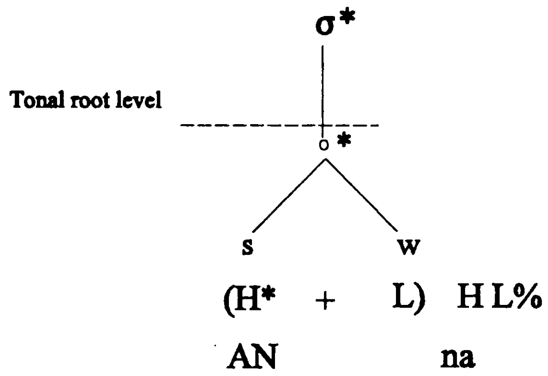
As can be seen below, the accented syllable is on a different tier from the tonal root level. This is indicated by the dashed horizontal line. What was formerly the Supertone tier has disappeared. Instead, there is only the tonal root node to which tones are directly associated by means of vertical lines. The dashed slanted line indicates that H has associated to the accented syllable late in the derivation, after the starred tone has been associated. Because tones are associated to two different root nodes, they constitute a cluster of tones. In this structure, H constitutes a leading tone and is part of a bitonal pitch accent. L* is part of the core.

(13)



Tonal root nodes can also show a branching structure such as the one illustrated below, which accounts for downstepping contours like the one in (5') above. A **branching** node is what constitutes a melodic unit.

(14)

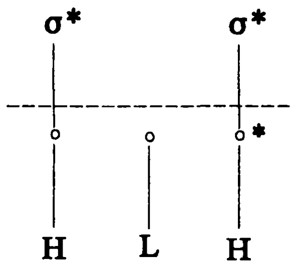


In the case of branching nodes, Grice adopts another idea, this time from Pierrehumbert & Beckman (1988), which is that nodes are labelled as strong or weak. However, Grice does not endorse the idea that nodes are freely labelled. Instead, she imposes some constraints on the structure, such that the strong node is always to the left, as seen above.

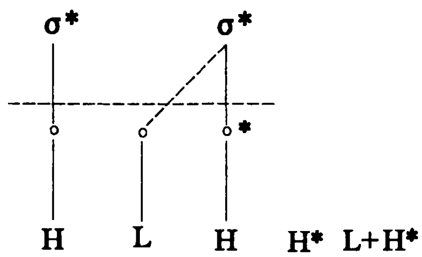
Notice that only bitonal pitch accents whose internal structure is that of a branching root node can trigger downstep on the following node. So, for example, the branching root node in (14) is a trigger of downstep; as a consequence of this, the following H phrase accent is interpreted on a lower level of pitch than the preceding H* (L is treated as a floating tone and is not interpreted). On the other hand, pitch accents like the one in (13) do not count as triggers of downstep, as their tones are associated to two independent tonal root nodes[9].

Another constraint in Grice's model is that starred nodes in an accented syllable occur to the right. For instance, the structure in (13) is well-formed. Also, according to this constraint, no structures can be generated in which the starred node occurs to the left; for example in a sequence such as the one in (15), L is associated to the second starred node, rather than to the first one. Otherwise, a violation of this constraint would result. This association is shown in (16).

(15)



(16)



The former constraint, in addition to the one mentioned above in relation to the strong branch of a node being to the right, together with the Obligatory Contour Principle (OCP)

[9]In her paper, Grice uses parentheses around those bitonal pitch accents which belong to a branching root node.

which operates at every level within the pitch accent, ensure that there is a high degree of restrictiveness in the model.

Let me now briefly outline and illustrate the inventory of pitch accents which can be generated under these constraints.

(17) Two monotonal pitch accents.

```
(a)   σ*                      (b)  σ*
      |                            |
------|---              -----------|---   Tonal root level
      σ*                           σ*
      |                            |
      H                            L
```

(17') Examples:

(a) the first accented syllable in *Marianna* (which is the same contour as in (5')(b)).

```
MA  ri      ANna
L*  (H+L)+H*  H  L%
```

(b) the accent in *millions*

```
There are MILlions of inter MEDiate...
          H*           (H+L)+H*
```

(18) Two bitonal pitch accents with a branching tonal root node.

```
(a)   σ*          (b)   σ*
      |                 |
------|---        ------|---
      σ*                σ*
     / \               / \
    /   \             /   \
   s     w           s     w
   H     L           L     H
```

(18') Examples:

(a) the accented syllable of *Anna*.

AN na

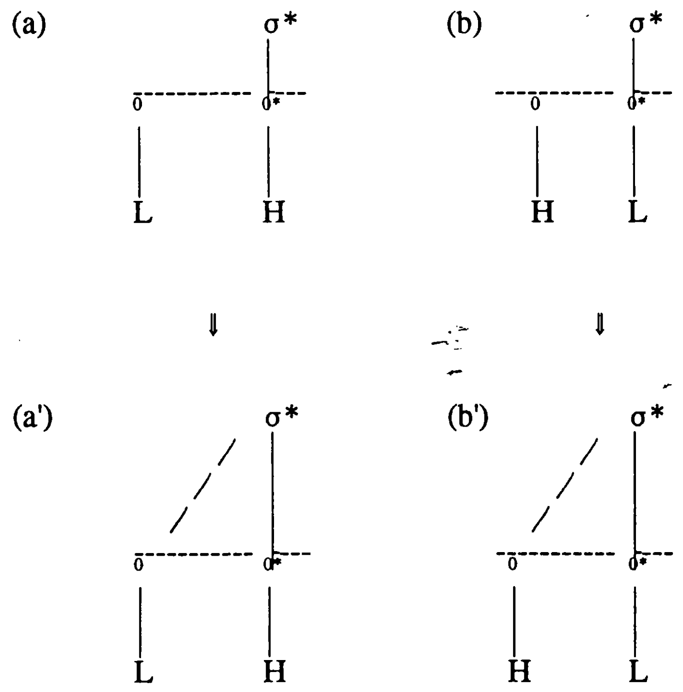
(H*+L) H L%

(b) the accented syllable of *really*.

I REALly don't...

(L* + H)

(19) Two bitonal pitch accents with two tonal root nodes. They constitute a sequence of tones, which can potentially become clusters. If they do become a cluster (as shown below in (a') and (b'), respectively), then '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' (Grice 1995: 217).



(19') Examples:

(a) the accented syllable of *car*

I had to get my CAR registered...

L + H*

(b) the accented syllable of *abysmal*

I t's a BYS mal

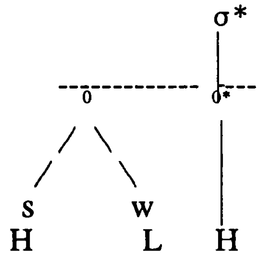
H +L* L L%

(20) Two underlyingly tritonal[10] (but phonetically bitonal) pitch accents with two tonal root nodes.

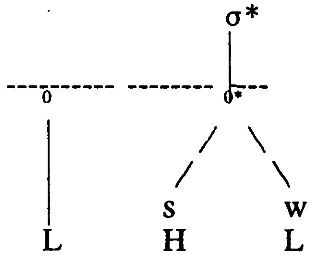
---

[10]Tritonal structures have already been used by Gussenhoven (1983) and Lindsey (1983) — the former where other analysts use a combination of a bitonal pitch accent and a boundary tone, and the latter, where other analysts use a combination of a bitonal pitch accent and a phrase accent tone. However, in Lindsey (1985) this analysis is rejected in favour of another analysis based on the use of features.

(a)



(b)



(20') Examples:

(a) see the second pitch accent in *Marianna* above.

(b) the accented syllable in *Marianna* (which is the same contour as in (5'(c')):

Mari AN na

L + (H*+L) H L%

Observe that the **underlyingly** tritonal pitch accents can potentially be interpreted as tritonal pitch accents. This would unnecessarily increase the number of representations, which could also turn out to introduce some degree of ambiguity. Grice is aware of this danger, and responds by introducing the constraints which I mentioned above and also by stipulating that the surface form of the pitch accent must be binary (that is why in (14) L is floating and is not interpreted phonetically). In the following section, I shall discuss the problems that remain unsolved from her previous model in the framework of the present version of her model.

2.1.3.1 Grice 's solutions to problems from (1992)

Remember from the previous section that in Grice'92 there was insufficient motivation for stipulating that stepping is limited to the core of pitch accents. Additionally, there were cases in which tones within the pitch accent trigger could also undergo stepping. In order to put forward a unified account of stepping, in the 1995 version of her model, Grice states that a **branching** node (that is, a melodic unit) constitutes the structure which triggers downstep in the following node or the following phrase accent.

An example of the former is shown in (17') in which (H+L) triggers downstep on the following starred tone H*. Notice also, that by having a branching node which triggers stepping on the following node, Grice solves the problem which she reported in relation to the structure in (5')(b'). Recall that she had to stipulate that a tone within the pitch accent trigger can also undergo stepping; this was seen as an arbitrary and undesirable contribution to her analysis, which has now been solved.

An example of the latter is illustrated in (18')(a) where (H*+L) triggers downstep on the following H phrase accent. In addition to all this, the reader must observe that there is no downstep-effect in (19')(b) because the bitonal pitch accent does not constitute a melodic unit.

Another issue related to early peak contours pointed out by Grice'95, is that there are cases in which a pitch accent can be downstepped, and yet it is not preceded by a bitonal pitch accent. More specifically, here she refers to a contour such as the one illustrated in (21), where the accented syllable of *toMAtoes* is downstepped with respect to preceding unaccented syllables (in this case to).

(21) The to MA toes haven't arrived y e t

I discussed this type of contour in (9)(b). However, there, the downstepped pitch accent was preceded by a **bitonal** pitch accent and, therefore, there was a clear motivation for the second pitch accent to undergo stepping. Now, the **first** accented syllable is downstepped without a pitch accent acting as a trigger for such a step. This finding in the data constitutes a problem for Pierrehumbert and Beckman (1988) and Grice'92 (see discussion in relation to (9)(b)). Nevertheless, Grice'95 overcomes this problem by having the structure included in (20)(a) in her inventory of representations, so that the contour in (21) is accounted for by the following combination of tones:

(22) The to MA toes haven't arrived y e t

(H+L)+H*

The pair of tones in parentheses belong to a branching node, which is the structure that triggers downstep on the following H* tone.

The last problem included in (12)(c) above — namely, that the model in Grice'92 presents a highly elaborate structure which could generate different representations which, in some contexts, receive the same phonetic interpretation — is overcome in Grice'95. This is achieved by the exclusion of the Supertone node and also by formulating the constraints outlined in §2.1.3.

The benefits of the revised model can be summarised as follows:

(23) (a) A higher degree of constraint is achieved by a reduction in the number of levels in tonal structure.

(b) The model can account for early peak, calling and stepping contours in an elegant way by proposing a phonological distinction between melodic units and tonal clusters.

Nevertheless, there are still some unsolved problems, which I discuss in the following section.

2.1.3.2 Difficulties in Grice (1995)

(24) (a) Tritonal structures may open the way to ambiguities in the representations of some contours.

(b) The purpose of adhering to binarity is not immediately clear and it raises a series of questions.

(c) Ordering in association of tones (see (19)).

(d) It is stipulated that an item triggers downstep on another item to its right. It could just as easily be stated that an item triggers upstep on another item to its left.

(e) Like Pierrehumbert's, the representation proposed by Grice for (3)(a) fails to produce a sequence of equal pitch accents.

I shall examine each one of these in turn.

Tritonal structures

Although tritonal pitch accents elegantly account for pitch patterns which involve downstep, it is highly unde sirable that a model should show such elaborate structures, as it may open the way for ambiguities in the representation of some contours. Consider, for example, a structure such as (L+H)+L* which is predicted to be perfectly possible in this model (and yet it is not included in the pitch accent inventory by Grice[11]), since it abides by all the conditions outlined above: the starred accent is to the right, it obeys the OCP and L may be treated as a floating tone so that only two tones will be interpreted. This structure could be used to account for the contour in (19'), which I reproduce below with both this structure and the one already proposed by Grice:

(25) I t's a BYS mal

H +L* L L%	(Grice'95)
L% (L+H) + L* L L%	(alternative representation)

The initial boundary tone accounts for low pitch and H from (L+H) for the rising movement. Due to the fact that there is a branching node (L+H) in phonological representation, the following tone is downstepped. However, because L* is already rather low, its effect is not clearly manifested in phonetic interpretation (the rest of the representation remains the same).

Obviously, these cases, in which a contour can be accounted for in different ways, allow for a high degree of indeterminacy in phonological representation, and for this reason they should be avoided. This matter clearly deserves some attention in the future.

Binarity

The idea that tones in pitch accents must be in a binary relation is taken very seriously by Grice, so much so that she even uses two different ways of showing binarity in phonological representations: by means of stars and by means of strong/weak labelling[12].

---

[11]See discussion about tritonal accents below.

[12]See chapter 2 where I introduced arguments against strong/weak labelling.

In addition to this, she stipulates that tritonal pitch accents in phonological representation show a floating tone which never surfaces, so that the surface interpretation of the tritonal pitch accent results in a binary pitch accent.

The problems which arise as a consequence of adhering to these stipulations are as follows:

(26) (a) the use of stars and strong/weak labelling may be misunderstood as implying that the type of relations between a pair of items are of a different nature depending on which diacritic is chosen;

(b) in unstarred branching nodes, strong/weak labelling is meaningless, and therefore, redundant. The same result is achieved if the labelling of some nodes is reversed. In addition, this labelling does not have any manifestation in phonetic interpretation (e.g. (20)(a), (unlike B&P'86's use, which does));

(c) it is not made explicit why bitonal (but not tritonal) pitch accents are the only ones which are allowed to appear in phonetic interpretation.

(d) it is not clear which tone in tritonal structures counts as a floating tone; whether it is a tone in a particular structural position (that is, starred or unstarred; in a branching node or non branching node) or whether it is H or L; another possibility would include a combination of the previous two options.

Given (26)(a), (b), and (c) above, I argue that Grice's use of binarity is somewhat incoherent and therefore, that this needs further clarification.

Ordering in association of tones

With reference to (24)(c), implicit in the claim that unstarred tones associate to syllables in the second stage of derivation — after starred tones have been associated — is the idea that there are different stages in the construction of phonological representations. The main argument against this proposal is that, logically, we may expect to find languages in which the association of unstarred tones precedes the association of starred tones.

However, this other possibility of association is not even contemplated by other authors and this may be an indication that it does not constitute a feasible option. This issue requires further investigation.

Target to the right of trigger

In relation to (24)(d) it must be pointed out that this problem is not exclusive to Grice's model. It has long been present in other models for tone languages (Clements (1983), Goldsmith (1990), Pulleyblank (1986) and for intonation languages (P'80, B&P'86). The core of the problem lies in the fact that it is possible to state an equally logical account which would run as follows: an item triggers upstep on another item which occurs to its left — in addition to other conditions which are not immediately relevant to the present issue. If this alternative is available, then it poses some fundamental questions as to the explanatory adequacy of current models.

Before I arrive at any conclusions and their repercussions for a phonological model for intonation, let us consider whether the option to upstep an item to the left which I mentioned earlier is, in fact, viable. For this, let us consider the contour and structures below (which are the same as in (17')):

(27) There are MILLions of inter MEDiate LEVels

L%	H — H*	H — $\underline{\text{H*}}$	H — $\underline{\text{H*}}$	L L%	(Grice'92)
	H*	(H+L)+$\underline{\text{H*}}$	(H+L)+$\underline{\text{H*}}$	L L%	(Grice'95)

It is not difficult to state the condition for upstep here. It suffices to claim that a pitch accent can trigger upstep on another pitch accent to its left. It might be the case, however, that for other contours, other stipulations may be necessary in order to make generalisations about the occurrence of upstep.

Nevertheless, the fundamental point which I want to raise here is that current models have yet to come close to producing an explanatorily adequate account of stepping, and that a model is desperately needed which can explain this phenomenon.

Unequal accents in stepping contours

Finally, as stated in (24)(e), note that the structure proposed by Grice for the contour in (3) is

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

which falls into the same pattern as the one originally proposed by Pierrehumbert (in as much as the first pitch accent is different from the rest) which is

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

Recall that in §2.1.2.2 Grice herself rejects Pierrehumbert's proposal, as it renders the account of stepping rather incoherent. Thus, Grice'95 displays no advantage over Pierrehumbert in this respect.

2.2 Stepping related to metrical structure

In this section I shall analyse and discuss Ladd's (1993) account of stepping by means of metrical structure. In order to set the context for Ladd's model, first I outline his earlier account of stepping (1983) and how this is negatively evaluated by other authors and myself. In response to their criticisms, he is forced to revise his proposal and offers the one put forward in the (1993) version.

Soon after the publication of Pierrehumbert's first work, Ladd (1983) launches an attack on her use of L as a trigger for downstep in the sequence H*+L H. One of the criticisms is directed at the fact that the manifestation of L in the F0 contour is far from clear; it does not appear as a well-defined F0 valley, as shown by the curved line in the schematised F0 diagram in (4). Instead, the F0 contour at this point closely resembles the movement shown by the dashed line in the same figure. Since there is insufficient evidence available for positing the presence of L in phonological representation, this approach is quickly rejected.

Another issue pointed out repeatedly by Ladd (1983:724; 1993:111) is that by having L as a trigger for downstep, similarities between members of a natural class of contours are highly obscured. For instance, it is expected that a sequence of falling accents in a non-downstepping and stepping contour would show similar tonal structures, given that the same type of (falling) pitch accents are involved. However, as Ladd points out, the tonal structure of the PAs varies considerably from one representation to the other: from a sequence of H* L- L% in the non-stepping pattern to a sequence of L-+H* in the stepping pattern. In addition, the sequence of low tone boundaries present in the former structure is absent in the latter. Hence, Ladd repeatedly rejects Pierrehumbert's analysis of stepping, claiming that it

(30') obscures the functional similarity between a downstepping contour and an otherwise similar series of non-downstepping pitch accents; at the same time it obscures the functional similarity shared by all downstepping contours as against all non-downstepping contours. (1993:111).

As a first alternative, Ladd (1983) then suggests that an SPE-type feature [downstep] is associated to those pitch accents which are downstepped, so that the sequence L+H* now looks like (30).

```
(30)  H* L    H* L
              |
              [downstep]
```

By doing it in this way, the nature of the pitch accents involved in the stepping structure remains unchanged from those which are not involved in a stepping structure. This means that the difference between (30) and a non-stepping sequence of accents is found **only** in that the former structure has [downstep], whereas the latter would presumably lack such a feature.

B&P'86 reject Ladd's feature-based proposal, and maintain their use of a bitonal accent as a trigger for downstep. In relation to Ladd's approach, one of the arguments they put forward[13] is that this feature enjoys too much freedom, as far as its distribution is concerned. For instance, it could be attached to any tone whatsoever, with the result

[13]For a full discussion of this issue, the reader is refered to the main sources, B&P'86 and to Grice'95. The former emphasizes the model's capacity for overgeneration; in addition, it highlights the ambiguity which can arise in the analysis of a given contour with a variety of representations.

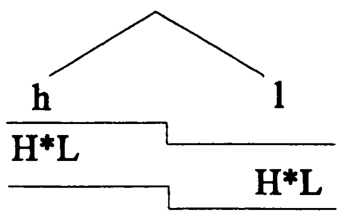
that it can even generate unattested pitch patterns. For example, it could be attached to the first H in a phrase as an indication that this tone has undergone downstep. But in English, there is no contrast between H and H [downstep] initially.

Another point is that Ladd is not very specific in his use of features when he allows for downstep to be represented by the symbol !. On the one hand, this may be understood as if this feature had a privative value, that is, presence/absence of !. On the other hand, no other feature from his set of features can be represented in this privative way. Hence, it seems as if he combines two phonological approaches which actually yield quite different phonological representations and are therefore rather indeterminate as to the nature of phonological oppositions.

In view of all these problems, Ladd (1988, 1990, 1993) abandons [downstep] altogether, and he adopts a completely different model which closely resembles that used in the analysis of tone languages by Huang (1980), and Clements (1983). In short, he proposes that downstep is due to a particular metrical relationship between tones. He represents this by means of an arboreal structure, whose nodes are labelled *high* and *low*, as shown in (31) below . These represent the register specification for two successive pitch accents.

A condition for the occurrence of downstep is that the node to the left is labelled *h*, and consequently, the one to the right is *l*. Only under these circumstances, *l* triggers a lowering of the register. In (31) I illustrate a sequence of falling pitch accents in which the second one is downstepped with respect to the first. The horizontal lines are merely a notational device to illustrate how the register is lowered; they are not part of phonetic interpretation.

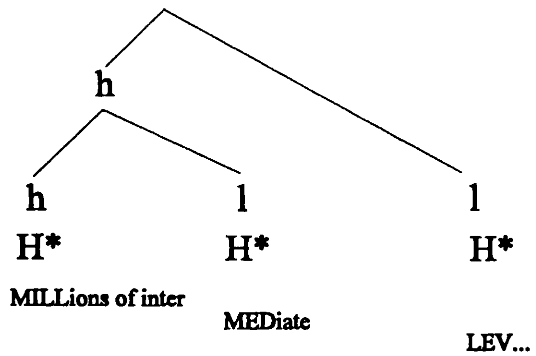
(31)



On the other hand, a metrical structure in which *h* is **preceded** by *l* does not trigger downstep.

As stated by Clements (1990), one of the advantages of this representation is that the manifestation of the cumulative effect of stepping, which I mention in the introduction to this chapter, is straightforwardly captured: there are nested structures, which parallel the representation of subordination shown by means of nested tree-structures in syntax. As an illustration, I reproduce the phonological structure for part of the contour in (3')(c') in (32) below:

(32)



By using structures and principles which have been used successfully in areas of syntax, a clearer and more direct link can be drawn between syntax and intonation. A further benefit of Ladd's proposal is that the representation of stepping does not obscure the pitch accent's identity. Recall that he addressed this as a problem in Pierrehumbert's analysis.

However, there are some disadvantages with Ladd's proposal. Some of them are summarised by Grice in the following quotation.

(33) 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 [...] 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 (Grice1995:207).

By (a), Grice is refering to contours like the ones in (17) which she accounts for by stipulating that only the second tone (H*) of the tritonal pitch accent (H+L)+H* undergoes stepping, and not the entire pitch accent. By (b) she is referring to early peaks followed by a stepped down pitch accent with no other pitch accent preceding, as in (21) above.

In addition Ladd's proposal suffers from other problems, which are related to phonological issues. Clements (1990) points out that Ladd's use of *h* and *l* to show register and *s* and *w* to show prominence relations are redundant and therefore, that one pair should be rejected. He suggests that, given that Ladd's model can account for prominence relations at the level of interpretation, *s* and *w* can be dispensed with in phonological representation.

Ladd (p.c.) admits that the *h-l*-type of representation is not fully formalised. For instance, there is no explanation as to why the order *h-l* should trigger downstep, whereas the reverse order does not.

Finally, in the context of a restrictive phonological model of intonation, having recourse to tree structure and labelling of its nodes, as I claimed in chapter 2, is too costly.

3 Conclusion

The proposals which I have outlined in this chapter leave a series of questions unanswered:

(34) (a) an item to the **right** of the trigger undergoes downstep, but an equally logical stipulation could be that an item to the **left** of the trigger undergoes upstep. There is nothing in their models which could help us to choose between these options.

(b) Pierrehumbert and her followers share the idea that a bitonal pitch accent is the factor underlying any account of stepping. This stipulation is highly arbitrary.

(c) no explanation is offered as to why in contours like (3) the pitch pattern of the first step is copied onto the following downstepping steps.

In addition to this, the above approaches to stepping inconveniently enrich phonological representation by increasing the number of tiers. Grice argues in favour of a tone tier, and the tonal root tier. Ladd contends that a register tier and a tonal tier are necessary to make generalisations about the phenomenon of stepping.

In the following chapter, I put forward my analysis of stepping and attempt to provide a solution to the unanswered questions in (34).

Chapter 6. Empty domains in phonological structure

1 Introduction[1]

This chapter develops some of the notions of phonological structure presented in chapter 4, where I put forward a **restrictive** model of intonation characterised by a single melodic unit, (T), which can associate to the boundaries of prosodic domains. I also proposed that prosodic constituents are organised according to the principles of licensing. Here I shall examine the phenomenon of stepping, which supports the conclusion that a phonological model for intonation can remain maximally constrained by referring to well-established phonological principles.

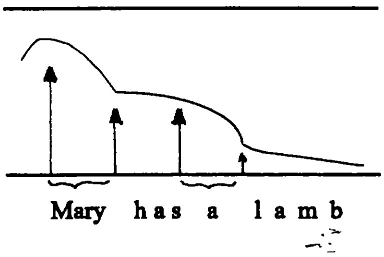
In the preceding chapter I discussed and evaluated current accounts of stepping, and the main argument against them has been that they all have the tendency to enrich phonological structure by increasing the number of tiers in the representation. On the other hand, here, I preserve the model's restrictive capacity by proposing that stepping should be accounted for by referring to what is already available in the representation — specifically, the presence of toneless boundaries.

I begin this chapter by reproducing the stepping contour illustrated in (3)(a) from the previous chapter. The reason for this is two-fold: first, to illustrate the phonological structure which I employ in my account of this contour; and second, to discuss how to formalise such a representation within the framework which I have already presented in chapter 4.

I reproduce (3)(a) in (1) below, in which the actual step is highlighted by a pair of arrows. The utterance is now *Mary has a lamb.*

[1]In this chapter, I omit L0 from prosodic structures in those instances where the illustration of a particular point is not directly related to this level. I also exclude tonal information for the same reasons.

(1)



Provisionally, let us assume that the phonological representation for this example is the one included in (2), in which *Mary, has a* and *lamb* constitute three independent domains at L1; T is associated to their left boundary.

(2)

$$\Big[\ \underset{\mathrm{T}}{[\text{Mary}]_{L1}} \quad \underset{\mathrm{T}}{[\text{has a}]_{L1}} \quad \underset{\mathrm{T}}{[\text{lamb}]_{L1}}\ \Big]_{L2}$$

This structure accounts for a series of falling pitch accents in a straightforward fashion. However, it is clear that this does not constitute the best representation for this contour. The present state of the structure misses the fundamental fact that the second and third pitch accents are shifted downwards in phonetic interpretation. I investigate here the factor I assume to be present in phonological representation which is responsible for this pulling-down effect on the interpretation of these pitch accents.

I argue that the aforementioned effect is due to the presence of toneless boundaries which affect the interpretation of the following T[2]. This proposal follows from the idea I have established above, that, in my model, toneless boundaries constitute the only structural factor which can have the phonetic result of pulling the F0 downwards. This, I illustrate below:

---

[2]At this stage of the discussion, the fact that only T's to the right of empty domains are downstepped is taken for granted. Nevertheless, as the chapter develops, I shall investigate whether or not this state of affairs constitutes a direct reflection of events in phonological structure.

(3)

$$\Big[\ \underset{\substack{|\\ \mathrm{T}}}{[\text{Mary}]}_{L1}\ [\ \ ]_{L1}\ \underset{\substack{|\\ \mathrm{T}}}{[\text{has a}]}_{L1}\ [\ \ ]_{L1}\ \underset{\substack{|\\ \mathrm{T}}}{[\text{lamb}]}_{L1}\ \Big]_{L2}$$

Note that I show empty domains with toneless boundaries in (3). Let me justify the appropriateness of this structure before I go on to describe its phonetic interpretation in more detail below.

Consider, for example, the alternative of adding a pair of boundaries to each L1 domain, as in (4).

(4)*

$$\Big[\ \underset{\substack{|\\ \mathrm{T}}}{[\,[\text{Mary}]\,]}_{L1}\quad \underset{\substack{|\\ \mathrm{T}}}{[\,[\text{has a}]\,]}_{L1}\quad \underset{\substack{|\\ \mathrm{T}}}{[\,[\text{lamb}]\,]}_{L1}\ \Big]_{L2}$$

This option is highly undesirable for reasons which I have already stated in chapter 4. Specifically, by increasing the number of boundaries in this fashion, we introduce the possibility of indeterminacy and overgeneration in phonological structure. In addition to this, it appears that the insertion of boundaries is rather arbitrary: they could be inserted anywhere in the representation, and moreover, in any quantity. Another argument against (4) is that it arbitrarily increases the number of levels in prosodic structure. Thus, in view of all these imperfections, I am forced to reject this alternative. On the other hand, as I shall illustrate throughout this chapter, by pursuing the structure in (3), these shortcomings are avoided.

Let us now briefly focus on the phonetic interpretation of the representation in (3), in order to illustrate its descriptive adequacy. Recall that, as I have stated above and also in chapter 4, toneless boundaries are interpreted as a down-pushing force on F0 by Gravitation Effect. According to this, toneless boundaries have the effect of pulling down the F0 contour. However, due to the fact that this domain is empty and that the boundaries are so near to each other, there is greater accumulation of GE, and consequently, the interpretation of T is greatly affected; so much so, that it is shifted downwards. For example, the second T in (3) is interpreted at a level lower than the corresponding T in (2).

So far, I have introduced the reader to the components of phonological structure which I propose are necessary for an account of stepping. Nevertheless, the actual organisation of these domains — especially the empty ones — into phonological constituents remains to be explained. This is the topic in the following section.

The last point I shall address here relates to empty domains, and more specifically to their encoding in the phonological representation of intonation. In other words, I would like to consider at what stage empty domains are encoded in the process of assembling the different parts of a sentence.

Without entering into a discussion of any technical depth (it must be borne in mind that it is not the aim of this thesis to investigate this area), I suggest that both empty and filled domains are encoded in the sentence simultaneously, and not at different levels of construction. In order to illustrate this, let us draw an analogy between the addition of a qualifying adjective and an empty domain into a sentence. For instance, let us assume that a speaker produces the sentence *Peter has a dog* in a given situation, and later he realises that he wants to be more specific and he says *Peter has a **big** dog*. The point I want to clarify here is that the speaker plans his message as he goes along, and that, in this process, he encodes all the different domains in the sentence. It is not the case that *big*, for example, is added to the second sentence once the speaker goes through the process of constructing the first one. In fact, they are completely independent items. In addition, the speaker chooses to encode *big* before *dog*, and this choice is under the control of general principles of grammar which dictate the distribution of qualifying adjectives in a sentence.

In linguistic terms, I understand this as follows: *Peter has a dog* and *big* constitute two independent strings which are put together to generate a composite representation. Notice that this example implies that one lexical item is inserted in another lexical item. Obviously, this is far from what is actually intended. In order to achieve a better understanding of the idea I am trying to illustrate, let us think in terms of stress:

given a stress pattern like (* . .), this can be superimposed on the word *vanity*. Similarly, *big* is not inserted in *Peter has a dog*, but instead, as I have just stated, it is **superimposed**.

In relation to empty domains, this is understood as follows: they are encoded in the intonation domain as this is being constructed, so that, according to some general conditions imposed by the grammar, the speaker can choose their location in the sentence as he goes along. Consider the following example, in which I show empty domains by means of @[3]: *Mary has a lamb* and @ count as two autonomous objects. The former is imposed over the latter in order to generate a stepping structure. That is, there is no insertion, but instead it is a superimposition, just like the case of stress patterns I mentioned above.

The motivation for treating *Mary @ has a @ lamb* as a derived form is to preserve the generative spirit of the model. Had *Mary @ has a @ lamb* been treated as a lexical representation, then we would have ended up with a model which simply lists representations, rather than one which **generates** them. Obviously, we do not wish to make this claim.

2 Phonological representation of stepping

In GP, it is claimed that a set of established principles apply throughout phonological structure; namely, the principles of licensing. I have argued that these are also applicable at the level of intonation; more specifically, in chapter 4, I have defended the view that the organisation of constituents into Onset and Nucleus below the foot domain, is reflected in intonation by the organisation of prosodic domains above the foot similarly into onset and nucleus. Inherent in this statement is the idea that phonological representation is unified by the fact that the same principles apply at all levels.

In order to understand how empty domains fit into constituent structure, I shall refer to the idea in GP that many languages show Onset-Nucleus sequences, in which

[3]For the purposes of clarity, I use @ to show an empty domain in text only. In phonological representation, the domains shall remain empty, thus [ ].

specific Nuclei can be empty[4]. Thus, I shall claim that the distributional patterns controlling empty Nuclei at the syllable level are also present at the level of intonation, and hence, that empty domains at the latter level correspond to empty nuclei.

As an illustration of empty Nuclei in GP, consider the representation of the word <fam(i)ly> (Harris (1994)), in which the medial vowel may or may not be interpreted:

(5)

O	N	O	N	O	N
|	|	|	|	|	|
x	x	x	x	x	x
|	|	|		|	|
f	æ	m		l	i

The case of non-interpretation is captured in phonological representation by the fact that the position in the second Nucleus remains devoid of melodic content[5].

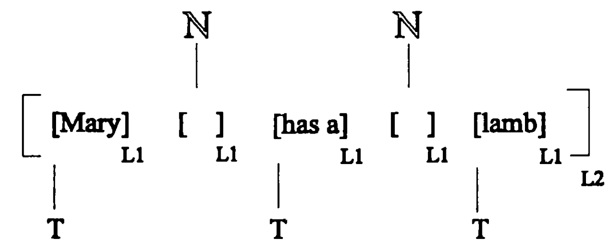
We are now in a position to label empty domains in intonation as empty nuclei, in keeping with the nature of empty positions at the syllable level. Hence, the structure in (3) now looks as follows:

---

[4]There are cases in which an Onset can be empty, as in, for example, the phenomenon of *h-aspiré* in French (Harris 1994:179). But these instances of empty Onsets are marginal cases.

[5]There are certain conditions which restrict the occurrence of empty Nuclei in phonological structure, but these are irrelevant at this stage. I shall examine them later in this chapter.

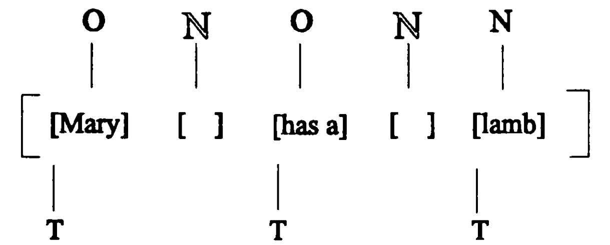
(6)



Empty domains ([ ]) are analysed as being licensed by empty nuclei (*N* is a label which I use for the purposes of clarity. It has no theoretical status different from *N*). Nevertheless, constituent structure remains incomplete until all the domains are identified.

Recall that, as I stated above, phonological structure is characterised by a sequence of onsets and nuclei. Therefore, it is safe to identify [*Mary*] and [*has a*] as onsets since they are themselves followed by nuclei (which have already been motivated). Let us now assume that [*lamb*] is the nucleus. Thus, the phonological representation of *Mary @ has a @ lamb* looks as follows:

(7)



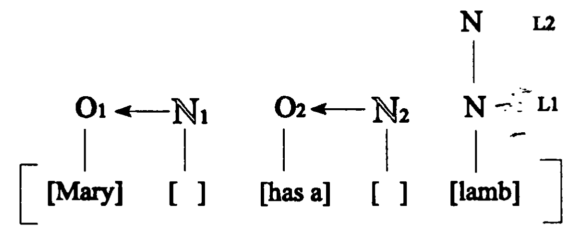
So far, I have introduced the reader to the phonological structure which accounts for the phenomenon of stepping by resorting to well-established principles in GP. However, the conditions which constrain the occurrence of empty categories in the structure remain to be explained. This is the topic of the following section.

2.1 Empty nuclei and the principles of licensing

The occurrence of a nucleus between the two onsets in (7) is justified by the well-established idea that an onset must be licensed by a nucleus in phonological representation

(Kaye 1990, and Harris 1994). This is illustrated in (8), where directionality of licensing at this level goes from right to left:

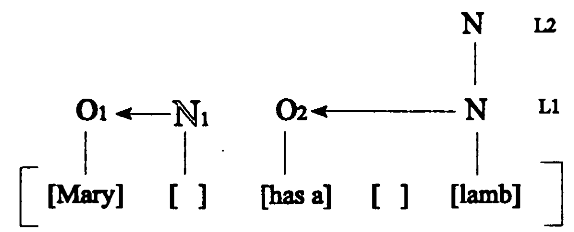
(8)



However, note that the presence of N_2 generates a sequence of two adjacent nuclei at L1[6]. In this respect, it could be argued that N_2 is redundant in the representation, since the task of licensing O_2 might as well be performed by N. Let us examine the reasons for rejecting this argument by analysing the consequences of omitting one of the relevant nuclei from the representation.

The possibility of omitting N is immediately rejected on the grounds that this position is the ultimate licensor of the whole structure[7]. Let us then focus on the other alternative, that of the omission of N_2, as I illustrate in (9).

(9)*



The crucial point to observe here is that, if N_2 is omitted from the representation, the phonetic effects of [ ] would similarly be absent. This is clearly not the case, since

[6]The situation in which two nuclei are adjacent to each other is common in GP analyses to account for a sequence of two vowels in a language in which there is not sufficient motivation for allowing the possibility of branching struture. For example, Japanese provides a case in point.

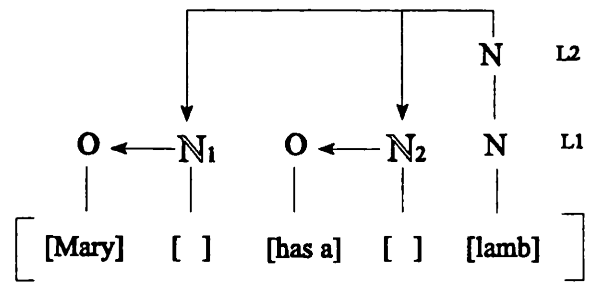
[7]The reason for treating *lamb* as the ultimate licensor is that it counts as the only obligatory constituent in the structure.

[ ] is manifested in phonetic interpretation; it displays a pulling-down effect in the following tonal pattern[8]. Thus, this possibility is also rejected.

I conclude that the above argument constitutes further evidence in support of the structure illustrated in (8). (Nevertheless, it must be made clear that, had there not been [ ] in the representation, this structure would have been well-formed. It accounts for another type of contour; one in which *lamb* is not under the effect of a down pulling force).

Of immediate relevance here is the question of how empty nuclei are themselves licensed, for, in order to be manifested in phonetic interpretation, all units must be integrated into a licensing relation. Given the principles of licensing which I have presented thus far, I conclude that empty nuclei are licensed by the ultimate licensor, which is the only constituent which possesses enough licensing potential to license empty domains. I show this below, which is a reproduction of (8);

(10)



Before we turn our attention to the conditions on the distribution of empty nuclei in phonological structure, I must clarify the following point: note that at L2, N can license two constituents in the representation, (N_1 and N_2); this clearly violates some very general principles of licensing, specifically with respect to the locality condition imposed on any binary licensing relation[9]. The licensing of N_1 by N is not strictly local, as it is interrupted by the presence of N_2.

[8] In order to prevent [ ] from being stranded, we could stipulate that prosodic constituents may show a branching structure, and thus, [ ] would be part of a branching constituent, either the onset to its left or the nucleus to its right. However, this line of argument has far reaching consequences in terms of the unnecessary enrichment of the structure. Therefore, it is not pursued any further in this work.

[9] See §3 in chapter 4 and further discussion below.

The argument I advanced in favour of the recognition of empty nuclei is a theory-internal one. These empty nuclei must be present in the structure in order to license the onsets (by onset licensing). This led us to examine how empty nuclei are licensed, and then to conclude that they must both be licensed simultaneously by the remaining unlicensed constituent, which is the filled nucleus. Let us now investigate whether there is any independent support for this type of licensing.

Recall that, in chapter 4, we encountered a similar situation, but at a different level in the prosodic hierarchy. Specifically, I claimed that constituents are unbounded and that this type of pattern is well-attested at the stress level in languages like French. Given that I pursue GP's idea of a unified phonological representation, it should not strike us as surprising to find that unbounded constituents are also present at levels other than L0. Hence, the structure in (10) is still within the principles which govern the model: N can license $\mathbb{N}_1$ and $\mathbb{N}_2$ in the intonation domain. It must be made clear though, that the ideas of a strict one-to-one relation between a head and its dependent position, and of locality, must be loosened at the level of intonation.

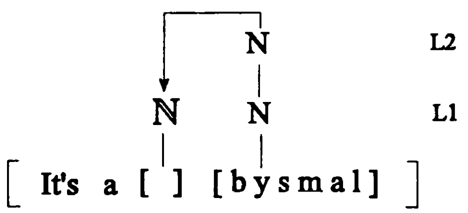
Another point in favour of empty nuclei in the account of downstep comes from the fact that their presence elegantly captures a fundamental feature of this phenomenon. In §1 of chapter 5, I claimed that a well-attested characteristic of downstep is that its levels are not restricted in number. Basically, there can be as many steps as the speaker succeeds in accommodating within his speaking range. This fact is reflected in phonological representation by allowing for an alternating pattern of onsets and empty categories.

Inherent in the acceptance of an unlimited number of empty nuclei in phonological representation is the risk that their occurrence may be abused in phonological analyses. The question of whether or not there is a principled restriction on their distribution will be answered throughout the development of the present chapter. Nevertheless, in this section, I shall demonstrate that there is, indeed, an array of constraints which prevent the empty nucleus from being licensed anywhere we wish in the representation. Let me pursue the discussion about empty nuclei and their distribution in phonological structure a little longer, before I turn to analyse licensing between onsets in the following section.

Let us focus on the constraint which prevents the occurrence of an empty nucleus after [*lamb*][10]. Notice that, in this hypothetical situation, N is not licensed, as the only possible licensor (which is N) occurs to its left. This is due to the fact that the directionality of licensing at this level goes from right to left. In other words, the empty nucleus would fall outside the scope of any type of licensing relation, rendering this structure ill-formed.

In view of the above circumstances relating to licensing, it is logically possible for the model to generate a structure in which the empty nucleus occurs to the **left** of a nucleus (as in (10) above) and in which there are no other constituents preceding the empty nucleus[11]. This structure appears as follows:

(11)[12]



The condition which allows a nucleus to remain empty in the location illustrated above (that is, to the left of a domain at L1, like [*bysmal*]), requires that some material be present to the left of the empty category. Note that this material need not be included in an L1 domain. Moreover, in view of the fact that it is possible to find the mirror image of this representation (see §5 and 5.1), it appears that the presence of that material counts as another factor which sanctions the presence of empty nuclei.

Let us now examine the type of data which are accounted for by this structure. Remember that in chapter 4, §4, I stated that the falling set of nuclear tones (the putative high fall and low fall nuclear tones) constitute a single phonological category, and that the

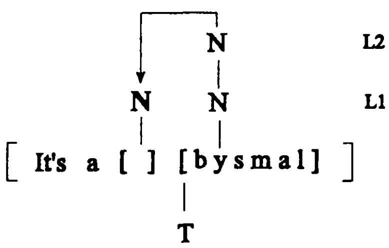
[10]But see §5.1, where I shall revise this condition.

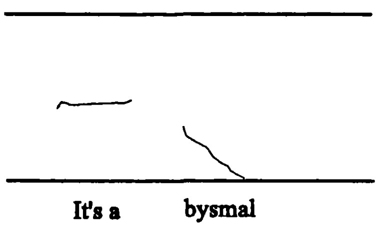
[11]In §4.1, I introduce further conditions on the occurrence of empty nuclei. Briefly, I shall claim that, minimally,one boundary of an empty nucleus must participate in a licensing relation with the boundary of another domain which is associated to T. This condition also allows for a straighforward representation of former high prehead followed by (low/high) rising nuclear tones.

[12]This structure remains incomplete (and as such it is ill-formed) until tone is associated.

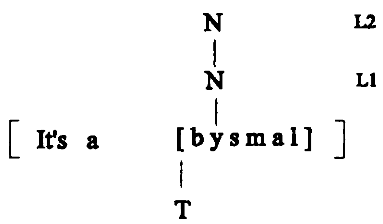
difference between them at the phonetic level arises as a consequence of the presence/absence of toneless boundaries to the left of the structure which represents a falling pattern. In this chapter, we are in a position to formalise this statement by claiming that these toneless boundaries correspond to the boundaries of empty nuclei. Thus, the phonological representations which I propose for this pair of contours do not involve distinctive phonological categories, but rather, the distribution of empty nuclei. I illustrate both representations and their respective stylised intonation contours below:

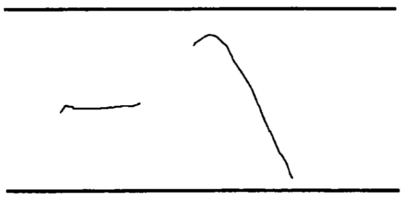
(12) 'low fall'





(13) 'high fall'





Notice that both structures show exactly the same tonal association: T is associated to the left boundary of [*bysmal*]. However, the difference between what was formerly treated as a low fall and a high fall is now captured in terms of the presence or absence, respectively, of an empty nucleus in their phonological representations. The presence of toneless boundaries in phonological representation is reflected phonetically by the extra force of GE pulling down the F0 contour (as in the first contour, but not in the second).

The use of empty nuclei in the above structures counts as independent and further evidence in support of their existence in phonological structure. More importantly, empty nuclei allow us to capture the generalisation that the contours in (1) and (12) are related, in the sense that they are both accounted for by a phonological structure which is characterised by the presence of toneless boundaries which belong to an empty nucleus. Thus, this highlights a further advantage of this model over previous proposals.

Finally, recall that in chapter 5, I claimed that current accounts of stepping tend to enrich phonological structure by allowing for bitonal pitch accents in their models. However, here, I have presented the case for differences in falling nuclear patterns providing additional support for the presence of empty nuclei in phonological representation, and thus, keeping the structure constrained.

2.2 Inter-onset licensing

Until now, I have established the following licensing relations between constituents in my model:

(14) (a) inter-nuclear licensing: N licenses other Ns (which are empty).

(b) inter-constituent licensing: N licenses O.

Following from these, and specifically in view of the fact that two constituents of the same type can participate in a licensing relation (as in (14)(a)), logically, we may raise the question of whether or not licensing is also found between onsets.

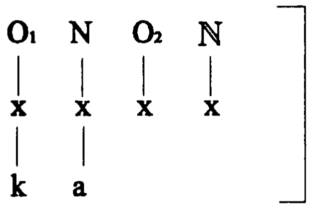
Evidence provided by the behaviour of some consonants in Japanese (Yoshida 1990), has led phonologists (Nasukawa, in preparation) to propose that, indeed, inter-onset licensing is a necessary notion in phonology to account for certain phenomena involving consonants; specifically, for gemination. For example, consider the output representation for the word *katta* ('win' + past suffix):

(15)

O_1	N	O_2	N		O_3	N
|	|	|	|		|	|
x	x	x	x		x	x
|	|	|			|	|
k	a	t			t	a

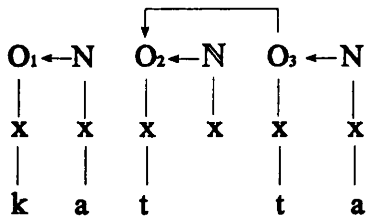
Note that O_2 and O_3 share the same melodic content 't', whereas O_1 does not. The observation that O_2 and O_3 each license the same type of melodic unit (together with the fact that this type of observation is widespread in other languages) leads Nasukawa to assume that this pair of onsets forms a domain in which there is gemination. On the other hand, O_1 falls outside this span, and therefore, is excluded from the effects of gemination. In other words, in the lexical representation of *kat*, the melodic identity of the rightmost consonant remains unspecified, as I show below (the square bracket indicates the right edge of this morpheme):

(16)



In order to account for gemination, Nasukawa proposes that there is inter-onset licensing between O_2 and O_3, and that the melodic content of the latter is determined by that of the former. However, O_1 does not participate in this licensing relation, and therefore, it is outside the scope of gemination. I illustrate this below:

(17)



One of the conditions for inter-onset licensing is that an **empty** nucleus occurs between the relevant onsets[13]. This follows from the well-established idea that each onset must be licensed by a nucleus.

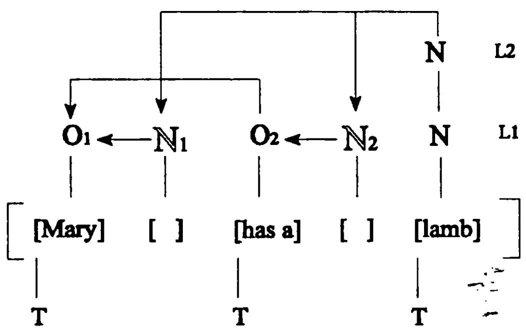
Having illustrated a proposal which accounts for gemination in Japanese, I now draw a parallel between this phenomenon and that of stepping in English. Specifically, I argue that the characteristic phenomenon of pitch-pattern copying in stepping terraces arises as a consequence of inter-onset licensing in phonological representation.

Let us turn to the structure in (10) in order to illustrate and examine the manifestation of inter-onset licensing in the field of intonation. For this, I reproduce (18) below, where I show tonal structure and inter-onset licensing.

---

[13]O_1 occurs to the left of a **filled** nucleus, and therefore, it falls outside the span of inter-onset licensing.

(18)



The directionality of licensing between onsets follows from phonological licensing at L1, which operates from right to left. Inter-onset licensing defines a *tone-harmony* domain between O_1 and O_2. In this domain, the tonal pattern of O_1 (which is the licensee) is determined by the tonal pattern of O_2 (which assumes the role of the licensor). In terms of the analogy with the case of Japanese, this corresponds to the events in the domain of gemination.

One of the advantages of this account of stepping over those presented in chapter 5 is that the copying phenomenon, or tone harmony, is straightforwardly accounted for by inter-onset licensing. In addition, this account identifies the source from which tonal patterns arise; this is accomplished through the directionality of inter-onset licensing. On the other hand, in previous models, researchers simply stipulate that the first pitch pattern is copied throughout the structure.

In view of the above tonal structure — in which the same pitch pattern is carried beyond the scope of inter-onset licensing into the domain of the filled nucleus — it may be predicted that another licensing relation could exist between the onsets and this nucleus. In fact, this prediction is correct only to the extent that N licenses O_1 and O_2 by virtue of being the ultimate licensor. However, this relation is too distant to have any immediate effect which could be seen as similar to that of tone harmony. Evidence which supports this claim is provided by the existence of well-attested pitch contours in which tone harmony ceases just before the last pitch pattern, that is, before the ultimate nucleus, (for example, the contour which is described in O'Connor and Arnold as an emphatic high head followed by a high fall).

2.3 Triggers and targets of stepping

An issue which remains unsolved from the previous chapter has to do with the identification of those units which count as a trigger or as a target of downstep. In chapter 5 I argued against the stipulation made by previous models that a unit to the **right** of the trigger undergoes downstep. Furthermore, I claimed that, in the context of their models — and after performing the necessary adjustments — it is equally possible to analyse this phenomenon in the following way: a unit to the **left** of the trigger undergoes upstep. Hence, we arrived at the conclusion that those models are suboptimal because they allow this indeterminacy.

In this chapter, we are in a position to offer a straightforward identification of targets and triggers of downstep: empty nuclei act as triggers, and those onsets to their right act as targets. This conclusion is reflected in the phonetic interpretation of toneless boundaries; note that GE pulls down the F0 contour, and that this has an effect on the interpretation of following tones. Further evidence in support of this claim is provided by the fact that the interpretation of the first onset in phonological structure lies beyond the scope of GE[14].

The observation that empty nuclei act as triggers, and onsets to their right as targets, appears to be reflected in phonological structure by the fact that such empty nuclei and onsets do not participate in a direct licensing relation. It must be borne in mind, though, that this is only an approximation to an account of the 'directionality' of this phenomenon, because there is evidence which casts doubt on the exact nature of the licensing relation. For example, in §2.1 above, [*bysmal*] undergoes downstep due to the presence of an empty nucleus to its left, and yet both constituents participate in an immediate licensing relation. In order to resolve this conflict, I must examine a wider range of data. I shall do so in §6.

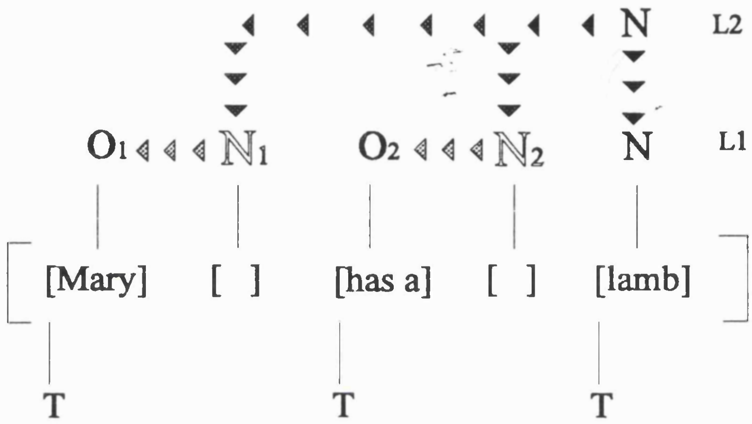
3 Licensing Inheritance

The main topic in this section relates to the manifestation of LI in those phonological structures which show empty nuclei, and to the effect of LI on the phonetic interpretation

[14]Note that this statement does not exclude the possibility of an initial pitch accent undergoing downstep. In fact, as I have already illustrated in the case of *It's abysmal* in §2.1 it is perfectly acceptable for an initial pitch accent to be downstepped.

of the structures as a whole. In (19) I illustrate the manifestation of LI in the above structure.

(19)



The ultimate licensor licenses both nuclei and onsets. However, notice that these constituents stand at different levels in the licensing path (shown by the arrows). Nuclei are immediate licensors, whilst onsets are licensed through licensed nuclei. Consequently, the amount of licensing potential stored by the former units is smaller than that stored by the latter. This differential in licensing potential, graphically illustrated by the fading arrows, is also manifested in phonetic interpretation.

Given that, as I claimed in chapter 4, §7 and 8, pitch specifications of units nearer to the ultimate licensor are enhanced and, by comparison, those farther away are attenuated, we expect the pitch specifications of $\mathbb{N}_1$, $\mathbb{N}_2$, and N to be more radical than those of the onsets. Since $\mathbb{N}_1$ and $\mathbb{N}_2$ show no melodic content, whereas N does, it is impossible to confirm or disconfirm this expectation.

Nevertheless, it is possible to analyse the implications of this claim in the case of N and onsets, since they do show the same melodic content. As the F0 contour of this structure illustrated in (1) shows, the above expectations do not correspond to what is actually observed. The pitch specifications of N are far from being enhanced. They are better described as attenuated.

In order to find an explanation for this situation, I remind the reader of the fact that this effect was absent from the structures illustrated in §7 and 8 from chapter 4, in which empty nuclei had not been introduced. This observation suggests that the reason for the manifestation of the phenomenon described above has to do with the presence of empty nuclei. In other words, the conditions have now changed. Thus, we may assume that the pulling down interpretation of a set of toneless boundaries prevents the interpretation of N from being the same as if they were absent.

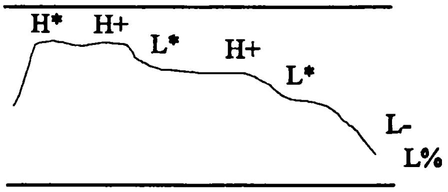
So far in this chapter, I have presented an alternative account of downstep based on the model which I introduced in chapter 4. For these purposes, I have further developed the model's original assumptions: by introducing the idea of empty nuclei, I have produced a coherent account of the above phenomenon. As an illustration of this, I have based most of the examples on the stepping down sequence of falling pitch accents over *Mary @ has a @ lamb*, but I have not illustrated sequences of other pitch accents[15]. As I stated in the previous chapter, the manifestation of downstep is also present in other contours, such as early peak and calling contours. I devote the rest of this chapter to a discussion and illustration of the aforementioned contours.

4 Stepping contours

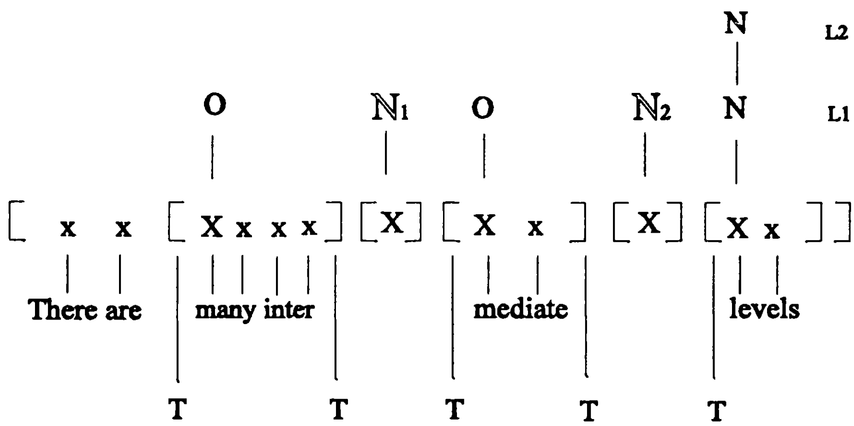
I begin this section by illustrating the phonological representation for the contours in chapter 5, §2.1.1, which I reproduce below:

[15]In P'80, a sequence of (H*+L-) pitch accents in a stepping fashion is different from the sequence (L-+H*) under the same conditions (see chapter 5, §2.1.1 for their corresponding stylised F0 contours). However, in my model, they are accounted for by the same phonological structure.

(20)



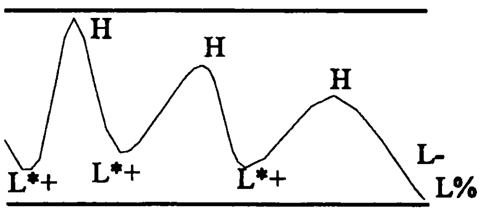
(20')



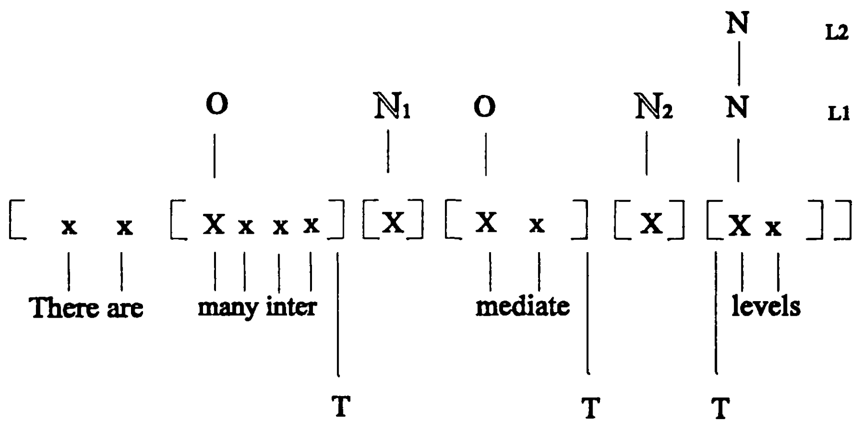
The phonological structure for *There are many inter- @ mediate @ levels* differs from that for *Mary @ has a @ lamb* only in the fact that the former shows T associated to both boundaries of its onsets, whereas the latter shows T associated to their left boundary. Otherwise, both metrical structures are identical. With regard to the tonal distribution in the nucleus of the structure in (20'), notice that T is associated to the left boundary exclusively, and not — as we may have expected — to both boundaries. Recall that T is banned from appearing on the right boundary of the nucleus constituent, since that boundary counts as a licensee and does not inherit enough licensing potential to bear tone.

Another stepping down contour and its proposed phonological structure are illustrated below. As can be seen in (21'), T is associated to the right boundary of its onsets, and to the left boundary of its nucleus.

(21)



(21')



At first glance, the tonal distribution in the nucleus might seem somewhat controversial. However, as I clarify below, I have already presented evidence in support of this representation in chapter 4. Nevertheless, first let us examine what might constitute a challenge to this proposal.

Since *levels* is analysed as bearing (L*+-H) in Pierrehumbert's account, it might be expected that, in my account, the left boundary in the nucleus would remain toneless, and that the right boundary would bear T. As I claimed above in relation to (20'), the latter point is ruled out on the grounds of its incompatibility with LI. In relation to the former point, this would be acceptable only if there was T associated to L2's right boundary, so that the rising movement could be accounted for. That is, it would have to look as follows:

(22)

$$\begin{array}{ccc} & & T \\ & & | \\ [&]_{L1} &]_{L2} \end{array}$$

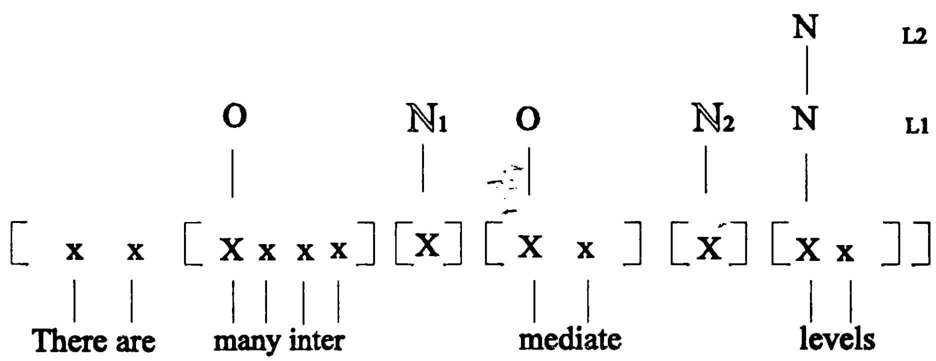
However, although (22) fully accounts for the rising movement, it results in a less than optimal representation. This is due to the fact that it is not descriptively adequate. As illustrated in the above F0 contour, it is unquestionable that, after rising, the F0 falls at the end of the contour. This is in sharp contrast to the information provided by the structure in (22), which shows, simply, rising pitch. Obviously, this structure does not account for the falling part of this contour. Given these circumstances, the safest assumption to make is that L2 must remain toneless in phonological representation, in order to capture the fall in pitch. In line with this thinking, we are immediately forced to reject (22) as part of an alternative phonological representation to account for the last pitch movement of the F0 contour in (21), and to accept the association of T to the left boundary of *levels*. Hence, we are left with the structure in (21') as the most appropriate.

An additional argument in support of this structure is found in chapter 4, §(4). There, I rejected the representation of the rise-fall nuclear tone — which would have had the same phonological structure as the one illustrated in (17). Furthermore, following other intonologists, I claimed that the rising movement preceding the falling movement in the relevant F0 contours was a matter related to phonetic interpolation. After that, I stated that these contours are accounted for in phonological representation by the same structure as that required for falling movements; that is, (21'). Thus, in the context of this discussion, I conclude that the latter constitutes the phonological structure which accounts for the contour in (21).

4.1 Empty nuclei and licensing between boundaries

In view of the distribution of tone in the above phonological structures, we must confront the question of whether or not onset boundaries may remain completely toneless. In other words, given the present state of the model, it is logically possible to predict the generation of the following structure:

(23)*



However, the interpretation of this structure — a contour characterised by continuous low pitch with dips — is unusual. Hence, we would wish the grammar to disallow the generation of such contours. Besides, the effect that empty nuclei have on following constituents which show toneless boundaries, although intuitively predictable — to pull down their interpretation in the frequency range — is, nevertheless, rather bizarre. The reason for this claim is that, if a unit is already interpreted as low pitch, no matter what factor may force it to be interpreted yet lower in frequency, there is no available space in the lower frequency range for it to be interpreted. Given this observation, it follows that, ideally, structures like this one should not be generated by the model.

When models *à la* Pierrehumbert encounter the latter situation — in their terms, a bitonal pitch accent can potentially trigger downstep on a following L — they **stipulate** that L tones do not undergo downstep, and thus, they safely use these structures to account for other pitch contours. Although this solution may be valid, note that it is based on a stipulation, and that the offending structure can still be generated in their models. In contrast, as far as the present model is concerned, I argue that the structure in (23) is rejected by imposing further constraints in the distribution of empty nuclei. Let us now examine which these are.

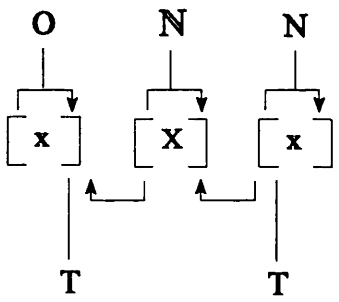
In chapter 4, I claimed that boundaries of prosodic domains participate in licensing relations, and that their licensing capacity controlled the distribution of tones. In this chapter the same principles apply, but additionally, they constitute another constraint on the occurrence of empty nuclei. This constraint is understood as follows: toneless boundaries of empty nuclei must license, or be licensed by, another boundary which bears tone. Let us formulate this constraint in terms of another licensing principle:

(24) **Empty nuclei and boundary licensing**

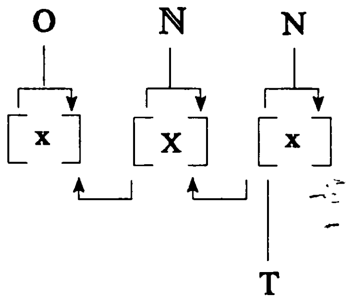
One of the boundaries of empty nuclei must participate in a licensing relation with another boundary which itself licenses T.

Given this condition, the structure in (23) above is rejected, since all the boundaries of empty nuclei violate this condition. Nevertheless, the well-formed structures which can be generated by the model are as follows:

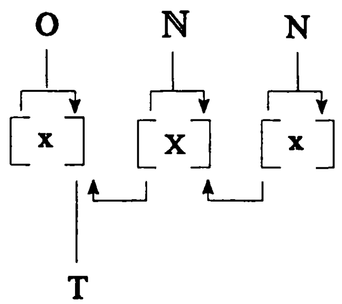
(25) (a) both nuclear boundaries participate in a licensing relation in which other boundaries license T.



(b) the nuclear **right** boundary participates in a licensing relation in which its licensor bears T.



(c) the nuclear **left** boundary participates in a licensing relation in which its licensor bears T[16].

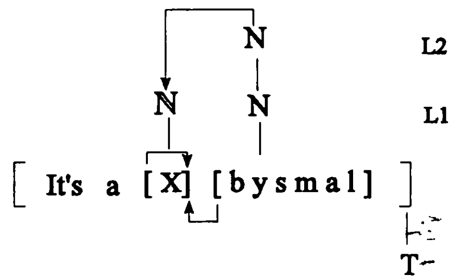


The solution I have adopted here not only excludes structures such as (23) from the model, but also allows us to impose more restrictions on the occurrence of empty nuclei.

The topic related to licensing relations between boundaries of empty nuclei and other constituents, leads us back to examine the occurrence of empty nuclei in initial position, in order to analyse what constitutes well-formed structures. In §2.1 I have stated that an empty nucleus may occur to the left of a nucleus whose left boundary is associated to T. In this position, this structure is sanctioned by empty-nuclei-and-boundary licensing. By the same token, notice that the following ill-formed structure results:

[16]In §7.1 I shall treat this structure as ill-formed, since the presence or the absence of [ ] yields two phonological structures which show the same phonetic interpretation, and therefore, are redundant.

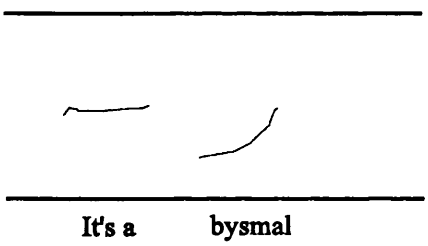
(26)*



As the arrows indicate, the right boundary of [*bysmal*] licenses the left boundary of [ ]. However, this structure is ill-formed because neither boundary of the empty nucleus performs the task of participating in a licensing relation with another boundary which itself licenses T. For this reason, I conclude that an empty nucleus is banned from appearing in this position, and thus, (26) results in being ill-formed.

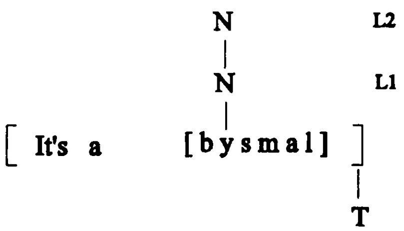
This situation has a direct bearing on the question of how to account for contours such as the following:

(27)



that is, a rising movement preceded by relatively mid pitch. The structure I propose for this type of contour is as follows:

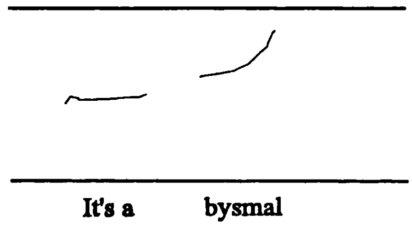
(27')



The only possible way to capture a rising movement over the last domain at L1 in this model is by the association of T to L2's right boundary. In addition, observe that L1's toneless left boundary already accounts for the drop in pitch. In other words, the step manifested in the F0 contour is predicted by the presence of precisely this toneless boundary.

However, it may be argued that there are other pitch contours in which the starting point of the rising movement is higher than the one illustrated above, and thus, that there is no manifestation of a step down in its phonetic interpretation. For instance, consider the following contour:

(28)



In view of this contour, the validity of the above representation in accounting for contours such as this is questioned.

A possible account of the difference between these contours in phonological representation might be that, in the former structure, an empty nucleus is the factor which pulls down the interpretation of the nucleus to its right (this is the structure in (26) above), whereas, in the latter structure, empty nuclei are absent, and consequently, this effect is absent, too. However, recall that this structure has already been rejected above by the principle of empty-nuclei-and-boundary licensing. Given this situation, the present state of the model forces us to treat those particular contours as *allotonic* interpretations of the same phonological structure[17].

---

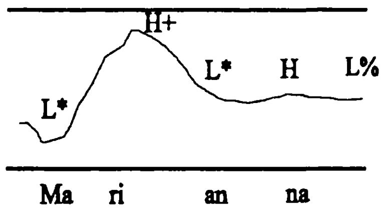
[17] I follow Gussenhoven (1983) on the representation of rising movements in nuclear position. However, for example Ladd (1983) maintains that the above contours are in contrast. If this statement proves to be true, the present state of the model may have to be revised with respect to this particular point.

I conclude the present discussion of this type of stepping contour by claiming that the proposed account has an advantage over that of Grice'95 and P'80[18] in as much as it shows a principled way of accounting for tonal harmony by means of inter-onset licensing. In addition, it shows an improvement on that of P'80, in the sense that the phonological identity of pitch patterns at each new level remains unaltered with respect to identical patterns found in a non-stepping structure[19]. Also, it avoids the tendency present in both Grice'92 and Ladd (1990) to elaborate tonal structure in order to be able to capture the downstep phenomenon. On the other hand, this is achieved in the present model, by having recourse to the idea of empty nuclei, while maintaining the simplest tonal structure. Finally, in the event of overgeneration, this model has recourse to general principles which exclude offending structures from being generated by its grammar. However, Pierrehumbert and her followers resort to a stipulation for the same purposes.

5 Calling contours

Below, I include the F0 stylised contour for the utterance *Marianna* with two pitch accents.

(29)

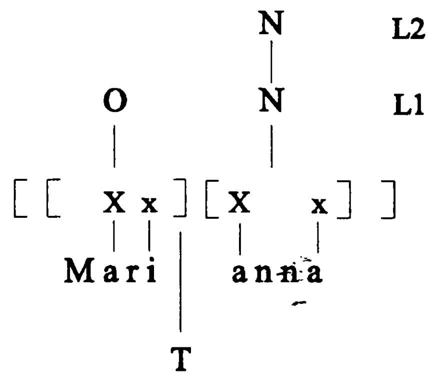


The phonological representation which accounts for this contour is as follows:

[18]See chapter 5, §2.1.2.2.

[19]See chapter 5, §2.2.

(29')[20]



T is associated to the right boundary of the onset. This is interpreted as a rising pitch movement. The interpretation of the toneless boundaries of the nucleus are responsible for pulling down the F0 contour, and consequently, given that Ts are absent from the rest of the structure, pitch remains level[21].

In the context of the present discussion, in which the main purpose is to produce a **unified** account of the phenomenon of stepping, it should not surprise the reader to encounter the above phonological structure in which a toneless boundary accounts for the step down. Recall that one of the fundamental characteristics of this model is that **toneless boundaries** (whichever domain they may belong to) encode the phonological information which is interpreted as pitch being pulled downwards by GE. In fact, this is the information conveyed by the left boundary of [*anna*]. An alternative representation, which may be seen as more in line with the representations presented thus far (with an empty nucleus just before [*anna*]), is rejected on the grounds that the presence of the empty nucleus would achieve the same result as the representation illustrated above, and therefore, would be deemed redundant. In view of the desire to avoid redundancy in phonological structures, the distribution of empty nuclei requires further revision. I shall return to this topic in §7 below, for a more thorough discussion of this matter.

So far, the present model improves on other models, in the sense that it avoids the shortcomings inherent in earlier approaches. For example, I claimed that the

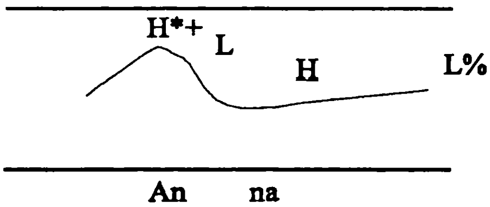
[20]Licensing relations in the phonological representations included in this section have already been discussed. Therefore, I shall not show them here, unless a point directly related to them is raised.

[21]See note 17 above.

interpretation of a toneless boundary remains unchanged, whereas in Pierrehumbert's account, L shows two completely different interpretations. As another example, in this model the trigger and the target of downstep are clearly defined, whereas in previous accounts, the target can be present either inside or outside the trigger (see discussion in §2.1.2). Unfortunately, this allows for some degree of indeterminacy.

Let us now concentrate on a similar type of calling contour, but this time mapped onto a phrase like *An @ na*, which was included in chapter 5. The F0 contour is as follows:

(30)



Note that we arrive at the following difference in metrical structure between *An @ na* and *Marianna*: in *An @ na*, there is only one licensor at L1, whereas in the case of *Marianna*, there are two licensors at L1. For the purposes of clarity, I illustrate the metrical structure of *An @ na* below:

(31)

```
[     x              ]  L2
      |
   [  x   ] [x] x       L1
      |         |
      An        na
```

The position which licenses *An-* constitutes the ultimate licensor, which also licenses the position which licenses *-na*. It must be emphasized that the above representation counts only as a preliminary version of an account of stepping manifested in this type of contour because, as I shall discuss shortly, it raises some questions as to a possible violation of the conditions which have already been imposed on the occurrence of empty nuclei. In addition, it introduces further problems for the conditions controlling the association of tone to the right boundary of the nucleus.

In order to put forward a complete version of the structure in (31) it is first necessary to address the following issues: on the one hand, what constitutes its melodic structure, and on the other, how this is organised into constituents. In the following paragraphs I discuss each of these in turn.

Recall that, in chapter 3 we encountered this type of contour, where I discussed the topic in relation to the treatment of spreading (the propagation of a specific tonal specification). There, I argued that spreading is accounted for by the fact that, in phonological representation, both boundaries of a given domain show the same tonal specifications or melodic content. This is manifested in phonetic interpretation as level pitch. Given this, and in view of the relatively high pitch in the F0 contour above, we predict that both boundaries of the domain *An-*, are associated to T:

(32)

```
[      X             ]  L2
       |
  [    X    ] [X] x     L1
  |    |    |     |
  |   An    |     na
  |         |
  T         T
```

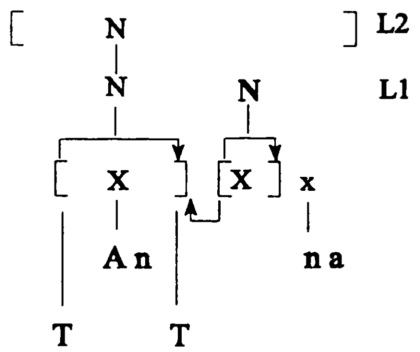
However, as the reader has doubtless anticipated, this proposal may encounter some difficulties. Namely, if the domain of *An-* constitutes the nucleus (as is implied in the above structure), then T cannot be associated to its right boundary. This is due to the fact that this licensee-boundary does not inherit enough licensing potential from its licensor-boundary for it to license any melodic content[22]. In order to propose a solution to this shortcoming, let us examine the constituent structure of the above representation.

In the light of the present framework, and more specifically, in view of the account of stepping presented so far, we would predict that the constituent structure which

[22] See chapter 4, §5.

accounts for this pattern consists of, minimally, two nuclei: an obligatory (filled) nucleus, as symptomatic of focus, and an empty nucleus, as a trigger for downstep. Straightforwardly, the former is identified as corresponding to the position which licenses *An-*, and the latter, to the position which licenses@. Initially, the outcome of this prediction is highly desirable, since it overcomes the problem of gathering enough licensing power on the right boundary of the nucleus for it to license T. This is achieved by the fact that the left boundary of the empty nucleus counts as an extra source of licensing potential which, added to the licensing potential inherited from the left boundary of the nucleus, renders the licensing of T possible. I illustrate this below:

(33)



This structure constitutes a solution to one of the problems which I raised above, specifically with regard to the danger of associating T to the right boundary of the nucleus. Note that, prior to this analysis, this boundary did not inherit enough licensing potential to license T. Nevertheless, now, the empty nucleus provides a solution to this problem. However, the other shortcoming, related to a possible violation of the conditions imposed on the occurrence of empty nuclei, remains unsolved. It is to this issue that I now turn.

In §2.1 I rejected the possibility of allowing an empty nucleus to the right of the nucleus which counts as the ultimate licensor, on the grounds that, in that position, the nucleus would not be licensed, and therefore, would not be allowed to remain empty. However, in the case under discussion, at least descriptive adequacy would be achieved, if the rightmost nucleus were allowed to remain empty. In addition, its left boundary would act as one of the sources of licensing power which ultimately licenses T on the boundary of the constituent to its left. In order to solve this conflict, we must search for a strategy which will succeed in the licensing of the rightmost nucleus; for this, I turn once again to GP and its application at the segmental level.

5.1 Final empty nuclei

In GP, the fact that, in a substantial number of the world's languages a string of prosodic constituents may end in an empty Nucleus, is well-established. Take, for instance, English, where a word like <pat> is analysed as follows:

(34)

```
O   N   O   N
|   |   |   |
x   x   x   x
|   |   |
p   æ   t
```

Note that the last audible segment in this word is consonantal, and thus licensed by an Onset position. However, as I have just stated, English is parametrically defined as a language whose prosodic structure allows for final Nuclei to remain empty. This parameter is stated as follows:

(35) **Final-empty-nucleus parameter**[23]

Final empty nucleus licensed? ON/OFF

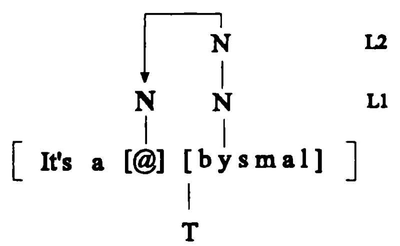
In the case of English, this parameter is turned [ON], and it sanctions the presence of the final empty Nucleus. This renders the above phonological representation of <pat> well-formed. The other logical possibility predicted by the Final-empty-nucleus parameter is that it may be set at [OFF]. In those circumstances, we expect to find languages in which

[23]There is also theory-internal motivation for the presence of an empty nucleus in this position. See Harris (1994) for a detailed discussion of this topic in relation to English.

the final Nucleus is not licensed to remain empty. According to Harris (1994), Luo is one of the languages which falls into this pattern.

Let us now return to the domain of intonation. As a preliminary representation of downstep in this type of phrase, I informally suggest that the final nucleus is licensed to remain empty, by virtue of being at the right edge of L2[24]. I have already illustrated this in (33). If this idea were to be pursued further, one of its advantages would be that it allows us to capture the generalisation that the phenomenon of stepping can manifest itself at both edges of the intonation phrase. Note that the aforementioned structure constitutes a mirror image of the events that are found at the left edge of the intonation phrase, which I reproduce below:

(12')



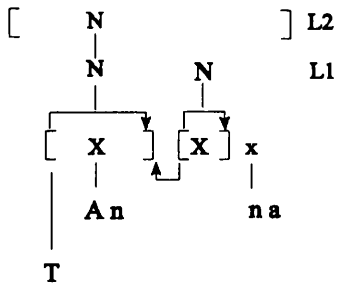
In §2.1 I stated that one of the conditions for an empty nucleus to occur (for example, in the above location) is that there must be at least one position to its left. According to this condition, (33) counts as a well-formed structure, as it remains completely faithful to its mirror image.

In addition to this, the occurrence of the empty nucleus in phonological structure is subject to the conditions stated under the heading of *Empty nuclei and boundary licensing* (§4.1): one of the boundaries of an empty nucleus must participate in a licensing relation with another boundary which itself licenses T. As can be seen in both (33) and (12'), the minimal task of the empty nucleus is accomplished, and therefore, these structures are well-formed. By the same token, a structure in which the right boundary

[24]Note that *-na* does not count for intonational purposes. Hence, [ ] constitutes the rightmost domain at both L1 and L2. The same can be said of *It's a-* in the structure that follows, only that, in this occas ion, [ ] constitutes the leftmost domain at both L1 and L2.

of *An-* remains unassociated, would be immediately ruled out by the condition just mentioned. I illustrate this structure below, for the purposes of clarity:

(36)*



In short, after having established the idea that nuclei can be licensed to remain empty at the left edge of the intonation phrase, it follows that the same circumstances can be met at the right edge[25].

A question of immediate concern here is related to the formalisation of this idea (that is, of the licensing of empty nuclei) in intonation. Note that in GP, this is captured in terms of a parameter which offers a choice between two settings. I have already illustrated the case of English, which selects the [ON] setting. Nevertheless, if the analogy between events above and below the foot is to be developed further, we expect to find the Final-empty-nucleus parameter also to be set [OFF]. In the field of intonation, this entails that it is possible to find languages in which the final nucleus is not licensed to remain empty, and hence, to show some melodic content. Even if this is possible, the detailed characteristics of these languages fall beyond the scope of the present work, and therefore, I shall leave their analysis for future research. Nevertheless, in order to render feasible the proposal of an empty nucleus at the edges of the intonation domain, at the present stage of this investigation, I suggest the following informal statement:

[25]An antecedent to this idea is present in Takahashi (in preparation). There, Takahashi proposes the **Edge parameter** which constitutes a revised version of the **Final-empty-nucleus parameter** in that an empty nucleus may be licensed at the right and/or left edge of a domain. This allows for a high degree of generalisation across a wide range of phenomena such as stress assignment in Latin, Aranda and English, and the occurrence of domain-initial nucleus in Parisian French and Tokyo Japanese.

(37) In a sequence of two nuclei, the one nearest to the edge is always empty, unless it is the ultimate licensor of the domain.

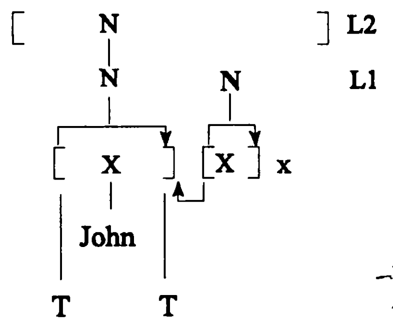
The main advantage of the above statement is that it allows us to make a generalisation about the occurrence of stepping in English in a straightforward manner, and also, across a wide range of data.

However, this solution must be adopted with caution, as it leaves many questions unanswered. For instance, the fact that the nucleus nearest to the edge can remain empty, needs further clarification. We could equally claim that the nucleus farthest from the edge can remain empty. It is not immediately clear whether or not this should be understood as an indication that an edge has the power to act as a source of licensing potential, and, if so, then, what the nature of this structural entity could be.

Another unsolved matter relates to the shape of the melodic content of this domain in those circumstances in which it is not licensed. In other words, what is the intonational equivalent to, for example, [ə], [ɯ], [e] (these are the phonetic shapes which unlicensed empty nuclei show in languages like English, Japanese and Spanish, respectively). This obviously constitutes a matter for further research in the future.

Let us now focus on another point which remains unsolved in the representation of calling contours. This relates to the phonological structure which accounts for phrases made up of a single syllable like, for example, *John* @. Note that the rightmost x position remains free of any melodic content, as I illustrate in the following representation:

(38)



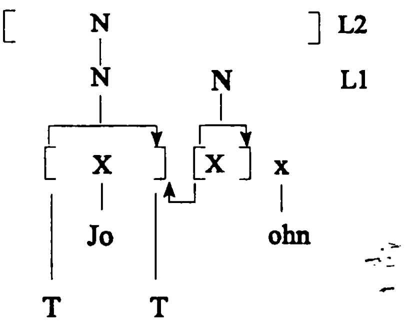
The situation in which timing positions can remain devoid of any melodic content is widespread in other areas of phonology[26], so it should not strike us as surprising to find that similar circumstances can manifest themselves in intonation. However, it is possible to foresee the difficulties that may result from such a representation, both in phonological representation and phonetic interpretation. I shall present a summary of these, since a detailed analysis of this type of utterance falls outside the scope of the present work.

As far as the former is concerned, it is not immediately obvious how to constrain the occurrence of a position devoid of melodic content. It seems appropriate to suggest that it occurs whenever the utterance is made up of a single syllable, but there is no phonological reason why this should be so. Furthermore, the fact that only one additional position occurs in the structure is completely arbitrary. Another point is that we may question whether or not this position is permanently devoid of content, and if not, then in which circumstances it can be filled.

With regard to the latter, the interpretation of this structure does not achieve descriptive adequacy. Note that, from the perception of this contour, the hearer receives the impression that the stepdown is encoded **within** the domain of *John*, and not, as our understanding of the above structure seems to suggest, at the **end**. This observation would lead us to claim that the last position is, in fact, filled with melodic content. This structure is represented as follows:

[26]One of the motivations for the existence of the timing tier in phonology is to capture the idea that the structural information and the melodic information of a phonological representation are completely independent, to the extent that even if a given position loses its melodic content, such a position can remain intact. See Goldsmith (1990) for examples of tonal stability after deletion of vocalic content.

(39)



The evaluation of these phonological representations will depend on the results of further investigation.

6 The directionality of downstep

In the discussion presented above, the target of a downstep occurs to the right of the trigger. In §2.3, I suggested that this phenomenon constitutes a reflection of particular circumstances found in phonological structure. More specifically, I claimed that an empty nucleus triggers downstep on an onset with which it does not share a direct licensing relation. However, given the evidence provided by phonological structures such as the following,

(40) (a) (19) in §2.3, in which the target shares an **indirect** licensing relation with the trigger,

(b) (12') in §5.1, in which the target shares a **direct** licensing relation with the trigger, and

(c) (33) in the previous section, in which the target and the trigger **do not share** an inter-constituent licensing relation,

it becomes obvious that the former suggestion ought to be revised. From the aforementioned structures, we observe that, independently from the distance between the target and the trigger participating in an inter-constituent licensing relation, it is always the unit to the right of the trigger which constitutes the target of downstep. Given this observation, I claim that there is no longer conclusive evidence to support the proposal that inter-constituent licensing counts as the phonological factor which determines the target of downstep.

Note that I do defend the idea of defining the phonological circumstances which control the phenomenon under discussion, but, due to the fact that I concentrate on data mainly from English, we only consider what counts as a partial view of the entire intonational spectrum. As a consequence, it might be the case that we have missed some decisive facts. In order to exclude this possibility, and to identify such circumstances, further research into this, and other languages is necessary, and will be part of future work.

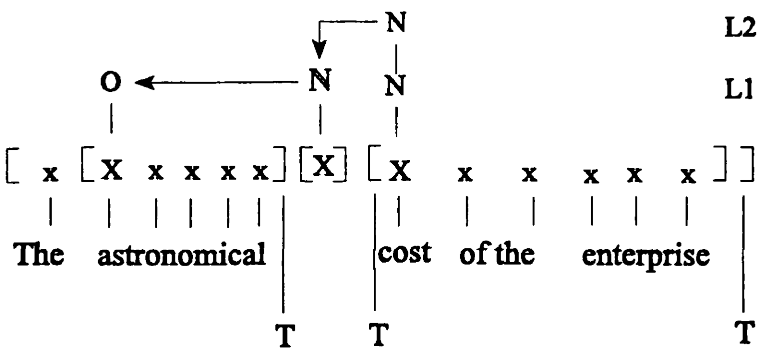
7 Early peak contours

The only contours which remain to be accounted for from the previous chapter, are those showing an early peak. For example, consider the following idealised pattern, which has been discussed at length by Grice'95:

(41) The AS Tr o no m i ca l COS T of the enterpri se

Recall that, in chapter 5, I illustrated how previous models of intonation accounted for these contours, and subsequently, how these accounts differed depending on the grammar of each model. For instance, B&P'86 account for the rising pitch on *-tronomical* and the step down on, and falling movement over *COST*, by means of a bitonal pitch accent with the starred tone on the right (H+L*). In contrast, the model in Grice'92 does not allow for a starred tone to occupy the position on the right in the core of a bitonal pitch accent; therefore, she accounts for that particular section of the contour like this: H—H* L (H* is the core unit in this pitch accent). After that, Grice'95 proposes the following combination of tones to account for the same section of the pattern under discussion: (H+L)+H*. The tones in parentheses account for the highest point in *-cal*, and the trigger for the step down on *COST* H* accounts for mid pitch on *COST*. This representation gains some advantage over the previous one in that, for example, the target and the trigger of the process are clearly differentiated, whereas before, this was not possible. Notice that the sequence H—H* constitutes both trigger — by virtue of being a bitonal pitch accent, and also target — by virtue of H* being the tone which undergoes downstep. The motivation for these changes resides in the need for models to remain constrained, and at the same time, to increase their capacity for generalisation. With the intention of achieving the same goal, I shall begin this section by presenting the phonological structure which I propose to account for the above contour. Following that, I shall discuss the phonological representations for the contours which remain unaddressed from the previous section.

(41')



In (41'), the ultimate licensor licenses its projection at L1 and the empty nucleus to its left. Subsequently, the empty nucleus licenses the onset to its left via phonological licensing. Although I have not illustrated boundary licensing in the above figure, it is understood that the structure abides by the principles of licensing which operate here, and therefore, is well-formed.

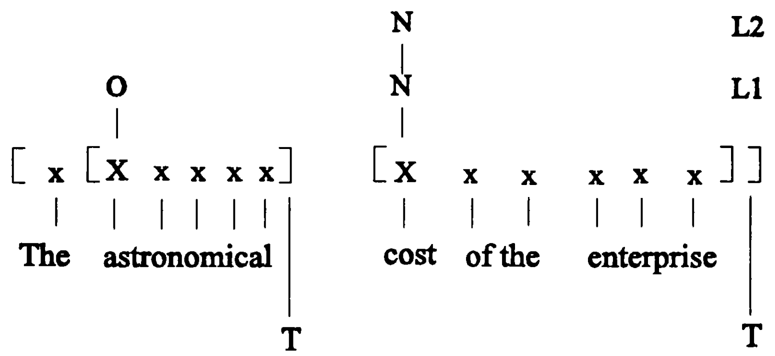
As far as the interpretation of this structure is concerned, it is straightforward: due to the absence of T in phonological representation, GE operates during the initial section of the contour, and its force gradually diminishes as the interpretation of the first T approaches. The toneless boundaries of the empty nucleus domain are interpreted as the switching off of the previous command to rise, and consequently, GE exercises its downpushing force. Inevitably, T to its right suffers the consequences of this force, and these are manifested in the fact that its interpretation results in a level lower than would have otherwise been expected, had there not been an empty nucleus present. After that, GE begins to apply its force over the last domain. This interacts with the presence of a toneless boundary to the right, followed by another boundary which bears T. In order to clarify the phonetic shape in this area, I remind the reader of one of the characteristics of GE, and also of the manifestation of LI in phonetic interpretation. Recall that in chapter 4, where I introduced the idea of GE developed by Cabrera-Abreu and Takahashi (1993), I argued that the longer the distance between two Ts, the greater the time for GE to exercise its force. This is precisely the situation found in *COST of the enterprise*, which, due to its greater length, allows for GE to apply its force, and therefore, the shape of the contour is manifested as a shallow dip. In addition, the right boundary of the domain which circumscribes *COST of the enterprise* counts as another pushing down factor, but this time weaker since, by LI (and the fact that it counts as a licensee), its tonal characteristics are attenuated. Finally, the association of T to the last boundary acts as the source of information which orders pitch to rise.

Another contour which illustrates an early peak is included below. However, this differs from the one in (41) in that pitch after the peak is rather low, as can be seen in the following example:

(42) The AS Tr o no m i ca l COST of the enterpri se

The phonological representation which I propose to account for this contour looks as follows:

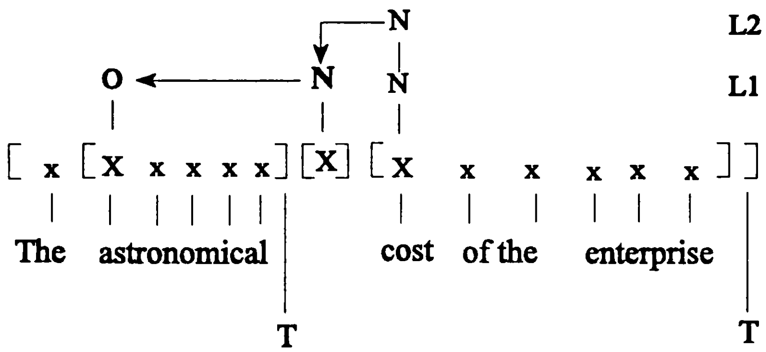
(42')



Phonologically, this structure differs from that in (41') only in the fact that, here, the empty nucleus is absent, and therefore, it is the presence of a toneless boundary that accounts for the low pitch in *COST*, rather than an empty nucleus.

A further reason for the different phonetic interpretations of both structures is found in the position that boundaries occupy in the licensing path. Note that, although there is a single toneless boundary present in (42'), it is nevertheless nearer to the ultimate licensor, and hence, its pitch specification is enhanced. On the other hand, the boundaries of the empty nucleus in (41') are farther away from the ultimate licensor; this renders the empty nucleus a weak position, and therefore, the pitch specifications of its boundaries are consequently attenuated, compared to the left boundary of *COST*. In view of these representations, and of the model hitherto presented, it is conceivable that an alternative well-formed phonological representation to (42') can be suggested, such as for instance, the representation below:

(43)*



The factor which differentiates the structure in (43) from that in (42') is that, in the former, an empty nucleus precedes the nucleus, whereas in the latter there is no empty nucleus. Obviously, this situation, in which two phonological structures can account for the same contour, is highly undesirable as it introduces problems of indeterminacy into the model, and it renders one of the structures redundant. Thus, the circumstances described here ought to be understood as a warning that the occurrence of empty nuclei in phonological structure is too powerful, and may have undesirable consequences.

7.1 More constraints on empty nuclei

In order to suggest a solution to this shortcoming, let us return to one of the conditions controlling the occurrence of empty nuclei in phonological representation. In §4.1, I introduced the constraint of *Empty nuclei and boundary licensing* which prevents the generation of ill-formed structures, such as a sequence consisting of an onset, an empty nucleus and a nucleus in which no boundary licenses T (see structure in (25)(c)) above. However, as we have just observed, this constraint is not restricted sufficiently to exclude the possibility of accounting for a given contour in more than one way (see also the case of *Marianna* in §5). Given this situation, it is obvious that this condition must be revised.

I suggest an alternative in the following terms:

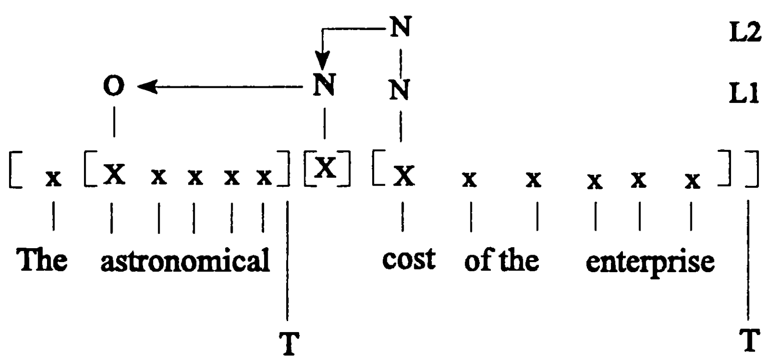
(44) **Empty nuclei and boundary licensing (revised)**

(a) one of the boundaries of the empty nucleus must participate in a licensing relation with another boundary which itself licenses T.

(b) the domain which licenses the empty nucleus must have, minimally, a boundary associated to T.

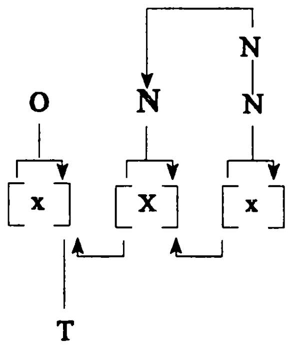
By (44)(b), we are now in a position to exclude the structures in (43) and (25)(c).

(45)*



The nucleus shows no T associated to its boundaries.

(25')(c)*



Given this proviso about the characteristics of the domain which licenses empty nuclei, we observe that it applies vacuously in the case of their occurrence at the right edge. This is due to the fact that, in this location, their licensor may constitute a structural unit other than a constituent — that is, an edge.

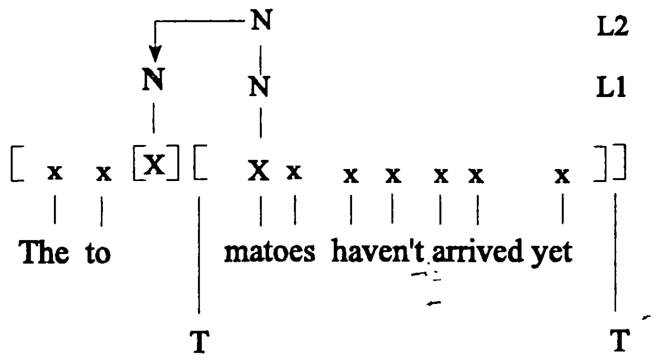
As a result of having encountered cases of redundancy in the model (those represented in the above structures, and the case of *Marianna*), the following point arises: it appears that toneless boundaries belonging to an empty nucleus may play a slightly different role from other toneless boundaries. In relation to this matter, it must be made clear that, although they all share the **same fundamental interpretation** of 'pulling down the F0 contour', those toneless boundaries which belong to an empty nucleus have the capacity to exercise an effect on the interpretation of following boundaries which are associated to T, whereas this capacity is absent from other boundaries. This may well be understood as an indication that the former show a global effect, whereas the latter show a local effect.

Let us now examine other contours from chapter 5, such as for example, the following,

(46) The to MA toes haven't arrived y e t

for which I propose this phonological representation:

(46)



In this structure, the interpretation of the empty nucleus is responsible for pulling down the interpretation of the following T. The resulting effect is a step down from *The to* to *matoes*.

As far as the account of all stepping contours is concerned, the present model boasts several advantages over previous models. These may be described as follows:

(47) (a) other models show two tones, and additionally, resort to further elaboration of tonal structure. This model shows a single tone and its tonal structure remains unchanged, and consequently, maximally restricted.

(b) Some accounts arbitrarily treat a bitonal pitch accent as the trigger for downstep, and thus establish a highly unnatural connection between them. This account treats toneless boundaries as the trigger for downstep, and thus succeeds in establishing a natural connection between them.

(c) Previous systems do not account for tone harmony, whereas the present system accounts for such phenomena by means of general principles of licensing.

This model is on a par with Grice'95 in that,

(48) (a) they can both account for the same type of contrasts;

(b) they capture the same type of generalisations: stepping can occur at both edges of the intonation domain;

(c) they clearly identify the target and the trigger of the phenomenon;

(d) they cannot explain why an item to the right of the trigger undergoes downstep.

8 Summary

In this chapter, I have examined the pivotal role played by toneless boundaries in the phonological representation of stepping contours, and how they participate in the organization of licensing relations in phonological structure. The account of this phenomenon by means of toneless boundaries, some of which belong to empty nuclei, allows us to establish a direct connection between the nature of the phenomenon and the factor which triggers its occurrence, without having recourse to the elaboration of tonal structure.

Licensing relations, which impose their effects on phonological structure, formally unify the account of stepping. They define the conditions under which empty nuclei may occur, and also offer a straightforward account of one of the basic features of stepping (*tone harmony*). Finally, their manifestation in phonological structure is directly reflected in their phonetic interpretation.

Chapter 7. Conclusion

I have tried to demonstrate, largely with reference to the intonation of English, that in terms of the principles of licensing, of a single tone, and of boundaries as TBUs, we can have a maximally restricted model of intonation, and yet we can account for well-attested F0 contours and also for widely attested intonational phenomena, like contour-tones, tone-spreading and stepping. Thus, the model I have developed here incorporates well-established ideas from Government Phonology, although an adequate account of stepping has necessitated some adjustments, especially to the conditions of locality associated with it.

I have revised earlier descriptions of intonation which account for intonation contours in terms of prehead, head and nucleus. The head and nucleus are formalised in terms of onset and nucleus. The idea underlying this terminology is that these are constituents which take part in a licensing relation, as in the analysis of non-tonal phenomena in the context of GP: the nucleus licenses the onset by inter-constituent licensing. Given this binary relation, there is little motivation for the notion 'prehead'. In the nuclear domain, the nuclear tones conventionally labelled fall, rise, and fall-rise can be generated. The possibility of generating a rise-fall is ruled out by boundary licensing and LI. In the onset domain, the model can also generate a high level configuration, in addition to allowing the same licensing relations to apply.

The definition of phonological structure in terms of constituents participating in licensing relations allows for a simple and general account of intonation contours.

Contour-tones are represented by assigning different tonal specifications to boundaries of prosodic domains. A falling movement is accounted for by the association of T to the left boundary, whilst the right boundary remains unassociated. A rising movement shows the opposite tonal association. Such tonal specifications are controlled by boundary licensing and LI, so that a licensor boundary possesses enough licensing potential to bear a privative tonal contrast, whilst a licensee boundary inherits a reduced amount of licensing potential, and hence, cannot support the presence of tone.

Whilst boundary licensing permits a principled association of tones, its definition temporarily prevents the model from generating rising movements in non-final position. This potential problem is solved by assuming the existence of 'coda-licensing', whereby, in the context of an onset followed by a nucleus, the leftmost boundary of the nucleus coda-licenses the rightmost boundary of the onset, and, as a consequence, transfers enough licensing potential for the rightmost boundary of the onset to bear presence/absence of tone.

Following P'80 and Pierrehumbert and Beckman (1988), I have also argued in favour of treating spreading as the interpretation of equal tonal specifications assigned to the TBUs of a prosodic domain. However the account I have presented differs from competing analyses in that the characteristic flat shape of the F0 contour arises as a consequence of Gravitation Effect in phonetic interpretation, rather than as a result of interpolation rules. The alternative I propose allows for a straightforward interpretation of phonological representations. In addition, by virtue of being a principle, not only does it have an application more general than that of rules, but also, avoids the arbitrariness inherent in the rule formalism.

This view stands in sharp contrast to earlier analyses which have recourse to distinctive features in phonological representation. Although these analyses achieve descriptive adequacy, they fail to present a non-arbitrary and restricted account of F0 contours, thus resulting in overpowerful models.

In relation to stepping, the account I have proposed here is characterised by its simplicity. I have avoided both the enrichment of tonal structure, a common feature of former models, and the introduction of tree structure. This has been achieved via a more efficient use of the resources already present in phonological structure. Specifically, I have exploited toneless boundaries of empty nuclei. I have also suggested that the tone-copying phenomenon frequently found in a sequence of stepping terraces correlates with inter-onset licensing in phonological structure, and that the repeated pattern in each onset is defined by the onset-licensor.

A fuller understanding of the fact that the target of downstep occurs to the right of the trigger would require further research into more languages which show this phenomenon. It remains to be discovered whether stepping is due to the absence of a licensing relation between target and trigger, or to some other phonological factor.

This account of stepping is highly beneficial for a restrictive model, since it also has the capacity for generalisation. Earlier models which had recourse to alternative means of accounting for high preheads plus low-falls on the one hand, and stepping on the other (represented by means of boundary tones versus bitonal pitch accents) are improved upon by the present model, which posits the occurrence of an empty nucleus in both cases.

The situation in which the empty nucleus occurs to the left of the nucleus is not an isolated case. In fact, it has a mirror image: an empty nucleus may also occur to the right of the nucleus. The occurrence of empty nuclei at both edges of the intonation domain allows us to put forward an elegant account of the behaviour of F0 contours at these locations. This is achieved by invoking an informal statement which derives from the final empty nucleus parameter. Such a statement ensures that, in a sequence of two nuclei, the one nearest to the edge is always empty, unless it is the ultimate licensor of the domain.

Although this proposal also accounts for calling contours, a formally complete account is left for future research. A better understanding of the encoding of empty nuclei in an account of stepping would require further research into its distribution. Although empty nuclei constitute lexical items which are imposed on other items, and this is constrained by licensing principles, it seems that such constraints are not restrictive enough to either sanction or rule out representations such as (38) and/or (39) in chapter 6.

This situation may be related, to some extent, to the loosening of the locality condition. Even though the model elegantly accounts for a series of stepping terraces, this is achieved at the expense of allowing a number of empty nuclei, all of which are licensed by a single licensor.

References

Anderson, J. and Ewen, C. (1987). *Principles of Dependency Phonology.* Cambridge: Cambridge University Press.

Anderson, S. (1985). *Phonology in the Twentieth Century.* Chicago: The University of Chicago Press, Ltd..

Archangeli, D. (1984). *Underspecification in Yawelmani Phonology and Morphology.* PhD dissertation: MIT.

Beckman, M. (1986). *Stress and Non-stress Accent.* (Netherlands Phonetic Archives 7.) Dordrecht: Foris publications.

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

Bolinger, D. (1958). A theory of pitch accent in English. *Word* 14, 109-149.

Cabrera-Abreu, M. (1994). An alternative model for intonation applied to English. *Speech, Hearing and Language; work in progress. UCL.* 8. 97-119.

Cabrera-Abreu, M. (1995). Tone-association in English. *UCL Working Papers in Linguistics.* 7. 437-462.

Cabrera-Abreu, M. and Takahashi, T. (1993). A single tone representation of intonation. *Speech, Hearing and Language; work in progress. UCL.* 7. 55-62.

Carr, P. (1993). *Phonology.* London: McMillan.

Carter, H. (1973). Tonal data in Comparative Bantu. *African Language Studies.* 14. 36-52.

Charette, M. (1991). *Conditions on Phonological Government.* Cambridge: Cambridge University Press.

Chomsky, N. (1981). *Lectures on Government and Binding.* Drodrecht: Foris.

Chomsky, N. (1986). *Knowledge of Language: its nature, origin and use.* New York: Praeger.

Chomsky, N and Halle, M. (1968). *The Sound Pattern of English.* New York: Harper and Row.

Clements, G. (1979). The description of terraced-level tone languages. *Language 55,3.* 536-558.

Clements, G. (1983). The hierarchical representation of tone features. In Dihoff, I. *Current Approaches to African Linguistics, Vol 1.* Dordrecht: Foris publications.

Clements, G. (1990). The status of register in intonation theory: comments on the papers by Ladd and by Inkelas and Leben. In Kingston, J. and Beckman, M. (eds). *Papers in Laboratory Phonology I.* Cambridge: Cambridge University Press.

Clements, G. and Goldsmith, J. (1984). *Autosegmental Studies in Bantu Tone.* Dordrecht: Foris publications.

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

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

Dikken, M. den and Hulst, H. van der (1988). Segmental hierarchitecture. In Hulst, H. van der and Smith, N. (eds.) *Features, Segmental Structure and Harmony Processes, Part I.* Foris: Dordrecht.

Durand, J. (1990). *Generative and Non-Linear Phonology.* Harlow, Essex: Longman.

Firth, J. (1937). The structure of the Chinese monosyllable in a Hunanese Dialect. *Bulletin of the School of Oriental Studies.* 8. 1055-1074.

García-Lecumberri, ML. (1995). *Intonational Signalling of Information Structure in English and Spanish.* PhD dissertation. University of London.

Goldsmith, J. (1976). *Autosegmental Phonology.* Inidana University Linguistics Club. Published by Garland Press, (1979).

Goldsmith, J. (1988). Prosodic trends in the Bantu Languages. In Hulst, H. van der and Smith, N. (eds.) *Autosegmental Studies on Pitch Accent.* Foris: Dordrecht.

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, University College London, University of London. Published by Tübingen: Niemeyer, (1995).

Grice, M. (1995). Leading tones and downstep in English. *Phonology 12.* 2. 183-234.

Gussenhoven, C. (1983). Focus, mode and the nucleus. *Journal of Linguistics*. 19. 377-417.

Gussenhoven, C. (1984). *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris publications.

Halle, M. and Vergnaud, J. (1987). *An Essay on Stress*. Cambridge, Massachusetts: MIT press.

Harris, J. (1990). Segmental complexity and phonological government. *Phonology* 7.2. 255-300.

Harris, J. (1992). Licensing Inheritance. *UCL Working Papers in Linguistics* 4. 359-406.

Harris, J. (1994). *Sound Structure of English*. Oxford: Blackwell.

Harris, J. and Lindsey, G (1995). The elements of phonological representation. In Durand, J. and Katamba, F. (eds) *Frontiers of Phonology*. Essex: Longman. 34-79.

Hayes, B. (1995). *Metrical Stress Theory*. Chicago: Chicago University Press.

Hayes, B. and Lahiri, A. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory, 9*. 47-96.

Hirst, D. (1983). Structures and categories in prosodic represention s. In Cutler, A. and Ladd, R. (eds). *Prosody: models and measurements*. Berlin: Springer.

Hirst, D. (1988). Tonal units as constituents of prosodic structure: evidence from English and French intonation. In Hulst, H. van der and Smith, N. (eds). *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris publications.

Hirst, D., Di Cristo, A. and Nishinuma, Y. (1994). Prosodic parameters of French: a cross-language approach. Paper presented at the *II International Symposium of the National Language Research Institute* in Tokyo.

Hobbs, J. (1990). The Pierrehumbert-Hirschberg theory of intonational meaning made simple: comments on Pierrehumbert and Hirschberg. In *Plans and Intentions in Communication*. Cohen, P., Morgan, J. and Pollack, M. (eds) Cambridge, Massachusetts: MIT press.

Hogg, R. and McCully, C. (1987). *Metrical Phonology: a coursebook*. Cambridge: Cambridge University Press.

House, J. (1995). Intonational stereotype: a re-analysis. In Windsor-Lewis, J. (ed). *Studies of English and General Phonetics*. London. Routledge.

Huang, C. (1980). The metrical structure of terraced-level tones. In Jensen, J. (ed). *NELS* (Cahiers Linguistiques d'Ottawa vol. 9), Dept. of Linguistics, University of Ottawa 257-270; expanded as: On the autosegmental and metrical nature of tone terracing. In Goyvaerts, D. (ed). (1985). *African Linguistics*. Ghent: Story-Scientia. 209-238.

Hulst, H. van der (1989). Atoms of segmental structure: components, gestures and dependency. *Phonology 6*. 2. 253-284.

Hulst, H. van der and Snider, K. (1993) (eds). *The Phonology of Tone*. New York: Mou ton de Gruyter.

Hulst, H. van der and Smith, N. (1988) (eds). *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris publications.

Hyman, L. (1982). Globality and the accentual analysis of Luganda tone. *Journal of Linguistic Research* 2.3. 1-40.

Hyman, L. (1986). The representation of multiple tone heights. In Bogers, K. Hulst, H. van der and Hous, M. (eds). *The Phonological Representation of Suprasegmentals*. Dordrecht: Foris publications.

Hyman, L. (1992). Register tones and tonal geometry. In van der Hulst, H. and Snider, K. *The Phonology of Tone: the representation of tonal register*. Mouton de Gruyter: Berlin.

Hyman, L. and Ngunga, A. (1994). On the non-universality of tonal association 'conventions': evidence from Ciyao. *Phonology 11*, 1. 25-68.

Johnson, M. and Grice, M. (1990). The phonological status of stylised intonation contours. *Speech, Hearing and Language: work in progress. UCL*. 4. 229-256.

Kaye, J. (1990). 'Coda' licensing. *Phonology 7*, 2. 301-330.

Kaye, J. (1991/2). Do you believe in magic: the story of s+C sequences. *SOAS Working Papers in Linguistics and Phonetics*. 2. 293-314.

Kaye, J. (1993). Derivations and interfaces. *SOAS Working Papers in Linguistics and Phonetics*. 3. 90-126.

Kaye, J., Lowenstamm, J. and Vergnaud, J. (1990). Constituent structure and government in phonology. *Phonology* 7, 2. 193-232.

Kiparsky, P. (1982). Lexical morphology and phonology. In The Linguistic Society of Korea, (eds). *Linguistics in the Morning Calm.* Hanshin, Seoul.

Ladd, R. (1980). *The Structure of Intonational Meaning, Evidence from English.* Bloomington: Indiana University Press.

Ladd, R. (1981). On intonational universals. In Myers, T., Laver, J., and Anderson, J. (eds). *The Cognitive Representation of Speech.* North-Holland Publishing Company.

Ladd, R. (1983). Phonological features of intonational peaks. *Language 59. 4.* 721-759.

Ladd, R. (1988). Declination "reset" and the hierarchical organisation of utterances. *Journal of the Acoustical Society of America.* 84. 530-544.

Ladd, R. (1990). Metrical representation of pitch register. In Kingston, J. and Beckman, M. (eds). *Papers in Laboratory Phonology I.* Cambridge: Cambridge University Press.

Ladd, R. (1993). In defense of a metrical theory of intonational downstep. In Hulst, van der H. and Snider, K. (eds). *The Phonology of Tone: the representation of tonal register.* Mouton de Gruyter: Berlin.

Leben, W. (1971). Suprasegmental and segmental representation of tone. *Studies in African Linguistics,* Supplement 2. 183-200.

Leben, W. (1973). *Suprasegmental Phonology.* PhD dissertation, MIT.

Liberman, M. (1975). *The Intonation System of English.* PhD dissertation, MIT: Massachusetts. Reproduced by the Indiana University Linguistics Club, Bloomington, 1978.

Liberman, M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry.* 8.2. 249-336.

Lindsey, G. (1983). *Units of English Intonation.* MA dissertation, University of California, Los Angeles.

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.

Lindsey, G. and Harris, J. (1990). Phonetic interpretation in generative grammar. *UCL Working Papers in Linguistics* 2. 355-369.

Lombardi, L. (1990). The non-linear organisation of the affricate. *Natural Language and Linguistic Theory*. 8. 375-425.

Mohanan, K. (1982). *Lexical Phonology*. Indiana University Linguistics Club.

Mtenje, A. (1985). Arguments for an autosegmental analysis of Chichewa vowel harmony. *Lingua*. 66. 21-52.

Nasukawa, K. (in preparation). *Harmony, Nasality and Voice in a Constrained-based model*. PhD dissertation. University College London. University of London.

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, Massachussetts: MIT Press.

Pierrehumbert, J. and Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In *Plans and Intentions in Communication*. Cohen, P., Morgan, J. and Pollack, M. (eds.). Cambridge, Massachusetts: MIT press.

Pulleyblank, D. (1986). *Tone in Lexical Phonology*. Reidel: Dordrecht.

O'Connor, JD. and Arnold, GF. (1969, 1973). *Intonation of Colloquial English*. London, Longman.

Selkirk, E. (1984). *Phonology and Syntax: the relation between sound and structure*. Cambridge, Massachussetts: MIT Press.

Snider, K. (1988). Towards the representation of tone: a three-dimensional approach. In Hulst, H. van der and Smith, N. (eds). *Features, Segmental Structure and Harmony Processes. Part I*. Foris: Dordrecht.

Takahashi, T. (in preparation). *The Theory of Phonological Licensing and Elements*. PhD disssertation. University College London. University of London.

Trubetzkoy, N. (1969). Principles of phonology. Berkeley: University of California Press. (Translation of *Grundzüge der Phonologie, Travaux du cercle linguistique de Prague* 7, 1939.).

Vanderslice, R. and Ladefoged, P. (1972). Binary suprasegmental features and transformational word-accentuation rules. *Language*, 48. 819-838.

Yip, M. (1980). *The Tonal Phonology of Chinese.* PhD dissertation, MIT.

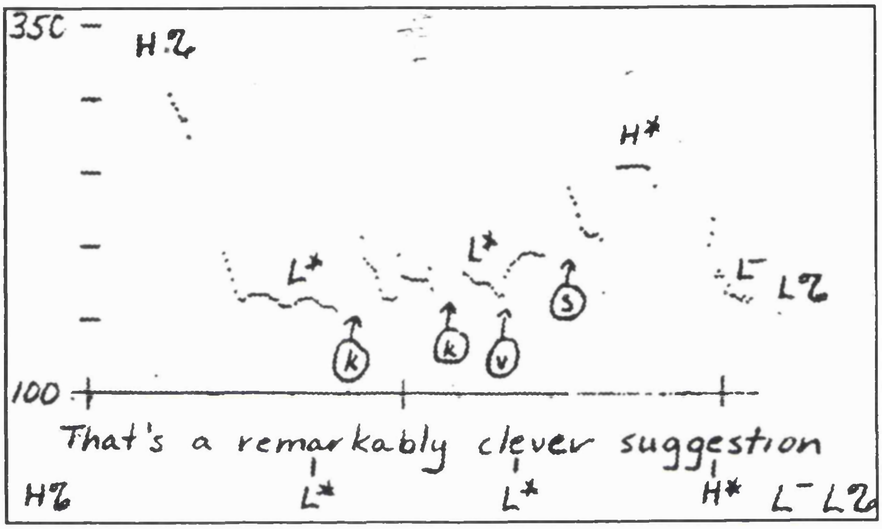
Yip, M. (1989). Contour tones. *Phonology 6.* 149-174.

Yoshida, S. (1990). *Some Aspects of Governing Relations in Japanese Phonology.* PhD dissertation, SOAS, University of London.

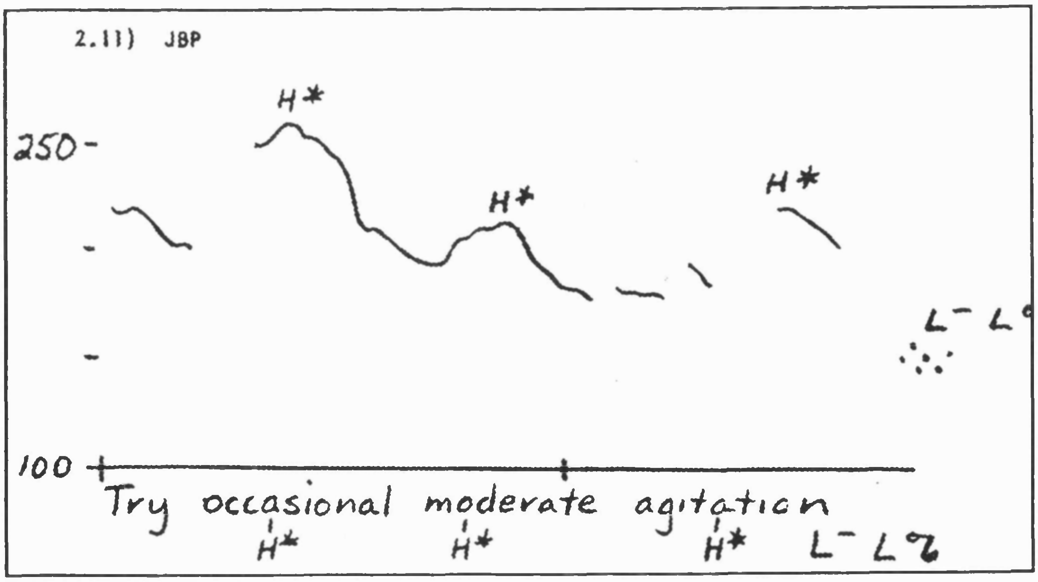
Yoshida, Y. (1991/1992) Pitch accent and tone harmony in Japanese. *SOAS Working Papers in Linguistics and Phonetics.* 2. 339-364.

Appendix to chapter 2, A valediction for L

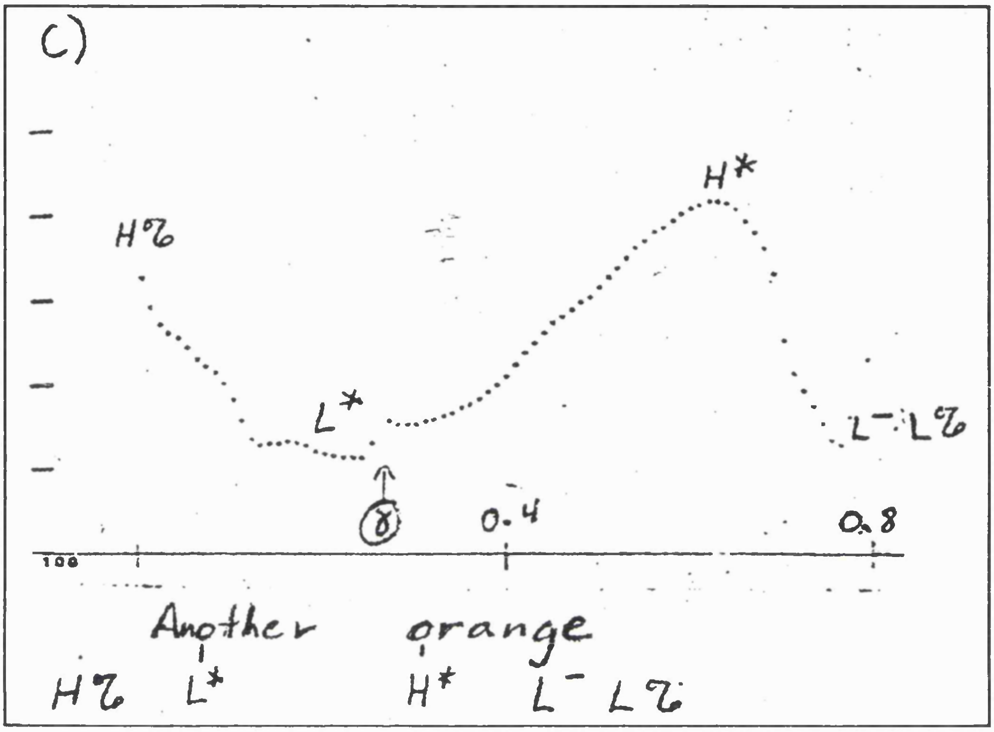
(1) 'That's a remarkably clever suggestion'. (Pierrehumbert 1980: 172).



(2) 'Try occasional moderate agitation'. (Pierrehumbert 1980: 172).

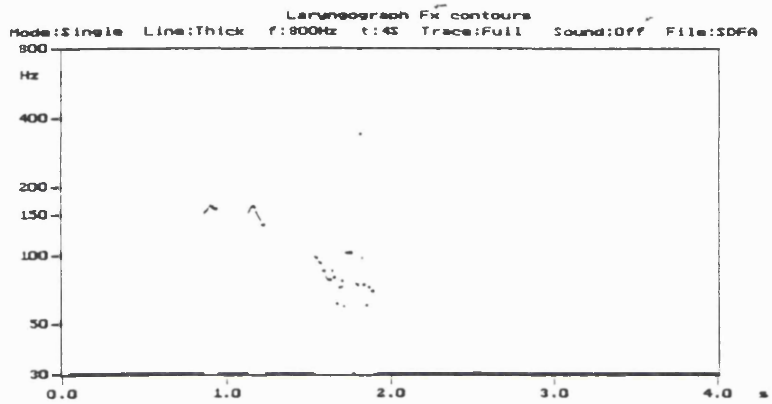


(3) 'Another orange' (Pierrehumbert 1980: 149)

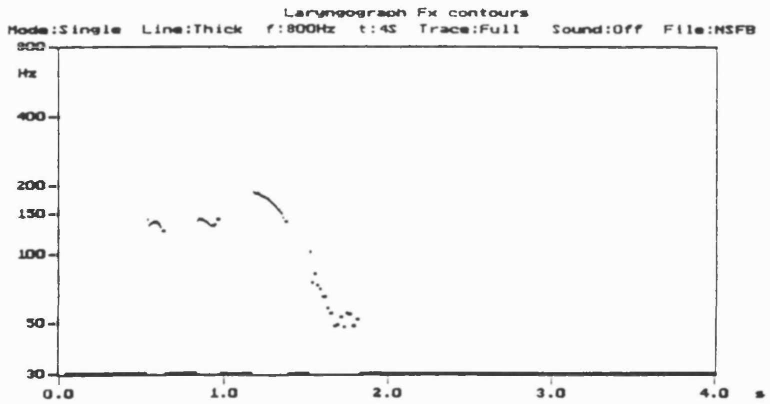


Appendix to chapter 4, The grammar of intonation

(1) Step-down + fall (SDFA) over the phrase *[It's a [@] [bysmal]].*



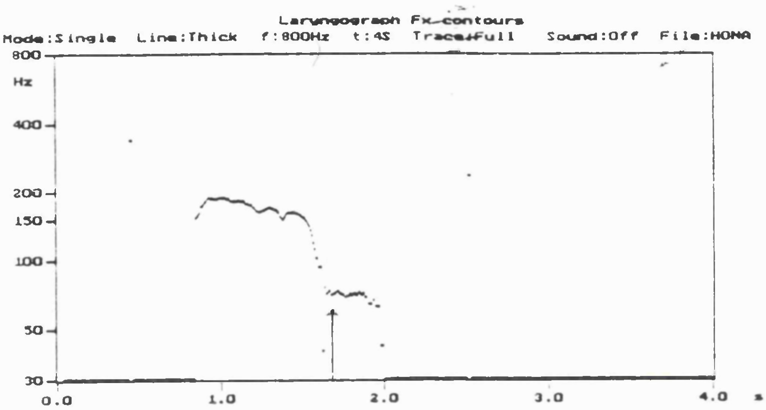
(2) Non-step + fall (NSFB) over the phrase *[It's a [bysmal]].*



(3) High onset + nucleus over the phrase *[[Mary has a] [lamb]]*[1].

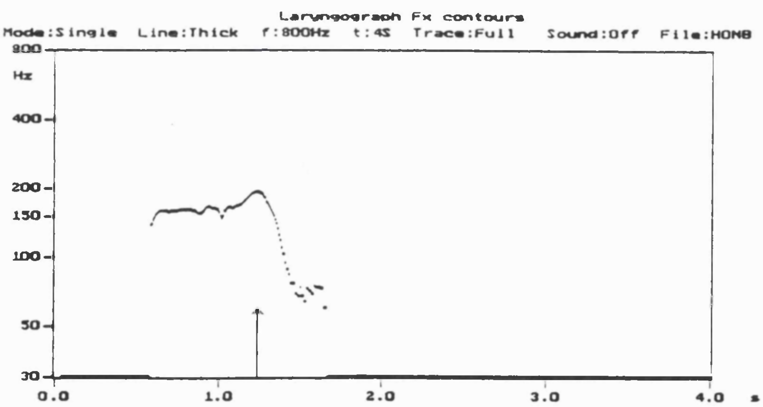
(a)

```
        T   T
        |   |
    [   [   ]   [   ]   ]
```



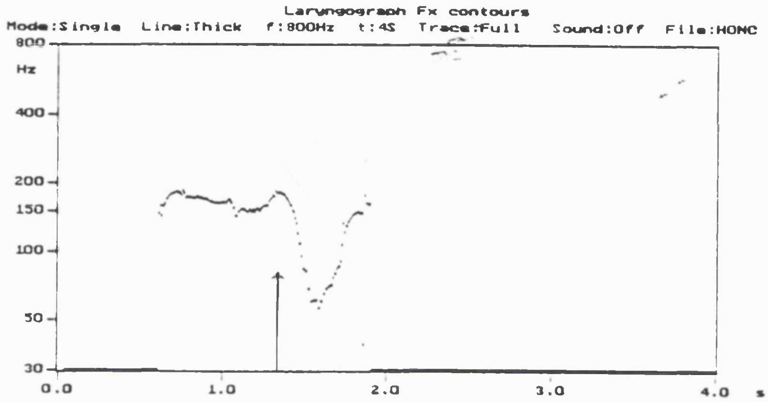
(b)

```
        T   T   T
        |   |   |
    [   [   ]   [   ]   ]
```

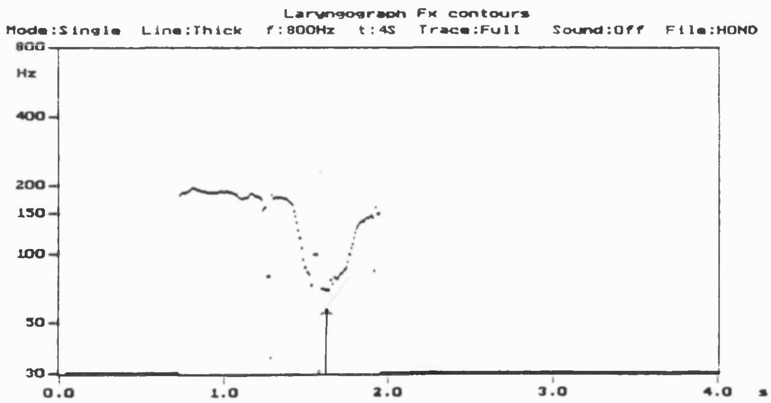


[1] The arrow s indicate the beginning of *lamb*. The utternace for the following F0 contours remains the same (*[[Mary has a] [lamb]]*).

```
(c)          T       T       T               T
             |       |       |               |
        [    [       ]       [       ]       ]
```


Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:HONC
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0
s

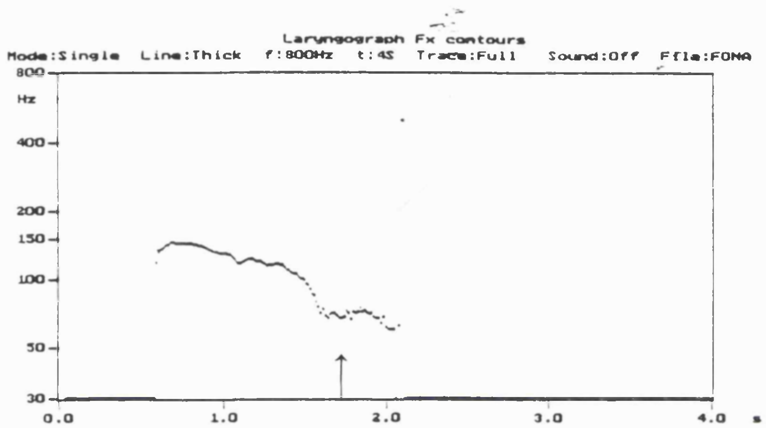
```
(d)          T       T                       T
             |       |                       |
        [    [       ]       [       ]       ]
```


Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:HOND
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0
s

(4) Falling onset + nucleus

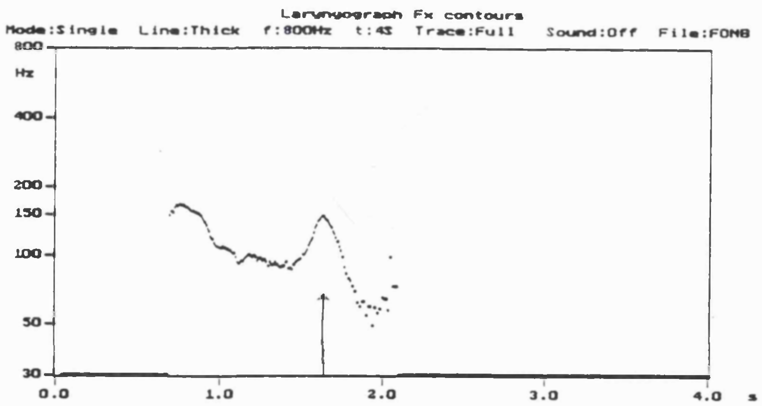
(a)

```
        T
        |
   [    [    ]    [    ]    ]
```



(b)

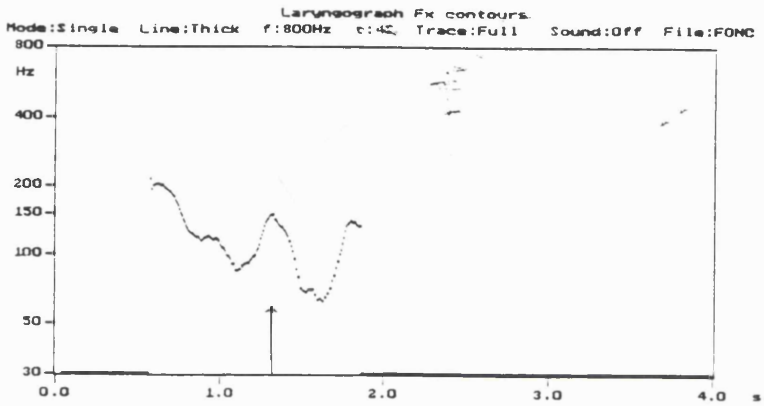
```
        T         T
        |         |
   [    [    ]    [    ]    ]
```



(c) T T T

| | |

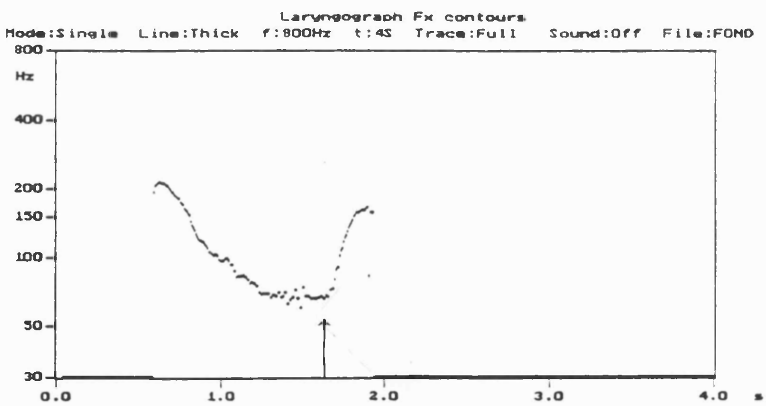
[[ ] [ ]]


Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:FONC
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0 s

(d) T T

| |

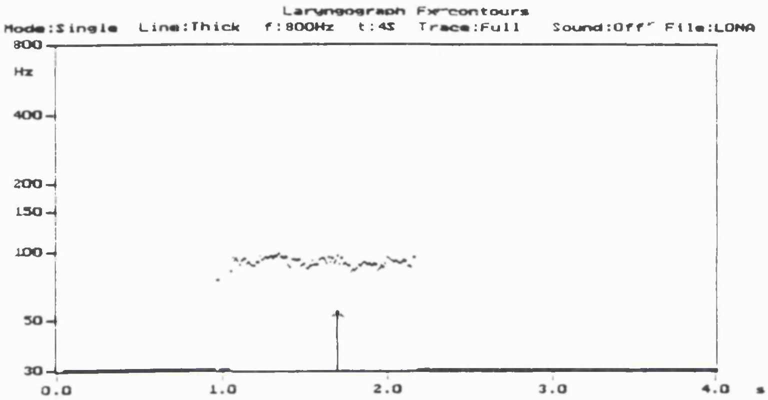
[[ ] [ ]]


Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:FOND
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0 s

(5) Low onset + nucleus

(a)

[[ ] [ ]]

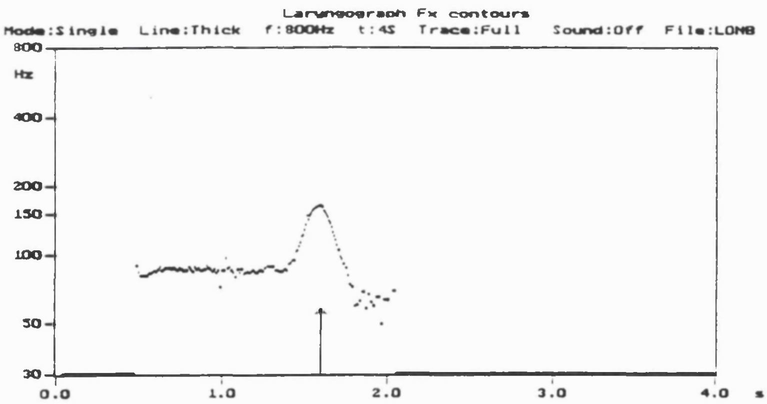


(b)

T

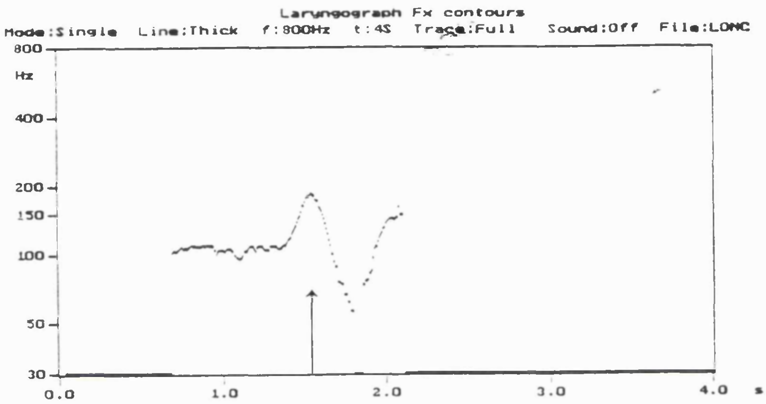
|

[[ ] [ ]]



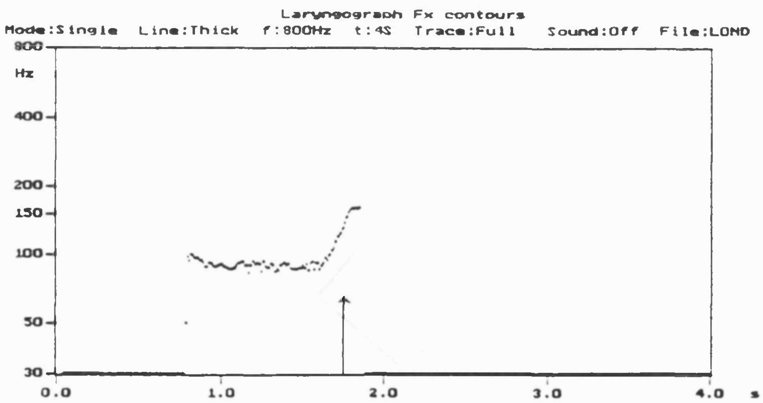
(c)

```
                  T         T
                  |         |
[    [    ]    [    ]    ]
```


Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:LONC
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0
s

(d)

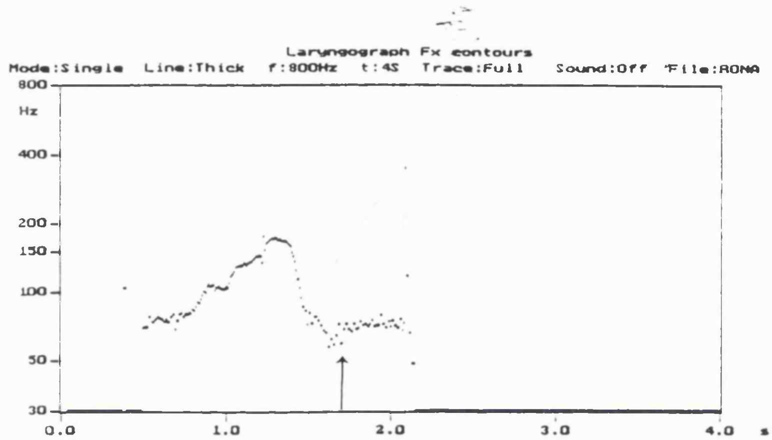
```
                            T
                            |
[    [    ]    [    ]    ]
```


Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:LOND
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0
s

(6) Rising onset + nucleus

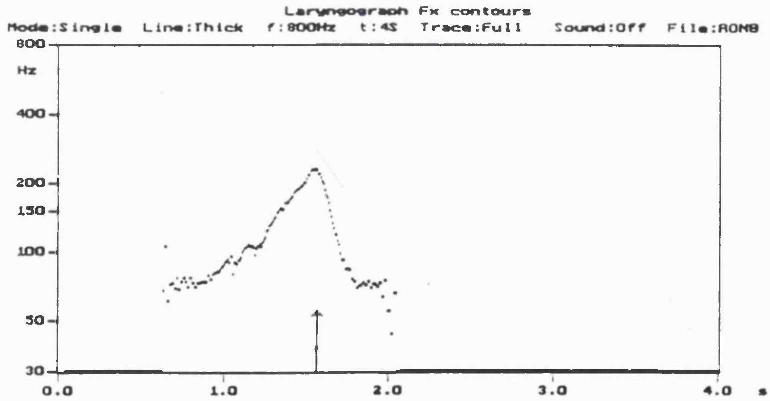
(a)

```
                T
                |
[      [        ]      [      ]      ]
```



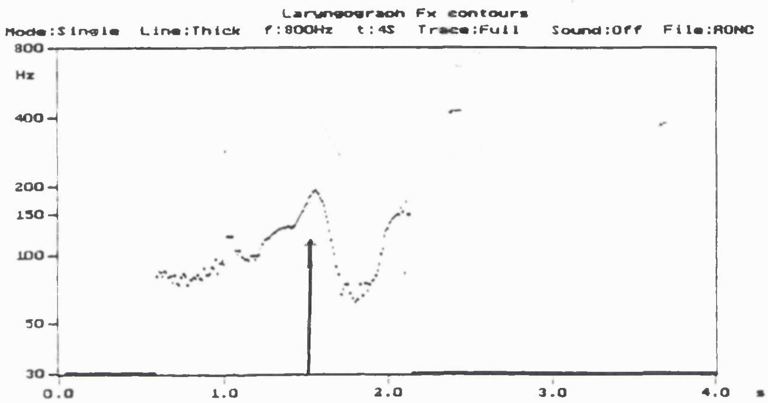
(b)

```
                T      T
                |      |
[      [        ]      [      ]      ]
```



(c)

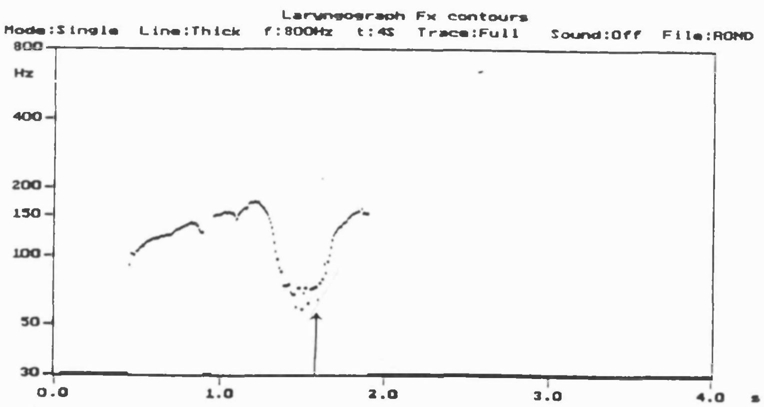
```
                T       T               T
                |       |               |
[       [       ]       [       ]       ]
```



Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:RONC
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0
s

(d)

```
                T                       T
                |                       |
[       [       ]       [       ]       ]
```



Laryngograph Fx contours
Mode:Single Line:Thick f:800Hz t:4S Trace:Full Sound:Off File:ROND
800
Hz
400
200
150
100
50
30
0.0
1.0
2.0
3.0
4.0
s

Appendix to chapter 5, The phenomenon of downstep

(1) 'Mari·Anna'. (Beckman and Pierrehumbert 1986: 277)