1 Introduction

This CD-ROM contains a revised issue of the previously published Vol. I of the Kiel Corpus of Read Speech (PHONDAT90 and PHONDAT92: Kiel-CD #1, 1994), providing the audio and the time-aligned segmental label files, as well as label files in which prosodic labels have been added to the segmental ones. There are several reasons for this revision: First, at the time CD #1 was published there was no established audio file format. We could therefore assume that users of the CD would have access to tools to convert raw data to their preferred format. Nowadays the situation is almost reversed which is why we choose the RIFF-WAV format on this CD. Second, it turned out that over 50 recordings contained loud clicks due to numerical overflow in the post-processing. Most of these could be corrected automatically but 4 recordings had to be copied from DAT and their label files realigned. Third, apart from adding files with prosodic labels, years of working with the segmental label files had led to numerous corrections, including some changes in the labelling conventions (see Section 4.2).

The speech material consists of isolated sentences and two texts. For the recording, the speaker was seated in a sound-treated booth and read the sentences through a window from a monitor that was placed in the adjacent room, where the operator controlled the orthographic presentation of the sentences one by one. The texts were presented in two or three parts, respectively, on boards positioned inside the booth. A Neumann U87 condenser microphone (cardioid settings) was placed approximately 30 cm from the speaker’s mouth. The microphone signals were amplified by a John Hardy M1 pre-amplifier and recorded on a SONY PCM 2500 DAT-recorder at a sampling rate of 44.1 kHz for PHONDAT90 and of 48 kHz for PHONDAT92, respectively, with 16 bit quantization. They were then digitally transferred to computer hard disk and downsampled to 16 kHz as well as high-pass filtered at 40 Hz.

Details concerning the make-up of the material, the recording and processing of the speech signals, as well as the old segmentation and labelling conventions applied to the data on Kiel-CD #1, can be found in Kohler (1992). The labelling conventions were subsequently further developed to be applicable also to spontaneous speech, cf. Kohler, Pätzold, Simpson (1995). Prosodic labelling (PROLAB) is based on the Kiel Intonation Model (KIM) and is described in Peters and Kohler (2004). Updates of both segmental and prosodic labelling are summarized in section 4.2.2 below. On the basis of canonical transcriptions, rule-generated from the orthographic text by Klaus J. Kohler, the segmental and prosodic labelling was carried out in xassp by phonetically trained students under the supervision of Klaus J. Kohler, Wim van Dommelen, Adrian Simpson, and Benno Peters. Cf. IPDS (1997); Kohler, Pätzold, and Simpson (1997).

Michel Scheffers serviced the speech processing software, integrated the data in the data bank and helped with updating Vol. I of the Kiel Corpus of Read Speech on this CD-ROM.
2 Directory structure

The CD-ROM has the following directory structure:

<CD-ROM>
+---ph90
  |   +---berlin
  |     |   +---k01
  |     |     |   +---k01be001.s[12]h
  |     |     |   +---k01be001.wav
  |     |     .
  |     |     .
  |     |     |   +---k01be100.s[12]h
  |     |     |   +---k01be100.wav
  |     |     +---k02
  |     |     +---k03
  |     |     +---k04
  |     |     +---k05
  |     |     +---k06
  |     |     +---k61
  |     |     +---k62
  |     |     +---k63
  |     |     +---k64
  |     |     +---k65
  |     |     +---k66
  |     +---marburg
  |         |   +---k07
  |         |     |   +---k01mr001.s[12]h
  |         |     |   +---k01mr001.wav
  |         |     .
  |         |     .
  |         |     |   +---k01mr100.s[12]h
  |         |     |   +---k01mr100.wav
  |         |     +---k08
  |         |     +---k09
  |         |     +---k10
  |         |     +---k11
  |         |     +---k12
  |         |     +---k61
  |         |     +---k62
  |         |     +---k67
  |         |     +---k68
  |         |     +---k69
  |         |     +---k70
  |     +---butter
  |         |   +---k22
  |         |     |   +---k22butt1.s[12]h
  |         |     |   +---k22butt1.wav
  |         |     |   +---k22butt2.s[12]h
  |         |     |   +---k22butt2.wav
  |         |     |   +---k22butt3.s[12]h
---k22butt3.wav
---k23
---k24
---k25
---k26
---k27
---k28
---k29
---k30
---k61
---k62
---k76
---k77
---k78
---k79
---k80
---nordwind
---k13
---k13nord1.s[12]h
---k13nord2.s[12]h
---k13nord1.wav
---k13nord2.wav
---k14
---k15
---k16
---k17
---k18
---k19
---k20
---k21
---k61
---k62
---k71
---k72
---k73
---k74
---k75
---restkorp
---k61
---k61cn001.s[12]h
---k61cn001.wav
---k61cn020.s[12]h
---k61cn020.wav
---k61ko001.s[12]h
---k61ko001.wav
---k61ko063.s[12]h
---k61ko063.wav
directories containing data banks for each speaker in PHONDAT92

**Key**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ph90</td>
<td>PHONDAT90</td>
</tr>
<tr>
<td>ph92</td>
<td>PHONDAT92</td>
</tr>
<tr>
<td>berlin</td>
<td>100 Berlin sentences</td>
</tr>
<tr>
<td>marburg</td>
<td>100 Marburg sentences</td>
</tr>
<tr>
<td>butter</td>
<td>The Butter Story in 3 parts</td>
</tr>
<tr>
<td>nordwind</td>
<td>The North Wind and the Sun in 2 parts</td>
</tr>
<tr>
<td>restkorp</td>
<td>20 CNET sentences</td>
</tr>
<tr>
<td></td>
<td>63 Kohler sentences</td>
</tr>
<tr>
<td></td>
<td>45 SEL sentences</td>
</tr>
<tr>
<td></td>
<td>25 additional SEL sentences</td>
</tr>
<tr>
<td></td>
<td>15 Schiefer/Sommer sentences</td>
</tr>
<tr>
<td></td>
<td>30 Tillmann/Kohler sentences</td>
</tr>
<tr>
<td>erlangen</td>
<td>100 Erlangen sentences</td>
</tr>
<tr>
<td>siemens</td>
<td>100 Siemens sentences</td>
</tr>
</tbody>
</table>

Each of the terminal nodes in the tree is a directory containing the signal and segmentation files of one speaker who was recorded speaking that particular part of the corpus. Two speakers (k61=kko and k62=rtd) produced the whole corpus.

### 3 File naming conventions

k05be005.s1h

```
   ^^^ speaker (PHONDAT90)
```

dlms001.s1h

```
   ^^^ speaker (PHONDAT92)
```

k05be005.s1h

```
   ^ sentence/text type:
     be Berlin sentences
     mr Marburg sentences
     butt The Butter Story
     nord The North Wind and the Sun
     cn CNET sentences,
     ko Kohler sentences
     s1 SEL sentences
     s2 additional SEL sentences
     sr Schiefer/Sommer sentences
     tk Tillmann/Kohler sentences
```

dlms001.s1h

```
   ^ e Erlangen sentences
       s Siemens sentences
```
4 File formats

4.1 Signal files

Signal files (*.wav): 16 bit, 16 kHz, Microsoft RIFF-WAV format

4.2 Label files

7-bit ASCII files, DOS format (CR/LF at line end). Each label file *.s1h or *.s2h contains three sections: orthographic, canonical and phonetically labelled. *restcorp* has only been phonetically labelled for speakers k61 and k62. Other speakers read varying sections of *restcorp*, for which only *.wav and *.s0 with orthographic text and canonical transcription are provided in the sub-directory *addcorp*.

4.2.1 Conventions in orthographic section

The orthographic text follows DUDEN before the spelling reform. Word-initial small letters are kept sentence-initially. TeX conventions are used to represent umlaut and "ß": "a, "A = ä, Å; "o, "O = ö, Ö; "u, "U = ü, Ü; "s = ß. Because of the TeX notation of umlaut no double inverted commas are allowed for quotations.

4.2.2 Conventions in label section

The segmental transcription is modified SAMPA, the prosodic annotation follows PROLAB: cf. Kohler et al. (1995) for segmental labelling; Peters and Kohler (2004) for PROLAB labelling. Both segmental and prosodic labelling conventions for the *Kiel Corpus of Read Speech* Vol. I are updated and summarized in this section. From orthographic representations of the sentences and texts, canonical, citation form word transcriptions were automatically generated in modified SAMPA, including punctuation marks. Subsequently, segmental and prosodic labelling was carried out in *xassp* (IPDS (1997); Kohler, Pätzold, and Simpson (1997)). This procedure produced a label file *.s[12]h for each sentence or text section from the canonical transcription by allocating one canonical symbol after another to a point on the time scale of the acoustic signal file, also modifying the series of segmental labels, as well as adding time-aligned prosodic labels. A segmental label referring to a signal portion is bounded by the time mark of the next label, i.e.
segmental labelling is strictly linear. Labels without duration are put on the same time point as the next label that is associated with a time extension.

4.2.2.1 General principles for segmental and prosodic labels

(1) All segmental and prosodic labels (including syntactic punctuation) are placed on one tier.
   • Segmental labels in lexical items are prefixed by
     ## at word boundaries, e.g. ##f in fort, fahren
     $# at word boundaries inside compounds, e.g. $#f in fortfahren
     $ in all other cases, e.g. $t in fort;
     the prefixes are in turn followed by % if the segment boundary is uncertain.
   • # prefix of non-verbal labels and syntactic punctuation marks, e.g. #p: (pause), #h: (breathing), #c: (beginning of sentence), #. (period)
   • & indicates prefixes of prosodic, and only prosodic labels, and is put after
     # in the default case, e.g. #&2^ (accent level 2 with medial peak)
     $ inside lexical items, e.g. in the case of stress shift, see PEAK Examples in 4.2.2.3 (4).

(2) Label prefixes ##, # and $ are obligatory in the label lines of the *.s1h files. In the running transcription section of these files, ## and # are replaced by two blanks, $ by one (see 4.2.3).

(3) Prosodic labels (as well as punctuation marks) are linked to a point in time, but have no duration. They receive the same time stamp as the subsequent label. So if there are several such labels in a row they are all put on the same time as the next label that is associated with a time extension. There are ordering conventions for prosodic labels, see section 4.2.2.3 B.

(5) Uncertainty as to the prosodic category to be labelled is marked by %, placed after the prosodic index & , e.g. &%PGn (uncertain prosodic phrase boundary). In the case of two prosodic features in one label, each may be marked by %, e.g. &%2^ or &2^% or &%2^%.

4.2.2.2 Segmental labels: modified SAMPA

The canonical transcription is based on a phonemic system of German which recognizes 46 units
24 vowels 20 oral, and 4 nasal vowels, which may occur in loan words from French
22 consonants 3 (x C Q) have non-phonemic status, using them avoids the need to include morphological information to account adequately for the differences in the intervowinal dorsal fricative in pairs such as Frauchen [fɾaʊəxn] and rauchen [ɾaʊəxn] or the presence vs. absence of glottalization in pairs such as vereisen [fɾeɐaiəzn] and verreisen [fɾeɐaɪəzn].

Vowels

<table>
<thead>
<tr>
<th>SAMPA</th>
<th>IPA</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>a:</td>
<td>a:</td>
<td>a k'am k'am kam Kamm</td>
</tr>
<tr>
<td>e:</td>
<td>e:</td>
<td>b'e:t b'Et Beet Bett</td>
</tr>
<tr>
<td>i:</td>
<td>i:</td>
<td>r'i:t r'It riet ritt</td>
</tr>
<tr>
<td>o:</td>
<td>o:</td>
<td>b'o:k b'Ok bog Bock</td>
</tr>
<tr>
<td>u:</td>
<td>u:</td>
<td>b'u:s@ b'Us@ Buße Busse</td>
</tr>
<tr>
<td>y:</td>
<td>y:</td>
<td>h'y:t@ h'Yt@ Hütte Hütte</td>
</tr>
</tbody>
</table>
E:  e:  k'E:z@  Käse
2:  9  œ:  h'2:l@  h'9l@  Höhle  Hölle

~a~  ă  rEsto:r'a~  Restaurant
E~  ě  t'E~  Teint
O~  ő se:z'O~  Saison
9~  ō  pa6f'9~  Parfum

aI  ai  tsv'aI  zwei
aU  aU  b'aUx  Bauch
OY  œv  n'Oyn  neun

@  œ  l'e:z@n  lesen
6  ë  l'e:z6  Leser
h'e:6  h'E6  her  Herr

Consonants

<table>
<thead>
<tr>
<th>Sampa</th>
<th>IPA</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b</td>
<td>b'ain</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
<td>d'ain</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>f'ain</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
<td>g'lOt</td>
</tr>
<tr>
<td>h</td>
<td>h</td>
<td>h'ain</td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>j'a:</td>
</tr>
<tr>
<td>k</td>
<td>k</td>
<td>k'ain</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>l'lOt</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>m'ain</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n'ain</td>
</tr>
<tr>
<td>p</td>
<td>p</td>
<td>p'ain</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
<td>r'ain</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>l'o:s</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>l'o:t</td>
</tr>
<tr>
<td>v</td>
<td>v</td>
<td>v'ain</td>
</tr>
<tr>
<td>z</td>
<td>z</td>
<td>z'ain</td>
</tr>
<tr>
<td>C</td>
<td>ć</td>
<td>z'lOt</td>
</tr>
<tr>
<td>X</td>
<td>x</td>
<td>z'Uxt</td>
</tr>
<tr>
<td>S</td>
<td>ŋ</td>
<td>S'ain</td>
</tr>
<tr>
<td>Z</td>
<td>ź</td>
<td>l'o:Z@</td>
</tr>
<tr>
<td>N</td>
<td>η</td>
<td>z'INt</td>
</tr>
<tr>
<td>Q</td>
<td>?</td>
<td>Q'aIns</td>
</tr>
</tbody>
</table>
Label Modifications

In segmental labelling, one of four decisions has to be taken:

1. the canonical symbol is aligned without modification
2. the canonical symbol has no corresponding signal portion, it is marked as deleted, and is put on the same mark as the next non-deleted label: e.g. $@-
3. the canonical symbol does not represent the signal portion adequately, it is marked as replaced by a phonoetically more appropriate symbol from the canonical list: e.g. $n-m
4. there is no canonical symbol on offer for the signal portion to be labelled, a label is inserted: e.g. $-t.

Biphonematic canonical symbol sequences of a monophthongal vowel and the vowelled realization 6 of post-vocalic /r/ are amalgamated to a single label, which is aligned with the beginning of the entire vocalic portion of the signal, e.g.

<table>
<thead>
<tr>
<th>Canonical</th>
<th>Label</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>E 6</td>
<td>$E6</td>
<td>Herr</td>
</tr>
<tr>
<td>e: 6</td>
<td>$e:6</td>
<td>Heer</td>
</tr>
<tr>
<td>2: 6</td>
<td>$2:6</td>
<td>hört.</td>
</tr>
</tbody>
</table>

This is done to avoid problems in attempting to segment such vocalic portions into two parts and to align the two symbols with the beginning of each. It follows from this procedure that in the canonical section of *.s[12]h files the two symbols are separated by a blank, but they are on one label line (see 4.2.3). However, if the sequence is disyllabic, the label is split up and the second part is marked by = on a separate label line, e.g. ##f $r $'y: $=6 früher

Replacements do not introduce symbols in addition to those contained in the canonical lists of vowels and consonants; they index changes to other categories within the vowel or the consonant system. In the case of vowels, only replacements by lax vowels, monophthongization or schwa are marked.

Insertions

In addition to the insertion of symbols from the canonical list, such as $-t in ##Q $a $l $-t $s, the following segment-related additional labels occur:

$-h Plosive release phase: burst and any aspiration (this label has a duration)

$-q Creaky voice
   (a) in addition to or instead of a glottal stop: ##Q $-q ##Q- $-q
   (b) to mark a morpheme boundary or the separation of vowels where a glottal stop is not symbolized in the canonical transcription:
       ##n $a $x $#-q $aI $n $#-q $a $n $d $6 nacheinander
       ##Q- $-q $E $v $E $n $t $-h $u: $-q $'E $l eventuell
   (c) to replace lenis or fortis plosives in nasal or lateral contexts:
       ##k $9 $n $t-q $- @ $n+ könnten
   (d) to insert glottalization instead of a velar plosive that is not represented canonically
       ##f $6 $b $'I $n $d-n $U $N $-kq ##n $a: $x+ Verbindung nach

In all these cases, the label has no duration and is put on the same time mark as the next non-deleted label.
Nasalization in the context of a deleted nasal consonant: placed after the latter if nasalization only affects the post-nasal sequence, or before it if it affects the preceding material:

$$Q\-U\-\-n\Rs+$$ $\text{uns}$

$$Q\-E\-@\Rs+\#n-\-\-C\Rs t+$$ $\text{es nicht.}$

The label has no duration and is put on the same time mark as the next non-deleted label.

Presence of phonetic correlates of deleted labels in preceding context:

$$v\sa:6-a\$s\'aI\Rs n\Rs-\Ma$\Rs l-\C\Rs-\\#Q-\Rs aI-\Rs n-m\#b\Rs s'I\Rs s\Rs C\Rs E\-@\Rs n$$

$\text{waarscheinlich ein bieschen}$

palatality of deleted syllable $-\text{lich}$ transferred to preceding nasal

The label has no duration and is put at the beginning of the deleted sequence to indicate a phonetic residue in the segment(s) preceding it. The inserted label and all subsequent deleted symbols are put on the same time mark as the next non-deleted segmental label with a duration.

Insertion of pause or breathing, which are not in the orthographic text and are therefore not automatically generated in the canonical transcription, but are produced in reading; these labels have durations.

Hesitational lengthening in reading, put after the lengthened segment on the same time mark as the next non-deleted segment.

“non-text”: deviates from orthographic text

with duration of its own, marks a stretch of speech that is not represented orthographically and is not segmented and labelled, e.g. false starts

$\text{speaker erroneously reads}$

$\text{then corrects herself and reads}$

$\text{the actual orthography}$

$\text{false start is labelled as}$

$\text{speaker erroneously replaces}$

$\text{then corrects himself by repeating}$

$\text{the canonical segments of}$

$\text{at the same time a deviant pronunciation of}$

$\text{is indicated by}$

$\text{preceding the word, see below}$

$\text{marks a word as deviant in reading, e.g. slip of the tongue, and is put at the beginning of}$

$\text{which is segmentally labelled; it occurs on the same time mark as the}$

$\text{in}$

$\text{the initial}$

$\text{probably triggered by the}$

$\text{subsequent initial oral}$

$\text{but}$

$\text{deviant pronunciation of}$

$\text{the following word}$

$\text{see above}$
instead of the written sequence *als schon* the speaker produces *schon als*; the canonical labels of *als* are labelled as deleted, and the deviant reading is indicated by a preceding #-n: ; the signal stretch corresponding to the actually pronounced *als* after *schon* is labelled as ##-n: with no segmentation

4.2.2.3 Prosodic labels: PROLAB

A  Labels for prosodic categories

(1)  **PROSODIC PHRASE BOUNDARIES**
- Syntagmatic prosodic phrasing is marked by phrase boundaries at the right edges of the prosodic chunks:
  `-PGn in the default case
- In the main PROLAB system, prosodic phrase boundaries are not subclassified according to strength of separation. n is a place holder for subsequent finer differentiation.

(2)  **SENTENCE ACCENT**
The lexical items within prosodic phrases receive an ACCENT specification on 4 levels, put before the segmental sequence
`&2` default accent
`&3` reinforced accent
`&1` partial deaccentuation
`&0` complete deaccentuation

E.g. `#&2 Max #&0 hat #&0 einen #&3 Brief #&1 geschrieben`. "Max has written a letter." with narrow focus on Brief (answering "what did he write?"); complete deaccentuation of the function words "has" and "a", strong prominence on "Max", and weak prominence on "written".

(3)  **LEXICAL STRESS**
Within words, one or more syllables are specified for STRESS.
- It is marked at the segmental level and linked to the symbol for the stressed vowel
  `- primary stress
  `'' secondary stress

E.g. `#&1 $a: $r $@ $n forfahren`
- Function words, identified by + postfixed to the last segment, by default do not get a lexical stress mark, e.g. `#m $I $t+ mit`.
- If function words receive a sentence accent, `''` is inserted before the accented vowel, e.g. `#m $`' $I $t+ 'mit`.
- The same applies in the case of stress shift in relation to the canonical transcription, e.g. `#v $ 'U $n $d $6 $b $' $a:6 wunderbar`.

(4)  **PITCH PATTERNS**
Each ACCENT > &0 receives a distinctive pitch pattern which is either
(rising-)falling – PEAK
or (falling-)rising – VALLEY
or FLAT
and is hooked to the environment of the stressed syllable of the accented word, more particularly to its stressed vowel.

It consists of two parts
- the SYNCHRONIZATION of the f0 maximum/minimum with the stressed vowel
• the CONTINUATION of the pattern beyond the synchronization point to concatenate with the next ACCENT pattern or to terminate at the next PHRASE BOUNDARY if there are no further accents in the prosodic phrase.

SYNCHRONIZATION is marked by an addition to the ACCENT label, CONTINUATION by the insertion of another prosodic label before the next ACCENT or PHRASE BOUNDARY label. In each case, the labels are specific for each of the three pitch patterns.

• PEAK
SYNCHRONIZATION

\&2) EARLY synchronization of the peak maximum before the stressed vowel

\&2^ MEDIAL synchronization of the peak maximum in the stressed vowel

\&2( LATE synchronization of the peak maximum late in, or after, the stressed vowel, similarly for \&1 and \&3

CONTINUATION

\&2. fall to low pitch, creating a large pitch dip between two accents or ending in the speaker's lowest voice range at the end of the prosodic phrase

\&1. fall to medium pitch between two accents or at the end of the prosodic phrase

\&0. no pitch dip between two accents (e.g. 'hat pattern' between two peaks) or ending level at the prosodic phrase boundary

\&2., or \&1., combined contours at phrase boundaries: fall is followed by low rise

\&2? or \&1.? combined contours at phrase boundaries: fall is followed by high rise

\&2; or \&1; pseudoterminal contours at phrase boundaries: very slight rise after fall, softens terminality without removing it altogether.

Examples

\#&0 am \&2^ Samstag \#&0 oder \#&1. \#&2. \#&2^ PGN
"on Saturday or Sunday" either – or alternatives

\#&0 am \&2^ Samstag \#&0 oder \#&0. \#&2) Sonntag \#&2. \#&2^ PGN
"on Saturday or Sunday" = "at the weekend"

\#&2^ tschiüß \&1., \#&2^ PGN "bye"

\#&2^ t-h $E6 $m $'i: $n $&0. $&2^ $v $' $"Y $n $S $&1. $&2^ $' $@ \&1. Ter'min'wün'sche "preferences for dates", pronounced with three accents: the first is hooked to the primary-stress syllable min of the canonical form and given a \#& ACCENT label preceding the segmental string; the other two require primary lexical stress positions to be inserted into the canonical form by $' ' , to which $& ACCENT labels can then be linked inside the lexical item.

• VALLEY
SYNCHRONIZATION

\&2) EARLY synchronization of the valley minimum before the stressed vowel

\&2^ LATE synchronization of the valley minimum in the stressed vowel

similarly for \&1 and \&3

CONTINUATION

&, low rising

&? high rising, ending in the speaker's high voice range
Examples


heute ist schönes warmes Sommerwetter
"today we have got beautiful warm summer weather"

he has written a letter

(5) DOWNSTEP and UPSTEP, RESET
In a sequence of peak contours of accent level 2 within a prosodic phrase, the maximum f0 is by default successively lowered. This is traditionally handled by DECLINATION over time. KIM treats it as structurally controlled DOWNSTEP from accent to accent, irrespective of the time that elapses between them. This default DOWNSTEP is not labelled. On the other hand, if the DOWNSTEP is interrupted within a prosodic phrase an UPSTEP is labelled by putting \| in between the prosodic prefix and the ACCENT marker, e.g. &|2^.

Across PHRASE BOUNDARIES there is by default f0 RESET, i.e. DOWNSTEP starts afresh. If this is not the case and DOWNSTEP continues across the prosodic phrase boundary it is marked as &=PGn.

(6) PRE-HEAD
Unaccented syllables before the first PEAK ACCENT in a prosodic phrase are the PRE-HEAD of the contour and may be low or high in relation to the PEAK maximum. A low PRE-HEAD is taken to be default and not labelled. A HIGH PRE-HEAD is labelled as one of two different types &HP1 flat f0 at level of PEAK maximum
&HP2 high f0 onset, gradually descending to level of PEAK maximum.

B Ordering of prosodic labels

Punctuation marks, labels for CONTINUATION, for PHRASE BOUNDARIES, for sentence beginning, for HIGH PRE-HEAD, AND for type of ACCENT are put in this fixed order on the same time mark as the word boundary segment, also at word boundaries inside compound words.

#[], . ?, !] punctuation mark
#&[0. 1. 2., ? etc.] CONTINUATION part of PITCH PATTERN
#&[PGn =PGn] PHRASE BOUNDARY
#c: sentence beginning
#-n: non-text
#&[HP1 HP2] HIGH PRE-HEAD
4.2.3 Structure and generation of label files (*.s12h)

Label files *.s1[12]h have the following structure:

a. file name

b. 7-bit orthographic representation of the sentence bounded by 'oend' at the beginning of the next line

c. canonical transcription of the sentence in SAMPA, bounded by 'kend' at the beginning of the next line

d. phonetic transcription of the utterance: segmental (modified SAMPA) in *.s1h, segmental and prosodic (PROLAB) in *.s2h, constructed from the labels placed by the segmenter, and bounded by 'hend' at the beginning of the next line

e. label lines with time stamp (sample number), label, time point in sec
   • the segmenter enters a prototype file *.s0 into xassp, having the sections a., b. and c., and label lines with imaginary sample numbers and time points 100 sec in the future
   • by placing each label, one after the other as offered by the program, on the time scale of the signal file the imaginary values are changed to real ones
   • subsequently section d. of the label file is constructed
   • the *.s0 files of the add(itional)corp(us), which has not been labelled, consequently do not contain section d.

Examples

k01be004.s0
"Uber die Felder weht ein Wind .
oend
  Q y: b 6+ d i:+ f 'E l d 6 v 'e: t Q aI n+ v 'I n t .
kend
hend
  1600000 #c: 99.9999375
  1600000 ##Q 99.9999375
  1600000 $y: 99.9999375
  1600000 $b 99.9999375
  1600000 $6+ 99.9999375
  1600000 ##d 99.9999375
  1600000 $i:+ 99.9999375
  1600000 ##f 99.9999375
  1600000 $'E 99.9999375
  1600000 $l 99.9999375
  1600000 $d 99.9999375
  1600000 $6 99.9999375
  1600000 ##v 99.9999375
  1600000 $'e: 99.9999375
  1600000 $t 99.9999375
  1600000 ##Q 99.9999375
  1600000 $aI 99.9999375
"Uber die Felder weht ein Wind.

Q y: b 6+ d i:+ f 'E l d 6 v 'e: t Q aI n+ v 'I n t .

k01be004.s1h

3225 #c: 0.2015000
3225 ##Q- 0.2015000
3225 $y: 0.2015000
3847 $b 0.2403750
4741 $6+ 0.2962500
5837 ##d 0.3647500
6472 $-h 0.4044375
6666 $i:+ 0.4165625
7570 ##f 0.4730625
9175 $'E 0.5733750
10460 $l 0.6536875
11506 $d 0.7190625
11657 $-h 0.7285000
11926 $6 0.7453125
13715 ##v 0.8571250
14870 $'e: 0.9293125
16309 $t 1.0192500
16814 $-h 1.0508125
16958 ##Q- 1.0598125
16958 $-q 1.0598125
16958 $aI 1.0598125
18948 $n+ 1.1841875
20015 ##v 1.2508750
21173 $'I 1.3232500
22161 $n 1.3850000
23749 $-h 1.4842500
24322 $-h 1.5200625
24730 #. 1.5455625

k01be004.s2h

"Uber die Felder weht ein Wind.

Q y: b 6+ d i:+ f 'E l d 6 v 'e: t Q aI n+ v 'I n t .

kend

c: &0 Q- y: b 6+ &0 d -h i:+ &2( f 'E l d -h 6 &0 v 'e: t -h &0 Q- -q aI n+ &0. &2) v 'I n t -h . &2. &PGn

hend
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>orthographic representation of lexical item</td>
</tr>
<tr>
<td>2</td>
<td>position of word in orthographic part of the label file</td>
</tr>
<tr>
<td>3</td>
<td>orthographic representation of an overlaying event, e.g. laughing</td>
</tr>
<tr>
<td>4</td>
<td>position of the overlaying event in orthographic part of the label file</td>
</tr>
<tr>
<td>5</td>
<td>canonical transcription of lexical item</td>
</tr>
<tr>
<td>6</td>
<td>transcription of pronunciation variant</td>
</tr>
<tr>
<td>7</td>
<td>date and time of label file inclusion in database</td>
</tr>
<tr>
<td>8</td>
<td>full filename of label file (including path)</td>
</tr>
<tr>
<td>9</td>
<td>basename of label file (i.e. no path suffix)</td>
</tr>
<tr>
<td>10</td>
<td>speaker and text file index</td>
</tr>
<tr>
<td>11</td>
<td>gender</td>
</tr>
<tr>
<td>12</td>
<td>format of labels (SAMPA notation)</td>
</tr>
<tr>
<td>13</td>
<td>sample frequency in Hz</td>
</tr>
<tr>
<td>14</td>
<td>start of word in seconds</td>
</tr>
<tr>
<td>15</td>
<td>end of word in seconds</td>
</tr>
<tr>
<td>16</td>
<td>(and all even-numbered fields following, i.e. 18, 20, ...) start times of labels</td>
</tr>
<tr>
<td>17</td>
<td>(and all odd-numbered fields following, i.e. 19, 21, ...) durations of labels</td>
</tr>
</tbody>
</table>


### 4.2.4 Data banks (*.kdb)

Each line of the data bank has the following field structure:

1. orthographic representation of lexical item
2. position of word in orthographic part of the label file
3. orthographic representation of an overlaying event, e.g. laughing
4. position of the overlaying event in orthographic part of the label file
5. canonical transcription of lexical item
6. transcription of pronunciation variant
7. date and time of label file inclusion in database
8. full filename of label file (including path)
9. basename of label file (i.e. no path suffix)
10. speaker and text file index
11. gender
12. format of labels (SAMPA notation)
13. sample frequency in Hz
14. start of word in seconds
15. end of word in seconds
16. (and all even-numbered fields following, i.e. 18, 20, ...) start times of labels
17. (and all odd-numbered fields following, i.e. 19, 21, ...) durations of labels

Speaker information

The columns in the table below contain information about the speakers as follows:

1. Time of recording (month/year)
2. Speaker abbreviation
3. Gender (f = female; m = male)
4. Age
5. German spoken (classification made by speakers themselves):
   - SH – Schleswig-Holstein
   - NI – Lower Saxony (Niedersachsen)
   - ND – North German (Norddeutsch)
   - NW – North Rhine-Westfalia (Nordrhein-Westfalen)
   - HE – Hessen
   - RP – Rhineland-Palatinate (Rheinland-Pfalz)
   - MX – mixed
6. type of files
7. number of files
   - a: additionalcorp, not labelled, only *.wav and *s0 (orthographic and canonical)
8. duration of recording

<table>
<thead>
<tr>
<th>Date</th>
<th>File</th>
<th>Gender</th>
<th>Age</th>
<th>Location</th>
<th>Classification</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/90</td>
<td>k01</td>
<td>m</td>
<td>25</td>
<td>SH berlin</td>
<td>100</td>
<td>4 min 35 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k02</td>
<td>f</td>
<td>25</td>
<td>SH berlin</td>
<td>100</td>
<td>4 min 29 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k03</td>
<td>m</td>
<td>28</td>
<td>SH berlin</td>
<td>100</td>
<td>4 min 04 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k04</td>
<td>f</td>
<td>26</td>
<td>SH berlin</td>
<td>100</td>
<td>5 min 40 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k05</td>
<td>m</td>
<td>28</td>
<td>SH berlin</td>
<td>100</td>
<td>5 min 42 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k06</td>
<td>f</td>
<td>23</td>
<td>RP berlin</td>
<td>100</td>
<td>4 min 41 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k07</td>
<td>m</td>
<td>23</td>
<td>SH marburg</td>
<td>100</td>
<td>5 min 04 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k08</td>
<td>f</td>
<td>24</td>
<td>SH marburg</td>
<td>100</td>
<td>5 min 30 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k09</td>
<td>m</td>
<td>27</td>
<td>SH marburg</td>
<td>100</td>
<td>4 min 55 sec</td>
</tr>
<tr>
<td>10/90</td>
<td>k10</td>
<td>f</td>
<td>24</td>
<td>SH marburg</td>
<td>100</td>
<td>5 min 08 sec</td>
</tr>
<tr>
<td>10/90</td>
<td>k11</td>
<td>m</td>
<td>24</td>
<td>SH marburg</td>
<td>100</td>
<td>5 min 20 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k12</td>
<td>f</td>
<td>22</td>
<td>SH marburg</td>
<td>100</td>
<td>4 min 11 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k13</td>
<td>m</td>
<td>25</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 46 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:53</td>
<td>3 min 23 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:53</td>
<td>3 min 09 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k14</td>
<td>f</td>
<td>26</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 39 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:53</td>
<td>3 min 14 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k15</td>
<td>m</td>
<td>25</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 49 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:53</td>
<td>3 min 08 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k16</td>
<td>f</td>
<td>23</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 43 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:53</td>
<td>3 min 57 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k17</td>
<td>m</td>
<td>27</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 45 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:53</td>
<td>3 min 00 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k18</td>
<td>f</td>
<td>28</td>
<td>NW nordwind</td>
<td>2</td>
<td>0 min 43 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:58</td>
<td>4 min 17 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k19</td>
<td>m</td>
<td>27</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 54 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:58</td>
<td>3 min 35 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k20</td>
<td>f</td>
<td>24</td>
<td>NW nordwind</td>
<td>2</td>
<td>0 min 47 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:58</td>
<td>3 min 02 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k21</td>
<td>m</td>
<td>26</td>
<td>SH nordwind</td>
<td>2</td>
<td>0 min 40 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:58</td>
<td>2 min 15 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k22</td>
<td>f</td>
<td>26</td>
<td>SH butter</td>
<td>3</td>
<td>1 min 17 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:41</td>
<td>2 min 29 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k23</td>
<td>m</td>
<td>24</td>
<td>SH butter</td>
<td>3</td>
<td>1 min 32 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:41</td>
<td>2 min 48 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k24</td>
<td>f</td>
<td>20</td>
<td>SH butter</td>
<td>3</td>
<td>1 min 17 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>restkorp</td>
<td>a:41</td>
<td>2 min 48 sec</td>
</tr>
<tr>
<td>Date</td>
<td>Code</td>
<td>Gender</td>
<td>Location</td>
<td>Duration</td>
<td>Rest</td>
<td>Notes</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>10/90</td>
<td>k25</td>
<td>m</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 31 sec</td>
<td>restkorp a:41 2 min 51 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k26</td>
<td>N</td>
<td>but 3</td>
<td>1 min 17 sec</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/90</td>
<td>k27</td>
<td>m</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 24 sec</td>
<td>restkorp a:41 2 min 24 sec</td>
</tr>
<tr>
<td>10/90</td>
<td>k28</td>
<td>f</td>
<td>N</td>
<td>butter 3</td>
<td>1 min 15 sec</td>
<td>restkorp a:41 2 min 11 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k29</td>
<td>m</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 42 sec</td>
<td>restkorp a:41 2 min 53 sec</td>
</tr>
<tr>
<td>10/90</td>
<td>k30</td>
<td>m</td>
<td>N</td>
<td>butter 3</td>
<td>1 min 23 sec</td>
<td>restkorp a:46 2 min 38 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k61</td>
<td>M</td>
<td>MX</td>
<td>compl.corp. 403</td>
<td>26 min 07 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k62</td>
<td>f</td>
<td>NI</td>
<td>compl.corp. 403</td>
<td>23 min 45 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k63</td>
<td>m</td>
<td>SH</td>
<td>berlin 100</td>
<td>4 min 30 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k64</td>
<td>f</td>
<td>N</td>
<td>berlin 100</td>
<td>5 min 04 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k65</td>
<td>m</td>
<td>MX</td>
<td>berlin 100</td>
<td>4 min 34 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k66</td>
<td>f</td>
<td>MX</td>
<td>berlin 100</td>
<td>4 min 31 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k67</td>
<td>m</td>
<td>RP</td>
<td>marburg 100</td>
<td>5 min 11 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k68</td>
<td>f</td>
<td>N</td>
<td>marburg 100</td>
<td>5 min 02 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k69</td>
<td>m</td>
<td>HE</td>
<td>marburg 100</td>
<td>5 min 07 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k70</td>
<td>f</td>
<td>SH</td>
<td>marburg 100</td>
<td>5 min 13 sec</td>
<td></td>
</tr>
<tr>
<td>09/90</td>
<td>k71</td>
<td>m</td>
<td>ND</td>
<td>nordwind 2</td>
<td>0 min 44 sec</td>
<td>restkorp a:53 2 min 51 sec</td>
</tr>
<tr>
<td>10/90</td>
<td>k72</td>
<td>f</td>
<td>NI</td>
<td>nordwind 2</td>
<td>0 min 52 sec</td>
<td>restkorp a:53 3 min 09 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k73</td>
<td>m</td>
<td>SH</td>
<td>nordwind 2</td>
<td>0 min 48 sec</td>
<td>restkorp a:53 3 min 29 sec</td>
</tr>
<tr>
<td>10/90</td>
<td>k74</td>
<td>f</td>
<td>SH</td>
<td>nordwind 2</td>
<td>1 min 03 sec</td>
<td>restkorp a:58 5 min 27 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k75</td>
<td>m</td>
<td>ND</td>
<td>nordwind 2</td>
<td>0 min 43 sec</td>
<td>restkorp a:58 3 min 54 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k76</td>
<td>f</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 36 sec</td>
<td>restkorp a:41 2 min 34 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k77</td>
<td>m</td>
<td>ND</td>
<td>butter 3</td>
<td>1 min 34 sec</td>
<td>restkorp a:41 2 min 39 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k78</td>
<td>f</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 18 sec</td>
<td>restkorp a:41 2 min 08 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k79</td>
<td>m</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 43 sec</td>
<td>restkorp a:41 3 min 37 sec</td>
</tr>
<tr>
<td>09/90</td>
<td>k80</td>
<td>f</td>
<td>SH</td>
<td>butter 3</td>
<td>1 min 45 sec</td>
<td>restkorp a:46 3 min 44 sec</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Code</th>
<th>Gender</th>
<th>Location</th>
<th>Duration</th>
<th>Rest</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/92</td>
<td>dlm</td>
<td>M</td>
<td>SH</td>
<td>erlangen/siemens 200</td>
<td>13 min 13 sec</td>
<td></td>
</tr>
<tr>
<td>04/92</td>
<td>hpt</td>
<td>M</td>
<td>SH</td>
<td>erlangen/siemens 200</td>
<td>15 min 01 sec</td>
<td></td>
</tr>
<tr>
<td>04/92</td>
<td>kko</td>
<td>M</td>
<td>MX</td>
<td>erlangen/siemens 200</td>
<td>17 min 23 sec</td>
<td></td>
</tr>
<tr>
<td>04/92</td>
<td>rtd</td>
<td>F</td>
<td>NI</td>
<td>erlangen/siemens 200</td>
<td>17 min 07 sec</td>
<td></td>
</tr>
<tr>
<td>04/92</td>
<td>uga</td>
<td>F</td>
<td>SH</td>
<td>erlangen/siemens 200</td>
<td>13 min 29 sec</td>
<td></td>
</tr>
</tbody>
</table>
6 References


URL: http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html
Queries and comments

Any queries or comments can be directed to:
Dr Michel Scheffers
Institut für Phonetik und digitale Sprachverarbeitung der
Christian-Albrechts-Universität zu Kiel
D - 24098 Kiel
Germany
Tel:  ++49 431 880 3321
Fax:  ++49 431 880 1578
e-mail: ms@ipds.uni-kiel.de
www:  http://www.ipds.uni-kiel.de

Limits of Liability

Great care has been taken in the preparation of Kiel-CD #5, however the Editors are not responsible for any damage arising out of the use of this CD-ROM or the data contained on it.

Copyright notice

The data on the CD-ROM are copyrighted. They are for research purposes only. No form of commercial use is allowed, nor may the data be sold or transferred, as a whole or in parts, to another party. The files on the CD-ROM or any part thereof may not be reproduced in any form without the written permission from

The Editors of Kiel-CD #5
Institut für Phonetik und digitale Sprachverarbeitung der
Christian-Albrechts-Universität zu Kiel
D - 24098 Kiel
Germany
© September 2007