

Chapter 9

Prosodic representations

Part I: Segment-to-Tone Association

Amalia Arvaniti

University of California, San Diego

9.1. Introduction

This section deals with the representation of intonation – i.e. the linguistically structured and pragmatically meaningful modulation of pitch – and the formal ways in which intonational elements are said to relate to the segmental string. Specifically, section 9.2 briefly reviews the experimental results that led to the development of the *autosegmental-metrical model of intonational phonology* (henceforth *AM*), the main principles of which are presented in section 9.3. Section 9.4 reviews experimental research that addresses several issues arising from the central tenet of *AM* that intonation involves the phonological *association* of tones with constituents of the prosodic hierarchy (phonological association gives rise to phonetic *alignment*, the temporal coordination between segments and tones; laboratory evidence for alignment is reviewed in D’Imperio, this volume). Section 9.5 presents an alternative model of “text-to-tune” association, the *Parallel Encoding and Target Approximation* model (Xu, 2005), and briefly reviews evidence against its main claim that pitch contours are fully specified on a syllable-by-syllable basis. Finally, section 9.6 discusses new experimental perspectives and approaches to the investigation of intonation (for a full review of experimental methods used in intonation research, see Prieto, this volume).

When studying intonation, one of the first questions that arise is how a tune is coordinated with the co-occurring segmental string. In order to be able to provide a satisfactory answer to this question, it is important to know what the structure of intonation might be. It is of course possible to reach the conclusion that such a structure (in the sense of a combination of discrete elements) does not exist: for instance, in their model OXIGEN, Grabe, Kochanski, and Coleman (2007) treat tunes as gestalts, as did some earlier researchers, such as Cooper, and Sorensen (1981).

This kind of approach to intonational structure is problematic for two reasons. First, if melodies are gestalts, their meaning should be unique and relatively constant across utterances. However, it has been repeatedly noted in the past 65 years at least that the same melody can be used with different meanings (for discussions, see Pike, 1945; Lehiste, 1970: 95 ff. and references therein; Arvaniti, 2007a; Ladd, 2008, chap. 1; for some recent evidence supporting this observation, see Baltazani, 2006a). Problems like these make it clear that viewing melodies as composites of smaller and meaningful elements is more likely to be successful in accounting for intonational meaning (for such an approach see Pierrehumbert and Hirschberg, 1990).

In addition, seeing melodies as gestalts is problematic because melodies do not simply shrink or stretch to fit the duration of the utterance with which they co-occur. Rather, parts of the melody appear to coordinate independently with parts of the segmental string (Arvaniti, 2007a, b; Ladd, 2008, chap. 2). Recent experimental evidence for this view is provided in Arvaniti, Ladd, and Mennen (2006a), and Arvaniti and Ladd (2009) who examine the intonation of Greek polar and wh-questions respectively and show that the shape of pitch contours is radically affected by the position of stressed syllables, the

location of the focal element, and the length of the utterance. These results cannot be accounted for if contour shape is seen as a primitive, but are compatible with the idea of discrete tones that can vary in their phonological association, and consequently in their phonetic realization, depending on the parameters mentioned above.

It is thus clear that melodies must be composed of some kind of primitives, the nature of which has preoccupied many researchers. Answers have varied extensively, from level tones (e.g. Pike, 1945; Trager and Smith, 1951), to configurations such as rises and falls (e.g. the IPO model, presented in detail in 't Hart, Collier, and Cohen, 1990), to elements that can span F₀ stretches of arbitrary length, such as the *head*, *pre-head*, *nucleus*, and *tail* of the British school of intonation (e.g. O'Connor and Arnold, 1973; Halliday, 1967, 1970). Choosing the right answer, however, is neither trivial nor a matter of taste, as the answer has important (and often empirically testable) consequences for our understanding of the relationship between the text and the tune. As a discussion of intonational primitives is beyond the scope of this chapter, in the remainder I simply assume that these primitives are static tones and combinations thereof (for supporting experimental evidence see Pierrehumbert and Beckman, 1988, and Arvaniti, Ladd, and Mennen, 1998; for a review of the issue of tonal primitives, see Arvaniti, to appear).

9.2. The relationship between text and tune: Early empirical evidence

The relationship between text and tune did not feature prominently in intonational models until researchers started examining acoustic data and the close connection between text and tune was uncovered. For example, IPO researchers ('t Hart et al., 1990) noted that some of the F₀ movements that constitute the primitives in their model showed tight temporal

coordination with stressed syllables (*prominence lending* pitch movements), while others tended to spread over several syllables and did not co-occur with either prominent words or stressed syllables (*non-prominence lending* pitch movements).

A significant influence on our current understanding of the relationship between text and tune has been Bruce (1977), a phonetic study of the Swedish lexical pitch accents, known as accent I, and accent II. Bruce showed that these accentual patterns could be successfully accounted for if both accents are seen as falls that differ in terms of their timing with respect to the accented syllable. By doing so, Bruce essentially distinguished the *phonological* connection between tonal elements and segmental structure from the coordination between the two in real time.¹ Following Ladd (1983), these two properties have become known as *association* and *alignment* respectively. Simplifying somewhat, we could say that in Swedish both accent I and accent II associate with the same syllable but align differently with it (for an analysis along these lines, see Bruce, 1987).

In addition, Bruce showed that the second pitch peak seen in words with accent II was not part of the pitch accent per se but the reflex of phrasal tones. Thus his study highlighted the distinction between tonal elements that co-occur with prominent syllables and tonal elements that co-occur with phrasal boundaries. It further showed that despite their different origin in the grammar, phrasal tones and lexical tones are part of the same representation and realized in a similar fashion, rather than forming distinct layers of tonal structure.

9.3. The autosegmental-metrical model of intonational phonology

Although the empirical research discussed above provided detailed evidence for the relationship between text and tune, models like that of the IPO or Bruce do not formalize their findings in terms of phonological representations. Such formalizations appeared in Leben (1973), Liberman (1978), and Goldsmith (1979). The combined insight of these early approaches culminated in Pierrehumbert's thesis (Pierrehumbert, 1980), which gave rise to the AM model. Pierrehumbert's dissertation is an early example of laboratory phonology as it combines a formal phonological analysis of intonation with instrumental and quantitative data.

Pierrehumbert proposed that English tunes are composed of high (H) and low (L) tones which are linearly ordered on an autosegmental tier and associated to strong nodes and edges of metrical trees. Thus, similarly to Bruce's model (and unlike early autosegmental accounts of English intonation), these H and L tones do not exhaustively represent the course of F₀. Phonetically, the reflexes of H and L tones are *tonal targets* (most often local pitch peaks and troughs respectively), with the pitch between them being generated by interpolation. Thus, both at the phonological and at the phonetic level melodies are underspecified.

A corollary of underspecification – in combination with the association properties of tones discussed in more detail immediately below – is that the number of tones and the number of *tone bearing units* (henceforth *TBUs*) need not match: in some instances, several TBUs may not be associated to tones, while in others, several tones may associate with the same TBU, giving rise to *tonal crowding* (Bruce, 1977). Experimental evidence for underspecification was first provided in Pierrehumbert and Beckman (1988) who showed that the F₀ of Tokyo Japanese accentual phrases with unaccented words can be

captured by positing just a H phrasal tone co-occurring with the second mora of the phrase and a L% boundary tone co-occurring with the right phrasal boundary. In their data these two landmarks are separated by increasingly more segmental material, leading to an increasingly shallower F0 slope between H and L%. This change of slope cannot be explained if all the syllables between the H and L% are specified for tone, as in earlier models of Japanese, but is compatible with the idea that the melody is underspecified both phonologically and phonetically. Similarly, a host of laboratory studies have demonstrated that tonal crowding results in controlled variability in the realization of tones, including tone truncation (Bruce, 1977; Grice, 1995a; Arvaniti, 1998; Grabe, 1998; Grabe et al., 2000; Arvaniti and Ladd, 2009), tonal undershoot (Bruce, 1977; Prieto, 1998; Arvaniti, Ladd, and Mennen, 2000, 2006a, 2006b; Arvaniti and Ladd, 2009), and the temporal re-alignment of tones (Silverman and Pierrehumbert, 1990; Arvaniti and Ladd, 2009).² None of these effects should be observed if melodies were gestalts or if all syllables were tonally specified, as such views of intonation predict uniform expansion and compression of their primitives.

As noted, Pierrehumbert's system also incorporated the distinction between tones that associate with prominent syllables, that is strong nodes in the metrical tree, and tones that associate with utterance edges. The former, known as *pitch accents*, are notated with an asterisk, e.g. H*; the latter, known as *boundary tones*, are notated with a percent, e.g. H%. Pierrehumbert also noted that English tunes included a stretch between the last pitch accent (or *nuclear accent*) and the following boundary tone, where F0 was clearly not a simple interpolation between these two tones. She analyzed these stretches as the reflex of *phrase accents*, unassociated (floating) tones with variable realization. Pierrehumbert also

proposed that in bitonal accents, such as L*+H,³ the starred tone is metrically strong and phonologically associated to the accented syllable, while the unstarred tone is a floating tone that precedes or follows the starred tone by a fixed amount of time.⁴ In Pierrehumbert (1980) this analysis is supported by quantitative data showing that the H of L*+H is located approximately 200 ms after the L* and its location does not correlate with the segmental structure of the accented syllable.

Phonological association is formalized in more detail in Beckman and Pierrehumbert (1986), and in Pierrehumbert and Beckman (1988) where new formalisms are introduced on the basis of experimental data from Japanese. Specifically, in Pierrehumbert and Beckman (1988) metrical trees are replaced by prosodic trees which represent both prominence relationships and constituency. These prosodic trees differ from those proposed by Selkirk (1984), and Nespor and Vogel (1986) by permitting limited extrametricality (as opposed to exhaustive parsing), and a language-specific number of prosodic levels (for a discussion see Frota, this volume).⁵

Crucially, tones may associate with any node in the tree, including phrasal nodes, and not just with specific TBUs (or with prosodic boundaries as in Hayes and Lahiri, 1991). This new formalization had several advantages. First, it provided a stricter formalization of the association of phrase accents which are now seen as phrasal tones that associate with the smaller of two phrasal constituents posited for English, the *intermediate phrase* or *ip* (the other being the *intonational phrase*, or *IP*).

In addition, the phonological association of tones to prosodic nodes of different levels is said to determine their scaling (i.e. differences in pitch level) as well. Simplifying considerably, Pierrehumbert and Beckman (1988) show that the scaling of Hs and Ls in

Japanese is determined by their association, with tones associated to lower constituents exhibiting less extreme scaling than those associated to higher constituents; the latter can also affect the scaling of neighboring tones. Similarly, downstep (*catathesis* in the terminology of Pierrehumbert and Beckman) is shown to apply in Japanese within but not across intermediate phrases. Formalizing tonal association along these lines provides a way to account for both local and long-distance effects on tonal scaling while retaining a linear representation of tones, that is without resorting to hierarchical representations or the notion of registers (for an alternative view of hierarchical representations in intonation, see Ladd, 2008, chap. 8 and references therein).

The analysis of Pierrehumbert and Beckman (1988) was successfully applied by Truckenbrodt (2002) to data from Southern varieties of German which exhibit downstep of all peaks within a non-final IP except for the nuclear peak which is scaled at the same level as the first peak in its phrase (upstep). Truckenbrodt accounts for this pattern by positing that pitch accents are properties of the phonological phrase (roughly equivalent to Pierrehumbert and Beckman's ip) except for the nuclear accent which may be associated with the IP instead; since prosodic levels determine pitch registers, the difference in association between the nuclear and prenuclear accents accounts for their difference in scaling.

As noted by Pierrehumbert and Beckman (1988), one drawback of allowing tones to associate with prosodic nodes of various levels is that in some instances phonological representations provide no ordering among tonal elements, such as lexical and phrasal tones, thus providing no guidance as to how these elements are to be produced in time. To address this issue, Pierrehumbert and Beckman formulate a constraint which is informally

stated as follows: “a substantive element that is associated to a node in the prosodic tree must either also be associated to the center [head constituent at some lower level of the tree] of the constituent or be realized somewhere at its left or right periphery” (Pierrehumbert and Beckman, 1988: 131). This distinction between *central* and *peripheral* tones regulates the temporal order of tones in phonetic realization.

9.4. Further empirical evidence for association

9.4.1. The internal structure of pitch accents

In the early autosegmental work, a crucial assumption that was largely accepted without comment was that the relationship between phonological association and phonetic alignment is straightforward: elements that associate in phonology co-occur in time (e.g. Goldsmith, 1979, chap. 1 and 3). This idea was challenged by a series of empirical results and has led to new proposals regarding association that I briefly review here.

One of the first studies showing that the relationship between phonological association and phonetic alignment is not simple was that of Silverman and Pierrehumbert (1990) who examined the phonetic realization of prenuclear H* accents in American English. Their results showed that the pitch peaks of these accents do not always co-occur with the accented vowel but may show *peak delay* the extent of which is determined by the segmental makeup of the accented syllable’s rhyme and its distance from the following accent. Empirical evidence for peak delay and its contrastive potential within a linguistic system, as well as its susceptibility to dialectal variation, have since been documented in many typologically unrelated languages with widely different prosodic systems (for a review, see D’Imperio, this volume).

Arvaniti et al. (1998, 2000) address a complication that arises from peak delay with respect to the temporal patterning of bitonal pitch accents. Specifically, Arvaniti et al. examined rising accents of Greek that can be plausibly analyzed as bitonal L+H (Arvaniti and Ladd, 1995) and show that neither is it the case that one of the tones aligns with respect to the other (as the analyses of Pierrehumbert, 1980, or Beckman and Pierrehumbert, 1986, would predict) nor are they both aligned with respect to the same segmental landmark (as a possible association of both tones with the same TBU would predict). Rather, the L tone occurs slightly before the onset of the accented syllable and the H tone occurs early in the first post-accentual vowel. Thus, the distance between the two tones is variable and positively correlates with the duration of the accented syllable (Arvaniti et al., 1998). The finding that tones may align stably with segmental landmarks gave rise to the “segmental anchoring” hypothesis soon afterwards (Ladd et al., 1999). In addition, these results served to consolidate the distinction between discrete phonological association with a specific structural position and gradient phonetic alignment with segmental material (Arvaniti et al., 2000). Finally, they showed that it is possible for grouped tones to be aligned independently of each other.

The independent alignment of grouped tones is formalized in Grice (1995b) who proposed two distinct representations for bitonal accents: accents with leading tones, e.g. L+H*, are represented as clusters in which both tones independently associate with the same TBU, while accents with trailing tones, e.g. L*+H, are represented as contours that associate with a given TBU as a group (this particular type of association echoes the representation of the HL lexical pitch accent of Japanese by Pierrehumbert and Beckman, 1988). Although Grice’s proposal that this difference in association is tied to the presence

of a trailing vs. a leading tone is not empirically supported cross-linguistically (e.g. Arvaniti et al., 1998, for Greek; Ladd and Schepman, 2003, and Arvaniti and Garding, 2007, for English), the notion of two association options with concomitant effects on phonetic alignment is certainly valid and has been supported by experimental data from Portuguese (Frota 2002). Taken all together, the results briefly reviewed here suggest that tones may align in different ways with the segmental string and with each other both within and across languages. In turn, these different modes of alignment can be formally represented as different types of phonological association.

9.4.2. Primary and secondary association

The complex nature of phonological association is addressed in detail in Pierrehumbert and Beckman (1988) who introduced the notion of *secondary association* to account for a series of data from Japanese. As mentioned, in Japanese each accentual phrase is associated with two tones, the phrasal H and the boundary L%. Typically, phrasal tones are realized at the boundary of the constituents with which they are associated, but the Japanese phrasal H co-occurs with the second sonorant mora of its accentual phrase if the first syllable is short (otherwise it co-occurs with the first sonorant mora), while the boundary L% co-occurs with the first mora of the *following* phrase if its first syllable is short and unaccented (otherwise it is realized at the right edge of its accentual phrase). Pierrehumbert and Beckman (1988: 126 ff.) account for this variability by positing that in addition to their primary associations with the accentual phrase node, the phrasal H and boundary L% have secondary associations to the second sonorant mora and the first mora

of the following accentual phrase respectively; these are realized when certain conditions are met.

The notion of secondary association was taken up by Grice, Ladd, and Arvaniti (2000) who examined a variety of tunes in English, German, Dutch, Standard and Cypriot Greek, and Standard and Transylvanian Hungarian and Romanian and showed that the variability in the phonetic realization of these tunes can be accounted for if we assume the existence of a phrase accent with both a primary and a secondary association. Phrase accents have a primary association to a phrasal constituent (the exact nature of which is left unspecified in Grice et al.) but also a secondary association to a specific TBU, such as the last metrically strong syllable of their phrase or the first syllable of the last word (independently of metrical strength). When this TBU is associated with another tone, the phrase accent is realized at the phrasal boundary, i.e. by means of its primary association; but when the TBU is available, the secondary association takes over instead. Quantitative support for this view of phrase accents has been reported for Cypriot Greek (Arvaniti, 1998), Standard Greek (Arvaniti et al., 2006a, Arvaniti and Ladd, 2009), French (Welby, 2004), Dutch (Lickley, Schepman and Ladd, 2005), and English (Barnes et al., 2006).

A different use of secondary association is developed in Prieto, D'Imperio, and Gili Fivela (2005). Prieto et al. present data from Catalan and Italian involving bitonal accents which are phonologically contrastive but show phonetically only small differences in peak alignment. They propose that these differences be represented by means of secondary association of one of the tones to a syllabic or segmental edge, with peak delay being the default. Face and Prieto (2006-7) further apply this idea to a three-way contrast of peak alignment in Spanish.

This use of secondary association differs from the use of the mechanism by Pierrehumbert and Beckman (1988), Grice et al. (2000), and others, where secondary association accounts for a discrete alternating pattern that does not affect meaning: in Prieto et al. (2005) secondary association is used *in conjunction* with the primary association, not in place of it, and thus does not involve an alternation in the surface realization of a tone. Arvaniti et al. (2006b) argue that this use is essentially a formalization of phonetic alignment rather than of association proper and propose alternative ways of representing the patterns uncovered by Prieto et al. that do not require the use of secondary association.

The notion of secondary association is also explored within OT by Gussenhoven (2000a, b) in his analysis of Roermond Dutch, a variety with a lexical tonal contrast. Simplifying somewhat, in Roermond Dutch phrasal tones show controlled variability in spreading that is determined by the location of the focal word (which carries a *postlexical* pitch accent) and (for some phrasal tones) the possible presence of a *lexical* H tone on the focal word. Gussenhoven accounts for these differences in the spreading of phrasal tones by drawing a distinction between *alignment* (in the OT sense) and *association*: phrasal tones are aligned with the right edge of their intonational phrase and also show alignment to the left of a preceding tone. This formalization of alignment does not entail association with a particular TBU, though phrasal tones may also have such an association; e.g. the L_i of Roermond Dutch declaratives associates with the focal word's second mora, if the word is bimoraic and its second mora is not associated with a lexical H. As noted by Gussenhoven (2000a, b), this analysis accounts for tonal spreading and more generally for

the durational aspects of tonal implementation without the need of positing additional targets in phonetic realization (but at the expense of additional alignment constraints).

9.4.3. Association and the temporal ordering of tones

In addition to the issues they raise about the role of secondary association, the Roermond data question a stable assumption about the realization of boundary tones, namely that they occur as close as possible to the boundary with which they are phonologically associated. This idea, which is largely supported by both qualitative and experimental results, is connected to the well-established notion of the linear ordering of autosegments and is an instance of the prohibition of crossed association lines (Goldsmith, 1979, Sagey 1988), the aim of which is to ensure the same left-to-right sequence among associated elements in each tier. (Maintaining temporal order is also the motivation behind the distinction of Pierrehumbert and Beckman, 1988, between central and peripheral tones, discussed in section 9.3.).

Gussenhoven's Roermond data provide rare evidence that F0 contours do not always reflect the expected order of tones. Specifically, in this variety boundary tones are realized *before* a lexical H tone, if this tone occupies the last mora of the intonational phrase. In order to account for this unusual pattern, Gussenhoven posits that $\text{ALIGN}_{\text{T}}\text{RIGHT}$, the constraint for the right alignment of Roermond Dutch boundary tones, is outranked by a similar constraint for the lexical tone, $\text{ALIGN}_{\text{LEX}}\text{RT}$, and suggests that other AM models cannot account for this pattern (for further discussion see also Gussenhoven, 2004, chap. 7).

Although Gussenhoven's contention appears to be *prima facie* correct – or at the very least, difficult to formalize except in *ad hoc* ways – ultimately the issues raised by the Roermond Dutch boundary tones and more generally by the timing patterns discussed in Gussenhoven (2000a, b) and Prieto et al. (2005) have to do with the interface between phonology and phonetics. If it is accepted that the phonological representation of intonation need not be faithful to surface facts any more than other phonological representations are, it is possible to represent tonal contrasts making sparing use of formalisms such as secondary association and alignment constraints, and to view differences in tonal timing, spread, and duration as the realm of phonetic realization, as advocated, among many, by Pierrehumbert and Beckman (1988), Silverman and Pierrehumbert (1990), Elordieta and Calleja (2005), Arvaniti et al. (2006b), Arvaniti (2007a), and Ladd (2008, chap. 5).

9.5. An alternative view of tonal association

Most of the experimental research reviewed above has been couched in terms of the AM model of intonational phonology. Results have led, as is inevitable, to revisions of Pierrehumbert's original model, and also to the development of alternatives, such as ToDI (*Transcription of Dutch Intonation*; Gussenhoven et al. 2003), and IViE (*Intonational Variation in English*; Grabe, 2001). These models deviate in some respects from the original AM model – e.g. both ToDi and IViE dispense with the phrase accent and make boundary tones optional – yet they retain the main assumptions of AM, that is, the notion that melodies are composed on tones, that tunes remain largely underspecified and that phonological structure mediates between F0 and meaning. This is so because by and large

the experimental data have proved compatible with AM and not easily amenable to alternative interpretations.

A radically different view is taken by Xu and colleagues in the Parallel Encoding and Target Approximation (PENTA) model (e.g. Xu, 2005; Xu and Xu, 2005, among many). PENTA rests on the idea that F0 directly encodes “communicative functions” (such as statement vs. question or the location of focus). Crucially, every syllable in an utterance is specified for F0, and F0 events are synchronized with syllables. Although PENTA is based almost exclusively on Mandarin, Xu and Xu (2005) have also applied it with some success to English declaratives with narrow focus.

Their results, however, are not incompatible with existing AM accounts, while results from several other studies are incompatible with main PENTA assumptions. For example, many studies show that the connection of focus with raised and expanded pitch range advocated by PENTA does not hold for all languages and melodies (Baumann, Grice, and Steindamm, 2006, for German; Arvaniti et al., 2006a, and Arvaniti and Ladd, 2009, for Greek; Chen and Gussenhoven, 2008, and Chen, to appear, for Mandarin;⁶ Pan, 2007, for Taiwanese; Gordon, 2008, for Chickasaw). Similarly, many studies show that both anticipatory and progressive coarticulation can affect tonal realization (among many, Silverman and Pierrehumbert, 1990, for English; Arvaniti et al., 2006a, b, and Arvaniti and Ladd, 2009, for Greek; Kochanski, Shih, and Jing, 2003, and Shih, Kochanski, and Yoon, 2007, for Mandarin). Finally, as Arvaniti et al. (2006a) and Arvaniti and Ladd (2009) amply demonstrate, PENTA does not have predictive power, as it is not conducive to appropriate generalizations about the meaning and structure of tunes. Thus, the experimental evidence so far argues against the full specification of F0 contours and for a

principled distinction between phonetic realization and an abstract phonological level of representation that involves tones and their association with the segmental string.

9.6. Promising new directions

Recently, there have been several new directions in experimental research on intonation. Two of these deal particularly with different aspects of association. First, recent laboratory research examines the connection between the F0 of an utterance and articulatory gestures, as recorded using electromagnetic midsagittal articulography or *EMMA* (Mücke et al., 2006; Bauman et al., 2007; D'Imperio et al., 2007; Mücke and Hermes, 2007; Prieto et al., 2007; Gao, 2008; Mücke et al., 2009). Research so far shows that this temporal coordination is in many cases tighter than that between acoustic events and F0. More recently this type of coordination has been formalized within Articulatory Phonology as a phasing relationship between articulatory and tonal events on the basis of data primarily from Mandarin (Gao, 2008). Gao's approach is further explored on the basis of German data in Mücke et al. (2009) who also note the similarity of the phasing relations depicted in gestural scores with the formal proposal of Grice (1995b) regarding the association of bitonal accents with their TBUs (see section 9.4.1). Such data are clearly valuable for understanding how intonation is phonetically implemented, but can also shed light on issues of phonological representation, such as the association of bitonal accents with the segmental string.

Finally, in the past decade or so, researchers have been testing experimentally the interaction of intonation with other components of the grammar both from formal perspectives and from the perspective of language processing and acquisition. One such

strand of research examines the mechanisms that result in the association of specific pitch accents and phrasal tones with specific words and boundaries (traditionally, these relationships were either taken for granted, or examined on the basis of impressionistic data alone; among many, Ladd, 1980; Bolinger, 1989; Selkirk, 1995; Steedman, 1991, 2000). For instance, German, Pierrehumbert, and Kaufmann (2006) tested the extent to which speakers can accent a stranded preposition that is seen as new information, Beaver et al. (2007) used production and perception experiments to examine second occurrence focus, and Swerts, Krahmer, and Avesani (2002) explored cross-linguistic differences in focus marking by means of intonation using Dutch and Italian semi-spontaneous speech. In addition, several studies have employed eye-tracking and ERPs (event-related potentials) to observe the effects of intonational and other prosodic cues in language processing and acquisition (Dahan, Tanenhaus, and Chambers, 2002; Salverda et al., 2007; Ito and Speer, 2008; see also Speer and Ito, 2009, for a review of the acquisition literature). These newer techniques together with well-established methods promise to provide us with a better understanding of the function of intonation in discourse and its interaction with other components of the grammar, but should also lead to a better understanding of intonational structure itself.

References

Arvaniti, Amalia (1998). 'Phrase Accents Revisited: Comparative Evidence from Standard and Cypriot Greek', in *Proceedings of the 5th International Conference on Spoken Language Processing*, vol. 7: 2883-2886.

- Arvaniti, Amalia (2007a). 'On the Relationship Between Phonology and Phonetics (Or Why Phonetics is not Phonology)', in *ICPhS XVI: 16th International Congress of Phonetic Sciences* (Special Session: Between Meaning and Speech: On the Role of Communicative Functions, Representations and Articulations), pp. 19-24.
- Arvaniti, Amalia (2007b). 'Greek Phonetics: The State of the Art', *Journal of Greek Linguistics* 8: 97-208.
- Arvaniti, Amalia, and Mary Baltazani (2005). 'Intonational Analysis and Prosodic Annotation of Greek Spoken Corpora', in S. Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, pp. 84–117. Oxford: Oxford University Press.
- Arvaniti, Amalia (to appear). 'Levels Versus Configuration and the Representation of Intonation', in M. van Oostendorp, C. Ewen, B. Hume, and K. Rice (eds.), *Companion to Phonology*. Wiley-Blackwell.
- Arvaniti, Amalia, and Gina Garding (2007). 'Dialectal Variation in the Rising Accents of American English', in J. Cole, and J. H. Hualde (eds.), *Papers in Laboratory Phonology* 9, pp. 547-576. Berlin, New York: Mouton de Gruyter.
- Arvaniti, Amalia, and D. Robert Ladd (1995). 'Tonal Alignment and the Representation of Accentual Targets', in *Proceedings of the XIIIth International Congress of Phonetic Sciences*, vol. 4, pp. 220-223.
- Arvaniti, Amalia, and D. Robert Ladd (2009). 'Greek Wh-Questions and the Phonology of Intonation', *Phonology* 26: 43-74.
- Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen (1998). 'Stability of Tonal Alignment: The Case of Greek Prenuclear Accents', *Journal of Phonetics* 26: 3-25.

- Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen (2000). 'What Is A Starred Tone? Evidence from Greek', in M. Broe, and J. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, pp. 119-13. Cambridge: Cambridge University Press.
- Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen (2006a). 'Phonetic Effects of Focus and "Tonal Crowding" in Intonation: Evidence from Greek Polar Questions', *Speech Communication* 48: 667-696.
- Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen (2006b). 'Tonal Association and Tonal Alignment: Evidence from Greek Polar Questions and Contrastive Statements', *Language and Speech* 49: 421-450.
- Baltazani, Mary (2006a). 'Intonation and Pragmatic Interpretation of Negation in Greek', *Journal of Pragmatics* 38: 1658-1676.
- Baltazani, Mary (2006b). 'Focusing, Prosodic Phrasing, and Hiatus Resolution in Greek', in L. Goldstein, D. Whalen, C. Best (eds.), *Laboratory Phonology 8*, pp. 473-494. Berlin: Mouton de Gruyter.
- Barnes, Jonathan, Stefanie Shattuck-Hufnagel, Alejna Brugos, and Nanette Veilleux (2006). 'The Domain of Realization of the L- Phrase Tone in American English', in *Speech Prosody 2006*. http://aune.lpl.univ-aix.fr/~sprosig/sp2006/contents/papers/PS3-11_0163.pdf
- Barnes, Jonathan, Nanette Veilleux, Alejna Brugos, and Stefanie Shattuck-Hufnagel (2008). 'Alternatives to F0 Turning Points in American English Intonation', *Journal of the Acoustical Society of America* 124: 2497.

- Baumann, Stefan, Martine Grice, and Susanne Steindamm (2006). 'Prosodic Marking of Focus Domains – Categorical Or Gradient?', in *Speech Prosody 2006*.
http://aune.lpl.univ-aix.fr/~sprosig/sp2006/contents/papers/PS3-09_0065.pdf
- Baumann, Stefan, Johannes Becker, Martine Grice, and Doris Mücke (2007). 'Tonal and Articulatory Marking of Focus in German', in *ICPhS XVI: 16th International Congress of Phonetic Sciences*, pp. 1029-1032.
- Beaver, David, Brady Z. Clark, Edward Flemming, Florian T. Jaeger, and Maria Wolters (2007). 'When Semantics Meets Phonetics: Acoustical Studies of Second-Occurrence Focus', *Language* 83: 245-276.
- Beckman, Mary E., and Jan Edwards (1994). 'Articulatory Evidence for Differentiating Stress Categories', in P. A. Keating (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, pp. 7-33. Cambridge: Cambridge University Press.
- Beckman, Mary E., and Janet Pierrehumbert (1986). 'Intonational Structure in English and Japanese', *Phonology Yearbook* 3: 255-310.
- Bolinger, Dwight (1989). *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford: Stanford University Press.
- Bruce, Gösta (1977). *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.
- Bruce, Gösta (1987). 'How Floating is Focal Accent?', in K. Gregersen and H. Basbøll (eds.), *Nordic Prosody IV*, pp. 41-49. Odense: Odense University Press.
- Campbell, Nick, and Mary E. Beckman (1997). 'Stress, Prominence, and Spectral Tilt', in A. Botinis, G. Kouroupetroglou, and G. Carayannis (eds.), *Intonation: Theory*,

Models and Applications (Proceedings of the ESCA Workshop on Intonation), pp. 67-70.

Chen, Yiya (to appear). 'Post-Focus F0 Range Suppression in Beijing Mandarin: Now You See It, Now You Don't', *WCCFL 28: Proceedings of the 28th West Coast Conference on Formal Linguistics*, 19-21 February 2010, USC.

Chen, Yiya, and Carlos Gussenhoven (2008). 'Emphasis and Tonal Implementation in Standard Chinese', *Journal of Phonetics* 36: 724-746.

Cooper, William, and John Sorensen (1981). *Fundamental Frequency in Sentence Production*. Heidelberg: Springer.

Dahan, Delphine, Michael K. Tanenhaus, and Craig G. Chambers (2002). 'Accent and Reference Resolution in Spoken-Language Comprehension', *Journal of Memory and Language* 47: 292-314.

D'Imperio, Mariapaola#

D'Imperio, Mariapaola, Robert Espresser, H el ene Loevenbruck, Caroline Menezes, No el Nguyen, and Pauline Welby (2007). 'Are Tones Aligned With Articulatory Events? Evidence from Italian and French', in J. Cole and J. I. Hualde (eds.), *Papers in Laboratory Phonology 9*, pp. 577-608. Berlin: Mouton de Gruyter.

Elordieta, Gorka, and Nagore Calleja (2005). 'Microvariation in Accentual Alignment in Basque Spanish', *Language and Speech* 48: 397-439.

Face, Timothy L., and Pilar Prieto (2006-2007). 'Rising Accents in Castilian Spanish: A Revision of Sp_ToBI', *Journal of Portuguese Linguistics* 5-6: 117-146.

Frota, Sónia (2002). 'Tonal Association and Target Alignment in European Portuguese Nuclear Falls', in C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology 7*, pp. 387-418. Berlin/New York: Mouton de Gruyter.

Frota, Sónia #

Gao, Man (2008). *Tonal Alignment in Mandarin Chinese: An Articulatory Phonology Account*. Yale University unpublished dissertation.

German, James, Janet Pierrehumbert, and Stefan Kaufmann (2006). 'Evidence for Phonological Constraints on Nuclear Accent Placement', *Language* 82:151-168.

Goldsmith, John A. (1979). *Autosegmental Phonology*. Outstanding Dissertations in Linguistics. New York and London: Garland Publishing.

Gordon, Matthew (2008). 'Pitch Accent Timing and Scaling in Chickasaw', *Journal of Phonetics* 36: 521-535.

Grabe, Esther (1998). 'Pitch Accent Realisation in English and German', *Journal of Phonetics* 26: 129-144.

Grabe, Esther (2001). 'The IViE Labeling Guide'.

<http://www.phon.ox.ac.uk/files/apps/IViE//guide.html> (accessed February 3rd, 2010).

Grabe, Esther, Greg Kochanski, and John Coleman (2007). 'Connecting Intonation Labels to Mathematical Descriptions of Fundamental Frequency', *Language and Speech* 50: 281-310.

Grabe, Esther, Brechtje Post, Francis Nolan, and Kimberley Farrar (2000). 'Pitch Accent Realization in Four Varieties of British English', *Journal of Phonetics* 28: 161-185.

- Grice, Martine (1995a). *The Intonation of Interrogation in Palermo Italian: Implications for Intonational Theory*. Tübingen: Niemeyer.
- Grice, Martine (1995b). 'Leading Tones and Downstep in English', *Phonology* 12:183-233.
- Grice, Martine, Stefan Baumann, and Nils Jagdfeld (2009). 'Tonal Association and Derived Nuclear Accents: The Case of Downstepping Contours in German', *Lingua* 119: 881-905.
- Grice, Martine, D. Robert Ladd, and Amalia Arvaniti (2000). 'On the Place of "Phrase Accents" in Intonational Phonology', *Phonology* 17:143-185.
- Gussenhoven, Carlos (2000a). 'The Boundary Tones are Coming: On the Non-Peripheral Realization of Boundary Tones', in M. Broe, and J. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, pp. 132-151. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos (2000b). 'The Lexical Tone Contrast of Roermond Dutch in Optimality Theory', in M. Horne (ed.), *Prosody: Theory and Experiment. Studies Presented to Gösta Bruce*, pp. 129-167. Amsterdam: Kluwer.
- Gussenhoven, Carlos (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos, Toni Rietveld, Joop Kerkhoff, and Jacques Terken (2003). 'Transcription of Dutch Intonation – Courseware', <http://todi.let.kun.nl/ToDI/home.htm> (accessed, February 3rd, 2010).
- Harrington, Jonathan, Janet Fletcher, and Mary E. Beckman (2000). 'Manner and Place Conflicts in the Articulation of Accent in Australian English', in M. Broe, and J.

- Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*, pp. 40-51. Cambridge: Cambridge University Press.
- Halliday, M.A.K (1967). *Intonation and Grammar in British English*. The Hague and Paris: Mouton.
- Halliday, M.A.K (1970). *A Course in Spoken English: Intonation*. Oxford: Oxford University Press.
- 't Hart, Johan, René Collier, and Antonie Cohen (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.
- Hayes, Bruce, and Aditi Lahiri (1991). 'Bengali Intonational Phonology', *Natural Language and Linguistic Theory* 9: 47-96.
- Ito, Kiwako, and Shari R. Speer (2008). 'Anticipatory Effects of Intonation: Eye Movements During Instructed Visual Search', *Journal of Memory and Language* 58: 541–573.
- Kochanski, Greg, Chilin Shih, and Hongyan Jing (2003). 'Quantitative Measurement of Prosodic Strength in Mandarin', *Speech Communication* 41: 625-645.
- Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, MA: The MIT Press.
- Ladd, D. Robert (1980). *The Structure of Intonational Meaning: Evidence from English*. Bloomington: Indiana University Press.
- Ladd, D. Robert (1983). 'Phonological Features of Intonational Peaks', *Language* 59: 721-759.
- Ladd, D. Robert (2008). *Intonational Phonology* (2nd edition). Cambridge: Cambridge University Press.

- Ladd, D. R., Dan Faulkner, Hanneke Faulkner, and Astrid Schepman (1999). ‘Constant “Segmental Anchoring” of F0 Movements Under Changes in Speech Rate’, *Journal of the Acoustical Society of America* 106: 1543-1554.
- Ladd, D. Robert, and Astrid Schepman (2003). “‘Sagging Transitions” Between High Pitch Accents in English: Experimental Evidence’, *Journal of Phonetics* 31: 81–112.
- Leben, William R (1973). *Suprasegmental Phonology*. MIT dissertation.
- Lieberman, Mark Y. (1978). *The Intonational System of English*. MIT dissertation, reproduced 1978, Indiana University Linguistics Club, Bloomington, IN.
- Lickley, Robin J., Astrid Schepman, and D. Robert Ladd (2005). ‘Alignment Of “Phrase Accent” Lows in Dutch Falling Rising Questions: Theoretical and Methodological Implications’, *Language and Speech* 48: 157-183.
- Mücke, Doris, Martine Grice, Johannes Becker, Anne Hermes, and Stefan Baumann (2006). ‘Articulatory and Acoustic Correlates of Prenuclear and Nuclear Accents’, in *Speech Prosody 2006*. http://aune.lpl.univ-aix.fr/~sprosig/sp2006/contents/papers/PS3-08_0143.pdf
- Mücke, Doris, Martine Grice, Johannes Becker, and Anne Hermes (2009). ‘Sources of Variation in Tonal Alignment: Evidence from Acoustic and Kinematic Data’, *Journal of Phonetics* 37: 321-338.
- Mücke, D., and Anne Hermes (2007). ‘Phrase Boundaries and Peak Alignment: An Acoustic and Articulatory Study’, in *ICPhS XVI: 16th International Congress of Phonetic Sciences*, pp. 997-1000.

- Mücke, Doris, Hosung Nam, Pilar Prieto, and Louis Goldstein (2009). 'Coupling of Tone and Constriction Gestures in Catalan and German', poster presented at PaPI 09 [Phonetics and Phonology in Iberia], Las Palmas de Gran Canaria, 17-19 June 2009.
- Nespor, Marina, and Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.
- O'Connor, J. D., and G. F. Arnold (1973). *Intonation of Colloquial English*. London: Longman.
- Pan, Ho-Hsien. (2007). 'Focus and Taiwanese Unchecked Tones', in C. Lee, M. Gordon, and D. Büring (eds.), *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation*, pp. 195-213. Dordrecht: Springer.
- Pierrehumbert, Janet B. (1980). *The Phonology and Phonetics of English Intonation*. MIT dissertation, reproduced 1988, Indiana University Linguistics Club, Bloomington, IN.
- Pierrehumbert, Janet B. (1981). 'Synthesizing Intonation', *Journal of the Acoustical Society of America* 70: 985-995.
- Pierrehumbert, Janet, and Mary Beckman (1988). *Japanese Tone Structure*. Cambridge, MA: The MIT Press.
- Pierrehumbert, Janet, and Julia Hirschberg (1990). 'The Meaning of Intonation in the Interpretation of Discourse', in P. Cohen, J. Morgan, and M. Pollack, (eds.), *Intentions in Communication*, pp. 271-311. Cambridge MA: The MIT Press.
- Pike, Kenneth L (1945). *The Intonation of American English*. Ann Arbor: The University of Michigan Press.
- Prieto, Pilar (1998). 'The Scaling of the L Tone Line in Spanish Downstepping Contours', *Journal of Phonetics* 26: 261-282.

Prieto, Pilar #

- Prieto, Pilar, Mariapaola D'Imperio, and Barbara Gili Fivela (2005). 'Pitch Accent Alignment in Romance: Primary and Secondary Associations with Metrical Structure', *Language and Speech* 48: 359-396.
- Prieto, Pilar, Doris Mücke, Johannes Becker, and Martine Grice (2007). 'Coordination Patterns Between Pitch Movements and Oral Gestures in Catalan', *ICPhS XVI: 16th International Congress of Phonetic Sciences*, pp. 989-992.
- Prieto, Pilar, and Francisco Torreira (2007). 'The Segmental Anchoring Hypothesis Revisited: Syllable Structure and Speech Rate Effects on Peak Timing in Spanish', *Journal of Phonetics* 35: 473-500.
- Sagey, Elizabeth (1988). 'On the Ill-Formedness of Crossing Association Lines', *Linguistic Inquiry* 19: 109-118.
- Salverda, Anne Pier, Delphine Dahan, Michael K. Tanenhaus, Katherine Crosswhite, Mikhail Masharov, Joyce McDonough (2007). 'Effects of Prosodically Modulated Sub-Phonetic Variation On Lexical Competition', *Cognition* 105: 466-476.
- Selkirk, Elizabeth O. (1984). *Phonology and Syntax: The Relationship between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, Elizabeth O. (1995). 'Sentence Prosody: Intonation, Stress and Phrasing', in J. Goldsmith (ed.), *The Handbook of Phonological Theory*, pp. 550-569. Oxford: Blackwell.
- Shattuck-Hufnagel, Stefanie, Marie Ostendorf, and Ken Ross (1994). 'Stress Shift and Early Pitch Accent Placement in Lexical Items in American English', *Journal of Phonetics* 22: 357-388.

- Shih, Chilin, Greg Kochanski, and Su-Youn Yoon (2007). 'The Missing Link Between Articulatory Gestures and Sentence Planning', in *ICPhS XVI: 16th International Congress of Phonetic Sciences*, pp. 35-38.
- Silverman, Kim, and Janet Pierrehumbert (1990). 'The Timing of Prenuclear High Accents in English', in J. Kingston and M. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, pp. 72–106. Cambridge: Cambridge University Press.
- Speer, Shari R., and Kiwako Ito (2009). 'Prosody in First Language Acquisition – Acquiring Intonation as a Tool to Organize Information in Conversation', *Language and Linguistics Compass* 3: 90–110.
- Steedman, Mark (1991). 'Structure and Intonation', *Language* 67: 260-296.
- Steedman, Mark (2000). 'Information Structure and the Syntax-Phonology Interface', *Linguistic Inquiry* 31: 649-689.
- Swerts, Marc, Emiel Krahmer, and Cinzia Avesani (2002). 'Prosodic Marking of Intonation Status in Dutch And Italian: A Comparative Analysis', *Journal of Phonetics* 30: 629-654.
- Trager, George L., and H. L. Smith (1951). *An Outline of English Structure*. Norman, OK: Battenburg Press.
- Truckenbrodt, Hubert (2002). 'Upstep and Embedded Register Levels', *Phonology* 19: 77-120.
- Welby, Pauline (2004). 'The Structure of French Intonational Rises: A Study of Text-To-Tune Alignment', in B. Bel and I. Marlien (eds.), *Speech Prosody 2004*,

International Conference; Nara, Japan, March 23-26, 2004, pp. 127-130. ISCA
Archive, <http://www.isca-speech.org/archive/sp2004>

Xu, Yi (2005). 'Speech Melody as Articulatorily Implemented Communicative Functions',
Speech Communication 46: 220-251.

Xu, Yi, and Ching X. Xu (2005). 'Phonetic Realization of Focus in English Declarative
Intonation', *Journal of Phonetics* 33: 159-197.

¹ Strictly speaking, Bruce saw association as a property of the systematic phonetic level, not of phonology, which does not feature in his model.

² Ladd (2008, chap. 5) cites German data which suggest that another way of dealing with tonal crowding may be the avoidance of tonal choices that would lead to it, but this strategy has not been experimentally investigated to my knowledge.

³ Bitonal accents have been notated in various ways; to give an example, L*+H-, L*+H, L*H- and L*H represent essentially the same entity.

⁴ In Pierrehumbert (1980) and in Beckman and Pierrehumbert (1986) the unstarred tones of some bitonal accents are not realized, but are used as a means to triggering various scaling changes, including downstep. The formal treatment of downstep is beyond the scope of this paper, but see Ladd (2008, chap. 3) for a discussion; Arvaniti and Baltazani (2005) and Grice, Baumann, and Jagdfeld (2009) present data from Greek and German respectively which do not support the idea that bitonal accents universally trigger downstep.

⁵ Prosodic structure is taken to be independent of intonation (Ladd 2008, chap. 7) and thus it is not discussed at length here. However, laboratory research has uncovered interesting interactions between the two. Research on the Rhythm Rule of English has shown that this is primarily an intonational phenomenon, to do with the choice of foot that associates with a pitch accent in a given word, rather than purely a change in metrical strength between feet (Shattuck-Hufnagel, Ostendorf, and Ross, 1994). Further, empirical research has shown that the presence of a pitch accent does not simply result in the “layering” of an F0 movement on particular syllables but may also affect their duration and articulation

(Beckman and Edwards, 1994; Campbell and Beckman, 1997; Harrington, Fletcher, and Beckman, 2000); in addition, nuclear accents may also block sandhi (Baltazani, 2006b).

⁶ Chen and Gussenhoven (2008) also report a lack of anticipatory coarticulation between tones under emphasis and preceding tones and point out that this result supports PENTA's assumptions about tonal coarticulation. However, the lack of coarticulation in this case could also be due to the presence of focus itself (as has been shown for vowel coalescence in Greek by Baltazani 2006b). Similarly, the results of Prieto and Torreira (2007), which are presented as evidence in favor of PENTA, are compatible with AM as well (see Ladd, 2008, chap. 5, for a discussion).