

AN INTRODUCTION TO INTONATION – FUNCTIONS AND MODELS

Martine Grice¹ & Stefan Baumann²

^{1,2}Ifl – Phonetik, Universität Köln

¹martine.grice@uni-koeln.de

²stefan.baumann@uni-koeln.de

This chapter provides an introduction to intonation in general, and is loosely based on an oral presentation given in the workshop “Non-native prosody: phonetic description and teaching practice” in Saarbrücken. Although intonation is particularly difficult for learners of a second language to master, it is seldom taught systematically. Although much of the early work on intonation was didactic in nature, recent studies have tended to be more experimental and/or theoretically rigorous. This has created a gap between intonation as it is used in teaching and intonation research, making it difficult for the results of such research to be of use to teachers of a second language. It is our aim to bridge this gap. We provide an overview of the main issues dealt with in current theoretical research, discussing the different forms intonation can take and the functions it can fulfill, the one of course dependent on the other. Reflecting the context of the workshop, examples are predominantly in German with English translations, accompanied where relevant by Italian equivalents.

We then present the two currently most widespread models of intonation, which will hopefully be useful for second language teachers and textbook writers for their own research and for preparation of course material. We also aim to facilitate reading of current primary literature on aspects of intonation, in particular on languages not dealt with here. With this, we hope that results from theoretical research will find their way into the classroom.

1. Intonation

The term ‘intonation’ has been defined in at least two different ways in the literature. A narrow definition equates intonation with ‘speech melody’, restricting it to the “ensemble of pitch variations in the course of an utterance” (‘t Hart et al. 1990: 10). The crucial role of pitch variations for the interpretation of utterances can be seen in the German example utterances (1) and (2), in which the pitch contour is represented as a line above the words spoken.



- (1) Sie hat ein Haus gekauft ('She bought a house.')

- (2) Sie hat ein Haus gekauft ('She bought a house?!')

The examples display exactly the same string of segments. They only differ in their intonation, making (1) a statement with a (rising-)falling contour, and (2) an echo question with a (falling-)rising contour.

Pitch can be modulated in a categorical way, with the *presence vs. absence*, or *type* of pitch movement, and in a gradient way, involving e.g. variations in the way a pitch movement is realised: the extent of the rise or fall, or the pitch range within which a pitch movement is realised. The two main tasks of pitch modulation are (1) highlighting, marking prominence relations (*Haus* is more prominent than *ein*), and (2) phrasing, the division of speech into chunks. However, it is not pitch alone which is responsible for these tasks. A broader definition of intonation includes loudness, and segmental length and quality, although languages differ in the extent to which they modulate these to achieve highlighting and phrasing. Like pitch, loudness, length and quality are auditory percepts. Their articulatory and acoustic correlates are given in table 1 below, adapted from Uhmann (1991: 109), (cf. also Baumann, to appear).

Table 1 Aspects of speech contributing to intonation in its broad sense

Perception	Articulation	Acoustics
pitch perceived scale: high – low	quasi-periodic vibrations of vocal folds	fundamental frequency (F0) measure: Hertz (Hz)
loudness perceived scale: loud – soft	articulatory effort, subglottal air pressure	intensity measure: decibel (db)
length perceived scale: long – short	duration and phasing of speech gestures	duration of segments measure: millisecond (ms)
vowel quality perceived scale: full – reduced	vocal tract configuration, articulatory precision	spectral quality measure: formant values in Hz

We now examine the two tasks of intonation, highlighting and phrasing, in more detail.

1.1. Highlighting

In languages like English and German, utterance level prominence is realised on a designated syllable either by means of increased loudness and length, and unreduced vowel quality (all contributing to stress) or by means of the above, accompanied by a pitch movement (accent). This is not the case for all languages. Some languages use pitch movement *without* the accompanying loudness, length and vowel reduction (or at least using them to a lesser degree). English and German are referred to by Beckman (1986) as ‘stress-accent languages’, in contrast to, e.g., Japanese, which is a ‘non-stress accent language’. Both pitch movements with stress in stress-accent languages, and those without stress in non-stress-accent languages are referred to as *pitch accents*. In what follows, we concentrate on pitch accents in stress-accent languages.

The notion of ‘stress’ applies to both word and utterance levels. We differentiate between ‘lexical stress’, also called ‘word stress’, denoting abstract prominences at word level, and ‘postlexical stress’, concrete prominences at utterance level.¹ Table 2 summarises the different levels of description.

Table 2 – Levels of description

Lexical stress	word level, abstract, potential for concrete prominence
Postlexical stress	utterance level, concrete prominence
Accent	utterance level, concrete prominence

The difference between stresses and accents entails a difference in the strength or degree of (postlexical) prominence. There are at least four different degrees of prominence at utterance level, as listed in table 3.

Table 3 – Degrees of prominence

No stress/accent	
Stress (equivalent to ‘force accent’ or <i>Druckakzent</i>)	A stressed syllable is louder, longer and more strongly articulated, with less vowel reduction than an unstressed syllable
Pitch accent	An accented syllable (i.e. a syllable bearing a pitch accent) has additional tonal movement on or near it
Nuclear pitch accent	the nuclear syllable is the last pitch accent in an intonation phrase, usually perceived as the most prominent one in the phrase

¹ These two meanings of stress follow the British school approach, e.g. Crystal (1969). For Bolinger (1964) on the

In (3) we provide an extended version of utterance (1) above. It might conceivably be produced with a nuclear pitch accent on *Haus* ('house'), a non-nuclear pitch accent on the first syllable of *schönes* ('beautiful'), and stress on the first syllable of *Lena* (and possibly also on –*kauft*). All other syllables can be thought of as unaccented. In this and later examples, pitch accents are indicated by capital letters, stresses by small capitals.



- (3) LENa hat ein SCHÖnes HAUS geKAUFT. ('Lena bought a beautiful house.')

1.2. Phrasing

Speech is divided into chunks delimited by means of intonation. These chunks have been termed *breath groups*, *sense groups*, *tone units*, *tone groups*, *phonological phrases* or *intonational phrases*, to name but a few (cf. Cruttenden 1986: 35 ff.). The most obvious indicators of boundaries between intonation units are (filled and silent) pauses. The longer the pause, the stronger the perceived boundary. However, there are many cases in which a boundary is perceived although a pause is missing. This effect is often due to an abrupt change in pitch across unaccented syllables, i.e. a jump up or down in pitch which cannot be attributed to the highlighting function of intonation.

It is often difficult to decide whether an intonation unit boundary is present or not, in particular when investigating spontaneous speech. In fact, transcribers across a number of approaches to intonation have often reported that they need to capture different levels of phrasing – in simple terms larger and smaller phrases. Although the British School originally had only one level of intonational phrasing (Crystal 1969, for instance), large scale corpus transcription using this model carried out by Gerry Knowles and Briony Williams led to the introduction of an additional level, the major tone unit, which was able to contain a number of (minor) tone units (Williams 1996a, b).

The autosegmental-metrical model of English intonation which contributed substantially to the ToBI framework (cf. section 3.2.) also makes a distinction between smaller, *intermediate phrases* and larger, *intonation phrases*. It is not clear whether there is a one-to-one correspondence between the two systems in terms of their phrasing, but it is possible to say that in many cases an intermediate phrase corresponds to a tone unit/tone group and the intonation phrase to a major tone group (see Roach, 1994 and Ladd, 1996 for attempts at converting between the British School and autosegmental metrical models).

The intuitive need for at least two different sizes of phrase can be felt when comparing utterance (3) above, which consists of only one phrase, with (4), which appears most naturally to be composed of two:

other hand, 'stress' is a strictly lexical feature, whereas 'accent' exclusively applies at the postlexical level.



- (4) Findest Du NICHT, dass Lena ein SCHÖnes HAUS gekauft hat?
 ('Don't you think that Lena has bought a beautiful house?')

The jump in pitch (and thus the phrase break) occurs between *nicht* and *dass*. Although the tonal break coincides with a syntactic break here, we stress that intonational phrases and syntactic phrases are independent, even if they of course often correspond.

Other instances of larger phrases containing more than one smaller phrase are lists, as in (5).



- (5) Lena hat einen ROTen, einen GELben und einen BLAUen Ball.
 ('Lena has a red, a yellow, and a blue ball.')

In lists usually all but the last phrase end at a relatively high pitch, either as in (5) above, or with a high level pitch. The high pitch indicates that there is still at least one more item to come. After it the pitch is reset (i.e. there is a jump down), marking the beginning of the next phrase. A jump either up or down is a strong cue for a phrase break (the boundary between two phrases).

1.3. Consequences of highlighting and phrasing for the segments of speech

In section 1.1. we claimed that sounds are more strongly articulated when they are stressed or accented. The strength of sounds is also affected by the position of the sound in the syllable and, in turn, of the syllable within the phrase. Below we outline what is meant by strengthening, both with respect to prominence and to phrasing, and describe another phrasal effect on the duration of sounds. An account of intonation cannot ignore these effects, as they are often consciously or unconsciously used as diagnostics for the intonational analysis itself. This is particularly the case for phrasing, where intuitions about levels of phrasing based on the pitch contour are often unclear.

If we take the sound /t/, at the beginning of a stressed syllable it is stronger than it would be at the beginning of an unstressed syllable: compare /t/ realisations at the beginning of 'tomorrow' and 'tomcat,' where /t/ in 'tomcat' is stronger (we hear greater aspiration and a longer closure). 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 longer and more aspirated.

The strengthening of segments at the beginning of phrases (domains) is referred to as domain initial strengthening (cf., e.g., Keating et al., 2003). Let us take the sound /t/ in English

again. It is pronounced at the beginning of a larger phrase with greater strength than at the beginning of a smaller one. Furthermore, connected speech processes such as assimilation occur to a lesser extent across large boundaries than across small 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).

At the ends of phrases there is a slowing down of the articulators, which is reflected in the signal as final lengthening. The larger the phrase, the greater the degree of final lengthening (inter alia, Wightman et al., 1992). Final lengthening leads to an increase in the duration of segments which is different from the increase obtained by stress and accent; the sounds are often pronounced less loudly and clearly than in stressed and accented syllables. Thus, final lengthening cannot easily be mistaken for accentual lengthening. Final lengthening has been found in a large number of languages, and is assumed to have a physiological basis, although there are language-specific, and even contour-specific differences as to the degree of final lengthening present. If a phrase break occurs across a sequence of unstressed syllables, those which are at the beginning of the second phrase are often pronounced very fast, this is referred to as *anacrusis*. Like an abrupt change in pitch, an abrupt change in rhythm is a strong cue for a phrase break.

Now that the highlighting and phrasing tasks have been discussed, we turn to which functions they are used to express.

2. Functions of intonation

In spoken language, intonation serves diverse linguistic and paralinguistic functions, ranging from the marking of sentence modality to the expression of emotional and attitudinal nuances. It is important to identify how they are expressed in the learner's native language, so that differences between the native and target languages are identified. It is particularly important to point out that many aspects of information structure and indirect speech acts are expressed differently across languages. Making learners aware of the existence of these functions will not only help them learn to express them, but will also help them to interpret what they hear in a more analytic way, thus reducing the danger of attributing unexpected intonation patterns as (solely) a function of the attitude or emotional state of the speaker.

We have seen that intonation analysis involves categorical decisions about whether there is stress or accent, and, if there is an accent, which type of pitch accent it is. It also involves decisions about whether a boundary is present, and if so which pitch movement or level is used to mark it. There are also many gradient aspects to intonation, such as variation in pitch height or in the exact shape of the contour (equivalent to allophonic variation in the segmental domain).

2.1 Lexical and morphological marking

Lexical and morphological marking does not belong to intonation proper but uses pitch, and to some extent also the other channels used by intonation. Categorical tonal contrasts at word level are characteristic of *tone languages*. Two quite different examples of tone languages are Standard Chinese, which has lexical contrasts such as the well-known example of the syllable *ma* with four different tonal contours, each which constitutes a distinct lexical item (*mother*, *hemp*, *horse* and *scold*), and the West African (Niger Congo) language Bini, which has grammatical tone: a change of tone marks the difference between tenses, e.g. low tone marking present tense and high or high-low tones marking past tense (see Crystal 1987: 172). Categorical tonal contrasts are also characteristic of so-called *pitch accent languages* which may also have lexical or grammatical tone. Both Swedish and Japanese are pitch accent languages. The difference between tone languages and pitch accent languages is that the former have contrastive tone on almost all syllables, whilst the latter restrict their tonal contrasts to specific syllables, which bear a pitch accent. However, it is difficult to draw a dividing line between these two language categories (cf. Gussenhoven 2004: 47).

In intonation languages (the most thoroughly studied of which are generally also stress accent languages) like English and German, pitch is solely a postlexical feature, i.e. it is only relevant at utterance level. All tone and pitch accent languages have intonation in addition to their lexical and/or grammatical tone, although the complexity of their intonation systems varies considerably.

2.2. Syntactic functions

As we have already pointed out, syntactic structure and intonational phrasing are strongly related, but do not have to correspond exactly. Intonation can be used to disambiguate in certain cases between two different syntactic structures. The attachment of prepositional phrases is often said to be signalled by intonation. For example, in (6), a phrase break after *verfolgt* tends to lead to the interpretation that it is the man with the motorbike which Rainer is following. A phrase break after *Mann* would tend to lead to the interpretation that Rainer is on his motorbike and is following a man whilst riding it. In the first case the prepositional phrase modifies the noun phrase (*den Mann*) and in the second it modifies the verb (*verfolgt*). This phrasing has the same effect in the English translation.

- (6) Rainer verfolgt den Mann mit dem Motorrad.
(‘Rainer is following the man with the motorbike.’)

However, it is often unnecessary to disambiguate between two readings, particularly if the

context is clear. It should therefore not be expected that speakers will make such distinctions all of the time. A study on Italian and English syntactic disambiguation (Hirschberg & Avesani 2000) showed this particularly clearly, not only for prepositional phrase attachments, as in (7a), but also for ambiguously attached adverbials, as in (7b) (adapted from Hirschberg & Avesani 2000: 93).

- (7a) Ha disegnato un bambino con una penna. ('lit. He drew a child with a pen')
(7b) Lui le aveva parlato chiaramente. ('lit. He to her has spoken clearly.')

The two readings of (7b) are either that it was clear that he spoke to her (the adverbial modifies the sentence) or that he spoke to her in a clear manner (the adverbial modifies the verb).

2.3. Information structure

An important linguistic function of intonation is the marking of information structure, in particular (a) the expression of givenness of entities within a chunk of discourse and (b) the division of utterances into focus and background elements. In both (a) and (b) we are dealing with a continuum rather than a dichotomy: entities are not simply given or new, but may have an intermediate status between the two extremes, just as an utterance might contain elements which are focussed to a greater or lesser degree. We deal with (a) and (b) in sections 2.3.1. and 2.3.2. respectively.

2.3.1. Givenness

Degrees of givenness can be expressed through the choice of words. A clearly new discourse element can be expressed with a noun and indefinite article, as in the underlined noun phrase in (8). A clearly given one can be expressed as a pronoun, as in (9).

- (8) Thomas isst einen Apfel. ('Thomas eats an apple.')
(9) A: Was ist mit dem Apfel passiert? ('What happened to the apple?')
B: Thomas hat ihn gegessen. ('Thomas ate it.')

An intermediate degree of givenness can be expressed by the use of a definite article, as in (10), where the word *Apfel* is considered to be more given than in (8), since it refers to a specific instance of an apple which has already been introduced into the discourse in some way.

- (10) Thomas isst den Apfel. ('Thomas eats the apple.')

Of course, degrees of givenness can also be expressed through intonation. For example, the word *Apfel* in (11) receives a pitch accent and is thus more prominent than the same word in the second turn (B) in (12). In B's turn *Apfel* is *deaccented*, which means that it does not receive an accent although it would be accented under default conditions, i.e. in an 'all-new' context such as in (11).

- (11) Thomas hat Hunger. Also isst er einen Apfel. ('Thomas is hungry so he eats an apple.')

- (12) A: Hast Du gesagt, dass Thomas mit einem Apfel jongliert?

B: Nein, er ISST einen Apfel.

('Did you say Thomas is juggling with an apple? No, he's eating an apple.')

(12) is similar to an example of Cruttenden's (in press) for English, given in (13).

- (13) A: Would you like to come to dinner tonight? I'm afraid it's only chicken.
B: I don't LIKE chicken.

Indian English, by contrast, does not deaccent, as in the example taken from Ladd (1996: 176), reproduced in (14).

- (14) If you don't give me that CIgarette I will have to buy a CIigarette.

Italian is similar to Indian English in that the nuclear pitch accent tends to go on the final lexical item regardless of whether it is given or not. In (15), the nuclear accent is on *casa* in both cases, whereas in English it would have gone on *outside* and *inside*.²

- (15) È un lavoro che si fa fuori CAsa o dentro CAsa?
(‘Is it a job which you do outside the HOME or inside the HOME.’)

Cruttenden (in press) refers to examples such as those in Italian and Indian English as having *reaccenting*. Not all types of accent are equally as strong, and therefore the context sometimes dictates not only whether an accent is present or not but also which type of accent may be used. The interested reader is referred to Baumann & Grice (to appear), where degrees of givenness are shown to be reflected in the type of accent used. A high accent is used for new information, and a step down in pitch onto the accented syllable for information which is not totally

² It is important to point out that this distribution of accents in Italian is only a tendency; it is quite possible to have a

given but, rather, *accessible*. No accent at all is used for totally given information.

2.3.2 Focus

The second aspect of information structure is the division of utterances into focus and background elements, based on the structure of the previous discourse and the intentions of the speaker. Although there is a relation between focus and newness on the one hand and background and givenness on the other, the two dimensions are generally orthogonal to each other. For example, an item in focus may be given within the discourse, as the name *Maria* in (16) B. Compare this to (17), where *Maria* is both in focus and new.

- (16) A: Liebst Du Maria oder Anna? ('Do you love Maria or Anna?')
B: [Ich liebe]_{background} [MaRIa]_{focus}
given given

- (17) A: Wen liebst Du? ('Whom do you love?')
B: [Ich liebe]_{background} [MaRIa]_{focus}
given new

Both of these structures represent so-called ‘narrow focus’, that is only one element is focussed. What is important is that this element is accented irrespective of its degree of givenness. In broad focus structures, where focus extends over a number of words, the relation between focussed elements and pitch accents is less direct. In many languages, larger focus domains are marked by only one or two pitch accents, a phenomenon called *focus projection* (cf. Selkirk, 1984, Uhmann, 1991). The preference as to which element receives the accent, and thus serves as *focus exponent*, is language specific. Ladd (1996) points out that many languages place the focus exponent on the argument rather than on the predicate. For example, in (18) the accent is placed on the argument, *Haus*, and the following predicate, *kaufen*, is left unaccented. This is the case even if the argument is followed by the verb, not only in German but also in English, as in (19).

- (18) Ich habe kein Geld übrig. Ich muss ein HAUS kaufen.
(‘I don't have any spare cash. I have to buy a HOUSE.’)

- (19) I don't have any spare cash. I have a HOUSE to buy.

As pointed out above, the tendency to accent the last lexical item is stronger in Italian than it

nuclear accent on *fuori* and *dentro* as well.

is in English or German. Thus, in (20) the final word is accented despite the fact that it is a verb, as in Ladd's (1996: 191) example.

- (20) Ho un libro da LEGgere. ('I have a book to read').

Another important influence on the accentability of words is their 'semantic weight'. In (21) B and C the noun phrases *meinen Anwalt* and *jemanden* are both arguments and in focus (i.e. part of the broad focus domain).³ However, *jemanden* is semantically 'light', since it is an unspecific pronoun, and thus does not receive an accent (cf. Uhmann 1991: 200).

- (21) A: Was haben Sie ihrer Aussage hinzuzufügen? ('Do you have anything to add to your evidence?')
B: Ich habe meinen ANwalt belogen. ('I lied to my lawyer.')
C: Ich habe jemanden beLOgen. ('I lied to someone.')

It is important to point out that there are differences even within a language as to where the nuclear accent is placed in broad focus contexts. One example of this is Greek, where the accent tends to be placed on the argument in statements but on the predicate in polar questions (Grice et al., 2000, Arvaniti et al., to appear, more on polar questions in 2.4. below).

2.4. Speech acts

Intonation is used to encode distinctions such as whether an utterance is intended as a request for information (Request) or as a request for the interlocutor to perform a particular action (Command). There are four major categories of communicative illocutionary 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 falling pattern, constituting an LHL sequence. Intonation plays a crucial role in distinguishing polar questions from, e.g., statements if there is no distinct interrogative syntax or question particle, such as in Italian. Even in German and English it is possible to ask a question using a fragment, as in (22), in which case intonation plays the major role in disambiguating the question from a statement, providing the context does not make it entirely clear that a question seeking confirmation is being asked.

³ They cannot be treated as entirely given, since they have not been mentioned in the immediately preceding context

(22) mit LEna?

('with LEna?')

Wh questions are usually accompanied with a falling intonation unless there is some additional paralinguistic meaning such as an element of insistence or politeness. In some cases, a syntactic Wh question in German can also be interpreted as a suggestion if uttered with a fall as in (23).



(23) Warum ziehst du nicht nach KaliFORnien?

('Why don't you move to California?')

2.5 Paralinguistic functions and iconicity of intonation

Intonation is often said to serve primarily an emotive function, implying an inherently iconic usage of pitch variations. Such fundamental iconicity further implies that the (paralinguistic) meaning differences in spoken language brought about by changes in pitch height are universally valid. This is, in principle, Bolinger's view when he claims that

intonation is part of a gestural complex, a relatively autonomous system with attitudinal effects that depend on the metaphorical associations of up and down – an elaborate scheme of iconism. It assists grammar – in some instances may be indispensable to it – but is not ultimately grammatical. (1985:106)

However, Bolinger (1985: 97 f.) relativises this claim by arguing that the iconicity of intonation is only 'symptomatic' in nature; pitch variations do not *directly* mirror the meaning they help to convey, as is the case – at least to a larger extent – with onomatopoeic expressions, such as *bang, smash* in English and *klatschen, gurren* in German (see Crystal (1987: 174 f.) for examples of sound symbolism in many languages).

Carlos Gussenhoven (2002, 2004) brought together research on the different factors affecting intonational form, which have led to claims of a universal form-function relation, and, crucially, showed how they interact. It is precisely the analysis of the interaction of the different factors which has explained apparent discrepancies in the form-function relation in cross-language comparisons. Gussenhoven claims that the form-function relations are based on three biological codes: the *frequency code*, the *production (phase) code* and the *effort code*. Each code has affective and/or informational interpretations and may have different linguistic manifestations in different languages.

(here: A), and are thus candidates for pitch accents.

According to the frequency code, which was introduced by Ohala (1983, 1984), size is suggested by pitch height: since a bigger larynx (including longer vocal folds) and a longer vocal tract produce lower frequencies, low pitch is associated with larger creatures and high pitch with smaller ones. The frequency code has affective interpretations along dimensions such as dominant~submissive or impolite~polite and more informational interpretations along dimensions such as certain~uncertain or – closely related – assertive~questioning, with low pitch attributed to the first pole and high pitch to the second (Gussenhoven 2004: 80 ff.). The most obvious linguistic manifestation of the frequency code is the distinction between statements and polar (yes-no) questions, which is a categorical manifestation of the assertive~questioning dimension. Polar questions are marked in a great number of languages by rising or high pitch (as in example (2) versus (1) above).

For many interpretations of the frequency code, it is the contour *endings* which are particularly important (cf. Ohala 1983, 1984, Gussenhoven 2004: 82). However, for a large number of languages it is not a final rise but rather an accentual rise which marks polar questions. This rise is often followed by a fall. A rising-falling contour is found in many Southern varieties of Italian (Bari, Palermo, Neapolitan; see Grice, D'Imperio, Savino and Avesani, 2005). This is illustrated in example (24), taken from a recording of Bari Italian (Grice et al., 2005: 370).



- (24) Lo mandi a MassimiLIAno? ('Will you send it to Maximilian?')

A similar contour is also found in varieties of Hungarian, Romanian and Greek (Grice, Ladd and Arvaniti, 2000), as well as in varieties of German, as shown in example (25) from a recording of a Palatinate dialect (Peters, 2004: 384). Note that the rise-fall is on the final unaccented syllable, in contrast to (24), where the rise is on the accented syllable.



- (25) Isch des e gute WIRTSfre::? ('Is that a good barkeeper (female)?')

The end of the contour is also important for the production code, which derives its interpretations from a gradual decrease in subglottal air pressure in the course of a breath group (Lieberman, 1967, Gussenhoven 2004). One consequence of the drop in subglottal pressure is a gradual lowering of pitch (along with intensity), throughout the phrase, referred to as *declination* (Cohen & 't Hart 1967). The central linguistic interpretation of this code is finality~continuation, marked by low versus high endings.

Many languages have distinct contours which they use to express non-finality, see for

example the contour in (5). However, as with questions, not all languages signal finality right at the end of a phrase. Palermo Italian, for instance, uses a rising type of accent instead (Grice 1995), although this rise is distinct from the question rise. A fall to low pitch can express varying degrees of finality, depending on the extent of the fall and the final pitch reached.

At the beginning of a phrase, the relation is reversed: an initial high accent often signals a new topic, whereas a relatively low accent at the beginning marks topic continuation (in German and English; cf. Wichmann et al. 2000), emulating an intake of breath and therefore increased subglottal pressure, leading to faster vibration of the vocal folds (producing higher pitch).

The third biologically determined code is the effort code, which is based on the physiological phenomenon that an increased effort in producing speech leads to greater articulatory precision. This is reflected by more pronounced and wider pitch movements (cf. Gussenhoven 2004: 85 f.). The primary informational function of this code in many languages is to express emphasis or importance achieved through gradient use of pitch height. Its most common categorical manifestation is accentuation used in the marking of focus (see section 2.3.2.) and the types of accent used to mark stages along the given~new continuum: As discussed in section 2.3.1., higher pitch is used for items which are new to the discourse, whereas a step down onto a lower pitch is used for items which are accessible to the hearer through context, but are not entirely given.

To sum up, a representative sample of prosodic functions and the means used to express them are shown schematically in Figure 1.

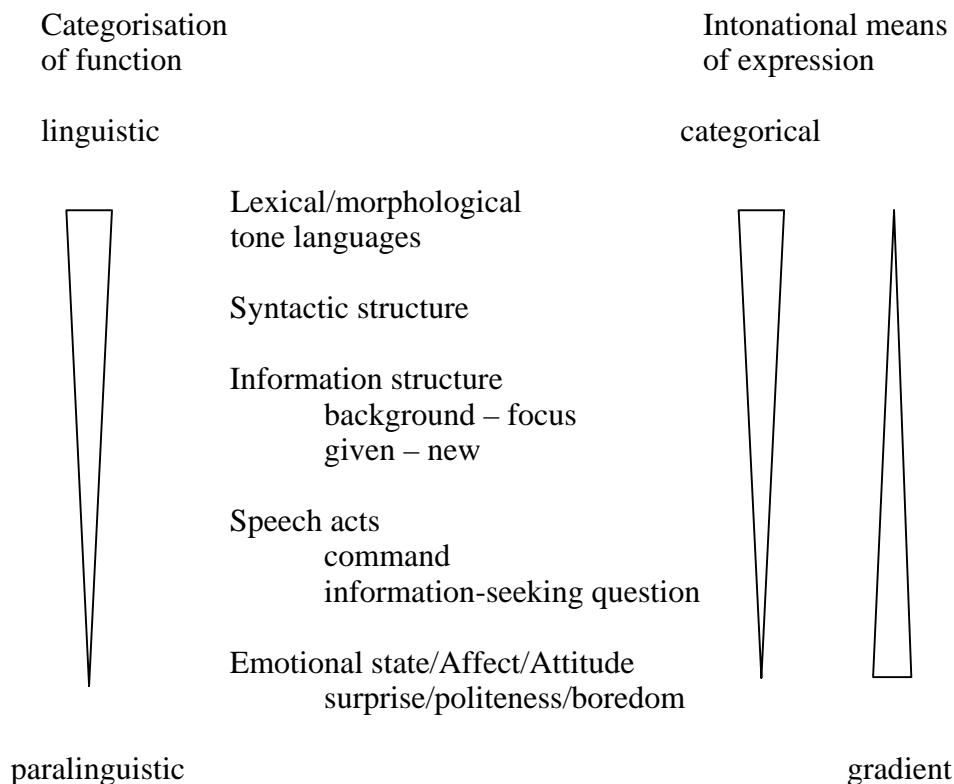


Figure 1 Functions of intonation and their intonational realisation

It should be clear from the figure (and from the discussion above) that although categorical means are employed to make lexical distinctions as well as distinctions pertaining to information structure and speech acts, it is not possible to state *either* that categorical means are used to express only linguistic functions, *or* that gradient means are used only for paralinguistic functions, although this is a widespread assumption. Therefore, anyone analysing the intonational forms of a language should keep an open mind when relating form to function. Furthermore, it should not be assumed that gradient means are universally valid, since different languages interpret pitch height in different ways.

3. Models of intonation

In the literature on intonation, pitch modulation is either captured as pitch *configurations* (as in the British School, cf. section 3.1.), such as rise, fall, rise-fall and so on, or as a sequence of *targets* (as in autosegmental-metrical models, cf. section 3.2.). Targets specify only specific points in the F0 contour, represented phonologically as ‘tones.’ H(igh) tones correspond to high targets, referred to as ‘peaks,’ L(ow) tones to low targets, referred to as ‘valleys’ or ‘troughs’. These tones can be combined into composite pitch accents, LH representing a rise, and HL a fall, or boundary tone combinations, e.g. LH representing a phrase final rise. In the British School, configurations such as rise or fall are the primitives (basic units), whereas in the autosegmental-metrical approach they are derived, the basic building blocks being the levels High and Low.

3.1. British School

British-style analyses (e.g. Crystal, 1969, Halliday, 1967, O'Connor & Arnold, 1973, and Tench, 1996, cf. also Kohler, 1991, for German), treat intonation in terms of dynamic pitch contours. The most important contour and the one by which tunes are classified is referred to as the ‘nuclear tone’. It starts at the ‘nucleus’ or ‘nuclear syllable’ (Halliday’s ‘tonic’), which is said to be the utterance’s most prominent syllable, and continues to the end of the phrase.

The nucleus represents the only obligatory part of a ‘tone group’. Maximally, a tone group consists of a ‘prehead’ (unaccented syllables before the first pitch accent), a ‘head’ (reaching from the first pitch accented syllable to – but not including – the nuclear syllable), a nucleus (last pitch accented syllable within the tone group) and a ‘tail’ (unaccented postnuclear syllables). Postlexical stresses (or *Druckakzente*), i.e. secondary prominences characterised by increased length and/or loudness but lacking an abrupt pitch movement (see section 1.1.), may occur within the prehead, the head, and the tail. Example (26) shows the structure of a tone group containing all possible parts (including a potential postlexical stress on *-kauf*):

(26)	Prehead	Head	Nucleus	Tail
Mag- da-	LE- na hat ein	HAUS	ge- kauf ^t	('Magdalena bought a house.')

The notation used in British-School analyses assigns a dot to every syllable, with stressed syllables larger than unstressed ones. Pitch accented syllables either represent turning points in a more or less smooth pitch contour (as the third syllable of *Magdalena* in (26)) or are characterised by a considerable pitch change within the syllable (as on *Haus* in (26)). The latter is indicated by a line. Due to the form of these symbols the notation has been called ‘tadpole’ notation. It has also been termed *interlinear*, since the transcription is placed between two lines indicating the upper and lower limit of a speaker’s pitch range. The usual method of transcription within the British School is to use tonetic stress marks for the nuclear contour, the pitch movement extending from the nucleus to the end of the phrase. This is called *intralinear* transcription, as in (27), where the diacritic indicates a high fall.

- (27) Magdalena hat ein `Haus gekauft. ('Magdalena bought a house.')

It is also possible to mark the beginning of the head and the direction the pitch takes during the head. Online material for practicing intonation within the British School is available at <http://www.eptotd.btinternet.co.uk/pow/powin.htm>.

3.2. Autosegmental-metrical models

The currently most widespread phonological framework for representing intonation is termed ‘autosegmental-metrical’, starting with the work of Pierrehumbert (1980), and treated in detail in Ladd (1996), in which the term was coined. The division of utterances into phrases and the assignment of relative prominence to elements within the phrase (phrasing and highlighting) represent the *metrical* aspect, which was first proposed by Liberman and Prince (1977). The association of the tones (grouped into accents – if the language has them – and boundary tones) with the metrical structure (in other words: the association of the tune with the text) represents the *autosegmental* aspect. The term autosegmental refers to the fact that the tune should be considered as reasonably autonomous with respect to the text – in fact they are represented as being on different tiers. A tune can thus be realised on a great many texts of different lengths and structures.

However, the tune has to be *anchored* to the text at strategic points – these are the associations between the two tiers.

The greatest advantage compared to the British School model is that tonal information can be precisely localised on single syllables and/or at the edges of phrases. In British School studies, the only direct connection between tones and text occurs on the nucleus. In most AM models, the nucleus does not have a special status. It is simply defined as the last fully-fledged pitch accent in a phrase, which means that there is no theoretical distinction between ‘prenuclear’ and ‘nuclear’ accents.

A widely used autosegmental-metrical framework for the description of intonation is the ToBI (‘Tones and Break Indices’) system, which was originally developed as a transcription system for American English, but has since become a general framework for developing intonation systems. There is a transcription system for Standard German, ‘GToBI’, which is based on speech data mainly from Northern German speakers (cf. Grice & Baumann 2002, Grice, Baumann & Benzmüller 2005 for an overview).

A (G)ToBI record consists of at least three different levels of description, which can be thought of as corresponding to autosegmental tiers. These tiers contain labels for text, tones, and break indices. The text tier provides an orthographic transcription of the words spoken, the tones tier mirrors the perceived pitch contour in terms of tonal events such as pitch accents and boundary tones, and the break index tier marks the perceived strength of phrase boundaries. Pitch accents are associated with lexically stressed syllables, indicated by a starred (*) tone placed within the limits of the accented word - generally at local F0 minima and maxima. Edge tones are assigned to phrase-final syllables, marked by ‘-’ or ‘%’ after the tone, signalling the edge of an intermediate (minor) phrase or a (major) intonation phrase, respectively (cf. section 1.2.).

As an example, the utterance in (26), which consists of a single intonation phrase, would be transcribed in GToBI as in (28).

- (28) MagdaLEna hat ein HAUS gekauft. (‘Magdalena bought a house.’)

L*	H*	L-%

The first (prenuclear) accent in the phrase is realised low in the speaker’s pitch range, the second (nuclear) one high, thus transcribed L* and H*, respectively. The tonal movement before and between these targets does not have to be transcribed, since no pitch minima or maxima are reached. Rather, the target points can be thought of as being joined up by quasi-linear ‘interpolation’. Finally, the falling nuclear movement is accounted for by the combination of a high accent and a low boundary tone (L-%). The combined notation of ‘-’ and ‘%’ stems from the fact

that the end of each intonation phrase necessarily coincides with the end of an intermediate phrase, since a hierarchical structure is assumed.⁴

The original ToBI model has been extended as a general framework for developing intonation systems for a large number of languages and varieties. Complete ToBI systems including online training materials are available for English, German, Korean, Japanese and Greek. These and other ToBI systems are described in detail in a book (Jun, 2005), and training materials as well as a number of related papers can be accessed from the ToBI homepage (<http://www.ling.ohio-state.edu/~tobi/>).

It is difficult to say which of these two models would work best teaching intonation to second language learners. The British School model is intuitively straightforward and has didactic origins. It is relatively easy to relate the transcription to an auditory impression. It is, however, very difficult to relate tonetic or interlinear transcriptions to F0 traces – something which might be a problem in an age where students have ever-increasing access to programmes which can estimate and display F0 contours. A further disadvantage of the British School model is that it is used less frequently than it used to be, so that research carried out for the purposes of preparing course materials must often be based on relatively old sources. Since pronunciation (including intonation) changes relatively quickly, both at a regional and standard level, this could be a problem, since any accompanying tapes will sound rather outdated and stilted.

The autosegmental metrical model is more helpful for students who might be interested in looking at F0 contours as well as listening. Further, a knowledge of this model is indispensable for anyone wishing to search the current literature for information on a specific language, or for communication amongst or with theoretical intonation researchers.

It must be stressed that both of the models are phonological in essence, and are therefore good for capturing the categories of the intonation system of a given system, but not suited for a detailed analysis into the finer phonetic details and gradient variation within a category. In other words, these models can be used for teaching what in segmental terms would be the 'phonemes' of a language, but not the allophonic variants.

4. Summary and Conclusion

In this paper we have provided an overview of the communicative functions attributed to intonation, starting out from the two main tasks intonation performs, i.e. highlighting and the division of utterances into chunks. In the languages we examined here, highlighting is achieved by means of stress and accent. However, not all languages have pitch accents and/or lexical stress, such

⁴ Due to the lack of a separate tonal target on the final syllable, an explicit symbol for tone immediately before the percentage '%' sign can be dispensed with. This notation is meant to increase the phonetic transparency of the contour, which used to be written as 'L-L%'.

as Korean (Jun, 2005), which uses phrasing to indicate narrow focus. All languages make use of phrasing of some kind.

Further, we have examined more specific linguistic and paralinguistic functions of intonation. At a clearly linguistic level, we have observed that intonation is not always used to disambiguate syntactically ambiguous structures but it can be in some languages in certain contexts (where disambiguation is necessary). As for information structure, givenness is expressed in some languages with deaccentuation, while in other languages there is no specific marking of givenness. Likewise, focus can be marked with certain types of accent. It is important to note, however, that not all languages use intonation to signal focus (e.g. Wolof; Rialland and Robert, 2001).

At the more paralinguistic level there appear to be more commonalities across languages but it is precisely these commonalities which lead to misunderstandings, since one language might interpret an utterance with high pitch as friendly (e.g. British English), whereas another might interpret the same utterance as emphatic (e.g. Dutch), a result which depends on the weighting of the frequency and effort codes (Chen, 2005).

Finally, we have outlined two influential models for transcribing intonation, the British School and the autosegmental-metrical approach. We have also provided links to further materials and exercises so that interested readers can hear examples in each model, and, in the autosegmental-metrical approach, in a number of languages.

Acknowledgements

We would like to thank Barbara Gili Fivela and Michelina Savino for their intuitions on Italian intonation, and Michelina Savino for providing the Italian recordings.

References

- Arvaniti, Amalia, D. Robert Ladd and Inneke Mennen to appear Tonal Association and Tonal Alignment: Evidence from Greek Polar Questions and Contrastive Statements, *Language and Speech*.
- Bach, Kent and Harnish, Robert M. 1979 Linguistic Communications and Speech Acts. Cambridge, Mass.: Blum-Kulka, S., House, J., and Kasper, G. 1989. *Cross-cultural pragmatics: requests and apologies*. N.J.: Ablex.
- Baumann, Stefan to appear The Intonation of Givenness – Evidence from German. PhD thesis, Saarland University.
- Baumann, Stefan and Grice, Martine to appear The Intonation of Accessibility. Special Issue ‘Prosody and Pragmatics’ *Journal of Pragmatics*.
- Beaver, D., B. Clark, E. Flemming, T. F. Jäger. and M. Wolters 2004 When semantics meets phonetics: Acoustical studies of second occurrence focus. Stanford University. Ms.
- Beckman, Mary E. 1986 *Stress and Non-Stress Accent*. Dordrecht: Foris.
- Bolinger, Dwight 1958 A Theory of Pitch Accent in English. *Word* 14, 109-149. Reprinted in Bolinger 1965, 101-117.
- Bolinger, Dwight 1964 Intonation: Around the Edge of Language. *Harvard Educational Review* 34:282-296.
- Bolinger, Dwight 1985 The Inherent Iconism of Intonation. In: John Haiman (ed.), *Iconicity in Syntax*, Amsterdam and Philadelphia: John Benjamins. 97-108.

- Büring, Daniel 1996 On (De)Accenting. Talk presented at the SFB 340 conference in Tübingen, October.
- Büring, Daniel 2003 Focus Projection and Default Prominence. Ms. University of California, Los Angeles.
- Chen, Aoju 2005 *Universal and language-specific perception of paralinguistic intonational meaning*. Utrecht: LOT.
- Chomsky, Noam and Halle, Morris 1968 *The Sound Pattern of English*. New York: Harper and Row.
- Cohen, Antonie and 't Hart, Johan 1967 On the Anatomy of Intonation. *Lingua* 19, 177-192.
- Cruttenden, Alan 1986 *Intonation*. Cambridge: Cambridge University Press.
- Cruttenden, Alan in press The Deaccenting of Given Information: a Cognitive Universal? In: G. Bernini (ed.), *The Pragmatic Organisation of Discourse*. The Hague: Mouton de Gruyter.
- Crystal, David 1969 *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- Crystal, David 1987 *The Cambridge Encyclopedia of Language*. Cambridge University Press.
- Grice, Martine 1995 *The intonation of Interrogation in Palermo Italian – Implications for Intonation Theory*. Tübingen: Niemeyer.
- Grice, Martine and Baumann, Stefan 2002 Deutsche Intonation und GToBI. *Linguistische Berichte* 191, 267-298.
- Grice, Martine, Baumann, Stefan and Benzmüller, Ralf 2005 German Intonation in Autosegmental-Metrical Phonology. In: Sun-Ah Jun (ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. 55-83.
- Grice, Martine and Savino, Michelina 2004 Information Structure and Questions: Evidence from Task-oriented Dialogues in a Variety of Italian. Tübingen: Niemeyer.
- Grice, Martine, D'Imperio, Mariapaola, Savino, Michelina and Avesani, Cinzia 2005 Strategies for Intonation Labelling across Varieties of Italian. In: Sun-Ah Jun (ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. 362-389.
- Grice, Martine, Ladd, D. Robert and Arvaniti, Amalia 2000 On the Place of Phrase Accents in Intonational Phonology. *Phonology*, 17, 143-185.
- Gussenhoven, Carlos 1983 Focus, Mode, and the Nucleus. *Journal of Linguistics* 19, 377-417.
- Gussenhoven, Carlos 2002 Intonation and Interpretation: Phonetics and Phonology. *Proceedings 1st Int. Conference on Speech Prosody*, Aix-en-Provence, 47-57.
- Gussenhoven, Carlos 2004 *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Halliday, M.A.K. 1967a *Intonation and Grammar in British English*. The Hague: Mouton.
- Halliday, M.A.K. 1967b Notes on Transitivity and Theme in English, Part 2, *Journal of Linguistics* 3, 199-244.
- Hart, Johan 't, Collier, René and Cohen, Antonie 1990 *A Perceptual Study of Intonation: An Experimental-Phonetic Approach*. Cambridge: Cambridge University Press.
- Hirschberg, Julia and Avesani, Cinzia 2000 Prosodic Disambiguation in English and Italian. In: Antonis Botinis (ed.), *Intonation. Analysis, Modelling and Technology*. Dordrecht: Kluwer Academic Publishers. 87-95.
- Jun, Sun-Ah 2005 *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.
- Jun, Sun-Ah 2005 Korean Intonational Phonology and Prosodic Transcription. In: Sun-Ah Jun (ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. 201-229.
- Keating, Patricia, Cho, Taehong, Fougeron, Cecile and Hsu, C. 2003 Domain-initial articulatory strengthening in four languages. In: *Phonetic Interpretation* (Papers in Laboratory Phonology 6), edited J. Local, R. Ogden, R. Temple, Cambridge University Press, 143-161.
- Kohler, Klaus 1991 Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics. *AIPUK* 25, 115-185.
- Ladd, D. Robert 1980 *The Structure of Intonational Meaning: Evidence from English*.

- Bloomington: Indiana University Press.
- Ladd, D. Robert 1996 *Intonational Phonology*. Cambridge: Cambridge University Press.
- Liberman, Mark and Prince, Alan 1977 On Stress and Linguistic Rhythm. *Linguistic Inquiry* 8, 249-336.
- Lieberman, Phillip 1967 *Intonation, Perception, and Language*. Cambridge, MA: MIT Press.
- O'Connor, J.D. and Arnold, G.F. 1973 *Intonation of Colloquial English*. London: Longman.
- Ohala, John J. 1983 CrossLanguage Use of Pitch: An Ethological View. *Phonetica* 40, 1-18.
- Ohala, John J. 1984 An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice. *Phonetica* 41, 1-16.
- Partee, Barbara H. 1999 Focus, quantification, and semantics-pragmatics issues. In: P. Bosch and R. van der Sandt (eds.), *Focus: Linguistic, Cognitive, and Computational Perspectives*. CUP, 213-231.
- Peters, Jörg 2004 *Regionale Variation der Intonation des Deutschen. Studien zu ausgewählten Regionalsprachen*. State doctorate thesis (Habilitationsschrift), University of Potsdam.
- Pierrehumbert, Janet B. 1980 The Phonetics and Phonology of English Intonation. PhD thesis, MIT. Bloomington: Indiana University Linguistics Club.
- Rialland, Annie and Robert Stéphane 2001 The intonational system of Wolof. *Linguistics* 39: 893-939.
- Roach, Peter J. 1994 ‘Conversion between prosodic transcription systems: “Standard British” and ToBI’, *Speech Communication*, 15, 91-99.
- Rochemont, Michael 1986 *Focus in Generative Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Searle, John 1969 *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Selkirk, Elisabeth 1984 *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Stechow, Arnim von and Uhmann, Susanne 1986 Some Remarks on Focus Projection. In: Werner Abraham & S. de Meij (eds.), *Topic, Focus, and Configurationality*. Amsterdam: John Benjamins.
- Tench, Paul 1996 *The Intonation Systems of English*. London: Cassell.
- Uhmann, Susanne 1991 *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.
- Wichmann, Anne, House, Jill and Rietveld, Toni 2000 Discourse Constraints on F0 Peak Timing in English. In: Antonis Botinis (ed.), *Intonation. Analysis, Modelling and Technology*. Dordrecht: Kluwer Academic Publishers. 163-182.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. and Price, P. J. 1992 Segmental durations in the vicinity of prosodic phrase boundaries. *JASA* 92, 1707-17.
- Williams, Briony 1996a The status of corpora as linguistic data. In: G. Knowles & P.R. Alderson (eds.), *Working with Speech: The Computational Analysis of Formal British English Speech*. London: Longmans.
- Williams, Briony 1996b The formulation of an intonation transcription system for British English. In: G. Knowles & P.R. Alderson (eds.), *Working with Speech: The Computational Analysis of Formal British English Speech*. London: Longmans.