Phonology of intonation

By Carlos Gussenhoven

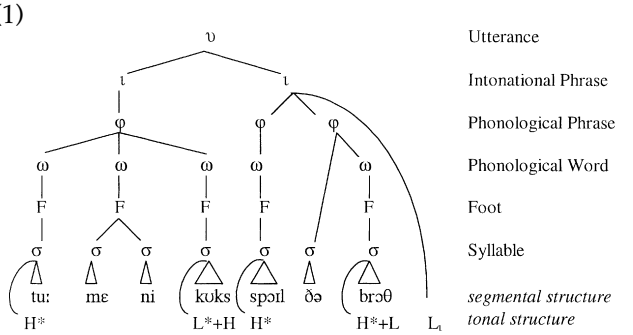
1. Introduction

Pierrehumbert's (1980) thesis marked the beginning of a new period in intonational research by presenting a model which separated the phonological representation from its phonetic implementation, thus allowing a characterization of the notion 'possible prosodic structure' as distinct from an account of the phonetic details of intonation contours. In this state-of-the-article, I sketch the historical background of this model, its role in redefining the phonetics-phonology interface, and what I see as the main developments. Introductions to the model are Ladd (1996), who labelled it the Autosegmental-Metrical (AM) model, Shattuck-Hufnagel & Turk (1996) and Beckman (1996). I leave treatments in other frameworks, notably Hirst & Di Cristo (1998), out of account.

The model is autosegmental because it has separate tiers for segments (vowels and consonants) and tones (H,L). It is metrical because it assumes that the elements in these tiers are contained in a hierarchically organised set of phonological constituents, as depicted in (1), to which the tones make reference in several ways. Among themselves, tones are organized into pitch accents and boundary tones. Tones may or may not be associated with Tone Bearing Units, whose nature varies across languages and ranges from the sonorant mora, as in Japanese, to the accented syllable, as in English. Representation (1) gives a possible pronunciation of the English proverb *Too many cooks spoil the broth*. The pitch accents have a starred tone to indicate their association with the accented syllable, and can be monotonal (e.g. H*) or bitonal (e.g. L*+H), while the boundary tone L_ι comes with the Intonational Phrase (henceforth ι) and associates with the appropriate node, as in Pierrehumbert & Beckman (1988). Example (1) illustrates a commonly adopted set of prosodic constituents, but other constituents have been postulated in intonation, notably the Intermediate Phrase and the Accentual Phrase, while for Japanese, a mora-tier would also be relevant to intonation. A non-crucial assumption is that boundary tones are optional, the first ι in (1) not having any.

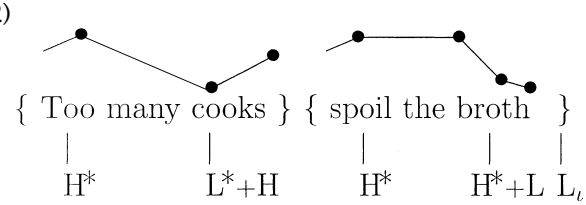
As a surface representation, (1) is a theory of a mental construct for a speaker of British English, who also possesses a phonetic implementation module which, among other things, translates every tone into an F_0 target. The phonetic 'alignment' of the target with the segmental tier, its timing, will to some extent be language-specific, as will its 'scaling', its F_0. Over

(1)



and above the effects of these implementation rules, the speaker's psychological condition and communicative purpose will influence the overall pitch range, so that the number of pronunciations of (1) is infinite. In the schematic implementation (2), the targets are given as bullets, which are connected by line segments that represent the F_0 interpolations between them. The boundary tone is attached to the bracket, as per convention (cf. Hayes & Lahiri, 1991).

(2)



My concern here is with substance (representations, contours), not with functions (focus, intonational meaning). A complete analysis of an intonational system will comprise a phonology as well as the morphology. Researchers usually feel more at ease with the phonological analysis, where the smallest units are given by the model. The elements in a morphological analysis are not given *a priori*, however. In principle, the ι-wide contour could be a single morpheme, any sequence of tones in the contour could be, or each tone could be a morpheme. A whole-contour proposal occurs in Liberman & Sag (1974), where the 'contradiction contour', for instance, consists of a two-accent sequence, a position which

Carlos Gussenhoven, Centre for Language Studies, University of Nijmegen, Erasmusplein 1, 6525 HD Nijmegen, The Netherlands, c.gussenhoven@let.kun.nl

has not been followed (Bolinger, 1986, 245). Morphemes consisting of pitch accents plus boundary tones occur in proposals for English intonational meaning (cf. Cruttenden, 1997) and Hayes & Lahiri (1991) for Bengali. Pierrehumbert & Hirschberg (1990) propose an analysis in which (virtually) every tone is a morpheme, a position defended in Bartels (1997). Recently, Dainora (2001, 2002) advanced the argument that the relatively high predictability of transitions between pitch accents and following tones in American English suggests that these elements are not morphemically independent. Of course, the same phonological analysis may be compatible with a large number of morphological analyses, just as different views of the morphological structure of an English word like *replicate* ([[*re* [*plic*]] *ate*], [[*replic*] *ate*], [*replicate*]?) do not compromise the assumption that its surface phonology is ['rɛplɪ,keɪt].

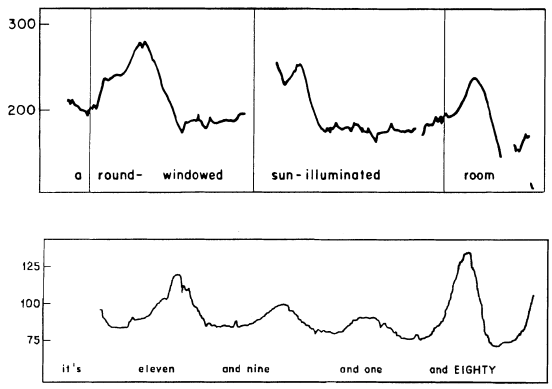
Other aspects that are left out of consideration are the distribution of pitch accents as a results of clash resolution within the phonological phrase (Bolinger, 1965; Bolinger, 1986; Vanderslice & Ladefoged, 1972; Gussenhoven, 1986; Gussenhoven, 1991; Shattuck-Hufnagel, 1989; Ladd & Monaghan, 1987), for which empirical support was provided by Horne (1990), Vogel, Bunnell, & Hoskins (1995), Ostendorf, Price, & Shattuck-Hufnagel (1995), Shattuck-Hufnagel, Ostendorf, & Ross (1979), Shattuck-Hufnagel (1995). This intonational phenomenon had been interpreted as shifts of stress in other work, e.g. Liberman & Prince (1977). Finally, although accounts of intonation in an Optimality theoretic framework have meanwhile appeared, both for phrasing (Selkirk, 2000; Truckenbrodt, 1999) and for the melody (Gussenhoven, 2000a; Gussenhoven, 2000b), this topic equally falls outside the scope of this article.

In section 2, I sketch the precursors of the various elements in the model, while section 3 gives the most important developments.

2. Background

The most innovative aspect of Pierrehumbert's account, which was consolidated as a general theory of tone and intonation by Pierrehumbert & Beckman (1988), was probably the concept of a phonetic 'target' as distinct from a phonological tone, and the principled separation of phonological representations and phonetic implementation that it involved. Most elements in the model can be traced back to earlier positions. First, the idea that there are separate tiers for tones and segmental phonemes was implicit in many descriptions of intonation, inasmuch as pitch features were not considered to be part of the featural composition of segments in the British tradition of intonation description (e.g. O'Connor & Arnold, 1973) or in the description of Dutch by 't Hart, Collier, & Cohen (1990). Also, Goldsmith's (1976) autosegmental model had earlier been applied to English intonation by Goldsmith (1980) (which began as an unpublished MIT paper in 1974), Liberman (1975) and Leben (1975).

Second, the idea that speech reflects a phonological representation consisting of hierarchically organized constituents and that at least one of these constituents, the Intonational Phrase (*ι*), was intonationally defined, had been current in prosodic research at least since Selkirk (1978). In the revised theory of Beckman & Pierrehumbert (1986), an additional intonationally defined constituent was introduced, the Intermediate Phrase (ip), ranked immediately below the *ι*. In this way, two degrees of depth became available for an intonational boundary. Two situations given by Beckman & Pierrehumbert (1986) are illustrated in (3) and (4), where the square brackets enclose ip's and the curly brackets the *ι*. In (3), the two adjectives are considered to be followed by just an ip-boundary, because the disjuncture with what follows is less complete than that observed for a full-fledged *ι*-boundary. In (4), the ip-boundary after *nine* is motivated by the high F_0 peak on *eighty*, which is due to an interruption of the downstepping pattern shown by the preceding F_0 peaks on *one* and *nine*. The F_0 contours are given in panels (a) and (b) of Figure 1, respectively.



(3) { [A round-windowed] [sun-illuminated] [room] }

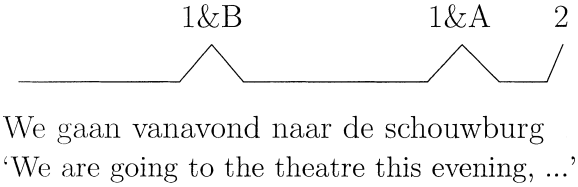
(4) { [It's eleven and one and nine] [and eighty] }

Analyses of West Germanic languages, including English, that do without ip and the 'phrase tone' that has been related to it (see below) are common (van den Berg, Gussenhoven, & Rietveld, 1992; Féry 1993; Grabe, 1998a; Gussenhoven, Terken & Rietveld 1999). In those analyses, the first internal ip-boundary of (3) would be an *ι*-boundary, while the second as well as that in (4) would not be an intonational boundary; the exemption from downstep would be attributed to the fact that it is the last ('nuclear') pitch accent of the *ι* (Ladd, 1983, 735; Gussenhoven, 1983b).

Third, the separation into pitch accents, symbolized T*(T), and boundary tones, symbolized T%, harks back to Trager & Smith Jr.'s (1951) juncture phonemes (# 'falling', – 'sustained', | | 'rising'), which existed by the side of the pitch phonemes. A division between pitch accents and boundary tones *avant-la-lettre* can also be recognised in the work in the 1970s and 1980s by Hans 't Hart, René Collier and Antonie Cohen on

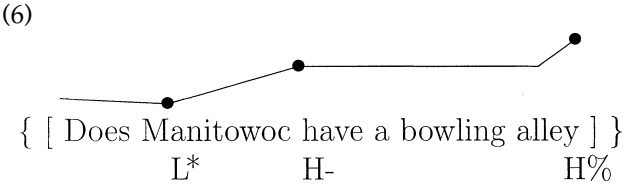
Dutch ('t Hart, Collier, & Cohen, 1990), as argued by Ladd (1996). They divided up the pitch movements of Dutch into 'accent-lending' and 'non-accent-lending', but stopped short of equating the latter with boundary features. An example is given in (5), a non-final contour, where '1' is an accent-lending rise, 'B' a non-accent-lending fall, 'A' a steep accent-lending fall, and '2' a non-accent-lending rise. (When two accent-lending movements appear on the same syllable, only one accent is produced.)

(5)



The chief motivation in Pierrehumbert (1980) for ending ι with two tones, a phrase accent, T-, and a boundary tone *tout court*, T%, was that in many contours two targets can be identified after the last ('nuclear') pitch accent. For instance, in (6), the L* pitch accent is followed by a high target at the end of the accented '*Manotowoc*, as well as a final high target, for which H-and H% were postulated, respectively. I reproduce her contour 2.29 as (6). This example would be natural in a conversation where someone had just asked the speaker if he knew of any towns with bowling alleys.

(6)



In Beckman & Pierrehumbert (1986), T was reanalysed as a boundary tone of the ip. As a result, ip's ended in T-, and ι's in T-T% in the new analysis, since the right edge of every ι coincides with that of an ip. This analysis was transferred to the practically oriented transcription system for American English that was derived from the Pierrehumbert model, *Tones and Break Indices* (ToBI), cf. Beckman & Ayers, 1994, see also below).

Fourth, the notion that a contour is an interpolation between levels, such that only the beginning and end points of a movement result from a target and thus from a tonal specification, was inherent in the analyses by Pike (1945) and Trager & Smith Jr. (1951), who used numbers to indicate pitch levels. Pike's (7), for instance, shows interpolations by means of dashes. In his system, '1' was the highest pitch phoneme, '4' the lowest. The notion of a starred tone was present in that accented levels, which begin a 'primary contour', are marked with the degree symbol. In this case, the representation translates quite readily into Pierrehumbert's (8). Other theories identified pitch *movements* as the basic elements instead of the interpolations between *level* pitches, although the idea of non-specification was present to the extent that some movements were stretchable, their duration being determined by the length of the segments over which they were pronounced, like the 'rising head' of O'Connor & Arnold (1973) or the gradually rising pitch movement '4' of 't Hart, Collier, & Cohen (1990).

(7) I wanted to do it, but I couldn't
4- °2- -4-3/4- °2- -4 //

(8) I wanted to do it, but I couldn't
H* L-H% H*L-L%

Fifth, the idea that the string of tones contained lexical and intonational tones forms the hallmark of Bruce (1977), who isolated the contribution of the lexical tones of Stockholm Swedish from that of the intonational tones, representing them as a string of pitch levels that were timed with the stresses and phrase ends much as in an autosegmental description (Pierrehumbert, 2000; Ladd, 2000). Ladd (1983a) characterized this type of description as a Tone Sequence model, to distinguish it from descriptions that superimpose accentual contours on phrasal intonation contours, termed Contour Interaction models by Ladd, as represented by Gårding (1983), Thorsen (1978, 1983) and Vaissièrre (1983), as well as by Fujisaki's (1983) model. The integration of lexical and intonational tones played an important role in the description of Japanese in the work by Pierrehumbert & Beckman, and later in that of Norwegian, varieties of Basque and Dutch dialects.

Sixth, the idea that there are only two tones, H and L, was also part of Bruce's thesis. Earlier, Liberman (1975) had described the intonation of American English with the help of four tones, H, L, raised H and raised L, thus staying closer to the earlier descriptions. At that point, descriptions were still vulnerable to Bolinger's (1951) criticism that four-level transcriptions of English intonation, like (7), were arbitrary, because 2–4 would not be discretely different from, say, 3–4 or 1–4.

2.1 The 1986 model for American English

Beckman & Pierrehumbert's (1986) revised model includes six pitch accents, H*, L*, H*+L, L*+H, L+H* and H+L*. An optional initial boundary-%H precedes the ι, and the ip and ι are obligatorily closed by boundary tones, as explained above. The grammar can be given as in (9), where parentheses include optional elements, accolades alternative options, and subscripts stand for '*n* or more occurrences', as usual. The part enclosed between the outermost (…) indicates the ip, of which there must be one or more, containing one or more pitch accent.

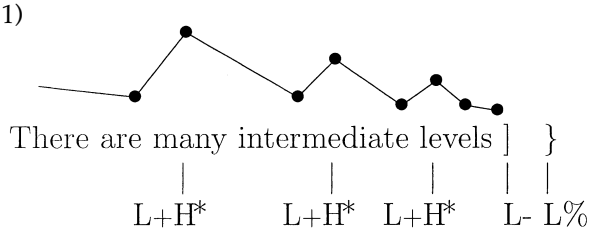
(9) The tonal grammar of Beckman & Pierrehumbert (1986)

$$(\%H)\left(\left(\left\{\begin{array}{l} H^* \\ L^* \\ H^*+L \\ L^*+H \\ L+H^* \\ H+L^* \end{array}\right\}\right)\left\{\begin{array}{l} H^- \\ L^- \end{array}\right\}\right)\left\{\begin{array}{l} H\% \\ L\% \end{array}\right\}$$

Perhaps more than Bruce's (1977) 'pitch rules', Pierrehumbert's implementation rules made it clear that the generation of an infinite number of F_0 values between the highest and lowest pitches allowed the distinction between phonological representations and phonetic contours to become very clear. The representations are there to describe what is a possible linguistic expression and thus characterize the contrasts of the language. For this purpose, two tones suffice, for English and for many other languages that have been described since. The phonetic implementation will create contextually appropriate targets for each tone. Downstep provides a suitable illustration. In (11), each non-initial H* is downstepped relative to the preceding H*. In Beckman & Pierrehumbert (1986), downstep applies to H* and H- after a bitonal pitch accent. Or again, a mid tone at the end of an Intonational Phrase is obtained by the combined working of DOWNSTEP, which lowers H- after a bitonal pitch accent, and UPSTEP (12), which raises L% to the level of a preceding H-, and raises H% above the level of a preceding H-. In the half-completed fall (13), therefore, the H- is downstepped and the L% is upstepped. UPSTEP is also responsible for the extra-high H% in (13b).

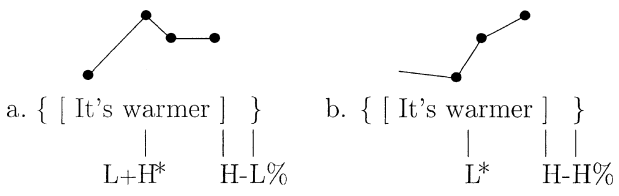
(10) PB DOWNSTEP: H→ !H/T*T . . . ___ . . . T%
(Implementation)

(11)



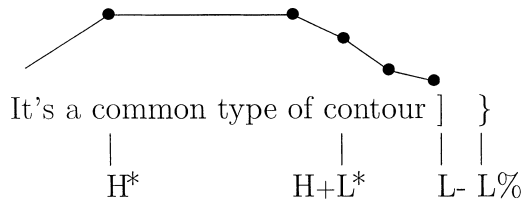
(12) PB UPSTEP: T% → raised T% / H- ___
Implementation

(13)



In addition to the two implementations rules given in (10) and (12), and the abstract downstep-trigger L of H*+L (see section 3.2), a final implementation rule was that L* of H+L* was realised as a downstepped H, H+L* being equivalent to H+!H*. This pitch accent is used to described the spreading plus-downstepping pattern illustrated in (14). This fairly abstract system has in practice been replaced with the practically oriented ToBI transcription system, but may still be seen as the 'real' Pierrehumbertian analysis behind the newer system. ToBI did away with the abstract downstep trigger, and noted downstepped H's directly with !, following a suggestion by Ladd (1983b). Earlier H+L* thus also became H+!H* in the new system.

(14)



In addition to downstep, two further phenomena are involved in F_0 downtrends (Pierrehumbert, 1980; Liberman & Pierrehumbert, 1984; Ladd, 1984; Pierrehumbert & Beckman 1988). First, there is 'declination', a time-dependent, gradual F_0 lowering, associated with one or more *ι*'s but otherwise context-independent. Second, there is 'final lowering', loosely associated with the last syllable(s) of the Utterance (*υ*). Downstep may be seen as a grammaticalization of declination, just as many tone languages have grammaticalizations of final lowering. Earlier, these three concepts tended to be collapsed under a single notion, usually simply referred to as 'declination'.

Questions like 'How do you pronounce H in English' will thus require a lengthy consideration of all the contexts in which H can occur. The difference between discrete ('digital') representation and gradient ('analogue') implementation as drawn for intonation by Pierrehumbert (1980) was extended to the phonology-phonetics interface generally by Pierrehumbert (1990), Keating (1990), Cohn (1990) and others. This work has crucially contributed to our understanding of speaker control in phonetic implementation (Kingston & Diehl, 1994), phonetic underspecification (see below), and the understanding of the place of paralinguistic meaning in intonation. To take up the latter point briefly, implementation rules that apply as a function of phonological context (e.g. DOWNSTEP, UPSTEP) are to be distinguished from structurally contextless variation signalling paralinguistic meanings (Gussenhoven, 2002). While English contrasts early peaks (H*) with late peaks (L*+H) (Pierrehumbert & Steele, 1989), the height of the peak is gradiently variable and correlates positively with degrees of urgency. Such variation may also correlate with other attitudinal meanings, like surprise. Differences in meaning are therefore no longer criterial for phonological contrasts (Ladd & Morton, 1997; Gussenhoven, 1999). Haan (2002)

shows that this type of variation may be a function of sentence categories like 'declarative question', wh-question, and yes-no question, each of which appears to some extent to have its own phonetic profile in Dutch.

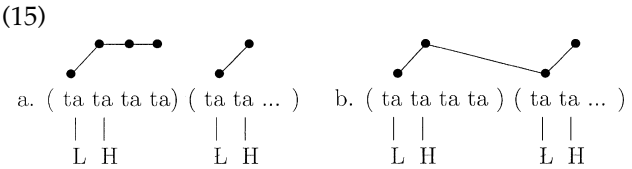
3. Some developments

Pierrehumbert's model has been applied to a number of different languages, like Japanese (Pierrehumbert & Beckman, 1988), Bengali (Hayes & Lahiri, 1991), varieties of Latin-American Spanish (Sosa, 1991), German (Uhmann, 1991; Féry, 1993), Palermo Italian (Grice, 1995b), European Portuguese (Frota, 1998), varieties of Korean (Jun, 1993), Basque (Elordieta, 1997; Hualde, Elordieta, Gaminde, & Smiljanić, 2002), French (Post, 2000; Jun & Fougeron, 2000), Bern Swiss German (Fitzpatrick-Cole, 1999), and European Spanish (Face, 2002), among others. This section discusses some noteworthy developments.

3.1 Phonetic underspecification

From Pierrehumbert (1980) onwards, a distinction has generally been made between interpolation, the creation of F_0 values in the phonetic implementation between the targets of phonological tones, and specification through spreading. The point was particularly clearly made in Beckman & Pierrehumbert (1986, 263), Pierrehumbert & Beckman (1988). Japanese Accentual Phrases (henceforth α), often no more than word-sized prosodic constituents, are either accented or unaccented. Both of these typically begin with a pitch rise, while accented α's are characterized by a subsequent sharp fall, from the accented syllable. Unaccented α's lack this fall, and after the initial rise the pitch slowly descends. Earlier descriptions described this slow descent as fully high pitch, and assumed that it was to be explained by a spreading H. This is shown schematically in (15a), after Poser (2001), which represents an α plus the rise of a following α. In (15b), the non-spread, phonetically underspecified situation is given, after Pierrehumbert & Beckman (1988), with an interpolation between the high target in the first α and the low target in the next. The evidence that (15b) is the superior theory comes from the dependence of the slope between the high and low targets on their distance. Pierrehumbert & Beckman showed that the longer the first α was, the less steep its sloping F_0. By contrast, theory (15a) would predict that the slope remains high up till the last syllable of the first α, regardless of its distance from the first α's beginning.

(15)

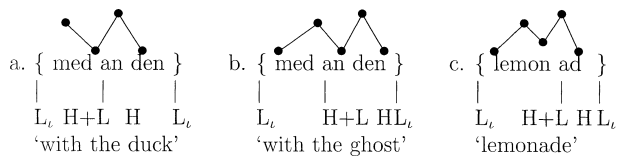


3.2 Unassociated tones

Pierrehumbert (1980) applied the Africanists' distinction between associated and unassociated lexical tones to intonational tones. There are two different uses to which unassociated (or 'floating') tones have been put. The first is based on the role that they have as a trigger for downstep in Bantu languages, while remaining without a target themselves (Stewart, 1966; Clements & Ford, 1980). Thus, a disyllable may have three tones, HLH, of which the first H is realized as high pitch on the first syllable, and the second as a downstepped !H on the second syllable, L itself obtaining no target. Similarly, Pierrehumbert (1980) describes one of the downstepped patterns of English by H*+L H*, whereby the bitonal pitch accent causes the second H* to be downstepped, even when there was no low target in between. The nonrealization of L in H*+L needs to be specified in the implementation rules, and was specific to Pierrehumbert's description of American English.

Generally, however, floating tones *are* realized in intonation. They differ from associated tones in that the timing of their targets is not attributable to a specific point in the segmental tier, but is rather with reference to the targets of other tones. Typically, while one tone in a pitch accent, T*, associates with the accented syllable, the leading or trailing T receives a target which occurs some fixed distance before or after that of T*. As a result, the H of L*+H as used on *Rigamarole* will occur on *–ma-*, the second syllable after the accented *Rig-*, while in *Stein*, it will be realized on the same syllable as L* (Beckman & Pierrehumbert, 1986). Bruce (1987) showed that the focus-marking tone of Stockholm Swedish, which is a single H, is floating, and is pronounced after the associated lexical tone complex HL. Since the lexical tone contrast depends on whether the L (Acc 1) or the H (Acc 2) associates with the stressed syllable, the focal H is pronounced later in Accent-2 words than in Accent-1 words. This is shown in the minimal pair in (16a, b), where the target for focal H falls in the [nd] in 'the duck', but in [ɛ] in 'the ghost'. In (16c), where Accent 1 occurs on a final syllable, H occurs inside the stressed syllable (Bruce, 1987; Bruce, 1990). (The L-boundary tones define the ι; the initial L_ι in (16a) is truncated for lack of segmental space; see also below.) Because the focal H does not associate, the term 'pitch accent' has been avoided for this tone. Pierrehumbert & Beckman (1988, 251) have in fact suggested it may be a boundary tone which is pronounced early.

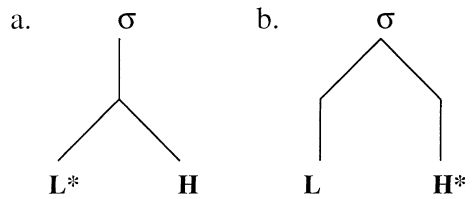
(16)



3.3 Bitonal pitch accents

A decision to designate one of the two tones in a bitonal pitch accent as T* may be based on nothing more than that its target is closer to the accented syllable. Not all pitch accents appear to be timed in this asymmetrical fashion, however. Work on Standard Greek has shown that the distance between the two targets of the prenuclear LH pitch accent depends on the duration of the stressed syllable, since the low target occurs just before the syllable and the high target just after it (Arvaniti, Ladd, & Mennen, 2000). As a result, the rise is longer when the syllable has more consonants; a further implication is that it is no longer clear which is the starred tone, L or H. This pitch accent could be interpreted as a branching structure, which as a unit associates with the accented syllable, as the authors suggest. The internal structure of the bitonal pitch accent had earlier been discussed by Grice (1995b), who recognised a 'cluster' by the side of a 'contour'. This distinction requires that a tonal node is introduced between the syllable and the tones. In the contour, the two tones are gathered under the tonal node, a representation originally proposed by Yip (1989) for tone complexes in southeast Asian languages, where they are underlying, while in the cluster the tones are directly dominated by the syllable, a representation Yip used for tone in African languages, where contour tones are typically derived. Grice claimed that in English, pitch accents with trailing tones, like L*+H are contours (cf. (17a)), but that pitch accents with leading tones, like L+H* are clusters (cf. (17b)). One of her arguments is that leading tones tend to be truncated when the accent is *ι*-initial, while *ι*-final trailing tones are not. To return to the Greek case, a structure like (17a) could be provided with a star for the whole complex to serve as the prenuclear LH (Arvaniti, Ladd, & Mennen, 2000).

(17)

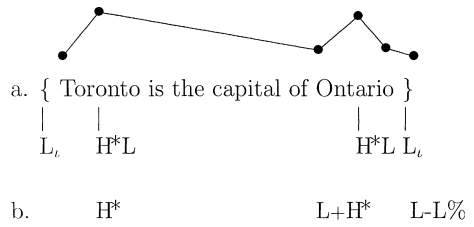


Timing characteristics were used by Frota (2002) to decide between an analysis of the European Portuguese focal and non-focal pitch accents as either bitonal H*+L and H+L* or as HL combinations of boundary tone and single-tone pitch accent. The evidence leads her to conclude that both are bitonal pitch accents, but that only the targets of H*+L are characterized by a constant interval.

Detailed phonetic studies of the timings of targets have on the one hand revealed cross-linguistic and cross-varietal variation, and on the other a tendency for targets to be coupled to 'segmental landmarks', like the syllable offset, as in Mandarin lexical tones (Xu, 1998), or the CV-boundary (Ladd, Faulkner, Faulkner, & Schepman 1999). They can be sensitive even to the tenseness of the vowel (Ladd, Mennen, & Schepman 2000). Evidence for the influence of the length of the onset and the sonorant status of consonants is provided in Prieto van Santen, & Hirschberg (1995) and Rietveld & Gussenhoven (1995), while the role of constraints like the speed of F_0 movements and the time difference between implementation and articulatory effect are discussed in Xu (2002).

The assumption that the targets of bitonal pitch accents are close together is not generally shared. The timing of trailing tones in English was made dependent on the distance to the next accent in Gussenhoven (1983a, 1988, 1999), Féry (1993) and Grabe (1998a). Part of the motivation was functional, as illustrated by (18a), after Gussenhoven (1983a), which has two occurrences of what appears to be the same neutral pitch accent. The target of the trailing L of the non-final H*L is timed rightmost, and bounded by the following associated tone. The trailing L of the final pitch accent is however constrained so as to occur immediately after the target of its T*. Because the timing of the trailing tone is context-dependent, the '+' is avoided, as it suggests that the tones are always realized close together. The representation of contour (18) in Pierrehumbert's theory is (18b), which has different pitch accents, H* and L+H*. A second argument for assuming (18a) is that both the timing and the scaling of the low target before the second peak is not precise, and could be higher and earlier with no appreciable perceptual difference, which suggests that L's rightward drift is somewhat imprecise. This is not what is suggested by the pitch accent L+H*, however. A third argument is that the right-moving trailing tone, or 'displaced' tone, to use Grabe's (1998a) term, also occurs in non-final L*H and H*LH (Gussenhoven, 1983a). The rightward displacement was termed a 'partial linking' in Gussenhoven (1983a), where 'linking' referred to the coherence of the two pitch accents, and was seen as a step towards 'complete linking', the deletion of the trailing tone.

(18)



3.4 Boundary tones

Boundary tones mark the edges of prosodic constituents. Prosodic phonology holds that speech is

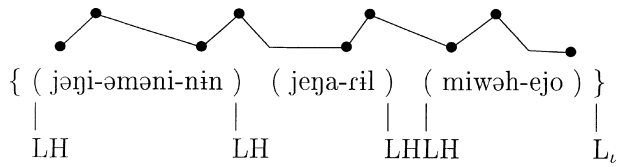
produced in batches of segments that are hierarchically ordered: within any such batch except the lowest a smaller batch can be identified. The gestural integration correlates inversely with rank: syllables are highly integrated articulations, while Utterances may contain noticeable pauses. In (1), which represents a widely adopted view of this hierarchy (Selkirk, 1978; Nespor & Vogel, 1986; Hayes, 1989) the syllable (σ) is included in the foot (F), the foot in the phonological words (ω, also 'prosodic word'), the phonological word in the phonological phrases (ϕ) or Accentual Phrase (α), depending on language, the phonological phrase in the intonational phrase (ι) or Intermediate Phrase (ip), depending on language or analysis, and the intonational phrase in the Utterance (υ).

The reality of these prosodic constituents is apparent from a number of phenomena. The context for segmental processes like assimilation is often defined by the boundaries of specific prosodic constituents (e.g. Nespor & Vogel, 1986). Second, their boundaries reveal themselves through lengthening at the end (Wightman, Shattuck-Hufnagel, Ostendorf, & Price 1992; Gussenhoven & Rietveld, 1992) and consonantal strengthening at the beginning (Fougeron & Keating 1997; Fougeron, 2001; Cho & Keating, 2001). For instance, in *Tiptoe through the tulips*, the υ-initial [t] in *tip-* will be longer and have a more extensive articulatory contact than the ω-initial [t] in *tu-* which in its turn will be longer and stronger than the F-initial [t] of *-toe*. The presence of a pitch accent independently increases the durations of segments in, and to some extent near, the accented syllable (Beckman & Edwards, 1990; Cambier-Langeveld & Turk, 1999; Cambier-Langeveld, 2000), due to 'accentual lengthening'. Abstracting away from the presence of pitch accents and segmental effects on duration, *-toe*, the last syllable of a phonological phrase, will be longer than *tu-*, the first syllable of a prosodic word, due to final lengthening. Third, syntactic movement rules may be sensitive to the size of the constituents they manipulate (Inkelas, 1989; Inkelas & Zec, 1990).

In intonation, the prosodic structure may play a role in the context of phonological or phonetic rules (e.g. downstep, which is always confined to some prosodic constituent), in the rhythmic distribution of pitch accents, and in the presence of boundary tones, the topic of this section. One, sometimes two, higher prosodic constituents may come with boundary tones initially and/or finally. These are now reported so frequently that they may well be universal. Some languages appear to have *only* boundary tones, like Unangan, which has L_ϕ at the beginning and H_ϕ at the end of every ϕ, and the only function of intonational structure in this language is thus to signal phrasing (Taff, 1997).

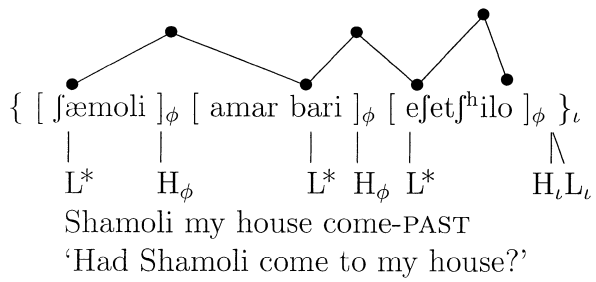
Boundary tones may be complex, i.e. consist of a tone sequence. Jun (1993) gives an analysis of Seoul Korean where every α has a string LH-LH, which is reduced to just an initial LH if the α has only three syllables (Jun, 1998). In the final α, any second occurrence of LH is preempted by one of the ι-final boundary tones H_ι, L_ι, L_ι H_ι, or $H_\iota L_\iota$. which express different intonational meanings. An example is (19). Formally and functionally, Korean is thus more complex than Unangan, but like Unangan lacks pitch accents.

(19)

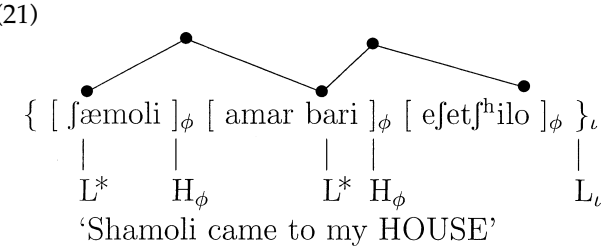


Bengali combines boundary tone with pitch accents, and closes the ι either by L_ι or by the boundary complexes $L_\iota H_\iota$, for continuation, and $H_\iota L_\iota$, for yes-no questions. This is illustrated in (20), where L* marks the accented syllables. Evidence that $H_\iota L_\iota$ both end the ι is provided by the contrast with L* $H_\phi L_\iota$, which contour is used to mark narrow focus with declarative intonation, as shown in (21), and is realized with a lower peak, as shown in (22). The F_0 peak signalling the question is always ι-final and high.

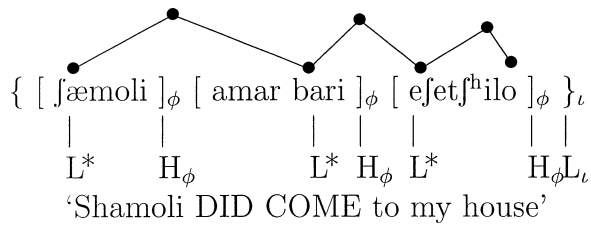
(20)



(21)



(22)

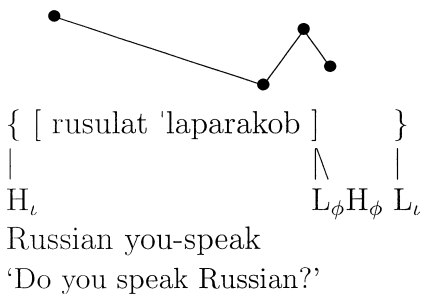


These examples also illustrate that more than one prosodic constituent may come with boundary tones, since in addition to H_ι, Bengali has H_ϕ. Evidence that

H_ϕ is a boundary tone is given by the pronunciation of *amar* 'my' in (20)–(22). The L* goes to the first syllable of the first lexical word in the ϕ, and thus skips the function word *amar*. Because *amar* is pronounced on a downward slope, H_ϕ cannot be interpreted as the leading tone of the following pitch accent, since this analysis would predict that *–mar* has high pitch.

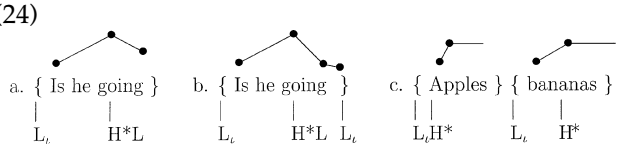
For Georgian, Bush (1999) reports a complex boundary $L_\phi H_\phi$, as in (23), which is part of a question intonation contour that begins with H_ι. As Bush points out, the L-tone is not a pitch accent, as it spurns the stressed syllable in *laparakob*. Neither is it a sequence of L_ϕ H_ι; one of his arguments is that the end of ι may contribute a L_ι in polite speech which could not be abstracted out of the contour if it was combined with the preceding H-tone. The Dutch dialect of Venlo has boundary tones of ι and υ (Gussenhoven & van der Vliet, 1999): all ι's end in H_ι or L_ι, while the utterance can have an additional H_υ, leading to four utterance-final contours, H_ι, L_ι, $L_\iota H_\upsilon$, and $H_\iota H_\upsilon$.

(23)



Finally, as will be clear from the Georgian and Venlo cases, boundary tones can be optional. While ι-initially, Pierrehumbert's analysis and the ToBI system that it gave rise to have an optional %H, mid and low pitch being unmarked, these analyses have *final* obligatory boundary tones at two ranks, the ip (L-, H-) and the ι (L%, H%), as opposed to a single-rank, optional boundary tones. That decision may have led to a tendency to assume obligatory boundary tones in other languages. In my description of standard Dutch, L% and H% were explicitly added in Gussenhoven (1991). Earlier they had been understood as part of the phonetic realization of 'basic' H*L and L*H in nuclear position (Gussenhoven, 1983a), where they contrasted with 'half-completed' versions of the same contours. Absence of L% and H% in these contours came to stand the half-completed realizations in a synthesis-by-rule programme (Gussenhoven & Rietveld, 1992). To exemplify in English, (24a) gives a half-completed fall, which is contrasted with the fall to low in (24b), while (24c) shows a sequence of high levels, one type of listing intonation. In Pierrehumbert (1980), these three contours are transcribed H*+L H-L%, H*L-L%, and H* H-L%, respectively.

(24)

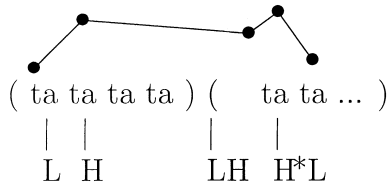


Likewise, Grabe (1998a) argues that an analysis of both German and English with an optional, single-rank boundary tone is to be preferred to a system with two constituents and obligatory boundary tones. Both the transcription system developed for Dutch, *Transcription of Dutch Intonation* (ToDI) (Gussenhoven, Terken, & Rietveld, 1999) and the partly phonetic transcription system developed for varieties of British English, IViE (Grabe, 2001) have optional boundary tones at a single rank only.

3.5 Secondary association

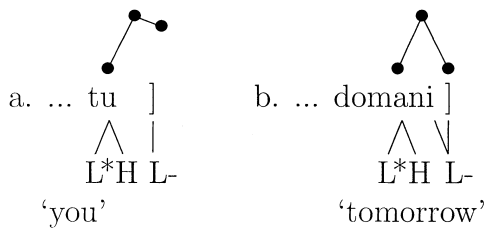
Pierrehumbert & Beckman (1988) introduced the concept of a boundary tone associating with a Tone Bearing Unit (TBU), the sonorant mora in Japanese. They used it to explain the low target at the beginning of the α in (15) and the first α in (25). A different situation exists when an α's first mora is associated with a H-tone, as in the second α of (25), where a pitch accent occurs on the first syllable. There is still a pitch rise, but it starts from a less low target. In (15b) and the first α in (25), therefore, they assumed that the low boundary tone is associated *both* with the prosodic constituent (the 'primary association') *and* with the mora in the first syllable ('secondary association'), while in the second α of (25) the L-tone is merely associated with the prosodic constituent. This difference in representation is translated by the phonetic implementation rules by a more precise, that is, lower target for the L-tone with the secondary association.

(25)



The notation of secondary association was applied by Grice (1995a) to account for the variation between the slight fall on ι-final accented syllables in Palermo Italian question intonation, as in (26a), and the deep fall observed when the accented syllable is non-final, as in (26b).

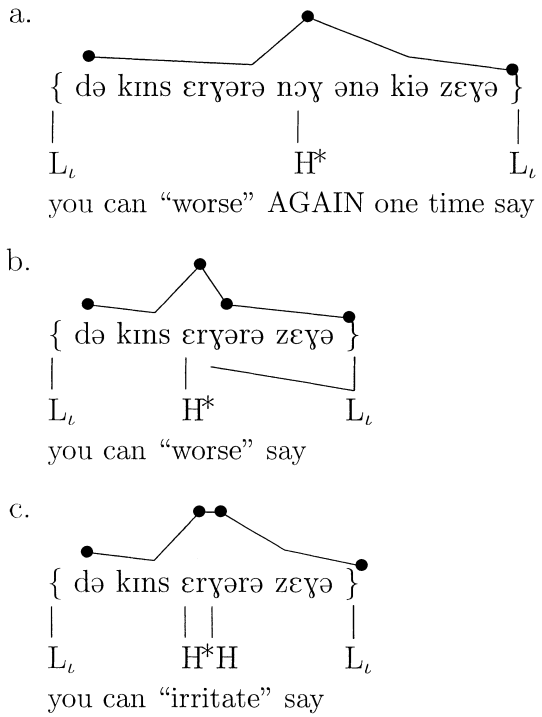
(26)



There is no necessary implication, however, that a tonal representation like (26a) inevitably leads to a phonetic implementation whereby the final L is not or is barely pronounced. Grønnum (1991) and Ladd (1996, 133) would describe a phonetic contour like (26a) as a case of 'truncation', and distinguish it from 'compression', i.e., a contour which reaches low pitch. Languages or language varieties may differ in that one is 'compressing' and another 'truncating', as shown by Grabe (1998b) for RP and northern standard German, respectively, and by Grabe, Post, Nolan, & Farrar (2000) for Cambridge English and Leeds English, respectively.

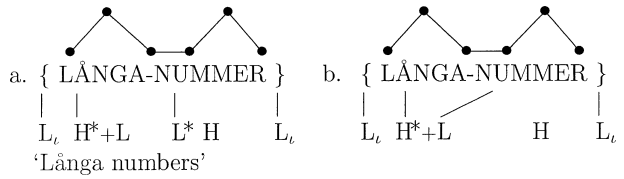
While the difference between 'compression' and 'truncation' might exclusively be accounted for by language-specific implementation rules that are sensitive to the availability of segmental material, confirmation of the moraic association of boundary tones was provided by data from Venlo Dutch. This dialect has a privative tone contrast on stressed syllables with two sonorant moras. There are thus three prosodic types of stressed syllable: those with one sonorant mora (27a), those with two but no lexical tone (27b) (also known as Accent 1), and those with two and H on the second mora (27c) (also known as Accent 2). An intonational pitch accent H* associates to the first mora of the stressed syllable of every focused word, and a declarative L_ι closes the ι. The fall for Accent 1 is completed inside the Accent-1 syllable in (27b), which is explained by the secondary association of L_ι with the sonorant mora in the accented syllable, a TBU which requires tone. Neither in (27a) nor in (27c) is there such a 'free' mora available, and as a result the falls in the latter two contours are slower (Gussenhoven & van der Vliet, 1999). Gussenhoven (2000a) gives data for all three conditions in the related dialect of Roermond. When H_ι occurs instead of L_ι in situations like (27b), there are clearly two targets, forming a high level stretch.

(27)



Not only boundary tones have been claimed to behave in this way, also the trailing L of the Swedish Accent 2 pitch accent has been described as associating with a stressed syllable some distance away from its H*. Compounds have Accent 2 on the first member, regardless of the underlying tone of the first and second members. The second member has a L-target on the stressed syllable, followed by the focal H. However, this LH cannot be equated with the L*H of Accent 1, because there is no leading H before the L (cf. (16a,c)). Instead of postulating a special L* pitch accent to mark second members of compounds, as in (28a), Gussenhoven & Bruce (1999) proposed that the trailing tone of Accent 2 is pronounced twice, once immediately after the target of H*, where it is timed as in English, and once in the next stressed syllable, where its target is due to an association.

(28)



3.6 Phonological adjustments

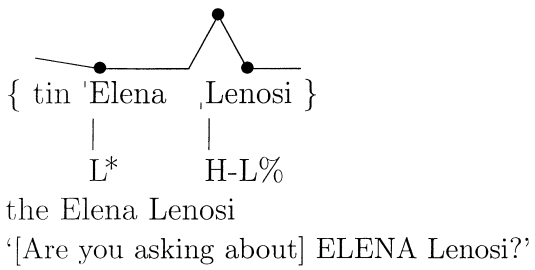
Pierrehumbert (1980) stated that, in English, the tone string arising from the compilation of pitch accent and boundary tones was at the same time the surface representation which offered itself to the phonetic implementation module. Likewise, Stockholm Swedish required no phonological adjustments (assimilation, deletions, insertions) of the underlying tone string. Cases of adjustments in intonational tone strings have since been reported, however. Bengali disallows adjacent like tones (OCP), and violation is solved by tone deletion (Hayes & Lahiri, 1991). Thus, when intonational H* is introduced before H_ι, H* deletes (Lahiri & Fitzpatrick-Cole, 1999). Assimilation of lexical or intonational H to L* occurs in the dialect of Roermond, if H follows L* in the same syllable (Gussenhoven 2000b).

3.7 Phrase accent

The term 'phrase accent' has a checkered history. Pierrehumbert (1980) applied it to the internal boundary tone T-, equating it with Bruce's (1977) 'sentence accent', the focal H of Stockholm Swedish. What they have in common is that they occur between the final boundary tone and the last T*(T), an intonational pitch accent in English and a lexical pitch accent in Swedish. Functionally, the Swedish 'sentence accent' is equivalent to the intonational pitch accents of English. In Beckman & Pierrehumbert (1986), the 'phrase accent' was reanalyzed as a boundary tone of the ip, as noted above, allowing for an analysis in

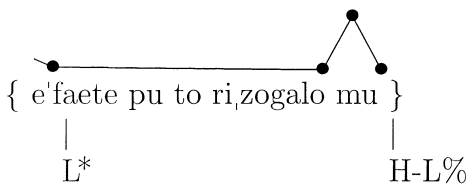
which lower-ranking ip's end T- and higher-ranking *ι*'s end in T-T%. Grice, Ladd & Arvaniti (2000) narrowed the meaning down to a boundary tone with secondary association. This analysis explains why the English phrase accent (T-) is not in fact pronounced at the boundary, but usually immediately after the last pitch accent, and they therefore claim that T- associates with a stressed syllable inside the ip. They also suggest that the much-discussed issue of the difference between 'fall-rise' and 'fall-plus-rise' in descriptions of British English (cf. Cruttenden, 1997) can be explained in terms of 'phrase accents'. While the dependence of T- on stress may be less obvious in German of English, a comparison between Standard Greek and Cypriot Greek makes the difference between an edge-seeking tone and a stress-seeking 'phrase accent' clear. Both languages have a yes/no question contour, analyzed as L* H-L%. In Cypriot Greek, the H- remains close to the *ι*-boundary, with the peak usually falling in the last syllable, while in Standard Greek its target falls in the rightmost stressed syllable. In (29) for instance, H- associates with the unaccented, but main stressed syllable *Le-* in Standard Greek, while L% possibly does too, or else is timed after H-. (In (29), I place the L-target where it would appear to occur in the F_0 track.) The Cypriot Greek case is like that of Hungarian (Ladd, 1983b; Gósy & Terken, 1994; Varga, 2002) and Bengali (Hayes & Lahiri, 1991), and shown in (30), where the rightmost stress is [zo], but the F_0 peak is on the unstressed final [mu].

(29)



the Elena Lenosi
'[Are you asking about] ELENA Lenosi?'

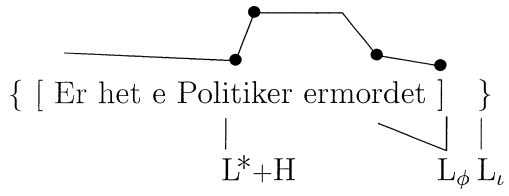
(30)



eat.2PL.PAST of the rice-pudding mine
'Have you eaten of my rice pudding?'

A third possibility is for the H-tone to have two high targets forming a plateau on post-focal stretches of speech, as in Roermond Dutch (Gussenhoven, 2000a) and Transylvanian Romanian (Grice, Ladd, & Arvaniti, 2000). In Bern Swiss German, a similar plateau is due to H of the L*+H pitch accent, which in broad focus continues until the last stressed syllable of the phrase, where a L_ι seeks a secondary association (Fitzpatrick-Cole, 1999). This is shown in (31).

(31)



4. Conclusion

While this brief survey may give one a feel for the cross-linguistic variation in intonational structure, the collection of languages of which fairly complete descriptions are available is still small. Future research is likely to bring many more tone systems to light couched in the autosegmental-metrical framework, with and without lexical tones. Many of these descriptions will have a practical bias in the sense that in addition to claiming to provide a tonal grammar, they are usable as transcription systems of spontaneous speech (cf. Jun (forthcoming)). More grammatically oriented accounts are likely to include a description of how information status ('focus') is expressed prosodically. Three aspects of the prosodic structure have been reported to be used for this purpose. First, there may be a requirement to align the beginning or end of the focus constituent with a prosodic constituent of a given rank, as happens in Japanese, Chichêwa, and Bengali (Pierrehumbert & Beckman, 1988; Kanerva, 1989; Hayes & Lahiri, 1991). In each of these cases, there is a different effect on the phonetics and phonology of the expression: the suspension of downstep (Japanese), the distribution of lexical tones and of vowel quantity (Chicheŵa), and the location of a boundary tone (Bengali, cf. (20)–(22)). Second, there may be a special pitch accent signalling narrow focus, as occurs in Bengali and European Portuguese (Frota, 1998). Third, postfocal words may be de-accented, as is most notably the case in West Germanic languages, but also in many other languages (cf. Ladd, 1996). Further research will undoubtedly refine this picture, as exemplified by Wolof, which lacks prosodic marking of focus and relies on verb morphology for this purpose (Rialland & Robert, 2001).

Acknowledgments

I am grateful to Martine Grice for her comments on an earlier version.

A Phonology of Intonation Bibliography

ARVANITI, A., LADD, D. and MENNEN, I. (2000) What is a starred tone? Evidence from Greek. See Broe & Pierrehumbert (2000), pp. 119–131.

BARTELS, C. (1997) *Towards a Compositional Interpretation of English Statement and Question Intonation*, PhD dissertation, Amherst: University of Massachusetts.

BECKMAN, M. E. (1996) The parsing of prosody. *Language and Cognitive Processes*, 11, 17–67.

BECKMAN, M. E. and AYERS, G. M. (1994) ToBI annotation conventions. <http://ling.ohio-state.edu/tobi/ame_tobi/ >.

BECKMAN, M. E. and EDWARDS, J. (1990) Lengthenings and shortenings and the nature of prosodic constituency. See Kingston & Beckman (1990), pp. 152–178.

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

BEL, B. and MARLIEN, I. (eds) (2002) *Speech Prosody* 2002. *An International Conference*, Aix-en-Provence. Laboratoire Parole et Language. CNRS and Université de Provence.

BOLINGER, D. L. (1951) Intonation – levels versus configurations. *Word*, 7, 199–200.

BOLINGER, D. (1965) Pitch accent and sentence rhythm, in I. ABE and T. KANEKIYO (eds) *Forms of English: Accent, Morpheme, Rhythm*. pp. 139–180. Tokyo: Hokuou.

BOLINGER, D. (1986) *Intonation and its Parts: The Melody of Language*. Stanford, CA: Stanford University Press. Appendix *Three kinds of vowels, two kinds of rhythm* was distributed by IULC, 1981.

BROE, M. and PIERREHUMBERT, J. B. (eds) (2000) *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, Cambridge: Cambridge University Press.

BRUCE, G. (1977) *Swedish Word Accents in Sentence Perspective*, Lund: Gleerup.

BRUCE, G. (1987) How floating is focal accent, in K. GREGERSEN and H. BASBØLL (eds) *Nordic Prosody IV*, pp. 41–49. Odense: Odense University Press.

BRUCE, G. (1990) Alignment and composition of tonal accents: comments on Silverman and Pierrehumbert's paper. See Kingston & Beckman (1990), pp. 107–114.

BUSH, R. (1999) Georgian yes–no question intonation, in *Phonology at Santa Cruz*, Vol. 6, pp. 1–11. Santa Cruz, CA: UC Santa Cruz.

CAMBIER-LANGEVELD, T. (2000) *Temporal Marking of Accent and Boundaries*. The Hague: Thesus (Subsidiary of Holland Academic Graphics).

CAMBIER-LANGEVELD, T. and TURK, A. E. (1999) A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, 27, 255–80.

CHO, T. and KEATING, P. A. (2001) Articulatory and acoustic studies of domain-initial strenghtening in Korean. *Journal of Phonetics*, 29, 155–190.

CLEMENTS, G. N. and FORD, K. C. (1980) On the phonological status of downstep in Kikuyu. See Goyvaerts (1980), pp. 309–357.

COHN, A. (1990) *Phonetic and Phonological Rules of Nasalization*, PhD dissertation, UCLA.

CRUTTENDEN, A. (1997) *Intonation*, (2nd edn), Cambridge University Press, Cambridge.

CUTLER, A. and LADD, D. R. (eds) (1983) *Prosody – Models and Measurements*, Berlin: Springer.

DAINORA, A. (2001) *An Empirically Based Probailistic Model of Intonation of American English*, PhD dissertation, University of Chicago, Chicago.

DAINORA, A. (2002) Does intonational meaning come from tones or tunes? Evidence against a compositional approach. See Bel & Marlien (2002), pp. 235–338.

ELORDIETA, G. (1997) Accent, tone and intonation in Lekeitio Basque, in F. MARTÍNEZ-GIL and A. MORALES-FRONT, (eds) *Issues in the Phonology and Morphology of the Major Iberian Languages*, pp. 4–78. Washington, DC: Georgetown University Press.

FACE, T. L. (2002) *Intonational Marking of Contrastive Focus in Spanish*, Munich: LINCOM Europa.

FÉRY, C. (1993) *German Intonational Patterns*, Tübingen: Niemeyer.

FITZPATRICK-COLE, J. (1999) The Alpine intonation of Bern Swiss German, in *Proceedings of the International Congress of Phonetic Sciences*, Vol. 14, pp. 941–944.

FOUGERON, C. (2001) Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29, 109–36.

FOUGERON, C. and KEATING, P. A. (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the American Acoustic Association*, 101, 3728–40.

FROTA, S. (1998) *Prosody and Focus in European Portuguese*, PhD dissertation, Garland, New York: University of Lisbon. 2000.

FROTA, S. (2002) Tonal association and target alignment in European Portuguese nuclear falls. See Gussenhoven & Warner (2002), pp. 305–418.

FUJISAKI, H. (1983) Characteristics of voice fundamental frequency in speech and singing. in MacNeilage, P.F., (ed.) *The Production of Speech*, pp. 37–47. New York/Berlin: Springer.

GÅRDING, E. (1983) A generative model of intonation. See Cutler & Ladd (1983), pp. 11–25.

GOLDSMITH, J. A. (1976) *Autosegmental Phonology*, PhD dissertation, MIT. Distributed by IULC, 1976, Published by Garland Press, New York, 1979.

GOLDSMITH, J. A. (1980) English as a tone language. See Goyvaerts (1980), pp. 287–308.

GÓSY, M. and TERKEN, J. (1994) Question marking in Hungarian: timing and height of pitch peaks. *Journal of Phonetics*, 22, 269–81.

GOYVAERTS, D. (ed.) (1980) *Phonology in the 80s*, Gent. Story Scientia.

GRABE, E. (1998a) *Comparative intonational phonology. English and German*, PhD dissertation, University of Nijmegen. Published in Max Planck Institute Series in Psycholinguistics.

GRABE, E. (1998b) Pitch realization in English and German. *Journal of Phonetics*, 26, 129–43.

GRABE, E. (2001) The IViE labelling guide. Version 3. <http://www.phon.ox.ac.uk/~esther/ivywed/guide.html>.

GRABE, E., POST, B., NOLAN, F. and FARRAR, K. (2000) Pitch accent realization in four varieties of British English. *Journal of Phonetics*, 28, 161–85.

GRICE, M. (1995a) *The Intonation of Interrogation in Palermo Italian: Implications for Intonational Theory*, Tübingen: Niemeyer.

GRICE, M. (1995b) Leading tones and downstep in English. *Journal of Linguistics*, 12, 183–233.

GRICE, M., LADD, D. R. and ARVANITI, A. (2000) On the place of phrase accents in intonational phonology. *Phonology*, 17, 143–85.

GRØNNUM, N. (1991) Prosodic parameters in a variety of regional Danish standard languages. *Phonetica*, 47, 188–214.

GUSSENHOVEN, C. (1983a) A semantic analysis of the nuclear tones of English. Distributed by Indiana University Linguistics Club (IULC). Bloomington, Indiana. Published as chapter 6 in Gussenhoven (1984).

GUSSENHOVEN, C. (1983b) Stress shift and the nucleus. *Linguistics*, 21, 303–39. Reprinted in Gussenhoven (1984).

GUSSENHOVEN, C. (1986) Review of Selkirk (1984). *Journal of Linguistics*, 22, 455–74.

GUSSENHOVEN, C. (1988) Adequacy in intonation analysis: the case of Dutch, in H. VAN DER HULST and N. SMITH (eds) *Autosegmental Studies on Pitch Accent*, pp. 95–121. Dordrecht: Foris.

GUSSENHOVEN, C. (1991) The English rhythm rule as an accent deletion rule. *Phonology*, 8, 1–35.

GUSSENHOVEN, C. (1999) Discreteness and gradience in intonational contrasts. *Language and Speech*, 42, 281–305.

GUSSENHOVEN, C. (2000a) The boundary tones are coming: on the non-peripheral natural of boundary tones. See Broe & Pierrehumbert (2000), pp. 132–151.

GUSSENHOVEN, C. (2000b) The lexical tone contrast of Roermond Dutch in Optimality Theory. See Horne 2000, pp. 129–167. Rutgers Optimality Archive 381.

GUSSENHOVEN, C. (2002) Intonation and interpretation: phonetics and phonology. See Bel & Marlien (2002), pp. 47–57.

GUSSENHOVEN, C. and BRUCE, G. (1999) Word prosody and intonation, in H. VAN DER HULST, (ed.) *Word Prosodic Systems in the Languages of Europe*, pp. 233–271. Berlin: Mouton de Gruyter.

GUSSENHOVEN, C. and RIETVELD, A. (1992) Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics*, 20, 283–303.

GUSSENHOVEN, C., TERKEN, J. and RIETVEL, T. (1999) Transcription of Dutch intonation – courseware. <http://lands.let.kun.nl/todi >.

GUSSENHOVEN, C. and VAN DER VLIET, P. (1999) The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99–135.

GUSSENHOVEN, C. and WARNER, N. (eds) (2002) *Laboratory Phonology 7*, Berlin and New York: Mouton de Gruyter.

HAAN, J. (2002) *Speaking of Questions: An exploration of Dutch Question Intonation*, Utrecht: Netherlands Graduate School of Linguistics.

HAYES, B. (1989) The prosodic hierarchy in meter, in P. KIPARSKY and G. YOUMANS (eds) *Phonetics and Phonology. Rhythm and Meter*, pp. 201–260. New York: Academic Press.

HAYES, B. and LAHIRI, A. (1991) Bengali intonational phonology. *Natural Language and Linguistic Theory*, 9, 47–96.

HIRST, D. and DI CRISTO, A. (eds) (1998) *Intonation System: Survey of Twenty Languages*, Cambridge, UK: Cambridge University Press.

HORNE, M. (1990) Empirical evidence for a deletion analysis of the rhythm rule in English. *Linguistics*, 28, 959–81.

HORNE, M. (ed.) (2000) *Prosody: Theory and Experiment. Studies Presented to Gösta Bruce*, Dordrecht/Boston/London. Kluwer.

HUALDE, J., ELORDIETA, G., GAMINDE, I. and SMILJANIĆ, R. (2002) From pitch-accent to stress-accent in Basque. See Gussenhoven & Warner (2002), pp. 547–584.

INKELAS, S. (1989) *Prosodic Constituency in the Lexicon*, PhD dissertation, Stanford University, Stanford CA.

INKELAS, S. and ZEC, D. (eds) (1990) *The Phonology-Syntax Connection*, Chicago: Chicago University Press.

JUN, S.-A. (1993) *The Phonetics and Phonology of Korean*, PhD dissertation, Ohio: Ohio State University.

JUN, S.-A. (1998) The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15, 189–226.

JUN, S.-A. (ed.) (2003) *Prosodic Typology and Transcription: A Unified Approach*, Oxford: Oxford University Press.

JUN, S.-A. and FOUGERON, C. (2000) A phonological model of French intonation, in A. BOTINIS (ed.) *Intonation: Analysis*, pp. 209–242. Dordrecht: Kluwer.

KANERVA, J. M. (1989) *Focus and Phrasing in Chichewa Phonology*, PhD dissertation, Palo Alto: Stanford University.

KEATING, P. (1990) Phonetic representations in a generative grammar. *Journal of Phonetics*, 18, 321–34.

KINGSTON, J. and BECKMAN, M. E. (eds) (1990) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Cambridge: Cambridge University Press.

KINGSTON, J. and DIEHL, R. L. (1994) Phonetic knowledge. *Language*, 70, 419–54.

LADD, D. R. (1980) *The Structure of Intonational Meaning: Evidence from English*, Bloomington: Indiana University Press.

LADD, D. R. (1983a) Peak features and overall slope. See Cutler & Ladd (1983), pp. 39–52.

LADD, D. R. (1983b) Phonological features of intonational peaks. *Language*, 59, 721–59.

LADD, D. R. (1984) Declination: a review and some hypotheses. *Phonology Yearbook*, 1, 53–74.

LADD, D. R. (1996) *Intonational Phonology*, Cambridge: Cambridge University Press.

LADD, D. R. (2000) Bruce, Pierrehumbert and the elements of intonational phonology. See Horne (2000), pp. 37–50.

LADD, D. R., FAULKNER, D., FAULKNER, H. and SCHEPMAN, A. (1999) Constant 'segmental anchoring' of f0 movements under changes in speech rate. *Journal of the American Acoustic Association*, 106 (3), 1543–54.

LADD, D. R., MENNEN, I. and SCHEPMAN, A. (2000) Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the American Acoustic Association*, 107(5), 2685–96.

LADD, D. R. and MONAGHAN, A. (1987) Modelling rhythmic and syntactic effects on accent in long noun phrases, in J. LAVER and M. JACK (eds) *Proceedings of the European Conference on Speech Technology*, Vol. 2, pp. 29–32. Edinburgh: CEP.

LADD, D. R. and MORTON, R. (1997) The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics*, 25, 313–42.

LAHIRI, A. and FITZPATRICK-COLE, J. (1999) Emphatic clitics and focus intonation in Bengali, in W. ZONNEVELD and R. KAGER (eds) *Phrasal Phonology*, pp. 119–144. Nijmegen: Nijmegen University Press.

LEBEN, W. R. (1975) The tones in English intonation. *Linguistic Inquiry*, 2, 69–107.

LIBERMAN, M. and SAG, I. (1974) Prosodic form and discourse function, in *Proceedings of the Chicago Linguistics Society*, Vol. 10, pp. 416–427.

LIBERMAN, M. Y. (1975) *The Intonational System of English*, PhD dissertation, MIT. Distributed by IULC, 1978. Published by Garland Press, New York, 1985.

LIBERMAN, M. and PIERREHUMBERT, J. (1984) Intonational invariance under changes in pitch range and length, in M. ARONOFF and R. T. OEHRLE (eds) *Language and Sound Structure*, pp. 157–233. Cambridge, MA and London, England: MIT Press.

LIBERMAN, M. Y. and PRINCE, A. (1977) On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249–336.

NESPOR, M. and VOGEL, I. B. (1986) *Prosodic Phonology*. Dordrecht: Foris.

O'CONNOR, J. and ARNOLD, G. (1973) *Intonation of Colloquial English*. London: Longman, Second edition.

OSTENDORF, M., PRICE, P. and SHATTUCK-HUFNAGEL, S. (1995) The Boston University Radio News Corpus. Technical report: Boston University ECS Technical Report ECS-95–001.

PIERREHUMBERT, J. B. (1980) *The Phonetics and Phonology of English Intonation*, PhD dissertation. MIT. Published by Garland Press, New York, 1990.

PIERREHUMBERT, J. B. (1990) Phonological and phonetic representations. *Journal of Phonetics*, 18, 375–94.

PIERREHUMBERT, J. B. (2000) Tonal elements and their alignment. See Horne (2000), pp. 11–36.

PIERREHUMBERT, J. B. and BECKMAN, M. E. (1988) *Japanese Tone Structure*, Cambridge MA: MIT Press.

PIERREHUMBERT, J. B. and HIRSCHBERG, J. (1990) The meaning of intonational contours in the interpretation of discourse, in P. COHEN, J. MORGAN and M. POLLACK (eds) *Intentions in Communication*, pp. 271–311. Cambridge MA: MIT Press.

PIERREHUMBERT, J. B. and STEELE, S. (1989) Categories of tonal alignment in English. *Phonetica*, 46, 181–96.

PIKE, K. L. (1945) *The Intonation of American English*, Ann Arbor, Michigan: University of Michigan Press.

POSER, W. J. (2001) *The Phonetics and Phonology of Tone and Intonation in Japanese*, PhD dissertation MIT 1984, Stanford: CSLI Publications.

POST, B. (2000) *Tonal and Phrasal Structures in French Intonation*, The Hague: Thesus (Subsidiary of Holland Academic Graphic.

PRIETO, P., VAN SANTEN, J. and HIRSCHBERG, J. (1995) Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429–51.

RIALLAND, A. and ROBERT, S. (2001) The intonational system of wolof. *Linguistics*, 39, 893–939.

RIETVELD, T. and GUSSENHOVEN, C. (1995) Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics*, 23, 375–85.

SELKIRK, E. O. (1978) On prosodic structure and its relation to syntactic structure, in T. FRETHEIM, (ed.) *Nordic Prosody. II*, pp. 268–271. Trondheim: TAPIR.

SELKIRK, E. O. (2000) The interaction of constraints on prosodic phrasing. See Horne (2000), pp. 231–261.

SHATTUCK-HUFNAGEL, S. (1989) Stress shift as the placement of phrase-level pitch markers. *Journal of the American Acoustic Association*, 86, S493.

SHATTUCK-HUFNAGEL, S. (1995) Pitch accent patterns in adjacent-stress vs. alternating stress words in American English, in *Proceedings XIIIth International Congresss of Phonetic Sciences*, Vol. 3, pp. 656–659.

SHATTUCK-HUFNAGEL, S., OSTENDORF, M. and ROSS, K. (1979) Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics*, 22, 357–88.

SHATTUCK-HUFNAGEL, S. and TURK, A. E. (1996) A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25, 193–246.

SOSA, J. M. (1991) *Fonética y Fonología de la Entonación del Español Hispanoamericano*, PhD dissertation, Amherst, MA: University of Massachusetts.

STEWART, J. M. (1966) The typology of the Twi tone system. *Bulletin of the Institute of African Studies 1.* Preprint.

't HART, J., COLLIER, R. and COHEN, A. (1990) *A Perceptual Study of Intonation: an Experimental-Phonetic Approach to Speech Melody*, Cambridge: Cambridge University Press.

TAFF, A. (1997) Intonation patterns in Unangan, in A. BOTINIS & G. KOUROUPETROGLOU and G. CARAYANNIS (eds) *Intonation: Theory, Models and Applications. Proceedings of an ESCA Workshop*, pp. 301–304. Athens (Greece): Department of Informatics, ESCA and University of Athens.

THORSEN, N. (1978) An acoustical analysis of Danish intonation. *Journal of Phonetics*, 6, 151–75.

THORSEN, N. (1983) Two issues in the prosody of Standard Danish. See Cutler & Ladd (1983), pp. 27–38.

TRAGER, G. L. and SMITH, H. L. Jr., (1951) *An Outline of English Structure*. Norman, OK: Battenburg Press, Reprinted by the American Council of Learned Societies, Washington, 1957.

TRUCKENBRODT, H. (1999) On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry*, 30, 219–55.

UHMANN, S. (1991) *Fokusphonologie: Eine Analyse Deutscher Intonationskonturen Im Rahmen der Nicht-Linearen Phonologie*. Tübingen: Niemeyer.

VAISSIÈRRE, J. (1983) Language-independent features. See Cutler & Ladd (1983), pp. 53–66.

VAN DEN BERG, R., GUSSENHOVEN, C. and RIETVELD, T. (1992) Downstep in Dutch: Implications for a model, in G. DOCHERTY and D. R. LADD (eds) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, pp. 335–359. Cambridge: Cambridge University Press.

VANDERSLICE, R. and LADEFOGED, P. (1972) Binary suprasegmental features and transformational word-accentuation rules. *Language*, 48, 819–38.

VARGA, L. (2002) *Stress and Intonation: Evidence from Hungarian*, Basingstoke, UK: Palgrave.

VOGEL, I., BUNNELL, T. and HOSKINS, S. (1995) The phonology and phonetics of the Rhythm Rule, in B. CONNELL and A. ARVANITI (eds) *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, pp. 111–127. Cambridge: Cambridge University Press.

WIGHTMAN, C. W., SHATTUCK-HUFNAGEL, S., OSTENDORF, M. and PRICE, P. J. (1992) Segmental durations in the vicinity of prosodic boundaries. *Journal of the American Acoustic Association*, 91, 1707–17.

XU, Y. (1998) Fundamental frequency peak delay in Mandarin. *Phonetica*, 58, 26–52.

XU, Y. (2002) Articulatory constraints and tonal alignment. See Bel & Marlien (2002), pp. 91–100.

YIP, M. (1989) Contour tones. *Phonology*, 6, 149–74.