Forschungsberichte des
Instituts für Phonetik und
Sprachliche Kommunikation
der Universität München
(FIPKM) 35 (1997) 127-137

THE PARTITUR FORMAT AT BAS

F. Schiel, S. Burger, A. Geumann, K. Weilhammer

Institut für Phonetik und Sprachliche Kommunikations,
Ludwig-Maximilians-Universität München

ABSTRACT

Most spoken language resources are produced and disseminated together with symbolic information relating to the speech signal. These are for instance orthographic transcripts, labeling and segmentation on the phonologic, phonetic, prosodic, phrasal level. Most of the known formats for these symbolic data are defined in a 'closed form' that is not flexible enough to allow simple and platform-independent processing and easy extensions.

At the *Bavarian Archive for Speech Signals* (BAS) a new format has been developed and used over the last few years that shows some significant advantages over other existing formats. This paper describes the basic principles behind this format, discusses briefly the advantages and gives detailed definitions of the description levels used so far.

GENERAL OVERVIEW

Most file formats containing segmental information on speech signals have the disadvantage that

- they are not easy to extend (without rewriting software that uses the existing format).
- they are not easy to process with UNIX standard tools.
- they mix different description levels (which leads to technical and conceptual problems)
- they were defined as adhoc solutions for very specialized problems and are not capable of being re-used in a different setup.

Therefore a new open format based on the SAM Label Format was developed, which circumvents most of the mentioned problems. In this format all levels of description may be annotated independently but are time aligned like the individual tiers of a score. Hence this format was called 'BAS Partitur Format' ('Partitur' = German for 'score').

In the future all BAS corpora will be distributed with the new BAS Partitur Format if they contain segmental information of any kind. The formerly used formats will be retained but not further updated.

A first draft of the BAS Partitur Format was published in [Att95].

The BAS Partitur Format has the following features:

- SAM compatible header structure
- Easy to extend and to process by simple UNIX commands
- Open format; extensions to the format can be implemented without alterations to the software that reads the older format
- Time-aligned independent description of a virtually unlimited number of different levels of the speech signal (see examples later in this paper).
- Symbolic links between the independent levels allow logical assignments aside from the physical time scale. These links are based on the word units of the utterance.

DEFINITION (Version 1.2)

A Partitur file name has the same prefix as the corresponding signal file (8 Bytes for Iso 9660 compatibility) but the extension *.par*. All contents are in 7-bit ASCII exclusively (to guarantee portability to all platforms). Each line starts with a three-byte label followed by a colon; this label defines synopsis and semantics of the ensuing line. The following units of the line are seperated by 'white spaces' (blank, tab).

The Partitur file is structured into a header and a body (like SAM description files). The header stretches from the beginning of the file to the label *LBD:*; the body from the label *LBD:* to the end of file where the last line has to be closed by a 'new line' or a 'CR + LF' (the final SAM label ELF:was omitted for the BAS Partitur Format since it prevents effective processing of the Partitur files).

The header contains SAM-compatible lines of general information. The following entries are compulsory:

```
LHD: Partitur file version
REP: Place of recording
SNB: Number of Bytes per Sample
SAM: Sampling Frequency in Hz
SBF: Byteorder (Intel 01, Motorola 10)
```

```
SSB: Bit Resolution
NCH: Number of Channels
SPN: Speaker ID
LBD:
```

Example:

```
LHD: Partitur 1.2
REP: Muenchen
SNB: 2
SAM: 16000
SBF: 01
SSB: 16
NCH: 1
SPN: PS1
LBD:
```

The following entries are optional; apart from these, other entries are tolerated as long as they do not conflict with compulsory and optional entries:

```
FIL: SAM File Type
TYP: Type of SAM Label File
DBN: Corpus Name
VOL: Number of Volume
DIR: Directory in Volume
SRC: Name of speech file
BEG: Begin of labeling sequence
END: End of labeling sequence
RED: Date of Recording
RET: Duration
RCC: Recording Conditions
CMT: Comment
SPI: Speaker Information
PCF: Name of Protocol File
PCN: Protocol Number
EXP: Name of Segmenter
SYS: Labelingsystem
DAT: Date of Labeling
SPA: SAM-PA Version
```

The body starts after the label *LBD:* and stretches to the end of file. It contains the different tiers of the BAS Partitur Format. Each tier is identified by a unique label. The order of tiers as well as the order of lines within a tier is not significant.

In the following sections the five basic classes of tiers are defined.

Tiers with symbolic relation *(class 1)*

A line of this tier contains:

- the label
- a comma-separated list of integers (*symbolic links*)
- a string with the label information

The symbolic links refer to a reference tier which numbers the word units beginning with zero. The label string has an internal synopsis which is defined in the tier definition.

Example:

```
TRL: 6,7 mit´m
```

In this example the word events 6 and 7 of an utterance are transliterated.

Tiers with time relation, time-consuming *(class 2)*

A line of this tier contains:

- the label
- two integers denoting the begin and duration of the event.
- a string containing the label information

The semantics of the integers is defined by the tier definition (possible are samples, millisecs, etc.)

Example:

```
PHN: 13456 3450 aU
```

In this example a phonemic segment labeled */aU/* stretches from sample *13456* for the next *3450* samples.

Tiers with time relation, not time-consuming *(class 3)*

A line of this tier contains:

- the label
- an integer denoting the time position of the event
- a string containing the label information

Example:

```
PRO: 13456 TON: P*; FUN: PA
```

In this exapmle the prosodic event labeled *TON: P*; FUN: PA* (GTobi, see [Gri96]) takes place at sample *13456* of the utterance.

Tiers with time and symbolic relation, time-consuming *(class 4)*

A line of this tier contains:

- the label
- two integers denoting the start and duration of the event.
- a comma separated list of integers (*symbolic links*)
- a string containing the label information

Example:

```
SAP: 13456 3450 9 aU
```

In addition to the example above this tier not only gives the starting point and the duration of the phonemic segment but also a pointer to the word unit where it belongs (word 9).

Tiers with time and symbolic relation, not time-consuming *(class 5)*

A line of this tier contains:

- the label
- an integer denoting the time position of the event.
- a comma separated list of integers (*symbolic links*)
- a string containing the label information

Example:

```
PRB: 13456 9 TON: P*; FUN: PA
```

Again, in this example the prosodic event is not only placed in time but also assigned to a word of the utterance (word 9).

Remarks

- If not otherwise noted, durational parameters are given in samples counting from the beginning of the digitized utterance
- An item may be refered to more than one word in the utterance (suprasegmental events, assimilation at word boundaries, phrases, etc.)
- If the symbolic link in a tier is not (or not yet) known, the symbolic link is set to *-1*.
- The same symbolic relation may occur in different lines of a tier (for example if more than one event can be assigned to the same word of an utterance).

DEFINITION OF TIERS

The following sections give an overview of the currently defined tiers in the BAS Partitur Format (version 1.2.2). Please keep in mind that this is an *open list* in the sense that new tiers can be defined whenever there is a need for it. If somebody would like to work with speech resources from BAS and to define a new tier for his or her specific problem, please contact the BAS to get a new label assigned. By doing this we can keep up a consistent documentation of the format and avoid conflicts between matching labels. The version of the BAS Partitur Format is incremented by one on the third digit whenever a new tier definition is added to it. In accordance to the basic principle this does not imply that any software has to be changed.

Canonical Pronunciation

- *Label:* KAN
- *Class:* 1
- *Synopsis:* (symbolic links) (transcript)

This tier is the **reference tier** for all other tiers that use symbolic links. It contains a list of the spoken words within the utterance annotated in extended German SAM-PA (see [Sam89] for a general definition of the SAM-PA and [Sam96] for a special description of the extended German SAM-PA as used in several German projects). Note that these forms are the phonologically expected citation forms, **not** the actually spoken form.

The segmentation of the whole utterance is done into word units, where everything counts as a word that is produced by the articulatory organs of the speaker and can be seen as *speech*. Following this definition hesitations are words, whereas laughing, coughs, etc. are not. This separation isn't always clear, but on the other hand the selection of word units is arbitrary as well. The main point here is a unique reference tier for symbolic relations in other tiers. Another problem is the reduction of words that are annotated in the orthographic form, e.g. "mit'm". In these cases the reduction is restituted (in this example /mIt de:m/). The reason for this lies in the fact that some of these reductions should later be automatically accessible.

Example:

```
KAN: 0  j´a:
KAN: 1  Qalzo:
KAN: 2  QE:m
KAN: 3  h´OYt@
KAN: 4  Qo:d6
KAN: 5  m´O6g@n
```

Orthography

- *Label:* ORT
- *Class:* 1
- *Synopsis:* (symbolic links) (lexical orthography)

The tier *orthography* contains the orthographic (lexical) strings corresponding to the units in the tier *canonical form*. Words are not capitalized at the beginning of an utterance or sentence within an utterance (except nouns of course). German 'Umlauts' and other letters not included within 7 Bit ASCII are written in LaTeX notation. This tier is used for easy lexical access; therefore no additional markers except lexical words are allowed. There is no punctuation in this tier. Lexical words include items that are contained in the KAN tier (e.g. hesitations, repairs, word fragments, etc.). This tier can be used to access customized pronunciation dictionaries, to create unique word frequency lists, etc.

Example:

```
ORT: 0  ja
ORT: 1  also
ORT: 2  <"ahm>
ORT: 3  heute
ORT: 4  oder
ORT: 5  morgen
```

Verbmobil Transliteration - VM I

- *Label:* TRL
- *Class:* 1
- *Synopsis:* (symbolic links) (transliteration)

The tier *transliteration* contains the orthographic transcript of the utterance according to the VM I conventions 3.0. The transliteration is segmented into the units of the tier *canonical pronunciation*. Therefore multiple references may occur (eg. if a reduced form of two words is written as one unit in the transliteration). Although especially defined for the German Verbmobil I project, this format has been used in many other resources of spontaneous speech as well. See [Koh94] (German only) or online [Trl96] for a detailed description of the VM I Transliteration format.

Example:

```
TRL: 0  <A>
TRL: 0  ja ,
TRL: 1  also
TRL: 2  <"ahm>
TRL: 3  <:<#Klicken> heute:>
TRL: 4  oder
TRL: 5  morgen .
```

Verbmobil Transliteration - VM II

- *Label:* TR2
- *Class:* 1
- *Synopsis:* (symbolic links) (transliteration)

The tier *transliteration* contains the orthographic transcript of the utterance according to the VM II conventions. A detailed definition of this format can be found in [Tr297] (German only). In contrast to the VM I format this new updated definition has the advantage of being fully parsable. Furthermore, with this format multi-party and multi-lingual dialogs may be transliterated. To denote overlapping speech parts between different speakers in a dialog, a new tier *SUP* was defined (see below).

Superimposed Speech - VM II

- *Label:* SUP
- *Class:* 1
- *Synopsis:* (symbolic links) (turn marker) (transliteration)

This is a very specialized tier to denote overlapping speech in multi-party recordings. The synopsis of the turn marker and the transliteration is defined for the VM II transliteration format (see above). The speech annotated in this tier stems from a different speaker who actively superimposes his speech on the speech of this partitur file. See [Tr297] (German only) for a detailed description of superimposed speech in the VM II format.

Example:

```
TR2: 0 ich
TR2: 1 w"urde
TR2: 2 vorschlagen ,
TR2: 3 da"s
TR2: 4 wir9@
TR2: 5 dann9@
TR2: 6 <:<#> hinfliegen:> ,
TR2: 7 <:<#> ich:>
TR2: 8 hab´
TR2: 9 jetzt <!1 jetz´>
TR2: 10 aber
TR2: 11 <:<#Rascheln> grade:>
TR2: 12 <:<#Rascheln> keine:>
TR2: 13 Unterlagen
TR2: 14 da . <#>
SUP: 4,5 g002acn2_028_AAK @9ja
```

In this example the utterance of another speaker (AAK, utterance "ja") is superimposed on the 4th and 5th word of the Partitur file (utterance "wir dann").

Phonetic Segmentation - PhonDat

- *Label:* PHO
- *Class:* 4
- *Synopsis:* (integer) (integer) (list of symbolic links) (label string)

This tier contains a totally time-consuming segmentation into broad phonetic units (extended German SAM-PA). The first number denotes the beginning of the segment in samples counted from the beginning of the speech file; the second number the duration of the segment in samples. The label string contains an additional relation to the canonical pronunciation (aside from the symbolic links to the tier *caninocal form*). The '-' sign denotes differences to the expected canonical pronunciation on a segmental level: a leading '-' sign means the following segment was inserted (e.g. /-a:/); a trailing '-' sign means the segment was deleted (e.g. /a:-/); a '-' sign between segment labels means that the canonical expected segment was replaced (e.g. /a:-E:/). This tier also contains prosodic and phrasal labeling and segmentation. The full conventions of labeling and segmentation for German are briefly described in [Pom92] or online in [Pho95].

Phonetic Segmentation - Verbmobil

- *Label:* SAP
- *Class:* 4
- *Synopsis:* (integer) (integer) (list of symbolic links) (label string)

In contrast to the *PHO* tier this segmentation is not stringently time-consuming. That is, there might be pauses in the signal that are not labeled (which happens frequently in spontaneous speech). Furthermore the conventions are different in some points to the *PHO* tier to simplify parsing and processing of the tier. *SAP* is an exclusively phonemic tier; there is no other information encoded here.

The conventions of labeling and segmentation are briefly described in appendix A.

Automatic Broad Phonetic Segmentation by MAUS

- *Label:* MAU
- *Class:* 4
- *Synopsis:* (integer) (integer) (symbolic links) (label string)

This tier contains an automatically generated broad phonetic segmentation in units of German SAM-PA. The segmentation is done fully automatically by the MAUS system ([Kip96]). The segmentation is totally time-consuming and the labeling has no direct relation to the tier *canonical form* as done in the tier *SAP*. (However, there are symbolic links to the words). The units are labeled in extended German SAM-PA as in the definition of the *SAP* tier (see appendix A). Additional labels are <nib> (non-speech event) and <p:> (pause). These labels always get the symbolic link -1 (no link).

Example:

```
MAU: 0 676 -1 <p:>
MAU: 677 7861 -1 <nib>
MAU: 8539 450 0 g
MAU: 8990 2436 0 u:
MAU: 11427 1740 0 t
MAU: 13168 958 1 d
MAU: 14127 1298 1 a
MAU: 15426 3820 1 n
MAU: 19247 303 2 n
MAU: 19551 1785 2 e:
MAU: 21337 624 2 m
MAU: 21962 636 2 n
MAU: 22599 501 3 v
```

Word Segmentation

- *Label:* WOR
- *Class:* 4
- *Synopsis:* (integer) (integer) (symbolic links) (word label)

This tier contains a segmentation of the utterance in word or word equivalents. The segmentation need not be without gaps. The label string may contain orthographic or pronunciation information (eg. in SAM-PA). A '-' at the end of the label string denotes a missing word in the reference of the tier *canonical from* (of course a missing word has zero duration); a leading '-' denotes an inserted word; a '-' between two words (*word1-word2*) denotes a replacement. The symbolic links give the relation to the tier *canonical form*. Note that inserted words have a symbolic link to the previous word in the reference tier.

Example:

```
WOR: 1245 13245 0 <"ahm>
WOR: 14490 10787 1 guten
WOR: 25277 5089 1 -<hm>          # insertion
WOR: 30366 8786 2 Tag
WOR: 39152 3089 3 ich
```

Dialog Act Segmentation

- *Label:* DAS
- *Class:* 1
- *Synopsis:* (symbolic links) (marker string)

This tier contains a segmentation in dialog acts according to the ongoing work of the 'Deutsches Forschungszentrums für künstliche Intelligenz' (DFKI), Saarbrücken, Germany. Each marker covers a portion of the speech signal that is denoted by the symbolic links to the reference tier *canonical form*. A description of the format can be found in [Jek95]) or online in [Das96].

Example:

```
DAS: 0,1,2,3,4,5          @m(REJECT_DATE BA)
                          @m(GIVE_REASON BA)
DAS: 6,7,8,9              @(SUGGEST_SUPPORT_DATE BA)
DAS: 10,11,12,13,14,15    @(REQUEST_SUGGEST_DATE BA)
```

Prosodic Segmentation - GTobi

- *Label:* PRB
- *Class:* 5
- *Synopsis:* (integer) (symbolic links) (marker string)

This tier contains the prosodic segmentation (by hand) according to GTobi defined by the Technical University of Braunschweig, Germany. A detailed description of the GTobi labeling format can be found in [Gri96] or online in [Prb96] (German only).

Example:

```
PRB:    54212    5   TON: H*; FUN: NA
PRB:    63269    7   TON: L+H*; FUN: EK
PRB:    76371    8   BRE: B3; TON: L-L%
PRB:    79967    8   TON: L*+H; FUN: PA
```

Literatur

[Att95] Atmanspacher, S. et al. (1995): Partiturformat für die Darstellung unterschiedlicher Repräsentationsebenen von gesprochener Sprache. Verbmobil Memo 90, September 1995.

[Gri96] Grice, Martine and Ralf Benzmueller (1995): Transcription of German Intonation using ToBI tones; The Saarbruecken System. Paper presented at Tutorial Workshop on Discourse and Dialogue Prosody, Stuttgart, February 1995, modified version also in Phonus 1, University of the Saarland, pp33-51.

[Jek95] Susanne Jekat, Alexandra Klein, Elisabeth Maier, Ilona Maleck, Marion Mast, Joachim Quantz (1995): Dialogue Acts in Verbmobil (Verbmobil-Report 65). Universitt Hamburg, DFKI GmbH, Universitt Erlangen, TU Berlin. April 1995.

[Kip96] A. Kipp, M.-B. Wesenick, F. Schiel (1996): Automatic Detection and Segmentation of Pronunciation Variants in German Speech Corpora; in: Proceedings of the ICSLP 1996. Philadelphia, pp. 106-109, Oct 1996.

[Koh94] Kohler,Lex,Pätzold,Scheffers,Simpson,Thon (1994): Handbuch zur Datenaufnahme und Transliteration in TP14 von VERBMOBIL - 3.0. IPDS, University of Kiel.

[Pom92] Pompino-Marschall, B. (1992): PhonDat - Verbundvorhaben zum Aufbau einer Sprachsignaldatenbank für gesprochenes Deutsch. FIPKM 30/1992, pp. 99-128.

[Das96] http://www.phonetik.uni-muenchen.de/Bas/BasDialogaktDok/vm-report-for-partitur_1.html

[Pho95] http://www.phonetik.uni-muenchen.de/Bas/BasFormatsPHOdeu.html

[Prb96] http://www.phonetik.uni-muenchen.de/Bas/BasProsodie.html

[Sam89] http://www.phon.ucl.ac.uk/home/sampa/home.htm

[Sam96] http://www.phonetik.uni-muenchen.de/Bas/BasSAMPA

[Trl96] http://www.phonetik.uni-muenchen.de/VMTraLexeng.html

[Tr297] http://www.phonetik.uni-muenchen.de/VMtrlex2d.html

APPENDIX A: CONVENTIONS OF THE SAP TIER FOR GERMAN

A.1 Symbol inventory

Consonants:

IPA-Number	SAMPA	Example(orth.)	Example(SAMPA)
102	b	bei	baI
104	d	du	du
128	f	verfahren	f6fa:r@n
110	g	Gast	gast
146	h	Hast	hast
153	j	ja	ja
109	k	Kahn	ka:n
155	l	Licht	lICt
114	m	Mann	man
116	n	neun	nOYn
101	p	Platz	plats
122	r	Rauch	raUx
132	s	las, Maß	la:s, ma:s
103	t	Torte	tO6t@
129	v	Vase, wann	va:se
133	z	lesen	le:z@n
134	S	Tasche	taS@
135	Z	Loge	lo:Z@
138	C	dich	dIC
140	x	Dach	dax
119	N	Junge	jUN@
113	Q	ich	QIC

Vowels:

IPA-Number	SAMPA	Example (orth.)	Example(SAMPA)
304 503	a:	Kahn	ka:n
304	a	kann	kan
302 503	e:	Beet	be:t
302	e	Meteor	meteo:6
303 503	E:	Käse	kE:z@
303	E	Bett	bEt
301 503	i:	riet	ri:t
301	i	Politik	politi:k
319	I	ritt	rIt
307 503	o:	bog	bo:k
307	o	Politik	politi:k

306	O	Bock	bOk
308 503	u:	Mus	mu:s
308	u	Kulisse	kulIs@
321	U	muš	mu:s
309 503	y:	Hüte	hy:t@
309	y	kyrillisch	kyrIlIS
320	Y	Hütte	hYt@
310 503	2:	Höhle	h2:l@
310	2	Ökonom	2kono:m
311	9	Hölle	h9l@
322	@	lesen	le:z@n
324	6	Leser	le:s6

Diphthongs:

IPA-Number	SAMPA	Example (orth.)	Example(SAMPA)
304 319	aI	zwei	tsvaI
304 321	aU	Bauch	baUx
306 320	OY	neun	nOYn

Diacritics, not used in the canonical form:

IPA-Number	SAMPA	Example
406	q	(Laryngealization, creaky) aq
424	∼	(Nasalization) a∼
404	H	(Aspiration) tH

Other Symbols not used in the canonical form:

IPA-Number	SAMPA	Meaning, Syntax
960	-	Modifications of the canonical form Elision @- Replacement O-@ Insertion -t
953	%	Uncertain boundary of a segment p Uncertain beginning p%< Uncertain end p%> Uncertain beginning and end p%<%>

The symbol inventory is in large parts the same as used in the Phondat 2 corpus (see [Pom92]). Also, the mode of segmentation is much the same as for the Phondat corpus.

A.2 Rules of Transcription

- TR1
  A transcription symbol is chosen mainly by the auditory judgement of the segment in its context. The context should have the size of at least one syllable.
- TR2
  A deviation from the canonical form is only quoted if another category is perceived, e.g. /I/ instead of /i:/ as in /fy:zIk/ oder /fy:zi:k/. More subtle variants induced by coarticulatory effects are neglected.
- TR3
  Only the symbols given above may be used.
- TR4
  If two adjacent homorganic segments merge (e.g. in 'hat den') the final segment is deleted (see TR5). If the auditory impression of a geminate remains, both symbols are preserved; the uncertain boundary can be marked by % , e.g. /kOm%>m%<al/.
- TR5
  In case of ambiguities concerning the elision of adjacent segments or assignment of symbols in reduced forms, the following rules have to be considered.

1. A postvocalic segment in an unaccented syllable can be elided.
2. Next, non-initial segments of an accentable syllable can be elided.
3. Finally, a word- or syllable-initial segment can be elided.

- TR6
  If the deviation from the canonic form affects entire words, and an erroneous canonic form is obvious, the canonic form is corrected.

A.3 Rules of segmentation

- SR1
  Segment boundaries are positioned at positive zero crossings in the oscillogram (changing polarity from negative to positive).
- SR2
  The placing should be checked in the spectrogram.
- SR3
  In transient areas between two succeeding elements, when the sound quality of both appears simultaneously, the border is set in the middle of the transient area, if no other criterion fits. Examples are /sf/, /sS/, possibly /l/ and /r/ in intervocalic position. The boundary then is marked as %, e.g. /s%>f%</.
- SR4
  Voiced (periodic) segments start with the first clearly noticable oscillation. The initial boundary matches the first significant zero crossing. Otherwise, segments judged as voiced can be devoiced initially. In this case, the devoiced portion is assigned to the voiced element (e.g. vowel) and the auditory judgment is crucial.
- SR5
  In fricatives with low intensity (especially /h/) the segment starts when the signal to noise ratio becomes positive. The position of the end boundary (with utterance final fricatives) follows the same rule. Utterance final exhalation noise should be separated from friction or aspiration.

A.4 The treatment of single sound classes

1. Stops
   (a) The phases closure and burst are both included in the stop segment. If the aspiration of /p,t,k/ or /b,d,g/ is clearly noticable and has a duration of more than 15 ms, H is inserted.
   (b) The beginning of utterance-initial stops or stops following a pause is, in most cases, uncertain, /p%</ .
   (c) In a non-homorganic stop cluster in which the first stop is unreleased (but audible!), the first half of the closure is labelled as the first stop, the remainder and the burst are labelled as the second stop. The boundary should be labelled as uncertain, e.g. /t%>p%</. The end of utterance-final unreleased stops is to be marked as uncertain.
   (d) The closure phase of a voiced stop starts after the last identifiable vowel oscillation.
   (e) Utterance final stops end after the burst. If necessary aspiration has to be inserted, but should be distinguished from exhalation noise.
   (f) The beginning of a voiced plosive with a preceding homorganic nasal is often difficult to determine. In these cases the low level amplitude phase of the nasal is considered as part of the stop. The burst is often only marked by a slight irregularity of the following oscillation.
   (g) Stops with an incomplete closure are treated as stops only if perceived as a stop. Otherwise it should be replaced by the corresponding fricative, or as appropriate.
   (h) The voiced - voiceless distinction keeps to the canonical proposal as far as possible. The canonical proposal is only rejected if a categorial change is audible and appears in the acoustic signal. E.g. /p,t,k/ is realised with a voice bar, /b,d,g/ is voiceless (and aspirated) in syllable initial or medial position.
   (i) The glottal stop is treated like other stops; therefore it consists of a closure and a burst. Usually the first oscillation is counted as burst phase. Utterance initial glottal stops begin with an arbitrary boundary. If no closure phase exists, the glottal stop is elided. In this case the adjacent vowel is often laryngealised.
2. AFFRICATES
   (a) Affricates (e.g. ts, pf, tS) are treated as two segments.
   (b) The boundary between the two segments is placed after the burst. Only if a distinct aspiration is intelligible, is an H inserted, and the fricative follows the aspiration.

3. FRICATIVES
   (a) The glottal fricative may be elided following other fricatives.
   (b) Succeeding homorganic fricatives should be kept apart. If no separation is possible, the rules TR4, and TR5 are applied, as for stops.
4. NASALS
   (a) Syllabic nasals are separated from an adjacent nasal, as far as audible. Internal structure or longer duration may serve as a clue.
   (b) A nasal might be realized by nasalization of an adjacent vowel. In this case the nasal is deleted and the vowel is nasalized, e.g. /a~/.
   (c) Devoiced nasals are not denoted any further.
5. RHOTICS
   The symbol /r/ summarizes the following articulations:
   - uvular trill
   - alveolar trill/flap
   - uvular fricative voiced/voiceless
   - velar fricative or glide

   Postvocalic 'r' is canonically represented as /6/, as in /hambu6k/ 'Hamburg'. If postvocalic /6/ is realized as a trill or fricative, it has to be replaced by /r/. If only the preceding vowel is lengthened, the (short) vowel is replaced by a long vowel and the /6/ is deleted, e.g. /dE:/ instead of /de:r/, or /va:/ instead of /va:6/. The diphthong quality of vowel and following /6/ is marked by an arbitrary boundary, e.g. /va:%>6%</
6. VOWELS
   (a) Long tense vowels are marked for length. Short tense vowels can be used, if necessary, but use should be limited.
   (b) Differing vowel quality should be noted as far as a categorial (perceptual) switch occurs.
   (c) If a vowel has diphthong quality, it may be replaced by /aI/, /OY/, or /aU/.
   (d) Devoicing is not marked.
   (e) Strong laryngealyzation/creaky voice is marked even if it does not replace a glottal stop.
   (f) Nasalization can be marked by a diacritic if it replaces an adjacent nasal, or in words like 'Restaurant' /rEsto:ra~/
7. DIPHTHONGS
   (a) If diphthongs have monophthong quality the deviation from the canonic form has to be notated.
   (b) Additive diphthong qualities are avoided. Only in rare cases can new diphthtong qualities be built up by combining two monophthongs with an intermediate arbitrary boundary, e.g. /o%<U%>/.

A.5 Reduced forms

The following forms are derived from the rules in TR4 and TR5.

'haben wir' realized as /h a m 6/

h	→	h
a:	→	a:-a
b	→	b-m
@	→	@-
n	→	n-
v	→	v-
i:	→	i:-
6	→	6

'Abend' realized as /Q a: m t/

Q	→	Q
a:	→	a:
b	→	b-m
@	→	@-
n	→	n-

t	→	t

'hast du einen' realized as /h a s n/

h	→	h
a	→	a
s	→	s
t	→	t-
d	→	d-
u:	→	u:-
Q	→	Q-
aI	→	aI-
n	→	n
@	→	@-
n	→	n-

'einen neuen' realized as /aIq n OY n/

Q	→	Q-
aI	→	aI-aIq
n	→	n-
@	→	@-
n	→	n-
n	→	n
OY	→	OY
@	→	@-
n	→	n