German Intonation in Autosegmental-Metrical Phonology

Martine Grice, Stefan Baumann and Ralf Benzmüller

3.1 Introduction

English, Dutch and German are often claimed to have very similar prosody and intonation, an observation that might be related to the fact that they all belong to the West Germanic language family. All three have a stress-timed rhythm with left-headed feet, and they all make use of a number of different pitch accents for highlighting information, and of edge tones for delimiting phrases. In broad focus contexts, they all place the nuclear pitch accent on the final argument in the intonation phrase.

Within the autosegmental-metrical framework there are essentially two major approaches to German intonation. On the one hand, there are accounts such as those by Féry (1993) and Grabe (1998), which follow Gussenhoven's (1984) analysis of Dutch. On the other there is GToBI, a consensus system developed by Martine Grice, Matthias Reyelt, Ralf Benzmüller, Anton Batliner and Jörg Mayer (Grice et al 1996; Reyelt et al 1996), which is closely related to the original English ToBI (Mainstream American English ToBI, henceforth MAE_ToBI: Beckman & Hirschberg 1994, Beckman & Ayers-Elam 1997, Beckman, Hirschberg & Shattuck-Hufnagel, this volume) and the analysis of English by Pierrehumbert (1980) upon which MAE_ToBI is based.

The major differences between these approaches are twofold. First, in Gussenhoven's model pitch accents, like feet, are always left-headed. This means that a given pitch accent cannot account for the pitch before the accented syllable to which it is associated. GToBI and MAE_ToBI, by contrast, have not only left-headed but also right-headed pitch accents. The latter account both for the pitch on the accented syllable

and for the pitch immediately before it, in which case the tone to the left is referred to as a leading tone. Second, Gussenhoven analyses nuclear contours as a combination of a pitch accent and an intonation phrase boundary tone. GToBI and MAE_ToBI instead postulate an additional tone after the final pitch accent in the phrase. This extra tone is referred to as the phrase accent.

At first glance, it might appear that the differences result from the fact that one group of researchers consider German to be like English, whereas the other take it to be like Dutch. However, since Gussenhoven also disputes the existence of leading tones and phrase accents in English, it becomes obvious that the differences are of a theoretical rather than a typological nature.

In this chapter we shall begin by looking at the relatively theory-neutral traditional literature on German intonation, often based on auditory impressions with a great deal of phonetic detail. We then go on to give an overview of the autosegmental-metrical literature which builds on Gussenhoven's work, and provide a detailed exposition of GToBI. Finally, we offer motivation for the analysis used in GToBI as compared to the other autosegmental-metrical models.

3.2 Accounts of German Intonation

Traditionally, intonation has been analysed either in terms of tonal configurations, i.e. pitch contours whose direction is important, or in terms of levels, where the pitch range is divided up into a number of discrete levels. In the latter, intonation contours are derived from sequences of these levels. We shall first examine the configurations-based accounts and then go on to look at early levels approaches, and at the more recent levels-based autosegmental-metrical accounts.

3.2.1 Configurations-based approaches

Early accounts of German intonation, such as those by von Essen (1964), Pheby (1975), Kohler (1977), and Fox (1984) are mainly auditory-based and didactically oriented, representing intonation patterns with a detailed interlinear transcription of the pitch of each syllable of an utterance. All of the above are akin to British-style analyses, e.g. Crystal (1969) and Halliday (1967), in treating intonation in terms of dynamic pitch contours, and in attributing particular importance to the nucleus (Halliday's 'tonic' and Pheby's 'Tonstelle') which is said to be the utterance's most prominent syllable. For Pheby and Fox, the contour by which tunes are classified starts at the nuclear syllable and continues to the end of the phrase. In the British School, this contour is referred to as the nuclear tone.

A somewhat different configurations-based approach has been proposed by Selting (1995). Her aim is to develop a description system for the analysis of spontaneous dialogues. Selting's model is auditory-based, partly supported by instrumental analysis. She distances herself from the British School in regarding intonation contours as holistic units with no special status assigned to the nucleus. For Selting, unlike for the early auditory studies of German, intonation does not reflect grammatical structure, but is used for signalling 'prosodically cohesive' units relevant for discourse organisation, e.g. in the construction of turns. Selting's is what can be termed an 'overlay approach' (Ladd 1996). She not only specifies the shape of individual pitch accents but also whether the upper and lower limits of the pitch range are globally declining, inclining or level.

Kohler treats pitch accents as pitch peaks which may be aligned in different ways with the text. In his more recent instrumentally based work (e.g. 1991), he shows that it is possible to differentiate between three types of peak (early, medial, and late) and to assign pragmatic interpretations to them: An early peak, where the peak is on the prenuclear syllable, marks a self-evident or established fact. A medial one, where the

peak is around the middle of the accented syllable, indicates a new fact. A late peak, occurring towards the end of the accented syllable or even on the following syllable, places emphasis on a new fact and/or represents greater involvement on the part of the speaker than is the case with a medial peak (1991:160).

3.2.2 Early levels-based approaches

Not all accounts of German have been configurations-based. In the early sixties, Moulton (1962), in the American structuralist tradition of Pike (1945) and Trager & Smith (1951) described the intonation of German in terms of distinct pitch levels (Moulton had three rather than the usual four pitch levels; he did not discuss level 4, the emphatic level). Like Trager & Smith, Moulton also had what is called 'terminal contours', indicating whether the pitch was rising, falling or sustained at the very end of the phrase, thus making the approach more a mixture of levels and contours than a strict levels approach. The first strict levels approach, Isačenko & Schädlich (1966), reduced the number of levels to two. They resynthesized utterances on high and low monotone pitch levels with a step up or down from one level to another, either before the accented syllable (which they refer to preictic, the ictus being the accented syllable) or after it (postictic). The preictic fall is equivalent to Kohler's early peak contour, the postictic fall to the medial or late peaks. We shall return to early peak contours in the discussion of GToBI pitch accents.

3.2.3 Autosegmental-metrical accounts

More recent levels-based approaches have been developed within the autosegmental-metrical framework. Autosegmental-metrical (AM) is a term coined by Ladd (1996) to refer to the approaches to intonation which developed following on from the seminal work of Pierrehumbert (1980). These approaches generally make use of

minimally two (H and L), maximally three (H, L and M) levels for the description of intonation. These may have a prominence-lending function, being grouped together into pitch accents. Pitch accents are generally either monotonal (e.g. H*) or bitonal (e.g. L*+H). The starred tone is said to phonetically align with the accented syllable, although recent research has shown that alignment is more complicated than this (Arvaniti et al 1999). If the unstarred tone precedes the starred tone, it is referred to as a leading tone. If it follows the starred tone, it is a trailing tone. As we have already seen, one of the differences between Pierrehumbert's model and GToBI on the one hand and the rest of the AM models for German on the other is that the former have both leading and trailing tones, whereas the latter have only trailing tones. We deal with this issue in more depth in section 3.4.1.

Tones may also have a delimitative function, acting as initial or final edge tones of intonationally relevant phrases. In the models surveyed below, tones are phonetically realised as coordinates on the frequency-time axis. However, the scaling of these tones when they are combined into accent or accent-edge tone clusters is not always transparent. As we shall see, the use of H and L tones differs considerably among the different accounts. Furthermore, scaling is affected by downstep, which lowers the pitch of certain H tones, or even upstep which raises the pitch of both H and L tones.

The AM models of German intonation include those of Wunderlich (1988), Uhmann (1991), Féry (1993) and Grabe (1998), the latter two in turn influenced by Gussenhoven's (e.g. 1984) account of the intonational systems of English and Dutch. Since GToBI is based on autosegmental-metrical theory, we shall briefly survey each approach,[1] list their inventories of pitch accents and boundary tones, and examine how they describe commonly occurring tunes, to pave the way for the comparison with GToBI in section 3.4.

We begin by looking at the work of Wunderlich (1988). Like most other German phonologists, he emphasises the 'grammaticalised' functions of intonation, especially sentence modality and focus-background structure (cf. e.g. Altmann, Batliner & Oppenrieder (1989), who provide extensive experimental data relating to such functions). Wunderlich's inventory of intonation patterns consists of single accents, accent-accent sequences and accent-boundary tone sequences. They are listed below, along with the contexts in which they occur, if any are given:

H*	peak accent (Gipfelakzent) - default accent
H* H L*	bridge accent (Brückenakzent) - multiple foci, contrast
%H L*	falling-to-low accent (Fallend-Tiefakzent) - exclamations
L* H%	low accent-to-rise (Tiefakzent-Steigend)
L* H (H%)	echo accent (Echoakzent) - echo questions
H* H	left bridge pillar (linker Brückenpfeiler) - beginnings of lists

Wunderlich claims that each pattern may have several functions, paying particular attention to the functions of the bridge accent. This accent type is said to be typical of German sentences with multiple foci: after the peak accent (H*) the pitch stays high until it sharply falls on or to the accented syllable (L*). A floating H tone accounts for the high pitch between the two accents. It surfaces as what appears to be at once a trailing H for the first and a leading H for the second accent. %H L*, where %H stands for a phrase initial boundary tone, is used in the second half of a bridge accent if there is a sentence-medial phrase boundary separating the two halves of the bridge. GToBI would transcribe the final accent in the bridge accent as having a leading H tone, which would correspond to both the floating H in H* H L* and the %H in %H L*.

Uhmann (1991) provides a book-length and therefore more detailed account of German intonation. Like Wunderlich, she restricts her investigation to the prosodic marking of grammatical functions, especially the relation between intonation and focus-background structure. Uhmann's inventory consists of an optional initial boundary tone L% or H% (the default being mid), an obligatory final boundary tone L% or H%, and four pitch accents L*, H*, L*+H, H*+L. The nuclear accent is always bitonal, prenuclear accents can be mono- or bitonal. Tonal targets before the accented syllables (i.e. leading tones) are not considered necessary.

Uhmann treats boundary tones as the phonological correlates of phrasing, and pitch accents as the phonological correlates of the focus feature. She assigns more or less distinctive meanings to the pitch accents: prenuclear L*+H functions as topic marker, L* highlights background constituents, H* highlights focus or background constituents, and H*+L represents the default focus accent.

She lists the following patterns and the sentence modality with which they co-occur:

H*+L	L%	declaratives, w-questions
L*+H	H%	echo questions, yes/no-questions
H*+L	H%	yes/no-questions (nuclear pitch accent marked)
L*+H	L%	w-questions (nuclear pitch accent marked)

The account proposed by Féry (1993) also deals with the influence of focus structure on German tonal patterns. Her inventory slightly deviates from Uhmann's: She has the same four bitonal pitch accents. The monotonal ones are derived by phonological rule rather than being underlyingly part of the inventory. Féry does not posit an initial boundary tone, although she does have an optional high (H%) terminal boundary tone which is used when the pitch at the boundary is considerably higher than the pitch of the

trailing tone of the nuclear accent. Following Gussenhoven (1984), she claims that all pitch accents are left-headed and underlyingly bitonal (H*+L or L*+H)[2], although her inventory has exceptions which will be discussed below. In a sequence of pitch accents (e.g. H*+L H*+L), a prenuclear accent can be linked to a nuclear accent, either partially, in which case the trailing tone of the prenuclear accent is associated with a syllable near the nuclear accent resulting in what appears to be a surface tritonal accent with a leading L tone (e.g. H* L+H*+L), or totally, in which case the trailing tone of the prenuclear accent is deleted, resulting in H* H*+L.

One aim of Féry's study is to give a comprehensive overview of the variety of tonal patterns occurring in German. She describes in detail two different types of hat pattern (Wunderlich's bridge accent). The first is exemplified in (1).

(1) H* H*+L

BALD ist sie DA (Féry 1993:150)

SOON is she THERE

'SOON she'll be THERE'

It is usually realised in one phrase. The second type of hat pattern surfaces as a sequence of a rise and a fall (L*+H H*+L), an example of which is given in (2).

(2) L*+H H*+L

geSCHLAfen hat KEIner von uns (Féry 1993:129)

SLEPT had NONE of us

'NONE of us had SLEPT'

Féry claims that this pattern is obligatorily realised in two intermediate phrases, and that it can be distinguished from the first type of hat pattern by the contexts in which it occurs. She claims that the first type has a very wide usage, whereas the second is restricted to topic-comment sentences, i.e. it consists of a topic accent (L*+H) and a focus accent (H*+L), or to several kinds of contrast, gapping, or clefts.

Féry postulates six different nuclear contours, giving the contexts in which they typically occur, or the meanings which they impart with respect to sentence modality, pragmatic interpretation and speaking style.

H*+L	declaratives, w-questions, wishes, imperatives
L*+H	progredient intonation, questions (e.g. echo questions), uncertainty/indignation
H*+L H%	questions, threats
L*+H+L	implying ‘of course’, slightly menacing
H+H*+L	TV-reporter style
H*+M	calling contour

However, the structure L*+H+L is problematic, since Féry has to take recourse to a bitonal trailing tone. Such a tone is not needed anywhere else in the system and one might consider whether the contour would have been better described as a L*+H followed by L%. She rejects this analysis, presumably because of the way the tones are distributed across the syllables; in terms of phonetic alignment, the rise-fall is not simply a mirror image of the fall-rise (which she represents as H*+L H%). GToBI, which, as we shall see, has not only trailing tones but also two different boundary tones, is able to capture a rising-falling pattern as L*+H L-(%), where H is a trailing tone. This is not parallel to the falling-rising pattern H* L-H%, where the alignment of the fall is

more variable, depending on the position of postnuclear stresses and is represented as an L intermediate phrase boundary tone (also referred to as a phrase accent, see sections 3.3.3 and 3.4.2 below). GToBI therefore correctly predicts that the alignment of tones with the text is different in each pattern. Furthermore, the self-evident or 'of course' meaning assigned to L*+H+L, which Féry also refers to as a late peak contour ('später Gipfel'; 1993:96), is not the same as the meaning of Kohler's late peak. In fact, as we have seen in 3.2.1, Kohler assigns 'self-evidence' to a completely different contour: the early peak.

The description of stylised contours, e.g. calls, poses another problem for Féry's model. She represents the calling contour as H*+M, thus adding a mid pitch level to her inventory. Although there are arguments for treating calling contours as distinct from other intonational phenomena since they are often chanted, and might therefore necessitate a more musical notation, GToBI accounts for calling contours in using the regular intonational inventory: the accented syllable is high, H*, and the step down represented as a H- phrase accent is downstepped (!H-).

Although Féry explicitly claims that her pitch accents are left-headed, that is, the starred tone is always on the left, she allows for an accent where it is evident that the pitch before the starred tone is relevant for the accent shape. This is the early peak accent H+H*+L. We shall see in section 3.4.1 that this contour is also represented in GToBI, albeit with a different sequence of tones.

3.3. GToBI

3.3.1 Preliminaries

GToBI is a set of conventions for labelling German intonation with the aim of being easy to learn, reliable, and adaptable for different labelling purposes. It is therefore not a strictly phonological description of German intonation. It was developed

between 1995 and 1996 by researchers from Saarbrücken, Stuttgart, Munich and Braunschweig with a view to facilitating the exchange of prosodically annotated data. A cross labeller consistency test with the consensus system was carried out and reported on in Grice et al (1996) and Reyelt et al (1996). Results showed that labellers were able to use GToBI consistently, most of them having learned it within a short period of time from printed training materials and accompanying sound files (Benzmüller & Grice 1997) and with little or no individual coaching. The GToBI training materials have been updated and are available via the GToBI home page http://www.coli.uni-sb.de/phonetik/projects/Tobi/gtobi.html. What is presented in this chapter is a slightly modified version of the original GToBI.

Part of the flexibility of GToBI is achieved by different levels of description - so called tiers. The consensus system comprises at least 3 tiers, containing labels for words, tones, and break indices. As a general principle, information is only encoded if it cannot be derived (automatically) from labels from other tiers or from the speech signal. Thus, only mismatches, or non-default correspondences are transcribed by hand.

The words tier provides an orthographic transcription of the words spoken. On the tones tier the perceived pitch contour is transcribed in terms of pitch accents and boundary tones, with symbols for pitch range modifiers such as downstep and upstep placed immediately before the affected tone. Phrase boundary strength information is recorded in the break index tier. Other information may be added in an optional miscellaneous tier. An example screen shot with speech waveform, labels for tones, words, break indices and miscellaneous information, along with the F0 contour is given in figure 3.1. The labels will be discussed below.

FIGURE 3.1

Much of the coded information is only interesting if it involves relating the tiers to each other. For instance, in the tone tier there is information as to which pitch accents are realised. Now, one of the functions of pitch accents is to highlight particular words. Information as to which words are highlighted by means of a pitch accent can only be gleaned from relating the position in time of each tone label to the positions of the word labels. Thus, a pitch accent label falling within the bounds of an annotated word is taken to highlight that word. The stressed syllable of the word must be identified separately, since the convention is not to mark stress in the orthographical representation.

We assume that most databases will be syllabified or even be annotated at the segmental level, in which case the syllable bearing a given accent can be found by relating the time stamp of the tone label to that of the syllables or segments. However, should a database have no such annotations we suggest explicitly marking cases where word stress is optional, as in the word for 'coffee' which may be pronounced '*Kaffee* or *Kaf'fee* (Duden 2000). As above, the IPA stress mark (') is inserted in the orthographical string before the stressed syllable. We also suggest marking cases where the pitch accent occurs on a syllable other than the primary lexical stress. In (3a) below, there are no stress marks, since the accent is on the lexically stressed syllable *zwan*. This is not the case in (3b), where the accent is shifted backwards onto *hun*.[3] The IPA stress mark is given in (3b) as it would appear in the label file.

(3) a hundertzwanzig

hundred and twenty

b 'hundertzwanzig Mann

hundred and twenty men

(adapted from Giegerich 1985:218)

Since the information as to which syllable is accented is gained only indirectly, it is important that the label for a pitch accent is placed within the bounds of the associated syllable. This is not a problem if the pitch accent peak or valley occurs within the syllable. In this case, the label is placed on the peak or valley. However, if the peak or valley occurs outside the syllable, we follow the MAE_ToBI conventions: the label has to be placed within the associated syllable and the actual peak or valley is marked with a '>' or '<' label, depending on whether it occurs before or after the associated syllable, respectively. In German the peak of a L+H* pitch accent is generally reached late in an accented syllable, often even later (especially if the accented syllable is short). This is also occasionally true for simple H* pitch accents. To avoid a situation where labellers have to decide on the location of syllable boundaries, the '<' label should be used in cases where the peak is somewhere in the vicinity of the syllable boundary (the same consideration holds for L* and valleys). An example of the '<' symbol can be found in figure 3.1.

Transcriber confidence as to the accuracy of individual labels is captured by a '?' flag after uncertain labels, and a '$' flag where the example is perceived to be a prototypical realisation of a given category.

3.3.2 Pitch Accents

The six basic pitch accents in GToBI are described below. The H and L tones are described as high or low relative to a speaker's pitch range which can be thought of as having a topline as an upper limit and a baseline as a lower limit. As a rule of thumb, tones which are perceived as high are roughly in the top three quarters of the range, those perceived as low in the bottom quarter.

- H* 'peak accent'

A canonical H* syllable is perceived as relatively high and may be preceded by a shallow rise.

- L+H* ‘rise from low up to peak accent’

Here, as in H*, the accented syllable is perceived as high. It is preceded by a syllable with a low pitch target which leads to a sharp rise in (or a jump up to) the accented syllable. The peak is often late in the accented syllable (cf. Adriaens 1991, Grabe 1998).

- L* ‘low accent’

The L* syllable is a local pitch minimum low in the speaker’s range. It may be preceded by a shallow fall.

- L*+H ‘valley accent plus rise’

Here a low target within the accented syllable is followed by a rise, starting late in the accented syllable and reaching its peak on the next syllable (or sometimes later). In contrast to L+H*, the perceived pitch of the accented syllable is low.

- H+L* ‘step-down from high to low accent’

The accented syllable is low with a valley clearly at or very near the bottom of the speaker’s range. It is preceded by a high pitch target which generally occurs on the syllable immediately preceding the accented syllable.

- H+!H* ‘step-down from high to mid accent’

As in H+L*, the accented syllable is preceded by a higher pitch on the immediately preceding syllable. However, the accented syllable is not low, but rather around the middle of the range between the ‘H+’ peak and the speaker’s baseline. If H+!H* is immediately followed by a low boundary tone there is a continuous fall from the preaccented syllable, through the accented syllable up to the next stressed syllable, if there is one, otherwise to the final syllable of the phrase.

The absence from the GToBI inventory of a H*+L accent will be discussed in detail in section 3.4.2.

The six basic pitch accents may additionally be scaled within a modified pitch range. The most common modification involves a lowering of the topline by a process of downstep, shifting the pitch of H tones downwards. When this occurs, the affected H tone is marked by a preceding '!' symbol. This is what happens in the H+!H* pitch accent. However, any H tone from the basic set of pitch accents can undergo downstep (e.g. !H*, L*+!H, and so on). It is important to note that within an intonation phrase all H tones following a downstep are scaled within the same reduced range. These following H tones are not especially marked for range, unless there is a subsequent step down. When there is more than one step down, i.e. when downstep occurs in sequence, then it is transcribed separately at each step, such as in the second phrase of figure 3.1, ...*WAAgerecht RECHTS von der GOLDmine* ('...in a horizontal line to the right of the gold mine'): 'H* !H* !H* L-%'. In the more phonologically oriented autosegmental-metrical models, downstep is either triggered automatically by a particular pitch accent type, as in Pierrehumbert's original English model (1980), or treated as an optional operation (Gussenhoven 1984). Since GToBI is more surface-oriented, it is simply flagged explicitly each time there is a step down in pitch.

When a new phrase is begun after a phrase containing a downstep, the pitch range is reset. In certain cases, a sequence of downsteps may be followed by a reset within a phrase, usually just before the nuclear accent. That is, after a sequence of steps down, there is a step up to the peak on the nuclear syllable. This happens in English (Ladd 1983:735, example 4b) and has also been attested in Southern German (Truckenbrodt 1998, to appear). GToBI makes use of an upstep '^' symbol to capture such cases, as exemplified in figure 3.3 (section 3.4.1 (i)) . Phrases may also contain sequences of pitch accents where each accent involves a step up in pitch (e.g. in emphatic speech). Such contours are transcribed by Selting (1995) as globally rising. In GToBI each accent of such a sequence is marked with a '^' symbol.

The explicit marking of upstep, which distinguishes GToBI from other autosegmental-metrical accounts, is not only used to indicate a step up within a sequence of pitch accents, it also describes a step up to a boundary tone, as will be shown in section 3.3.3 below.

3.3.3 Boundary Tones

GToBI differentiates two levels of phrasing: the (minor) intermediate phrase and the (major) intonation phrase. Each intonation phrase contains at least one intermediate phrase, and each intermediate phrase contains at least one pitch accent.[4]

The edge tones for these phrases determine the contour from the last tone of the last pitch accent until the end of the phrase. There are three intermediate phrase edge tones:

- L-

L- constitutes an F0 minimum low in the range.

- H-

H- has roughly the same F0 value as the peak corresponding to the most recent H tone in the phrase. Given enough distance, there is a plateau between a nuclear H tone and the end of the phrase.

- !H-

A H- tone can also be downstepped in relation to a previous accentual H tone. This most commonly occurs in calling contours.

Theoretically, there is the possibility of a fourth intermediate phrase boundary type, ^H- , although we have not to date found examples which would unambiguously distinguish it from H-.

The target for the intermediate phrase edge tone is often reached at a postnuclear stressed syllable (if there is one) and extends up to the beginning of the last syllable of the phrase. The tendency for the intermediate phrase boundary tone to align with

postnuclear stressed syllables is reported on in Grice and Benzmüller (1998) and is evidence for it being a phrase accent, as discussed in Grice, Ladd and Arvaniti (2000). Phrase accents are tones which function as edge tones but can also associate with stressed syllables or other tone-bearing units. In GToBI there is an option to explicitly transcribe this association with a separate L(*) or H(*) label, whereby the star in brackets denotes the secondary nature of the postnuclear prominence (see figure 3.4 in section 3.4.1 (i) for an example of usage).

An intonation phrase (IP) edge never occurs without a preceding intermediate phrase (ip) edge. Their tones are therefore listed below as combinations. The new GToBI presented here has simplified the boundary tone notation in order to make it phonetically more transparent. For example, in cases where the IP and ip boundary tones would represent the same pitch level, only one tone is transcribed: H-% in the new GToBI instead of the original H-L% (see below for explanation), and L-% instead of L-L%. The description of the canonical shapes given below assumes a distance of at least two syllables from the final pitch accent to the end of the phrase.[5]

- H-% 'plateau'

The main difference between H- and H-% is not tonal, but rather relates to perceived boundary strength, as encoded by the labels 3 and 4 in the parallel break index tier (see below). The similarity between the two contours is captured by the use of only one H tone.

There have been several different ways of describing such a plateau at phrase boundaries. Grabe (1998), for example, suggests the transcription 0% for a contour that more or less stays the same from the end of the last pitch accent to the boundary. The problem with this transcription is that the unmarked boundary tone does not directly encode whether the phrase ends low or high. Its value depends on which accent precedes it. The original GToBI transcription of a plateau was H-L% (with automatic

upstep on the L% tone). Since using an L tone to represent mid or high pitch was considered counter-intuitive and difficult to learn, the new GToBI transcription eliminates the L tone altogether. The combined label H-% has the advantage of directly encoding the phrase final pitch height without syntagmatic reference to preceding pitch accents. This makes the system easier to learn and more straightforward for database access.

- H-^H% 'plateau followed by sharp rise at the end of the phrase'

The upstepped H% component causes a sharp rise in the last syllable of the phrase, often to a point very high in the speaker's range.

- L-H% 'low followed by rise to mid at end of phrase'

This edge tone combination accounts for a final fall-rise contour if it is preceded by a H tone, and otherwise simply for a low stretch with a rise to mid on the last syllable.

- L-% 'low stretch which may be followed by drop to extra low'

The main difference between L- and L-% is perceived boundary strength. In addition, L-% is generally lower than L-. A final drop at the end of the phrase is possibly due to factors such as final lowering which is little understood in German and is not necessarily confined to the final syllable. We do not distinguish between several degrees of lowness at the IP boundary.[6]

- %H 'initial high boundary'

GToBI also provides the option of marking a high intonation phrase onset with a %H initial boundary tone. A mid or low tone is not explicitly marked.

The following boundary tone combinations are not in GToBI at present:

- L-^H%

In principle, this combination could be used to transcribe a low stretch with a rise to extra high on the final syllable. However, we do not have clear examples of this contour as distinct from H-^H%.

- H-L%

This combination was used in the original GToBI to describe a level contour (with automatic upstep of the L% after a H- phrase accent). Since upstep is now marked explicitly, H-L% could be used to describe a fall to low after a high plateau. Although this contour is not attested in Standard German, it has been reported in an East Phalian dialect (Kerckhove 1948:63) and in dialects of the Palatinate (Peters 2001a, 2001b).

The other three logical possibilities for boundary tone combinations are captured by simpler ones, already given above in the inventory: H-H% and H-^L% can be equated with H-%, and L-^L% would describe a contour very much resembling L-H%. Further nuances in the description of tonal movements at intonation phrase boundaries have not yet proved necessary in the sense that GToBI already captures all the IP phrase final contours reported to have distinct meanings or functions.

3.3.4 Break Indices

The break index tier is based on MAE_ToBI, where in the default case ‘3’ and ‘4’ coincide with intermediate phrase and intonation phrase boundaries respectively. GToBI does not explicitly mark the break index unless it deviates from this. In such cases, there are three options: the label ‘4-’ is used for cases where a phrase boundary is perceived, but where it is unclear as to the level of phrasing. GToBI distinguishes between two mismatches in tonal and rhythmic structure which are both encoded in MAE_ToBI with index 2: a rhythmic break with tonal continuity, ‘2r’, e.g. a rhetorical or hesitation pause; and a tonal break with rhythmic continuity, ‘2t’, e.g. a perceived boundary without a pause but with a tonal contour not attributable to the accents in the phrase. This often occurs in fast speech. GToBI does not include a simple break index 2 in its inventory. Break indices below level 2 are not dealt with.

3.3.5 Commonly occurring nuclear contours

Schematic representations and textual examples of commonly occurring nuclear contours are given in table 3.1, along with a suggested context in which such an utterance might be produced. In the schematic contours, extra heavy lines represent accented syllables, heavy lines postnuclear stressed syllables, and dotted lines the baseline of the speaker's pitch range. The line drawings provide a maximally long contour, assuming that the nucleus is followed by at least one postnuclear stress (the heavy line) and at least one other syllable after that. Most of the contour types have at least one example with a postnuclear stress. However, since the examples were chosen because they are representative in terms of their pragmatic interpretation and not because of their rhythmic structure, some do not correspond to the maximal contour. For instance, the rise (3a) has two example sentences: *Tauschen Sie auch **BRIEF**MARken?* ('Do you also exchange stamps?') and *Von wem ich das **HA**be?* ('From whom I have it?'). In the first, the extra heavy line corresponds to the nucleus, ***BRIEF***, the heavy line to the postnuclear stress *MAR*. In the second, the nuclear syllable ***HA*** is followed by only one syllable. Since there are not enough segments for the realisation of the rise-plateau-rise shape, the pitch simply rises directly from ***HA*** to the end of the phrase. In cases where the nuclear syllable is final in the phrase, especially if the coda contains voiceless obstruents, the contour may be truncated. This is particularly true in the case of falling contours (see Grabe 1998).

The contexts provided in the table contain pragmatic interpretations referring to specific examples; they should not be taken as abstract meanings for given contours. If syntactic information is given, then it is simply that the pattern may be regarded as neutral for a particular syntactic construction. It does not imply any more than this. We do not distinguish between linguistic and paralinguistic functions of intonation, since it has been shown that both types of function can be expressed by discrete means such as

the choice of pitch accent and boundary tones (Scherer, Ladd & Silverman 1984). We therefore include information as to speaker attitude or affect, where this helps to clarify the context in which an utterance might be spoken. We discuss each section of the table separately below.

Table 3.1

Fall - In the autosegmental literature there is only one type of fall. In GToBI there is a simple fall, represented as H* L-%, and a fall preceded by a sharp rise. The latter is represented with a leading L tone, thus L+H* L-%. Although this combination does not necessarily signal contrast, it may do so (especially with a wide pitch range), as in the example given in figure 3.2.

Figure 3.2

Rise-fall - The rise-fall, represented as L*+H L-%, differs in timing from L+H* L-%; in the former, the rise begins later in the accented syllable than in the latter. In the former the accented syllable sounds low whilst in the latter it is clearly high.

Rise - In the early autosegmental-metrical literature there are at most two different types of rise. In GToBI there are two starting on a low pitch, L* L-H% and L* H-^H%, where the endpoint of the second is higher than the first. There is also a rise where the accented syllable is mid, with or without a steeply rising onglide, i.e. (L+)H* H-^H%.

Level - Contours ending in a level or sustained pitch are barely mentioned in the literature. According to Féry, a L*+H rise can be followed by a level stretch which is not given any explicit transcription; it is assumed that the pitch of the trailing tone is continued until the end of the phrase. She therefore does not distinguish between rises and level nuclear patterns, claiming that 'As a matter of fact, rising tones and progredient intonation cannot be kept apart.' (1993:89). GToBI marks a level contour with or without a steeply rising onglide as L+H* H-% or H* H-% respectively.

Fall-rise - Generally GToBI represents fall-rises as (L+)H* L-H%. In principle, it is also possible to mark a 'high fall-rise' (where the pitch between the two peaks does not drop to low) as (L+)H* !H-^H%, although it is unclear whether this distinction is really necessary.

Early peak - GToBI has two early peak contours: H+!H* and H+L*. The former is the early peak contour referred to by Kohler with the meaning 'established fact', and also the one transcribed by Féry as H+H*+L. Von Essen claims that this pattern can signal finality even on unfinished parts of an utterance, and points out that it is often used in radio announcements. This matches Féry's claim that this pattern is frequently used by TV reporters. The schematic contour for H+!H* allows for a postnuclear stressed syllable which is not present in the example (selected from a spontaneously produced corpus). An example such as *Sie hätte ja **LÜ**gen KÖNnen* ('She even could have lied') would have the syllable *kön* on the second heavy line. The other early peak contour, H+L*, is what von Essen describes as signalling a fatalistic tone. It can also be used for soothing or polite requests, as in the example given in table 3.1. Early peak contours will be discussed in more detail in section 3.4.1(i).

Stylised step down - The stylised step down (or calling contour) is represented as (L+)H* !H-%. Here the phrase accent !H- occurs on a stressed syllable if there is one. The prominent syllable upon which the step down occurs is optionally marked with '!H(*)'. The multiple uses of this contour in German have been described at length by Gibbon (1976, 1998).

3.4 GToBI compared with other AM accounts

Table 3.2 provides correspondences between GToBI and the models of Wunderlich (1988), Uhmann (1991) and Féry (1993). The gaps in the table indicate that GToBI makes more distinctions than the other models. The increased expressivity of GToBI is due to a number of different factors. It has leading tones, thus enabling distinctions to be made between e.g. a plain fall and a fall with a preceding onglide. It allows for a phrase accent as well as an intonation phrase boundary tone, whereas the other models have only one edge tone. Each of these issues will be dealt with separately below.

Table 3.2

3.4.1 Leading tones

In GToBI it is possible to have either H or L as leading tones. This means that the pitch before an accented syllable may be transcribed as high, in which case the

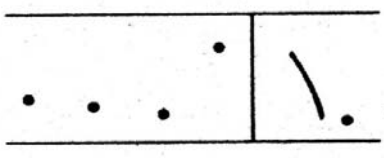
contour is referred to as an early peak, or low, in which case there is a rise up to the accented syllable, referred to as a rising onglide.

3.4.1 (i) Early peak contours

We have seen ample evidence for the existence of early peak contours, translatable into contours with a H leading tone.

Kohler's early peak, exemplified in (4a), provides us with such a case. It is contrasted with (4b), which he refers to as a medial peak.

(4) a

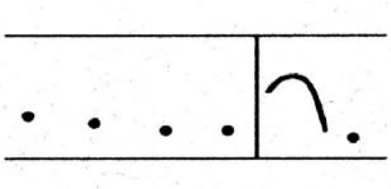


Sie hat ja ge- LO- gen (Kohler 1995:123)

She had actually LIED

'She actually LIED'

b



Sie hat ja ge- LO- gen (Kohler 1995:123)

She had actually LIED

'She actually LIED'

Moreover, Kohler performed perception tests which clearly indicated that high pitch on the preaccentual syllable (i.e. the one immediately before the accented syllable) is distinctive. That is, the early peak contour, which signals that information is old, is linguistically distinct from a medial peak signalling new information. GToBI captures this distinction by representing the early peak contour with a leading H tone, implying that the peak is *before* the accented syllable, and the medial peak contour with a H* (possibly with a L leading tone, see 3.4.1 (ii)), implying that the peak is *on* the accented syllable.

We have also seen that within the levels approach of Isačenko & Schädlich an early peak is represented as a 'preictic' fall, schematised in (5a) below, in contrast to a 'postictic' fall (5b).

(5) a preictic

die ↓ KINder

the children

b postictic

die KIN ↓ der

the children

(Isačenko & Schädlich 1966:60)

However, GToBI does not only have one early peak contour, but rather two: H+!H* and H+L*. Grabe (1998:89f.) argues that this distinction can be interpreted as one between total and partial downstep of her basic H*+L pitch accent (although she claims the distinction is gradual). As we have seen in 3.3.5, von Essen (1964) describes both types of early peak contour, attributing different usages to each of them.

Figures 3.3 and 3.4 illustrate spontaneous utterances of H+!H* L-% and H+L* L-% contours.

Figure 3.3

Figure 3.4

3.4.1 (ii) Rising onglides

Selting (1995) claims that a distinctive local pitch pattern begins on an accented syllable and is extended over the following unaccented syllables up to, but not including, the next accent. Thus the domain of the accent is related to the domain of the Abercrombian foot, or the 'Takt' (Kohler 1977, Pheby 1975), as it is referred to in German, and identical to Gussenhoven's (1990) 'Tonal Association Domain'. The domain of Selting's local pitch pattern is thus equivalent (at least for the final accent in a phrase) to the domain of the nuclear tone in the British School. This means that the pitch immediately prior to an accented syllable is excluded from the analysis of that accent. However, in the early configurations-based approaches there are descriptions of patterns in which the pitch of the immediately prenuclear syllable is important for the interpretation of the contour. In example (6) from Fox (1984:19), where the nuclear accent is on *KOMMST*, the British School is forced to analyse the nuclear tone as level. This would also be the case in Selting's approach.

(6)

Wenn du morgen KOMMST... (fahren wir zusammen)

If you tomorrow COME... (go we together)

'If you COME tomorrow... (we can go together)'

However, the movement from *gen* to *KOMMST* is clearly perceived as rising or as a jump up to the nuclear syllable. The specification of a level pitch on *KOMMST* is not enough to capture the essence of this contour. Fox states this clearly:

> "[...] an important characteristic of this pattern is the *jump up* to the high level pitch of the nucleus. The nucleus must always be at a higher pitch than the immediately preceding syllable. If the preceding head is high, its pitch must fall towards the end to allow for the jump up, hence the lower pitch given to *morgen* [...]"
>
> (1984:19f., italics as in original)

This jump up is also represented in the early levels-based approach of Isačenko & Schädlich as a preictic rise, as schematised in (7).

(7) die ↑ KINder

the children

(Isačenko & Schädlich 1966:60)

GToBI captures this tonal movement by the leading L tone in a L+H* pitch accent.

3.4.2. Levels of phrasing and phrase accents

Autosegmental-metrical accounts of German, as in much of the traditional literature, generally restrict the levels of phrasing to only one – the intonation phrase (e.g. Uhmann 1991, Grabe 1998). The AM exception is Féry (1993), who, like Pierrehumbert for English and GToBI for German, assumes intermediate phrases as well. Among the auditory approaches, von Essen's is the one which could be interpreted as allowing for an additional smaller level of phrasing. His unit of analysis is the

rhetorical phrase. He distinguishes between two types of rhetorical phrase: a major one with a nucleus (or 'Schwerpunkt') and a minor one without a nucleus. When an utterance contains more than one phrase, he claims it is the last one which contains a nucleus. This can be seen in example (8) below:

(8) Ich habe ge- TAN | was mir be- **FOH-** len war. (von Essen 1964:38)
I have done what me ordered was
'I DID what I was ORDERED to do.'

In a GToBI analysis, example (8) would be divided into two intermediate phrases (ip) forming one intonation phrase (IP), as in (9) below:

(9) [[Ich habe geTAN]ip [was mir beFOHlen war]ip]IP

However, although the analyses of von Essen and GToBI appear to be similar, there is one important difference: von Essen distinguishes the two types of phrase according to their pitch contours (progredient vs. terminal and interrogative). GToBI, by contrast, provides two different, hierarchically structured domains of phrasing, which are independent of specific pitch contours. The only restriction on contours at a given boundary is the number of tones available to capture them. At an intermediate phrase boundary the phrase accent is the only edge tone available. At an intonation phrase boundary there are always two edge tones: the phrase accent and the intonation phrase boundary tone. In the latter case there is one more tone available, which allows for a more complex tonal contour to be captured.

Féry dispenses with the intermediate phrase boundary tone, which she also refers to as the phrase accent. She claims that the phrase accent has two functions: to control the pitch movement between the nuclear accented syllable and the boundary tone of an intonation phrase, and to mark the boundary of an intermediate phrase (1993:79), and argues that the trailing tone of the phrase-final pitch accent takes over both of these functions. However, such an approach would make it difficult, if not impossible, to disambiguate between a sequence of prenuclear bitonal pitch accents (which Féry (1993:120) explicitly allows for) and a sequence of intermediate phrases, each with one bitonal pitch accent (which are also possible in her model).

The transcription of nuclear falls as H*+L, as in Féry and Uhmann, or as (L+)H* L- as in GToBI is still a controversial issue. Grice & Benzmüller (1998) found that the 'elbow' (the point at which the pitch reaches the baseline) after a medial peak accent differed according to the number of unstressed syllables between the nuclear syllable and the next postnuclear stressed syllable; the further away the stressed syllable, the later the baseline was reached. In fact, in 94% of fall-rises and in 91% of falls, the baseline was reached precisely on the postnuclear stressed syllable. This is taken as strong evidence for the analysis of those patterns as H* L-H% and H* L-% respectively, as opposed to H*+L H% and H*+L L%.

However, there are possibly contours which have not yet been investigated where the elbow is aligned differently, for example at a relatively constant distance from the H* peak, indicating that the L tone is part of a bitonal H*+L pitch accent (this could be the case in structures displaying narrow focus or contrast, as several examples in Uhmann (1991) suggest). We assume that there are dialectal differences in the position of the elbow. This is a common phenomenon in other languages, as is the case in Greek, Romanian and Hungarian where the question tune has the same tonal structure but has a different association of the phrase accent tone, depending on the dialect. As a

matter of fact, in an analysis of Bern Swiss German, Fitzpatrick-Cole (1999) has found evidence for the association of an edge tone (analysed as an intonation phrase boundary tone) to a lexically stressed syllable. In that dialect it appears that the tone associates with the final stressed syllable in the phrase rather than the immediately postnuclear one.

3.5 Prospect

GToBI is an annotation scheme, which should be regarded as a tool for research into the phonological structure of German intonation. Although its descriptive inventory is certainly richer than any other AM model of German intonation, the distinctions GToBI offers are all motivated by independent studies, either auditory or instrumental. This is hardly surprising since the aim of an annotation scheme is to be as flexible as possible in order to capture all relevant empirically observed patterns, even if it turns out later that part of the description is redundant, in that some of these patterns are derivative or make reference to gradient features rather than categorical ones.

Acknowledgements

We would like to thank Caroline Féry, Jürgen Trouvain and Jörg Peters for insightful comments on earlier drafts.

References

Adriaens, L.M.H. (1991), 'Ein Modell deutscher Intonation: eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenem Text', Ph.D. dissertation (Eindhoven

University of Technology).

Altmann, H., Batliner, A., and Oppenrieder, W. (eds.)(1989), *Zur Intonation von Modus und Fokus im Deutschen* (Tübingen: Niemeyer).

Arvaniti, A., Ladd, D.R. and Mennen, I. (1999), 'What is a starred tone?', in M.B. Broe, and J.B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon* (Cambridge: Cambridge University Press), 119-31.

Beckman, M.E. and Ayers-Elam, G. (1997), 'Guide to ToBI Labelling', Electronic text and accompanying audio example files available at http : //ling.ohio-state.edu/Phonetics/E_ToBI/etobi_homepage.html.

Beckman, M.E. and Hirschberg, J. (1994), 'The ToBI Annotation Conventions', Ms. and accompanying speech materials (Ohio State University).

Beckman, M.E., Hirschberg, J. and Shattuck-Hufnagel, S. (this volume), 'The original ToBI system and the evolution of the ToBI framework'.

Benzmüller, R. and Grice, M. (1997), 'Trainingsmaterialien zur Etikettierung deutscher Intonation mit GToBI', *Phonus* (Saarbrücken University), 3: 9-34.

Benzmüller, R. and Grice, M. (1998), 'The nuclear accentual fall in the intonation of Standard German', in *ZAS Papers in Linguistics: Papers on the conference "The word as a phonetic unit"*, (Berlin), 79-89.

Crystal, D. (1969), *Prosodic Systems and Intonation in English* (Cambridge: Cambridge University Press).

Duden (2000), *Aussprachewörterbuch* (Duden Volume 6), (Mannheim: Dudenverlag, Revised 4th edition).

Essen, O. von (1964)[2], *Grundzüge der hochdeutschen Satzintonation* (Ratingen: Henn, 1st edition 1956).

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

Fitzpatrick-Cole, J. (1999), 'The Alpine Intonation of Bern Swiss German', in *Proceedings of the ICPhS99* (San Francisco), 941-4.

Fox, A. (1984), *German intonation: An outline* (Oxford: Clarendon Press).

Gibbon, D. (1976), *Perspectives of intonation analysis* (Bern: Lang).

Gibbon, D. (1998), 'Intonation in German', in D. Hirst, and A. Di Cristo (eds.), *Intonation Systems. A survey of 20 languages* (Cambridge: Cambridge University Press), 78-95.

Giegerich, H.J. (1985), *Metrical Phonology and Phonological Structure. German and English* (Cambridge: Cambridge University Press).

Grabe, E. (1998), *Comparative Intonational Phonology: English and German* (MPI Series in Psycholinguistics 7) (Wageningen: Ponsen and Looijen).

Grice, M. and Benzmüller, R. (1995), 'Transcription of German Intonation using ToBI-Tones – The Saarbrücken System', *Phonus* (Saarbrücken University) 1: 33-51.

Grice, M. and Benzmüller, R. (1998), 'Tonal affiliation in German falls and fall-rises', Poster presented at the 5th Conference on Laboratory Phonology (York).

Grice, M., Ladd, D.R. and Arvaniti, A. (2000), 'On the place of phrase accents in intonational phonology', *Phonology*, 17.2: 143-85.

Grice, M., Reyelt, M., Benzmüller, R., Mayer, J. and Batliner, A. (1996), 'Consistency in Transcription and Labelling of German Intonation with GToBI', *Proceedings of the Fourth International Conference on Spoken Language Processing* (Philadelphia), 1716-9.

Gussenhoven, C. (1984), *On the Grammar and Semantics of Sentence Accents* (Dordrecht: Foris).

Gussenhoven, C. (1990), 'Tonal association domains and the prosodic hierarchy in English', in S. Ramsaran, (ed.), *Studies in the pronunciation of English* (London: Routledge).

Halliday, M.A.K. (1967), *Intonation and Grammar in British English* (The Hague: Mouton).

Isačenko, A.V. and Schädlich, H.J. (1966), 'Untersuchungen über die deutsche

Satzintonation‘, *Studia Grammatica*, VII: 7-64.

Kerckhove, M. van de (1948), ‘Intonationssystem einer Mundart‘, *Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft*, 52-65.

Kohler, K.J. (1977), *Einführung in die Phonetik des Deutschen* (Grundlagen der Germanistik 20), (Berlin: Schmidt, Revised 2nd edition 1995).

Kohler, K.J. (1991), ‘Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology, and semantics’, *AIPUK* (Kiel University) 25: 115-85.

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

Moulton, W.G. (1962), *The Sounds of English and German* (Chicago: University of Chicago Press).

Peters, J. (2001a), ‘Postnukleare Tonhöhengipfel in der Vorderpfalz und in Mannheim’, Ms. (University of Potsdam).

Peters, J. (2001b), ‘Frageintonation in der Pfalz. Eine Reanalyse des Güntherodt-Korpus’ Ms. (University of Potsdam).

Pheby, J. (1975), *Intonation und Grammatik im Deutschen* (Berlin: Akademie-Verlag).

Pheby, J. (1984)[2], ‘Phonologie: Intonation’ (ch. 6), in K.E. Heidolph et al (eds.), *Grundzüge einer deutschen Grammatik* (Berlin: Akademie-Verlag, 1st edition 1980), 839-97.

Pierrehumbert, J. (1980), 'The Phonology and Phonetics of English Intonation', PhD dissertation, (Bloomington: Indiana University Linguistics Club).

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

Reyelt, M., Grice, M., Benzmüller, R., Mayer, J. and Batliner, A. (1996), 'Prosodische Etikettierung des Deutschen mit ToBI', in D. Gibbon (ed.), *Natural Language and Speech Technology, Results of the third KONVENS conference* (Berlin, New York: Mouton de Gruyter), 144-55.

Scherer, K.R., Ladd, D.R. and Silverman, K.E.A. (1984), 'Vocal cues to speaker affect: Testing two models', *Journal of the Acoustical Society of America*, 76 (5): 1346-56.

Selting, M. (1995), *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation* (Tübingen: Niemeyer).

Shattuck-Hufnagel, S., Ostendorf, M. and Ross, K. (1994), 'Stress shift and early pitch accent placement in lexical items in American English', *Journal of Phonetics*, 22: 357-88.

Trager, G.L. and Smith, H.L. (1951), *An outline of English structure* (Norman, Oklahoma: Battenburg Press, Revised edition 1957, Washington, American Council of Learned Societies).

Truckenbrodt, H. (to appear), 'Upstep and Register Levels', *Phonology*.

Truckenbrodt, H. (1998), ‘On the Role of Prosody in Tonal Scaling’, Paper presented at Workshop on Focus, Intonation and Phrasing (Freudental, near Konstanz, Germany).

Uhmann, S. (1991), *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie* (Tübingen: Niemeyer).

Wunderlich, D. (1988), ‘Der Ton macht die Melodie - Zur Phonologie der Intonation des Deutschen’, in H. Altmann (ed.), *Intonationsforschungen* (Tübingen: Niemeyer), 1-40.

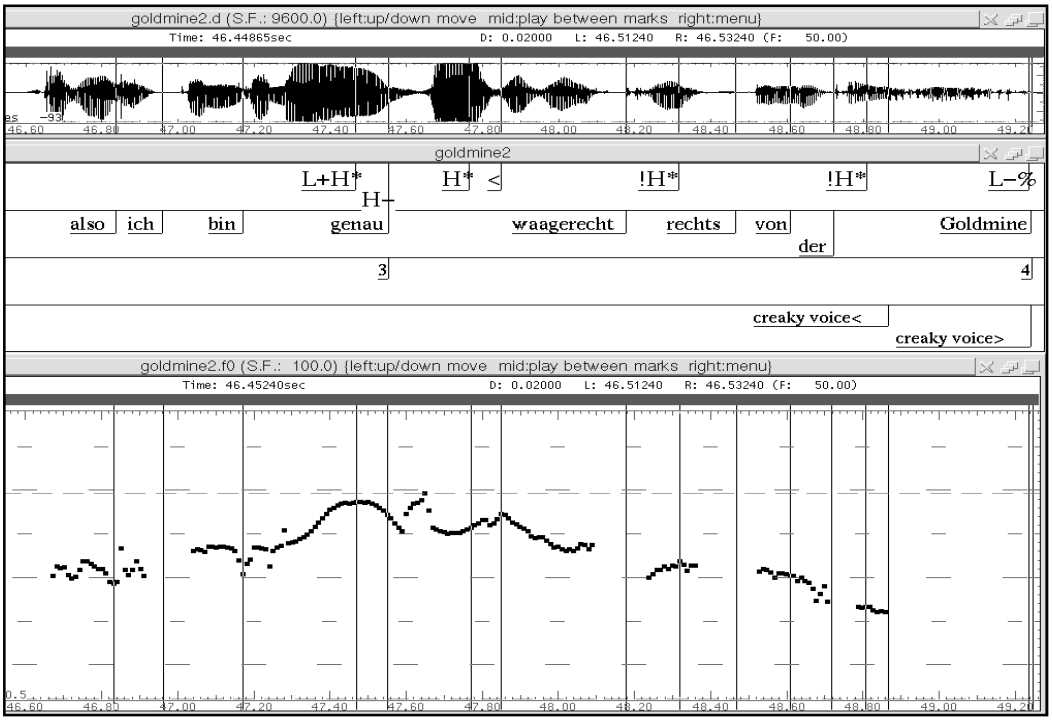


Figure 3.1 F0 contour and label tiers of the utterance 'Also ich bin genau waagerecht rechts von der Goldmine' (*Well, I'm exactly in a horizontal line to the right of the gold mine*) (adapted from Grice & Benzmüller 1995).

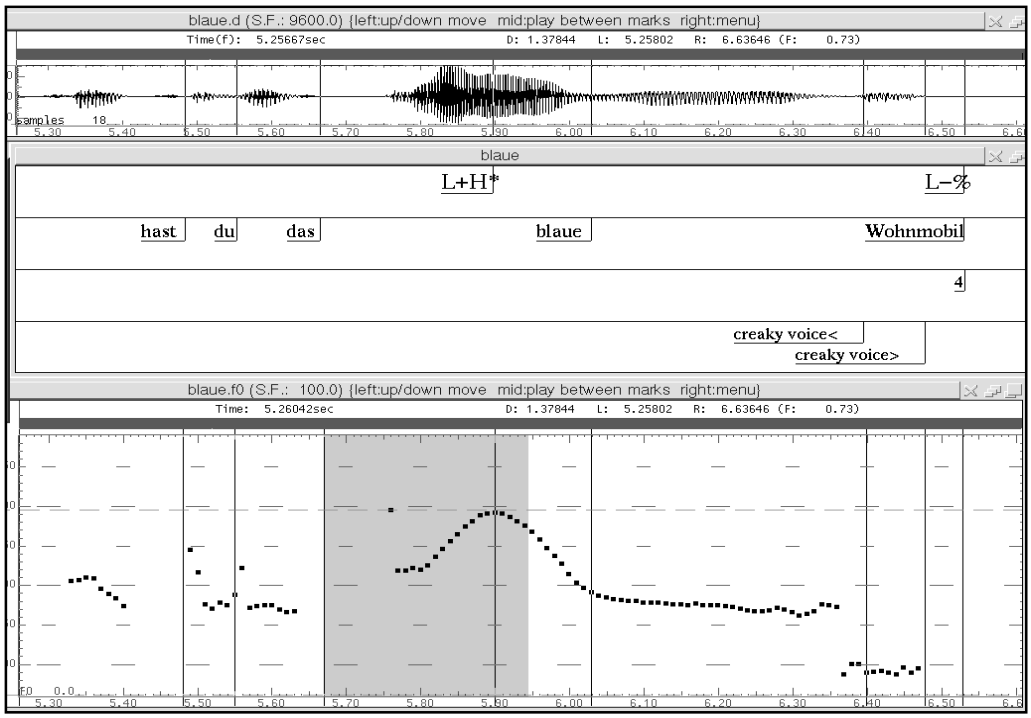


Figure 3.2 F0 contour of L+H* L-% on 'Hast du das **BLAU**e WOHNmobil?' (*Do you have the blue caravan?*) (adapted from Grice & Benzmüller 1995); for clarification purposes the shaded area marks the nuclear syllable *BLAU*.

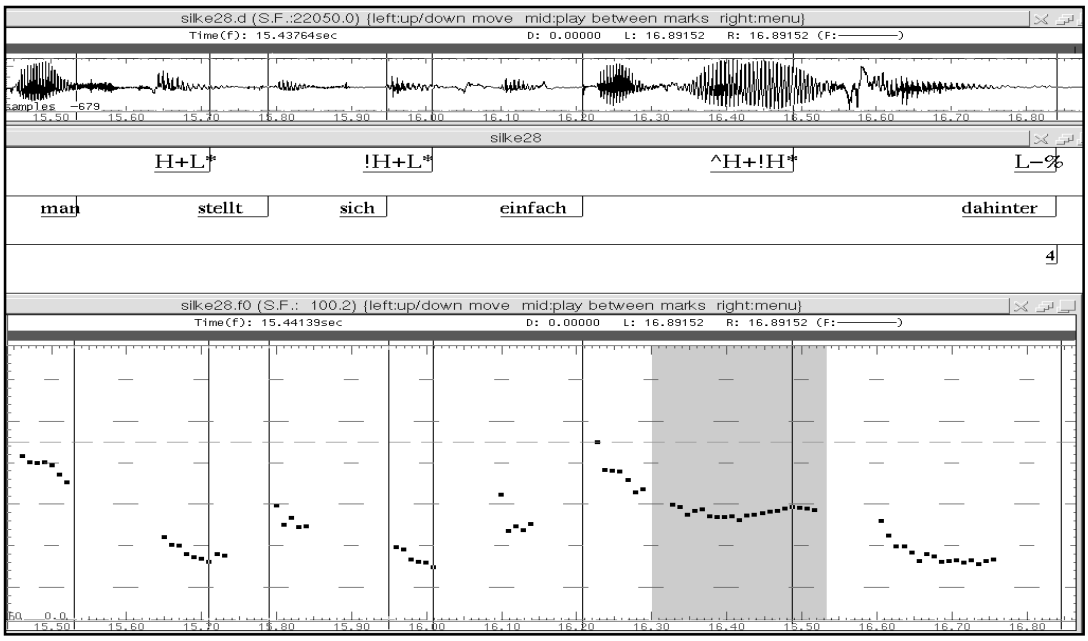


Figure 3.3 F0 contour of H+!H* L-%[7] on 'Man stellt sich einfach da**HIN**ter' (*You just queue up behind it*); the shaded area marks the nuclear syllable *HIN*.

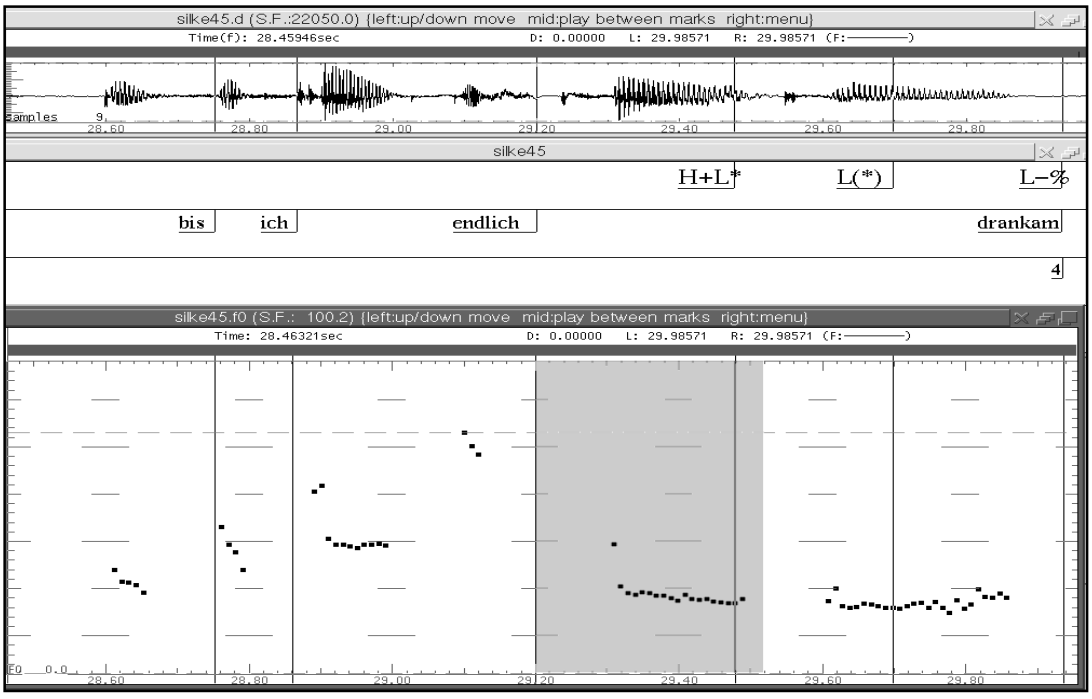
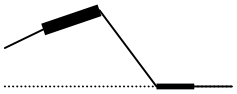
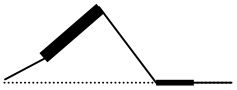
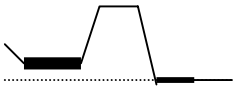
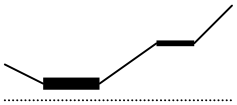
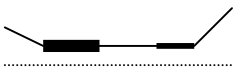
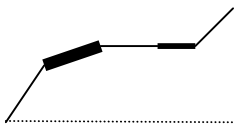
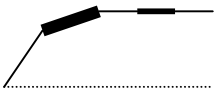
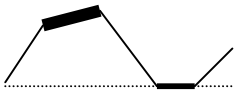
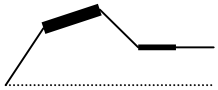


Figure 3.4 F0 contour of H+L* L-% on '... bis ich endlich **DRAN**KAM' (*...until it finally was my turn*); the shaded area marks the nuclear syllable *DRAN*.

Table 3.1 - Commonly occurring German nuclear contours and examples of their usage:

		GToBI	Schematic Contour	Context	Example
Fall	1a	H* L-%		Neutral statement Neutral W-question	Mein **ZAHN** tut WEH.[1] *My tooth hurts.* Wo hast du den **WA**gen gePARKT?[1] *Where did you park the car?*
	1b	L+H* L-%		Contrastive assertion	Schon der Ver**SUCH** ist STRAFbar![2] *Even to attempt is an offence!*
Rise-Fall (Late Peak)	2	L*+H L-%		Self-evident assertion Emotionally committed or sarcastic assertion	Das **WEISS** ich SCHON![6] *I already know that!* Der Blick ist ja **FA**belhaft![3] *The view is fantastic!*
Rise	3a	L* H-^H%		Neutral yes/no-question Echo question	Tauschen Sie auch **BRIEF**MARken?[1] *Do you also exchange stamps?* Von wem ich das **HA**be?[2] *From whom I have it?*
	3b	L* L-H%		Indignation Answering phone	**DOCH**! *It is!* **BEC**kenBAUer?[4]
	3c	(L+)H* H-^H%		Follow-up question	...oder ist Ihr **BRU**der HIER?[5] *...or is your brother in?*
Level	4	(L+)H* H-(%)		Incompleteness Ritual expression	**AN**dererSEITS...[6] *But then again...* Guten **MOR**gen! [3] *Good morning!*
Fall-Rise	5	(L+)H* L-H%		Polite offer	Mögen Sie **ROG**genBRÖTchen?[1] *Would you like rye rolls?*
Early Peak	6a	H+!H* L-%		Established fact	Hab' ich mir schon ge**DACHT**.[7] *That's what I thought.*
	6b	H+L* L-%		Soothing / Polite request	Nun er **ZÄH**le doch MAL![2] *Just tell me about it!*

Stylised Step Down	7	(L+)H* !H-%		Calling contour	**BEC**kenBAUer!

Examples are taken from [1]Féry (1993), [2]von Essen (1964), [3]Fox (1984) [4]Ladd (1996, adapted), [5]Moulton (1962), [6]Pheby (1984), and [7]Grice & Benzmüller (1995). Capitals in bold face indicate nuclear syllables, plain capitals postnuclear stresses.

Table 3.2 - German nuclear contours - three models compared with GToBI

		Wunderlich	Uhmann	Féry	GToBI
Fall	1 a	H* L	H*+L L%	H*+L	H* L-%
	1 b				L+H* L-%
Rise-Fall (Late Peak)	2		L*+H L%	L*+H+L	L*+H L-%
Rise	3 a	L* H H%	L*+H H%	L*+H	L*(+H) H-^H%
	3 b	L* H%			L* L-H%
	3 c				(L+)H* H-^H%
Level	4			L*+H	(L+)H* H-(%)
Fall-Rise	5		H*+L H%	H*+L H%	(L+)H* L-H%
Early Peak	6 a			H+H*+L	H+!H* L-%
	6 b	%H L* L			H+L* L-%
Stylised Step Down	7			H*+M	(L+)H* !H-%

Abstract

This chapter proposes a consensus system for the annotation of Standard German intonation within the framework of autosegmental-metrical phonology: GToBI. First, we provide a survey of existing studies of German intonation, including traditional auditory approaches as well as more recent phonological studies and instrumental analyses. We then go on to give a detailed exposition of GToBI, showing how the intonation contours considered to be distinctive in the surveyed works can be captured, and comparing GToBI to three earlier autosegmental-metrical approaches to German intonation. Finally, we discuss a number of theoretical issues, such as whether pitch accents need to be represented with leading tones or not, how many levels of phrasing are required, and the status and distribution of phrase accents.

Index Words

pitch accent
phrase accent
leading tone
trailing tone
GToBI
boundary tone
break index
early peak
late peak
rising onglide
intonation phrase
intermediate phrase
nucleus (= nuclear pitch accent)
bridge accent
topic accent
focus accent

upstep

downstep

calling contour

preictic

postictic

Key Words

intonation

German intonation

autosegmental-metrical phonology

pitch accent

phrase accent

leading tone

early peak

GToBI

ToBI

prosody

Footnotes

[1] From the two models based on Gussenhoven's system, we investigate in detail here the earlier of the two (i.e. Féry 1993).

[2] For clarity of exposition, a '+' sign is transcribed between tones belonging to a pitch accent. This notational convention taken from Pierrehumbert (1980) is not used by Féry.

[3] In their study of American English, Shattuck-Hufnagel et al. (1994) found that the perceived shift of prominence from the lexical stress to an earlier syllable in rhythmic clash contexts is due to substantial F0 movement rather than any durational increase on the prominent syllable. It is thus more appropriate to call this type of shift a 'pitch accent shift' rather than the traditional term 'stress shift'.

[4] Exceptions to this rule are the so-called 'intonational tags', which can be regarded as enclitic tone units without pitch accents.

[5] If the final accented syllable is closer to the boundary, then much of the shape is lost owing to lack of time for the realisation of the individual tones.

[6] This means that we do not distinguish between L-% and L-L% to transcribe the presence or absence of a final drop. For the moment, we leave open whether this distinction is functionally motivated for a transcription system of Standard German (cf. Peters (2001a) where the two transcriptions are used to express a difference in the tonal alignment of postnuclear falls in a regional variety of German).