GToBI - a phonological system for the transcription of German intonation

Stefan Baumann [1], Martine Grice [1] and Ralf Benzmüller [2]

Universität des Saarlandes, Saarbrücken [1]; G Data Software, Bochum [2]
baumann@coli.uni-sb.de, mgrice@coli.uni-sb.de, ralf.benzmueller@gdata.de

ABSTRACT

In this paper we present an updated version of GToBI, an annotation scheme for capturing aspects of the intonation of Standard German within the autosegmental-metrical framework. We concentrate on the tonal part of the system, providing an inventory of commonly occurring contours, and discuss issues which are still controversial within the framework, such as leading tones and phrase accents.

1. Introduction to GToBI

GToBI (German Tones and Break Indices) is a set of conventions for labelling aspects of the phonological structure of German intonation with the aim of being easy to learn, reliable, and adaptable for different labelling purposes. It is close to the English ToBI system (E-ToBI), which has its roots in autosegmental-metrical phonology. GToBI is able to capture distinctions drawn in the traditional auditory-based literature on German intonation (e.g. Moulton 1962, von Essen 1964, Kohler 1977, Pheby 1984, Fox 1984) as well as in later autosegmental-metrical studies (e.g. Uhmann 1991, Féry 1993, Grabe 1998), and has been applied to spontaneous and read corpora (Grice et al. 1996, Reyelt et al. 1996). It has recently been modified in order to make the system phonetically more transparent and to incorporate recent advances in intonational phonology.

GToBI consists minimally of three label tiers: tones, break indices, and words. For reasons of space only the first of these will be dealt with here. On the tonal tier the perceived pitch contour is transcribed in terms of pitch accents and boundary tones, with diacritics for pitch range modifiers such as downstep (´!´) and upstep (´^´) placed immediately before the affected tone. The tonal inventory comprises two monotonal (H*, L*) and four bitonal pitch accents (L+H*, L*+H, H+L*, H+!H*), and edge tones which are peripheral to minor (intermediate) phrases (L- or H-) and major (intonation) phrases (L% or H%). The intermediate phrase edge tone, or ´phrase accent´ (Grice, Ladd & Arvaniti (in press); Grice & Benzmüller 1998), may occur on postnuclear lexical stresses as well as at the phrase edge. In the new GToBI there is an option for capturing the location of the phrase accent explicitly (see section 3.2.).

The automatic upstep rule, which raises the pitch range after an H- phrase accent in the English ToBI and earlier GToBI, has now been dispensed with, and instead upstep is marked with a ´^´ diacritic. Furthermore, redundant boundary tone symbols have been deleted: If the contour following the phrase accent stays the same or slightly falls (due to final lowering) up to the end of the intonation phrase, an extra tone marking the level of the IP boundary is unnecessary. The new inventory is as follows:

L-% (low fall, formerly L-L%), L-H% (rise to mid), H-% (high level contour)[1], and H-^H% (high plateau with a final rise, formerly used without the upstep diacritic). The upstep diacritic is also used to indicate a step up within a sequence of pitch accents, for example on a high nuclear pitch accent which follows a prenuclear downstepping sequence (see Grice, Baumann & Benzmüller (to appear) for more details).

2. Commonly occurring nuclear contours

Schematic representations and textual examples of commonly occurring nuclear contours are given in table 2.1., along with a suggested context in which utterances might be produced. 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.

3. GToBI compared with other accounts

GToBI differs from other autosegmental-metrical (AM) accounts of German intonation both in the structure of its pitch accents and in the number of edge tones.

3.1. Pitch accent structure

GToBI allows for leading tones; 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.

3.1.1. Early peak contours

One type of early peak contour has been described by Kohler, exemplified in (1), corroborated by perception tests which clearly indicate that high pitch on the preaccentual syllable is distinct from high pitch on the accented syllable (medial peak).

(1)

Sie hat ja ge- | LO- gen (Kohler 1995:123)
She actually LIED (lit. *she had actually LIED*)

[1] 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.

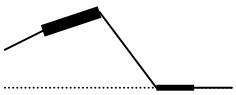
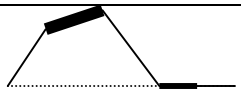
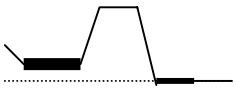
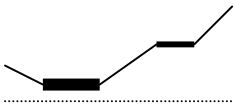
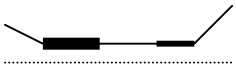
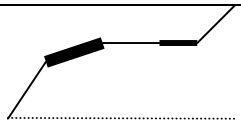
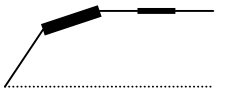
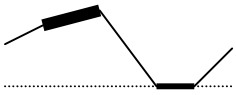
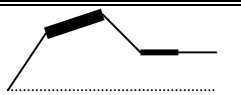
	GToBI	Schematic Contour	Context	Example
Fall	H* L-%		Neutral statement	Mein **ZAHN** tut WEH.[1] *My tooth hurts.*
			Neutral W-question	Wo hast du den **WA**gen gePARKT?[1] *Where did you park the car?*
	L+H* L-%		Contrastive assertion	Schon der Ver**SUCH** ist STRAFbar![2] *Even to attempt is an offence!*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

Table 2.1. Commonly occurring German nuclear contours and examples of use
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. In the schematic contours, extra heavy lines represent accented syllables, heavy lines postnuclear stressed syllables (if available), and dotted lines the speaker´s baseline.

The former signals that information is old, the latter signals new information. GToBI captures this distinction with a H leading tone which aligns with the prenuclear syllable

ge, as opposed to a starred tone, H*, on *LO*, which would be used in the representation of a medial peak.

This distinction had already been made by Isačenko & Schädlich, who represent the early peak as a ´preictic´ fall, as in (2a) below, in contrast to a ´postictic´ fall (2b).

(2a) preictic

die ↓ KINder
the children

(2b) 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*. The former is the early peak contour referred to by Kohler with the meaning ´established fact´. It is transcribed as H+H*+L by Féry (1993) who claims that it is often used by TV reporters. Von Essen (1964) also observes that this pattern is used by radio announcers and attributes to it a meaning of finality. The other early peak contour is labelled in GToBI as H+L*. It is the same contour that has been claimed to give a fatalistic tone (von Essen 1964). However, it can also be used for soothing or polite requests, as in the example given in table 2.1. The distinction between the two patterns has been interpreted by Grabe (1998) as a distinction between total and partial downstep of her basic H*+L pitch accent.

3.1.2. Rising onglides

A distinctive rising onglide can be observed in example (3) from Fox (1984:19), where the nuclear accent is on *KOMMST*.

(3)

Wenn du morgen KOMMST... (fahren wir zusammen)
If you COME tomorrow... (we can go together)
(lit. *if you tomorrow COME...*)

The movement from *gen* to *KOMMST* is clearly perceived as rising or as a jump up to the nuclear syllable. The specification of the nuclear pitch contour cannot be adequately described with a model which excludes the pitch immediately prior to an accented syllable from the analysis of that accent (as in British School accounts). 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 [i.e. the prenuclear stretch: B, G & B] 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 (4). GToBI captures this tonal movement by the leading L tone in a L+H* pitch accent.

(4)

die ↑ KINder
the children

(Isačenko & Schädlich 1966:60)

3.2. 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). The AM exception is Féry (1993), who, like Pierrehumbert (1980) for English and GToBI for German, assumes intermediate phrases as well, but unlike those models does not allow for a tone to be associated to its edge. Among the auditory approaches, von Essen's is the one which could be interpreted as allowing for an additional smaller level of phrasing. 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 (5), which in a GToBI analysis would be divided into two intermediate phrases (each containing a nucleus) forming one intonation phrase.

(5) Ich habe geTAN | was mir be**FOH**len war. (von Essen 1964:38)
I DID what I was ORDERED to do.
(lit. *I have done what me ordered was*)

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, which has at most three tones, consisting of a bitonal pitch accent and one edge tone (phrase accent), the pitch contour cannot be as complex as it can be at an intonation phrase boundary which may have four (bitonal pitch accent plus phrase accent plus intonation phrase boundary tone).

The transcription of nuclear falls as H*+L, as in Féry and Uhmann, or as (L+)H* L- as in GToBI is still controversial. GToBI offers the possibility of shedding light on the issue by allowing for the labelling of the point at which the pitch reaches the baseline[2], referred to as the ´elbow´ (Ladd 1996). Grice & Benzmüller (1998) found that the elbow 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 stress. This is taken as 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. Furthermore, we assume that there are dialectal differences in the alignment of the phrase accent. 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

[2] The label used is a separate L without diacritics. GToBI accordingly allows for the optional marking of a high phrase accent by a simple H label which is placed at the beginning of a high plateau.

association of the phrase accent tone, depending on the dialect. This can be to the nuclear syllable, one or more postnuclear stressed syllables, or a phrase-final or penultimate syllable. A case in point is Bern Swiss German (Fitzpatrick-Cole 1999) which also has an edge tone associated to a lexical stress. In our view this is also a phrase accent – but rather than being associated with the first postnuclear stressed syllable it is associated with the last in the phrase. Clearly more data from a range of dialects is needed to test these hypotheses.

REFERENCES

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

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

[3] Fitzpatrick-Cole, J. (1999). The Alpine Intonation of Bern Swiss German. *Proc. ICPhS99*. San Francisco, 941-944.

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

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

[6] Grice, M., S. Baumann & R. Benzmüller (to appear). German ToBI. To appear in Jun, S.A. (ed.) Intonation: Models and ToBI Labeling.

[7] Grice, M. & R. Benzmüller (1995). Transcription of German Intonation using ToBI-Tones – The Saarbrücken System. Saarbrücken: *Phonus* 1. 33-51.

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

[9] Grice, M., D.R. Ladd & A. Arvaniti (in press). On the place of phrase accents in intonational phonology. Provisionally accepted for *Phonology* 17.2.

[10] Grice, M., M. Reyelt, R. Benzmüller, J. Mayer & A. Batliner (1996). Consistency in Transcription and Labelling of German Intonation with GToBI. *Proc. IV ICSLP.* Philadelphia. 1716-1719.

[11] Isačenko, A.V. & H.J. Schädlich (1966). Untersuchungen über die deutsche Satzintonation. *Studia Grammatica* VII, 7-64.

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

[13] Ladd, D.R. (1996). *Intonational Phonology.* Cambridge: CUP.

[14] Moulton, W.G. (1962). *The Sounds of English and German*. Chicago: CUP.

[15] Pheby, J. (1984)[2]. Phonologie: Intonation. (ch. 6). In: Heidolph, K.E. et al (eds.), 1984. *Grundzüge einer deutschen Grammatik*. Berlin: Akademie-Verlag. 839-897.

[16] Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation.* PhD thesis, MIT. Bloomington: Indiana University Linguistics Club.

[17] Reyelt, M., M. Grice, R. Benzmüller, J. Mayer & A. Batliner (1996). Prosodische Etikettierung des Deutschen mit ToBI. In D. Gibbon (ed.) *Natural Language and Speech Technology, Results of the third KONVENS conference.* Berlin: Mouton de Gruyter. 144-155.

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