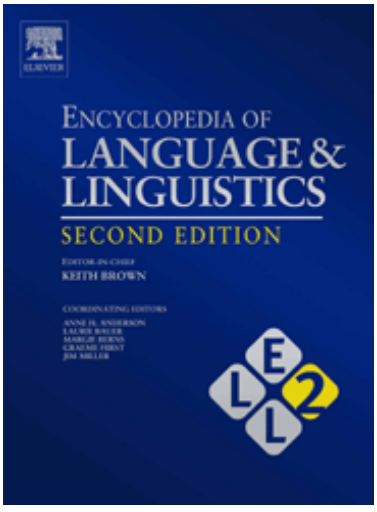

Grice M (2006), Intonation. In: Keith Brown, (Editor-in-Chief) *Encyclopedia of Language & Linguistics, Second Edition*, volume 5, pp. 778-788. Oxford: Elsevier.

Eades D (2004). 'Understanding Aboriginal English in the legal system: a critical sociolinguistics approach.' *Applied Linguistics 25(4),* 491–512.
Ehrlich S (2001). *Representing rape: language and sexual consent*. London: Routledge.
Ehrlich S (2002). '(Re)contextualizing complainants' accounts of sexual assault.' *Forensic Linguistics 9(2),* 193–212.
Gibbons J (2003). *Forensic linguistics*. Oxford: Blackwell.
Gumperz J (2001). 'Contextualization and ideology in intercultural communication.' In Di Luzio A, Gunthner S & Orletti F (eds.) *Culture in communication: analyses of intercultural situations*. Amsterdam: John Benjamins. 35–53.
Hoopes R (2003). 'Trampling *Miranda:* interrogating deaf suspects.' In Lucas (ed.). 21–59.
Jacquemet M (1992). '"If he speaks Italian it's better": metapragmatics in court.' *Pragmatics 2(2),* 111–126.
Jacquemet M (1996). *Credibility in court: communicative practices in the Camorra trials*. Cambridge: Cambridge University Press.
Liberman K (1981). 'Understanding Aborigines in Australian courts of law.' *Human Organization 40,* 247–255.
Lucas C (2003). *Language and the law in deaf communities*. Washington, DC: Gallaudet University Press.
Mather S & Mather R (2003). 'Court interpreting for signing jurors: just transmitting or interpreting?' In Lucas (ed.). 60–81.
Matoesian G (1993). *Reproducing rape: domination through talk in the courtroom*. Chicago: University of Chicago Press.
Montoya M (2000). 'Silence and silencing: their centripetal and centrifugal forces in legal communication, pedagogy and discourse.' *Michigan Journal of Race and Law 5,* 847–941.
Morrow P (1993). 'A sociolinguistic mismatch: central Alaskan Yup'iks and the legal system.' *Alaska Justice Forum 10(2),* 4–8.
Morrow P (1996). 'Yup'ik Eskimo agents and American legal agencies: perspectives on compliance and resistance.' *Journal of the Royal Anthropological Institute 2,* 405–423.
Trinch S (2001a). 'The advocate as gatekeeper: the limits of politeness in protective order interviews with Latina survivors of domestic abuse.' *Journal of Sociolinguistics 5(4),* 475–506.
Trinch S (2001b). 'Managing euphemism and transcending taboos: negotiating the meaning of sexual assault in Latinas' narratives of domestic violence.' *Text 21(4),* 567–610.
Trinch S (2003). *Latinas' narratives of domestic abuse: discrepant versions of violence*. Amsterdam: John Benjamins Publishing Co.
Walker A G (1993). 'Questioning young children in court.' *Law and Human Behavior 17(1),* 59–81.
Walker A G (1999). *Handbook on questioning children: a linguistic perspective* (2nd edn.). Washington, DC: ABA Center on Children and the Law.
Walsh M (1994). 'Interactional styles in the courtroom: an example from northern Australia.' In Gibbons J (ed.) *Language and the law.* London: Longman. 217–233.
Wodak-Engel R (1984). 'Determination of guilt: discourse in the courtroom.' In Kramarae C, Schulz M & O'Barr W M (eds.) *Language and power.* Beverly Hills: Sage. 89–100.

Intonation

M Grice, University of Cologne, Cologne, Germany

Introduction

Intonation is present in every utterance in every language. A well-formed utterance contains minimally one intonationally defined phrase, which is termed variously 'intonation phrase,' 'intonational phrase,' 'tone group,' 'tone unit,' or 'breath group.' Such a phrase entails tonal marking of some kind, at one or both of its edges and/or on at least one prominent element within it. The currently most widespread phonological framework for representing intonation is termed 'autosegmental-metrical,' starting with the work of Pierrehumbert (1980), and is treated in detail in Ladd (1996), in which the term was coined. It is generally claimed in this framework that the division of utterances into phrases and the assignment of relative prominence to elements within the phrase – leading to the **prosodic structure** – is the task of **prosody** (in the narrow sense of the term). This represents the **metrical** aspect, which was first proposed by Liberman and Prince (1977). The association of tonal entities – the tune – with the prosodic structure is the task of **intonation** (also in its narrow sense). This represents the **autosegmental** aspect, which has its origins in the work of Goldsmith (1979) and Leben (1975). However, this distinction between prosody and intonation is rather artificial, since the terms are often used interchangeably – not only in more traditional phonetic models such as the British School (Crystal, 1969; Cruttenden, 1997), but also within phonological models of intonation which embrace the autosegmental-metrical framework. In this

account, we use the term intonation, in its broad sense, to cover both aspects.

Phonetics and Phonology of Intonation

Phonetically, intonation is expressed through the following channels: (a) perceived pitch, (b) loudness, (c) vowel quality (e.g. reduction) and voice quality, and (d) relative length of segments, syllables, and words. Their approximate acoustic correlates are: (a) estimated fundamental frequency over time (F0), (b) relative intensity, (c) spectral quality (formant bandwidth and spectral tilt) and voice source, and (d) relative duration in, for example, milliseconds. For example, the word 'can' in the sentence 'John can cook' can be made prominent by virtue of a pitch obtrusion, along with it being louder, more clearly articulated, and longer than it would otherwise be: 'John CAN cook' Similarly, the word 'cook' can be clearly marked as phrase final (as it would be in 'John can cook,' as opposed to 'John can cook pasta,' where it is phrase medial) with a drop in pitch, a decrease in loudness, a creaky voice, and lengthening of the segments.

However, intonation's primary channel is perceived pitch, corresponding to fundamental frequency in the acoustic signal. By pitch obtrusion we mean that the pitch of the syllable entails a glide or jump up or down from the pitch of the syllables surrounding it. Such obtrusions are determined phonologically in terms of abstract discrete values assigned to tonal entities. These tonal entities can be regarded as either pitch configurations (Crystal, 1969; Tench, 1996; Kohler, 1991), or targets, as is the case in the autosegmental-metrical framework. In the latter framework, targets specify only specific points in the F0 contour, interpolation filling in the rest.

Targets are represented phonologically as 'tones,' which normally have one of two values, H(igh) or L(ow). H tones correspond to high targets, often referred to as 'peaks,' L tones to low targets, referred to as 'valleys' or 'troughs.' A further level, M(id), is used in a small number of accounts, often for transcribing intonation patterns that have a chanting nature ('stylized intonation' or the 'calling contour', Ladd, 1978). Since there are often more syllables in an utterance than intonational tones, the question arises as to how these tones are distributed. They are in fact anchored to prominent elements and edges of domains, or phrases. This anchoring corresponds to the two major tasks tones can carry out: lending of prominence to elements within phrases (words and parts of words, such as 'can' above), and marking the edges of those phrases (such as the edge coinciding with the edge of the word 'cook' above).

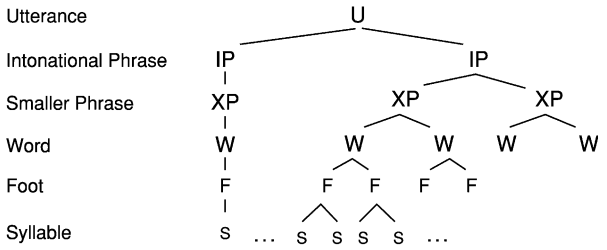

Figure 1 The Prosodic Hierarchy, adapted from Keating *et al.* (2003).

Phrasing and Prosodic Structure

Autosegmental-metrical phonology allows for multiple levels of phrasing. The phrases that are relevant for intonation differ from language to language. There is no universal prosodic structure, although some of the higher and lower levels of this structure appear to be common to a wide range of languages. Figure 1 provides a simplified prosodic structure, reflecting the part of prosodic structure that is currently accepted by most researchers in the field. It can be interpreted as follows. An Utterance (U) contains one or more Intonational Phrases (IP); an IP contains one or more Smaller Phrases (XP); an XP contains one or more Words (W), a W contains one or more Feet (F), which in turn contains one or more Syllables (s).

The Smaller Phrase (XP) represents the phrase or phrases between the Intonational Phrase and the Word. There is no common structure at this intermediate level. Depending on the language under investigation and on the theoretical bias of the investigator, an XP could be a Phonological Phrase (φ, or PP), an intermediate phrase (ip, note use of lower case), or an Accentual Phrase (α, or AP). For example, according to Gussenhoven (and others, 2004), English has the following prosodic constituents (for '>' read 'is higher in the prosodic hierarchy than'):

(1) Utterance (υ) > Intonational Phrase (ι) > Phonological Phrase (φ) > Phonological Word (ω) > Foot (Σ, or F) > Syllable (σ)

Pierrehumbert and Beckman (1988) and Beckman and Pierrehumbert (1986) provide a similar hierarchy for English, but with one crucial difference: in their account, between the Intonation(al) Phrase and the Prosodic Word there is an intermediate phrase in place of the Phonological Phrase.[1] We return to this difference in the next section, showing that although

[1]The term *Intonational Phrase* tends to be used in theories concerned primarily with the prosodic structure (e.g., Liberman and Prince, 1977; Hayes, 1995) and *Intonational Phrase* in those dealing primarily with tonal representations (e.g., Pierrehumbert and Beckman, 1988; Ladd, 1996). The two terms are equivalent.

the intermediate phrase and phonological phrase are roughly the same in size, they are not equivalent.

Prosodic structure is also involved in determining the pronunciation of sounds in a language. At the ends of domains there is a slowing down of the articulators, which is reflected in the signal as **final lengthening**, as would be the case in the word 'cook' above. The larger the domain, the greater the degree of final lengthening (*inter alia*, Wightman *et al*., 1992). More recently, research has concentrated on the beginnings of domains, and the strengthening of segments in this position: **domain initial strengthening** (see, e.g., Keating *et al*., 2003). For example, a sound such as /t/ in English is pronounced at the beginning of a given domain with greater strength (e.g., greater area of contact between tongue and palate, greater aspiration) than at the beginning of a domain lower in the prosodic hierarchy. Furthermore, connected speech processes such as assimilation occur to a greater extent across smaller boundaries than across greater ones. This resistance to assimilation is also considered to be due to initial strengthening in the sense that the segment preserves its identity, thus enhancing the contrast with adjacent segments (syntagmatic contrast), and possibly even enhancing a contrast with other segments which might occur in that position (paradigmatic contrast).

Another effect on the pronunciation of sounds is stress and accent. For instance, a /t/ at the beginning of a stressed syllable is stronger than one at the beginning of an unstressed syllable: compare /t/ realizations at the beginning of 'tomorrow' and 'tomcat,' where /t/ in 'tomcat' is stronger. Moreover, /t/ at the beginning of a syllable bearing a pitch accent is stronger than one at the beginning of a syllable which is stressed but bears no pitch accent: Compare initial /t/ in the word 'tomcats' in 'I like TOMCATS best' with 'Why not? I LIKE tomcats,' where the former /t/ is stronger.

Edge Tones

Some prosodic constituents are marked tonally, in that they have tones at one or both of their edges. For Pierrehumbert and Beckman and for Gussenhoven, the Intonation(al) Phrase has a **final edge tone**, also referred to as a **right edge tone**. For Pierrehumbert and Beckman, this final tone is obligatory; for Gussenhoven it is often present but may optionally be absent. Furthermore, Pierrehumbert and Beckman's account of English allows for an optional IP **initial edge tone** (**left edge tone**), but this initial edge tone can only be High; a phonetically low or mid pitch in initial position is not represented with a phonological tone.

The intermediate phrase in Pierrehumbert and Beckman's account has an obligatory final edge tone too, leading to a sequence of two edge tones at the end of Intonation Phrases: that of the final intermediate phrase in the IP, and that of the IP itself. The Phonological Phrase in Gussenhoven's account has no tones at either edge. The relative scarcity of edge tones in Gussenhoven's account has repercussions for the analysis of intonation patterns such as the one shown in Figure 2(A), and analyzed in two different ways in Figure 2(B), (C).

Whereas in Gussenhoven's model there is no edge tone after the word 'cooks,' and the rise is analyzed as part of the pitch accent, in the Pierrehumbert and Beckman style model the rise is achieved through an H-intermediate phrase edge tone (although a rising pitch accent, e.g., L* + H, is also possible). In Gussenhoven's model, the final fall to low on 'broth' is mainly due to the HL sequence in the pitch accent. In the Pierrehumbert and Beckman style model, it is due to an interpolation between the H*pitch accent and the following L- edge tone. Comparing the inventories and combinatory possibilities of both models, falls and fall-rises are always decomposed in Pierrehumbert and Beckman's model: H* L-(L%) for falls and H*L-H% for fall-rises; Gussenhoven captures these falling and falling rising contours with complex pitch accents: HL and HLH.

Returning to Figure 2 above, the **trailing** tones of Gussenhoven's model appear to play a similar role in specifying the contour as the ip edge tone of Pierrehumbert and Beckman. We return to the issue of ip edge tones below, where phonetic alignment details may provide a clue as to which account is more adequate.

Pierrehumbert and Beckman's (1988) account of Japanese provides the following hierarchy:

(2) Utterance > intermediate phrase > Accentual Phrase > Word > Syllable > Mora

In their account of Japanese, there are obligatory edge tones both initially and finally for two constituents: the Utterance and the Accentual Phrase. There is a difference, however, in the way these tones are associated. In Japanese, the edge tones are not simply associated at the edge of the phrase, but are additionally associated with a sonorant mora at or near the edges of each phrase. The section on **secondary association** examines this phenomenon in more detail.

Pitch Accents

We have just explored the properties of tones whose function is to mark edges – **edge tones** or **boundary tones**. Tones or combinations of tones associated with prominence are referred to as **pitch accents**. They can

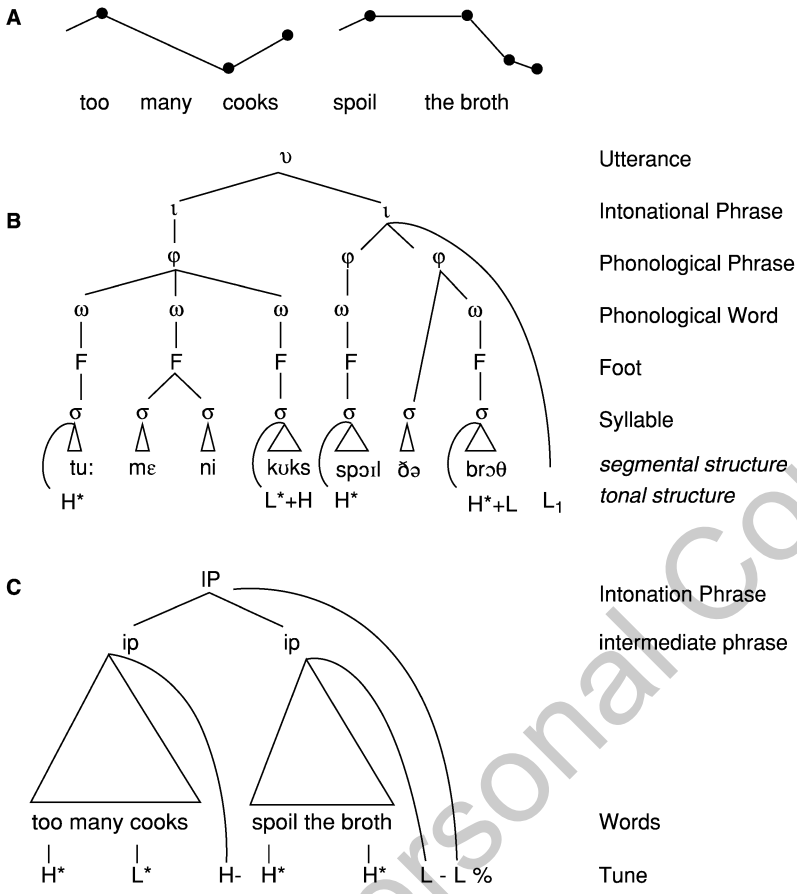

Figure 2 Stylized intonation contour for 'Too many cooks spoil the broth,' (A), and Gussenhoven's analysis (B), both reproduced with permission from Gussenhoven (2002: 271), and one way of analyzing the contour in a Pierrehumbert and Beckman style model, (C). Alternatively, it could be analyzed as having two IPs wih two edge tones after 'cooks.' Reproduced from Gussenhoven C (2002). 'State of the article: phonology of intonation.' *Glot International 6(9/10)*, 271–284, with permision from Blackwell.

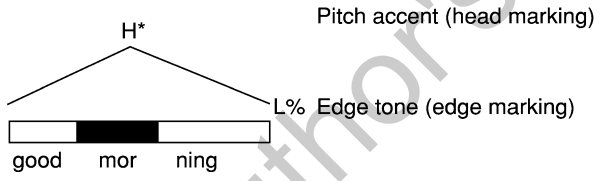

Figure 3 Anchoring of tones to a prominent element (H* to head) and to a phrase edge (L% to edge). Dark bar represents stressed syllable.

also be seen as the markers of the strongest elements, or **heads**, of phrases in the prosodic structure. The head of an IP is the strongest syllable of the Intonation Phrase. This is the syllable bearing the nuclear pitch accent.

Figure 3 provides a simple example where head marking and edge marking are clearly distinct.

We can observe that there is a pitch peak roughly aligned in time with the stressed syllable of 'morning.' This pitch peak marks the head of the phrase by means of a high pitch accent, H*. The final low pitch marks the final edge of the phrase with a low boundary tone, L%. Not all languages have pitch accents. Korean, for example, only has edge tones (Jun, 1996).

Pitch accents were originally analyzed in autosegmental terms as comprising one 'starred' tone and optionally one 'unstarred' tone. The star diacritic on a particular tone was used to indicate association with a **tone bearing unit** (TBU), often taken to be a vowel or syllable. The other tone was 'floating,' or unassociated. The TBU was also marked with a star. Figure 4 shows a typical association between the starred tone of a pitch accent and a TBU, as represented in earlier studies, e.g., Leben (1975).

In current autosegmental-metrical models the TBU does not need to be marked, since it is the strongest element, or head, of a constituent. However, the star notation has been maintained in the tonal representation. In autosegmental terms, the unstarred tone cannot strictly be considered to be unassociated, since this means that it would be deleted (e.g., by the process of **stray tone erasure**, a process whereby tones that have no association to another autosegmental tier are deleted). To resolve this problem,

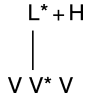

Figure 4 Association of a tone with a tone bearing unit (vowel), e.g., in the word 'tomato,' where the second syllable (and therefore the second vowel) is assigned a star.

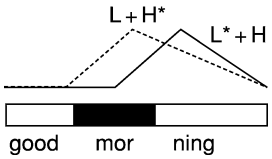

Figure 5 Bitonal pitch accents with schematic pitch contours.

autosegmental-metrical theory treats pitch accents as a sequence of tones packaged together, a relation symbolized by the + sign. It was proposed in Pierrehumbert and Beckman (1988), and elaborated on by Grice (1995), that the packaging of the two tones in bitonal pitch accents might be represented as a branching structure. Although the star notation was maintained as a convenient shorthand, the strong node of the branching structure represents the starred tone.

The issue as to how it can be determined which of the two tones has the primary association (i.e., which is the starred tone) has been the subject of much debate in recent years, backed up by research into the exact timing of tonal targets with landmarks in the acoustic signal, such as syllable or segment onsets and offsets.

Most studies on **phonetic alignment**, or tune-text synchronization, have concentrated on pitch accents involving a rise throughout the accented syllable. Figure 5 shows how two pitch accent types occurring in a given language would be analyzed in terms of starred tones. The leftmost trace (dashed line) shows a typical rise for L + H*, as it might be realized in English. Here the peak occurs during the accented syllable, and the beginning of the rise is at its left edge. The rightmost trace (solid line) shows a typical L* + H accent, also as it might be realized in English. In L* + H, the peak occurs later, and the rise begins well into the accented syllable.

Arvaniti *et al.* (1998) subjected to close scrutiny the tacit assumption that the **phonological association** of a tone with a tone bearing unit means **phonetic alignment** with that tone bearing unit (as is the case for H* and L* of the pitch accents in Figure 5). Since their findings for Greek revealed that both of the F0 events corresponding to the L and H of a bitonal LH pitch accent fell outside the lexically stressed syllable with which they were associated, it was clear that this assumption was too strong. Thus, exact synchronization with a landmark in the strong syllable appears to be unnecessary for a tone to be starred. However, the fact that the pitch accent investigated involved a separate anchor for each tone strengthened the arguments within autosegmental metrical theory that intonation should be modelled as targets between which there is interpolation, rather than as configurations.

Ladd *et al.* (1999) later investigated peak timing in English, and found that both tones of a bitonal pitch accent (corresponding to the F0 valley and the peak) were aligned in a consistent way to the same segmental landmarks, regardless of articulation-rate-induced changes in segment durations. However, although synchronization experiments have shown that the phonetic reflex of association cannot be taken to be simply an alignment in time with the segment or constituent with which they are associated, this line of research has yet to provide a definitive answer to the question asked by Arvaniti *et al.* (2000): "What is a starred tone?"

The transcription of the leftmost pitch accent as L + H* in Figure 5 is somewhat controversial, since it involves a starred tone preceded by another tone, a **leading tone**. Not all autosegmental-metrical models of intonation have leading tones, and those that do not would transcribe the contour as simply H*, arguing that there is no categorical difference between the contours in Figures 3 and 5. However, more recent studies, even those that tend not to make use of leading tones in general, allow for a H leading tone in 'early peak' contours (H + L* or H + !H*), where there is a clear peak on the syllable preceding the accented syllable (see, for example, Gussenhoven, 2004).

The issues of pitch accent structure and the status of leading tones are discussed at length in Grice (1995), Ladd (1996), and Gussenhoven (2004).

We have provided the principles behind current pitch accent analyses without an inventory of pitch accents used by each model for individual languages. For the inventory of pitch accents in English intonation according to Pierrehumbert's model, see the section on ToBI.

Secondary Association and Phrase Accents

'Secondary association' involves an association of a tone with an additional element further down the prosodic tree, typically with the smallest TBU in a given language, such as a syllable (e.g., English) or sonorant mora (e.g., Japanese and Roermond Dutch). Figure 6 shows a prosodic tree for Japanese from Pierrehumbert and Beckman (1988). Japanese has edge tones at the levels of the utterance and accentual phrase. Starting with the first four left-most tones in the tree (LHLH), the utterance initial L tone has a

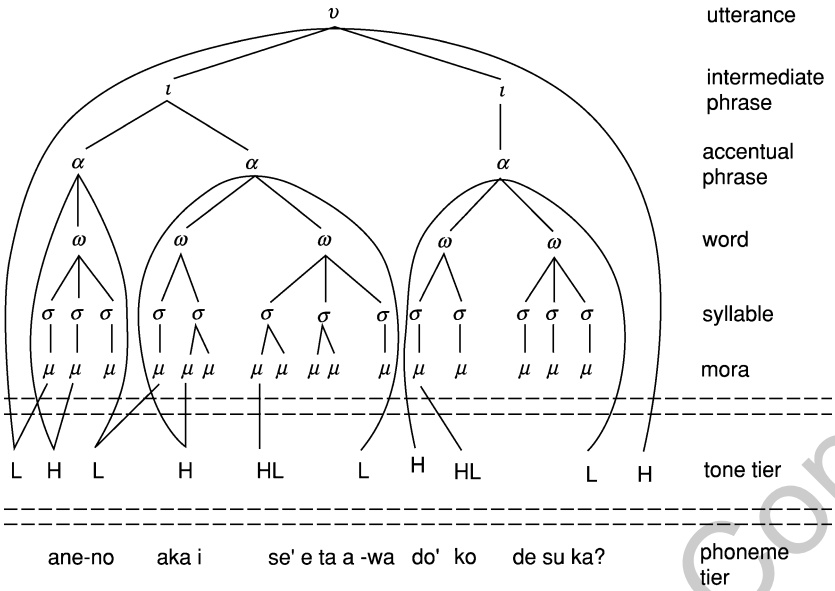

Figure 6 Prosodic structure and tonal association for the sentence ''Ane-no akai se'etaa-wa do'ko desu ka?'' (Where is big sister's red sweater?), from Pierrehumbert and Beckman (1988: 21) with kind permission from MIT Press. Japanese has lexical tones (HL on se' and do'); other tones are intonational.

secondary association to the first sonorant mora of the phrase (μ). The accentual phrase initial tone goes to the second mora (since the first is already taken by the higher level edge tone). The accentual phrase final edge tone, L, is secondarily associated with the first mora of the next accentual phrase, and the initial tone of that accentual phrase goes to the second mora. However, note that not all edge tones have a secondary association.

Edge tones for tonally demarcated phrases may also be secondarily associated to internal constituents other than the mora (or syllable). There can, for example, be a secondary association to a word or foot. This is then realized by an association to the strongest element of that word or foot (for many languages this means that there is an association with a stressed syllable). This is the case for the **phrase accent**, a tone or tonal structure that is primarily an edge tone, but is realized on stressed syllables some distance from the edge of the phrase. Grice *et al.* (2000a) pointed out that phrase accents can serve both head-marking and edge-marking functions (*see* Figure 3), either simultaneously or in linguistically conditioned alternation.[1]

[1]In fact, the term 'phrase accent,' sometimes also called 'phrase tone,' had been previously used to denote a tone of unclear status occurring between the nuclear pitch accent and the edge tone of the intonation phrase. Grice *et al.* redefine the term, giving it a clear definition and functionality. The phrase accent is not to be confused with the **phrasal accent**, which is another term used by researchers concentrating on prosodic structure for **nuclear accent** – the final accent in the phrase.

Figure 7 shows an example of the phrase accent H tone in the LHL question tune in Standard Greek. H alternates between being prominence lending, as in (A), and delimiting, as in (B). Observe that in (A) H co-occurs with a stressed syllable in a similar way to the H* pitch accent in Figure 3, whereas in (B) it is towards the end of the phrase.

Grice *et al.* (2000a) showed that when there is no postnuclear stressed syllable available as anchor for a phrase-accent, it occurs close to the phrase edge. It can therefore alternate between head marking (as in (A), when there is a postfocal stress) and edge marking (as in (B), when there is none). They show that this alternation is characteristic of phrase accents in many languages.

Their analysis of the English fall-plus-rise can be seen as additional evidence for a compositional analysis of nuclear contours: H* L-H%. L-has a secondary association to a stressed syllable after the nuclear one, as in (3) below, where there is a fall starting on the nuclear syllable, 'REA' of 'REASONABLE,' and the rise begins on the word 'costs'. The symbols before these two words are **tonetic stress marks** as used in the British School (see, e.g., Cruttenden, 1997), they are used iconically to symbolize a fall beginning high in the speaker's pitch range and a rise beginning low. The word bearing the nuclear accent is in upper case, the one bearing the phrase accent is in italics.

(3) It seems perfectly \ REASONABLE
to take the /*costs* into account.

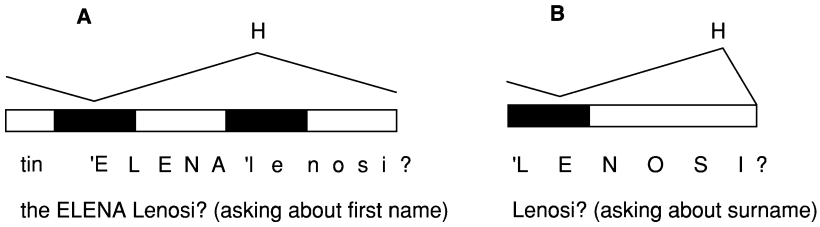

Figure 7 The phrase accent in Standard Greek yes-no questions. Dark bar represents focal (nuclear) stress, light bar represents postfocal (postnuclear) stress.

This type of alignment is not captured in Gussenhoven's model, where the fall-rise is accounted for as a complex HLH pitch accent.

The phrase accent (under its current definition as an edge tone with stress-seeking properties) has since become a central part of the intonation analysis of a number of different languages and varieties. Evidence for stress-seeking phrase accents has been found in regional varieties of German and Swedish. However, the exact definition of the phrase accent is still controversial; see Gussenhoven's survey on intonational phonology (2002).

Downstep and Upstep

It is, of course, not possible to capture the scaling of intonation contours using a simple model of High and Low tones. This problem was solved by introducing operators on H tones in the form of 'downstep' and 'upstep.' Downstep is present in almost all models; a downstepped H tone is scaled lower than a previous H tone and often marked with an exclamation mark before the downstepped tone, e.g., !H*. In Pierrehumbert's (1980) model, downstep was triggered by a L trailing tone of a previous pitch accent. Thus, a sequence H* + L H* is realized as a peak followed by a peak at a lower pitch. Ladd argued against this local conditioning for downstep and accounted for it using a metrical structure whereby a downstepped tone is dominated by a weak branch (see Ladd, 1996, for discussion). The phonological basis for downstep is still controversial.

Upstep is also a process that can be triggered by rule: in Pierrehumbert (1980) every high intermediate phrase edge tone, H-, triggers upstep on the following Intonation Phrase edge tone. However, there are models that allow for upstep on pitch accents as well as on edge tones; e.g., in the German autosegmental-metrical transcription system GToBI, it is marked with a '^' symbol, before the H (Grice *et al.*, 2004). Upstep is also present on pitch accents in other models, such as INSTINT (Hirst and Di Cristo, 2000).

ToBI

A theoretical study of prosody, whether based on introspection or on carefully controlled experiments, needs to be validated in the real world. An analysis of recordings of natural speech informs us as to which prosodic forms speakers actually produce and, depending on the context within which the recordings were made, gives more or less an indication of their functions.

Since annotating speech is a time-consuming process, and since current phonological models of prosody are not sophisticated enough to allow for (fully) automatic annotation, it was considered necessary to arrive at a consensus transcription and labelling system for capturing the most important aspects of intonation and junctural phenomena (tones and break indices respectively), so that recordings and the associated annotations could be shared across different sites.

Such an initiative was first taken for the intonation of English, from which the ToBI (Tones and break Indices) framework was conceived (Beckman and Hirschberg, 1994; Beckman and Ayers-Elam, 1997). The framework is defined as follows:

> ToBI is a framework for developing community-wide conventions for transcribing the intonation and prosodic structure of spoken utterances in a language variety. A ToBI framework system for a language variety is grounded in careful research on the intonation system and the relationship between intonation and the prosodic structures of the language. http://ling.ohio-state.edu/~tobi/

A ToBI transcription consists of a speech signal and F_0 record, along with time-aligned symbolic labels relating to four types of events. The two main event types are tonal, arranged on a **tone tier** and junctural, arranged on a **break index tier.** There is additionally a tier containing an orthographic transcription, the **orthographic tier,** and a further tier for the annotation of nontonal events such as voice quality or paralinguistic and extralinguistic phenomena, the **miscellaneous tier.**

Table 1 English ToBI pitch accents

H*	'peak accent'
L*	'low accent'
L + H*	'scooped accent'
L* + H	'rising peak accent'
H + !H*	'clear step down onto the accented syllable'

The original ToBI inventory for (Mainstream American) English has five basic pitch accents, as given in Table 1. The tonal part of the transcription is based on the work of Pierrehumbert (1980) and Beckman and Pierrehumbert (1986).

All of the H tones in the above inventory may be marked with a '!' diacritic, which indicates that they are downstepped relative to the immediately prior H tone. The downstep diacritic is obligatory in the H + !H* accent. The others, if downstepped would be transcribed !H*, L + !H*, L* + !H, and, in principle, !H + !H*. The prerequisite for using a ! diacritic is that there must be at least one H tone prior to the downstepped tone from which there can be a step down.

As outlined in the section on edge tones, two domains have an obligatory right edge tone: the 'intermediate phrase' and the 'intonation phrase.' Intonation phrases contain at least one intermediate phrase. The tones available at the right edge of the intermediate phrase are L- and H-. The right edge of an intonation phrase is automatically the right edge of an intermediate phrase. It is customary to label the sequence of tones at these two right edges together. Since there is also the choice of H or L tone at the intonation phrase boundary, there are four combinations to choose from: L-L%, L-H%, H-H% and H-L%. The '-' diacritic is used for intermediate phrase boundaries and '%' for intonation phrase boundaries.

The English ToBI system distinguishes between five **break indices** or levels of perceived juncture between words transcribed on the orthographic tier. They are numbered from 0 to 4. The lowest degree of juncture between two orthographic words is level 0, where the words are grouped together into a 'clitic group,' e.g., between 'did' and 'you' pronounced as 'didya.' Level 1 is the default boundary between two words in the absence of any other prosodic boundary. Levels 3 and 4 correspond to intermediate phrase and intonation phrase boundaries. Mismatches between the tone and break index tiers are expressed by a '-' diacritic (for instance, 4-, meaning that the perceived juncture is of the intermediate phrase boundary type when the tonal transcription contains an intonation phrase edge tone) and by break index level 2. Level 2 is used (*inter alia*) when the perceived juncture is even less than that for an intermediate phrase while the tonal transcription contains an edge tone.

Figure 8 gives an example of a ToBI transcription.

On the tone tier, (B), the first intonation phrase 'actually' has one pitch accent, H*, and has a falling-rising boundary, L-H%. The second is divided into two intermediate phrases, 'next week' and 'I am on vacation.' 'Next week' has one pitch accent, H*, and ends at a level around the middle of the range, transcribed as a downstepped H tone, !H-. In the second intermediate phrase, the transcriber was uncertain as to whether there was a pitch accent on 'I' or not; such uncertainty is marked with '*?' The nuclear accent is on 'vacation' and involves a rising accent with a leading tone, L + H*, followed by a fall to the bottom of the range, L-L%. The tones tier also contains an indication of the highest intonationally relevant F0 (ignoring spurious values) within the intonation phrase. This is transcribed as 'HiF0,' and the F0 value with which this label is time aligned is used as a rough indicator of pitch range.

On the break index tier, (D), instead of plain 4, 4- after 'actually' indicates that the perceived juncture at the phrase boundary is weaker than would be expected for an intonation phrase boundary. After 'next week,' the perceived juncture matches with the tonal transcription: an intermediate phrase boundary, 3. Break index 4 marks the end of the second intonation phrase. The default break index is 1, signalling a non-cliticized juncture between two words. This default juncture is transcribed after 'next' and between the words in 'I am on vacation,' indicating, among other things that the speaker did not produce 'I'm.' A rough indication as to the location of creak, and thus irregular F0 values, is given by a label on the miscellaneous tier (E).

ToBI systems for other languages have since been developed, building on the experience gathered for English but basing their categories on research into the structure of the language concerned.

Languages with a complete description of their specific ToBI system, with guidelines for use, training materials, and published transcriber consistency results are German (GToBI), Japanese (J_ToBI), and Korean (K_ToBI). Systems under development are Italian (IToBI), Greek (GrToBI), Serbo-Croation, Cantonese, Mandarin, and Spanish. These systems are all presented in a book outlining the ToBI framework and the ToBI systems in a number of languages (Jun, 2004).

ToBI annotated corpora are becoming important resources for investigations into the form and

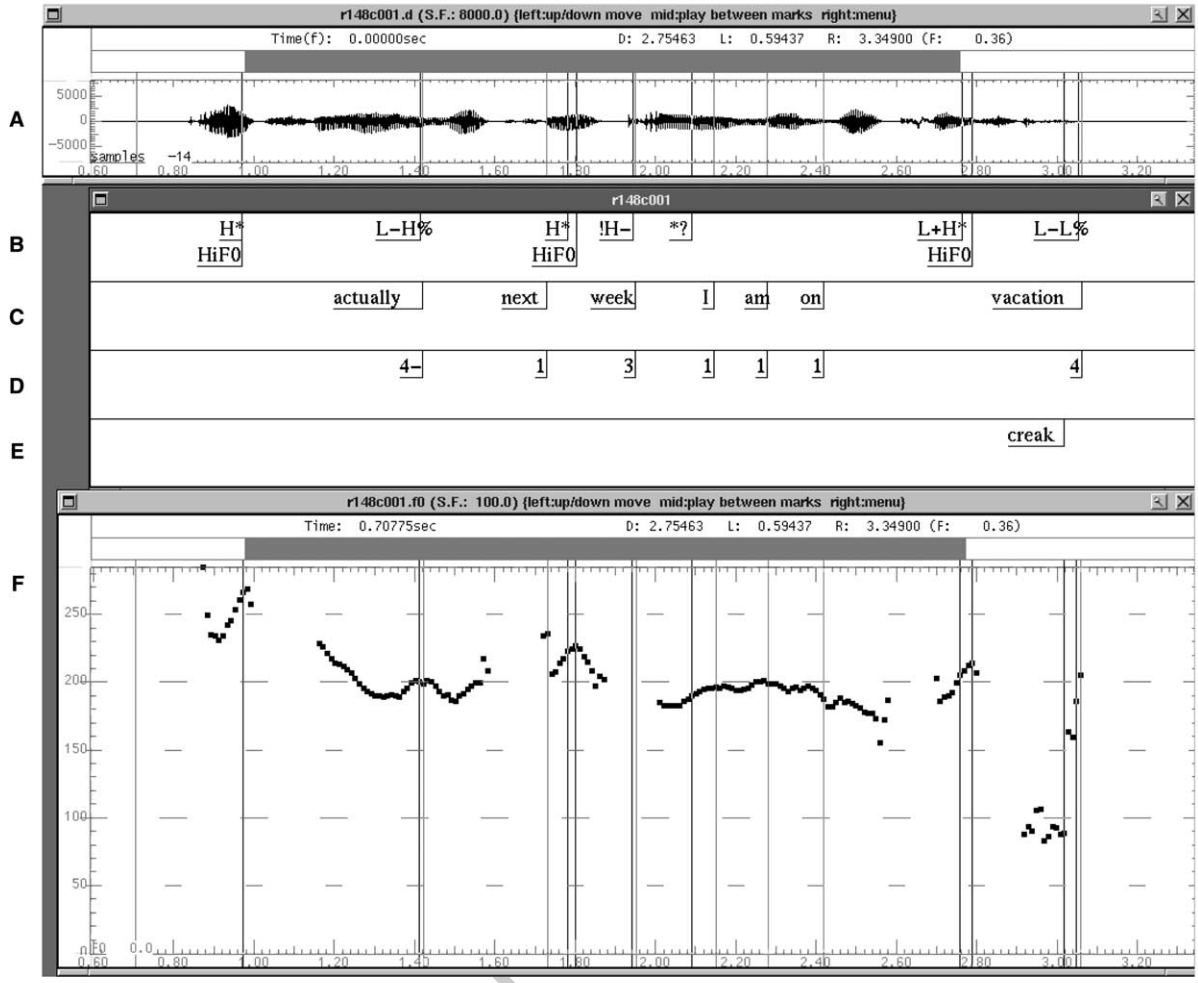

Figure 8 ToBI transcription: (A) speech waveform, (B) tone tier, (C) orthographic tier, (D) break index tier, (E) miscellaneous tier, (F) F0 track; reproduced with kind permission from Springer Science and Business Media from Grice *et al.* (2000b), transcription by Laura Dilley, MIT.

distribution of tonal and prosodic entities (Dainora, 2004). When combined with annotations at other levels, such as information structure and dialogue act categorization, they are an invaluable source of information as to prosodic functions (Baumann *et al.*, 2004).

An autosegmental-metrical system, also with training materials, and based on Gussenhoven's model of intonation, is ToDI (Tones in Dutch Intonation; Gussenhoven *et al.*, 1999). Some of the theory behind Gussenhoven's model is included in the IViE system for transcribing regional varieties of English (Nolan and Grabe, 1997), although the IViE system introduces a tier that captures fine acoustic-phonetic (and to some extent auditory-phonetic) timing details.

Functions of Intonation

A representative sample of functions that intonation can fulfill is shown schematically in Figure 9.

At the level of semantics, intonation is often used to mark **information structure** within sentences. According to Steedman (2000), within each **theme** (what the sentence is about, or sentence topic) and **rheme** (a comment about this topic), there is provision for a background-focus partitioning. Focus can be said to reflect an abstract notion of contrast between alternatives available in the discourse context (Rooth, 1992), and in many languages it is marked by means of pitch accents (while background is often marked by lack of accent) or by the type of pitch accent. In Korean, which has no pitch accents, focus is marked by means of phrasing.

At the pragmatic level, intonation is used to encode distinctions such as whether an utterance is intended as a request for information (e.g., Request) or as a request for the interlocutor to perform a particular action (e.g., Command). There are four major categories of **communicative illocutionary**

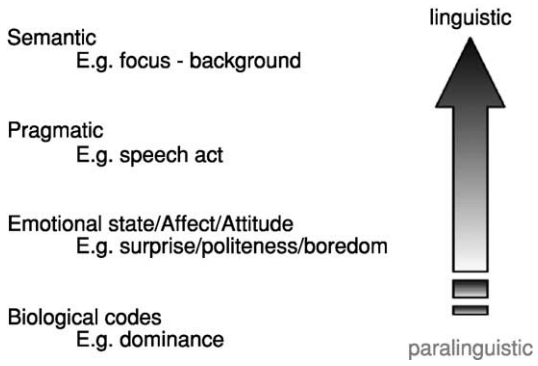

Figure 9 The functions of intonation.

acts: constatives, directives, commissives, and acknowledgments (Bach and Harnish, 1979; Searle, 1969), examples of which are statements, requests, promises, and apologies respectively. Much research has been carried out on **questions**, a special type of directive, and how they are marked intonationally. Although polar questions are often marked with a final rise (H% edge tone), there are a great many languages that have a rising pitch accent followed by a final fall, constituting an LHL sequence (see, for instance, the question tune described for Standard Greek in the section on secondary association and phrase accents). Intonation plays a crucial role in distinguishing polar questions from, e.g., statements if there is no distinct interrogative syntax or question particle (as in Italian).

Intonation is also used to signal **emotional states** of varying degrees of intensity, **speaker affect**, and **attitude**. Since these are generally considered to be external to grammar, and therefore paralinguistic, studies on their realization have concentrated on non-discrete aspects of intonation, such as pitch range, rather than on phrasing and prominence relations or pitch accent type. However, there are many cases where discrete means are used to signal paralinguistic meaning. For example, a final H% tone is often used to signal uncertainty or politeness (Cruttenden, 1997). A L* accent is used in 'scathing intonation' (Monaghan, 2000). Although the expression of these meanings has been grammaticalized, it is claimed that there is a universal basis to this means of expression in the form of **biological codes**, the most established of which is the **frequency code** (Ohala, 1994), whereby high pitch corresponds to submissiveness or friendliness and low pitch to dominance or aggression. Two further biological codes, introduced by Gussenhoven (2002) are the **effort code** and the **production code**.

To sum up, intonation is active at many different levels of communication, in areas deemed purely linguistic as well as those considered more peripheral to linguistic inquiry. However, since the intonational expression of many functional levels occurs simultaneously, it is not possible to understand the expression of one level without an taking into account the way the others are expressed.

See also: Autosegmental Phonology; Metrical Phonology; Phonological Phrase; Prosodic Aspects of Speech and Language; Speech Acts.

References

Arvaniti A, Ladd D R & Mennen I (1998). 'Stability of tonal alignment: the case of Greek prenuclear accents.' *Journal of Phonetics 26,* 3–25.

Arvaniti A, Ladd D R & Mennen I (2000). 'What is a starred tone? Evidence from Greek.' In Pierrehumbert J & Broe M (eds.) *Papers in laboratory phonology V.* Cambridge: Cambridge University Press. 119–131.

Bach K & Harnish R M (1979). *Linguistic communication and speech acts.* Cambridge, MA: MIT Press.

Baumann S, Brinckmann C, Hansen-Schirra S, Kruijff G, Kruijff-Korbayova I, Neumann S, Steiner E, Teich E & Uszkoreit H (2004). The MULI Project: annotation and analysis of information structure in German and English. In Lino M T, Xavier M F & Ferreira F (eds.). *Proceedings 4th International Conference on Language Resources and Evaluation (LREC 2004).* Portugal: Lisbon.

Beckman M E & Ayers Elam G (1997). *Guidelines for ToBI labelling.* Columbus, Ohio: Ohio State University.

Beckman M E & Hirschberg J (1994). *The ToBI Annotation Conventions.* Columbus, Ohio: Ohio State University.

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

Cruttenden A (1997). *Intonation* (2nd edn.). Cambridge: Cambridge University Press.

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

Dainora A (2004). 'Modeling intonation in English: a probabilistic approach to phonological competence.' In Goldstein L, Whalen D & Best C (eds.) *Papers in laboratory phonology* 8. Mouton de Gruyter.

Grice M (1995). 'Leading tones and downstep in English.' *Phonology 12.2.* Berlin: Mouton de Gruyter. 183–233.

Grice M, Baumann S & Benzmüller R (2004). 'German intonation in autosegmental-metrical Phonology.' In Jun S-A (ed.) *Prosodic typology: the phonology of intonation and phrasing.* Oxford: Oxford University Press. 55–83.

Grice M, Ladd D R & Arvaniti A (2000a). On the place of phrase accents in intonational phonology. *Phonology 17.2,* 143–185.

Grice M, Leech G, Weisser M & Wilson A (2000b). 'Representation and annotation of dialogue.' In Gibbon D, Mertins I & Moore R (eds.) *Handbook of multimodal and spoken dialogue systems: resources, terminology and product evaluation.* Dordrecht: Kluwer. 1–101.

Goldsmith J A (1979). *Autosegmental phonology.* New York: Garland Press.
Gussenhoven C (2002a). 'Intonation and Interpretation: phonetics and phonology.' In Bel B & Marlien I (eds.) *Proc. Speech Prosody 2002.* Aix en Provence: Laboratoire Parde et Language.
Gussenhoven C (2002b). 'State of the article: phonology of intonation.' *Glot International 6(9/10),* 271–284. Blackwell.
Gussenhoven C (2004). *The phonology of tone and intonation.* Cambridge: Cambridge University Press.
Gussenhoven C, Terken J & Rietveld T (1999). *Transcription of Dutch Intonation – courseware.* http://lands.let.kun.nl.
Jun S-A (1996). *The phonetics and phonology of Korean prosody: intonational phonology and prosodic structure.* New York: Garland Publishing Inc. NY.
Jun S-A (ed.) (2004). *Prosodic typology: the phonology of intonation and phrasing.* Oxford: Oxford University Press.
Keating P, Cho T, Fougeron C & Hsu C (2003). 'Domain-initial articulatory strengthening in four languages.' In Local J, Ogden R & Temple R (eds.) *Phonetic interpretation.* (Papers in Laboratory Phonology 6) Cambridge University Press. 143–161.
Kohler K J (1991). 'Terminal intonation patterns in single-accent utterances of German: phonetics, phonology, and semantics.' *AIPUK* (Kiel University) 25, 115–185.
Ladd D R (1978). 'Stylised intonation.' *Language 54(3),* 517–540.
Ladd D R (1996). *Intonational phonology.* Cambridge: Cambridge University Press.
Ladd D R, Faulkner D, Faulkner H & Schepman A (1999). 'Constant "segmental anchoring" of F0 movements under changes in speech rate.' *JASA 106(3),* 1543–1554.
Leben W R (1975). The tones in English intonation. *Journal of Linguistics 2,* 69–107.
Liberman M & Prince A (1977). On stress and linguistic rhythm. *Linguistic Inquiry 8,* 249–336.
Monaghan A (2000). Null focussed item? Linguist List, Subject 11.667. 23 March 2000, cited in Gussenhoven (2004).
Nolan F & Grabe E (1997). 'Can ToBI transcribe intonational variation in the British Isles?' In Botinis A, Kouroupetroglou G & Carayannis G (eds.) *Proceedings of the ESCA tutorial and research workshop on intonation: theory, models and applications.* Athens, Greece.
Ohala J J (1994). 'The frequency code underlines the sound symbolic use of voice of pitch.' In Hinton L *et al.* (eds.) *Sound symbolism.* Cambridge: Cambridge University Press. 325–347.
Pierrehumbert J & Beckman M E (1988). *Japanese tone structure.* Cambridge, MA: MIT Press.
Pierrehumbert J (1980). 'The phonetics and phonology of English intonation.' Ph.D. diss., MIT.
Rooth M (1992). 'A theory of focus interpretation.' *Natural Language Semantics 1,* 75–116.
Searle J (1969). *Speech acts: an essay in the philosophy of language.* Cambridge: Cambridge University Press.
Steedman M (2000). 'Information structure and the syntax-phonology interface.' *Linguistic Inquiry 31(4),* 649–689.
Shattuck-Hufnagel S & Turk A (1996). 'A prosody tutorial for investigators of auditory sentence processing.' *Journal of Psycholinguistic Research 25(2),* 193–247.
Tench P (1996). *The intonation systems of English.* London: Cassell.
Wightman C, Shattuck-Hufnagel S, Ostendorf M & Price P (1992). 'Segmental duration in the vicinity of prosodic phrase boundaries.' *Journal of the Acoustical Society of America 91,* 1707–1717.

Intonation and Syntax

I Fougeron, Université de Lille 3, France

As an integral part of any linguistic mechanism, intonation relates both to phonetics (with respect to the nature of its units) and to syntax (with respect to its area of application). Intonation is a suprasegmental tool, concerning only semantic units.

What Is Syntax?

It is necessary to distinguish between 'formal' and 'communicational' syntax. The former concerns anâlysis of specific grammatical (syntactic) structures: sentences (*propositions* in French). A sentence is a particular set of links between its constituents (Perrot, 1978). As for communicational syntax, it deals with the message such structures convey. Its area is the utterance, i.e., a communication unit endowed with prosodic features and specific word order.

A sentence is usually a syntagma, i.e., a binary structure whose terms are linked by a 'determining' (T′) – 'determined' (T) relation. A sentence is necessarily a predicative syntagma: the speaker assigns a T′ to a T. A sentence, in as much as it provides a structural basis for the utterance, is the most convenient way to communicate. But a single word or a word group, a syntactic structure or an interjection can, if the extralinguistic situation demands it, becomes an utterance and assumes the function of a communication unit, provided that it is actualized, i.e., adapted to the concrete situation, endowed with an appropriate word order and specific intonation. In the whole language