

The Kiel Corpus of Spontaneous Speech Vol. IV German: VIDEO TASK SCENARIO (Kiel-DVD #1)

Editors Klaus J. Kohler, Benno Peters, Michel Scheffers

1 Introduction

The *Kiel Corpus of Spontaneous Speech* in the APPOINTMENT-MAKING SCENARIO (German, Kiel-CD #2-4, 1995 – 1997, updated and reissued with prosodic annotations as Kiel-CD #6) was recorded with a technique that opened a channel for only one speaker in a dialogue and blocked the other. This made it impossible to study dialogue interaction. The technique also introduced a certain artificiality into unscripted communication. Therefore, Benno Peters looked for a new method to elicit spontaneous speech with overlapping dialogue turns. The result is the German LINDENSTRASSE corpus in the VIDEO TASK SCENARIO presented on Kiel-DVD #1. A description of the material and the strategy used to elicit it can be found in Peters (2001).

For the recordings, the two dialogue partners were seated in separate quiet, sound-treated rooms. They were not able to see each other and communicated via headphones, and microphones (Sennheiser MD421-U-4) placed on tables in front of them. This recording environment ensures high quality speech signals and good channel separation. Signals in each recording channel were digitized at 16kHz straight to disc yielding a stereo file for each dialogue (I*_dual.wav). From this stereo file two mono files were derived, one for each channel, which were also summed, generating a mono file of the complete dialogue (I*_mono.wav).

Benno Peters did the data collection; the recorded data were then orthographically transliterated by advanced students of phonetics under his supervision. Klaus J. Kohler generated the canonical transcriptions. The segmental and prosodic labelling was subsequently carried out in *xassp* by phonetically trained students under the supervision of Benno Peters and Klaus J. Kohler. Cf. IPDS (1997); Kohler, Pätzold, and Simpson (1997).

The DVD-ROM contains the transliteration and audio files as well as time-aligned segmental + prosodic label and commentary files of 6 overlapping German dialogues (4 female and 2 male speaker pairs) in the VIDEO TASK scenario (LINDENSTRASSE). In addition there are expanded label files with turn markers and the associated KielDat data banks for each of the 12 speakers.

The transliteration and the segmental labelling of the signals follow the conventions applied to the *Kiel Corpus of Spontaneous Speech* in the APPOINTMENT-MAKING SCENARIO, and are laid down in Kohler, Pätzold, and Simpson (1995). Prosodic labelling (PROLAB) is based on the Kiel Intonation Model (KIM) and is described in Peters and Kohler (2004), updated and summarized in English in section 4.2.2.1 below. Some additions introduced in the labelling of the LINDENSTRASSE corpus are also listed in sections 4.2.2/3.

The commentary files give descriptive phonetic comments at certain time points, beyond the systemic segmental and prosodic labelling categories.

Michel Scheffers serviced the speech processing software, integrated the data in the data bank and helped with the editing of the DVD-ROM.

2 Directory structure

The DVD-ROM has the following directory structure:

<DVD-ROM>

```
| +---l01
| | +---chan_cje          channel of speaker CJE
| | | +---l01cje.wav      complete signal file of speaker CJE
| | | +---l01cje.s1h      complete label file of speaker CJE
| | | +---l01cje.tur      complete label file of speaker CJE with markers for turn transitions
                           between the two dialogue partners
| | | +---l01cje.kdb      data bank of speaker CJE
| | | +---sect_cje        signal (*.wav) and label, commentary (*.s1h, *.kom) files of speaker
                           CJE, appr. 30-sec sections of the CJE channel of l01_dual.wav,
                           bounded by silence in both channels; numbering follows the number-
                           ing in the complete transliteration file (l01.trl), see 4.2.1

| | | | +---l01cje00.wav
| | | | +---l01cje00.s1h
| | | | +---l01cje00.kom
| | | | +---l01cje03.wav
| | | | +---l01cje03.s1h
| | | | +---l01cje03.kom
      .
      .
      .
| | +---chan_tev          channel of speaker TEV
| | | +---l01tev.wav      signal file of speaker TEV
| | | +---l01tev.s1h      label file of speaker TEV
| | | +---l01tev.tur      label file of speaker TEV with turn markers
| | | +---l01tev.kdb      data bank of speaker TEV
| | | +---sect_tev        section files of speaker TEV
| | | | +---l01tev01.wav
| | | | +---l01tev01.s1h
| | | | +---l01tev01.kom
| | | | +---l01tev02.wav
| | | | +---l01tev02.s1h
| | | | +---l01tev02.kom
      .
      .
      .
| | +---chan_mix          both speakers of dialogue l01
| | | +---l01_mono.wav    on one channel
| | | +---l01_dual.wav    on two channels
| | | +---l01.trl         complete transliteration file

| +---l02
| | +---chan_aha          channel of speaker AHA
      ...
| | | +---sect_aha
      ...
| | +---chan_tlu          channel of speaker TLU
      ...
```

+---sect_tlu	
...	
+---chan_mix	both speakers of dialogue 102
...	
+---103	
+---chan_ape	channel of speaker APE
...	
+---sect_ape	
...	
+---chan_kpa	channel of speaker KPA
...	
+---sect_kpa	
...	
+---chan_mix	both speakers of dialogue 103
...	
+---104	
+---chan_kth	channel of speaker KTH
...	
+---sect_kth	
...	
+---chan_che	channel of speaker CHE
...	
+---sect_che	
...	
+---chan_mix	both speakers of dialogue 104
...	
+---105	
+---chan_uho	channel of speaker UHO
...	
+---sect_uho	
...	
+---chan_jko	channel of speaker JKO
...	
+---sect_jko	
...	
+---chan_mix	both speakers of dialogue 105
...	
+---106	
+---chan_mpi	channel of speaker MPI
...	
+---sect_mpi	
...	
+---chan_tra	channel of speaker TRA
...	
+---sect_tra	
...	
+---chan_mix . . .	both speakers of dialogue 106

3 File naming conventions

The names of the transliteration files, signal files and label files are constructed according to the following conventions:

```
101xyz00.slh
^          dialogue scenario (l = Video Task LINDENSTRASSE)
101xyz00.slh
^^        dialogue number
101xyz00.slh
^^^       speaker
101xyz00.slh
^^        section number: the last two digits of the numbering in the
          101.trl file for speaker xyz of the 101 dialogue pair
101xyz00.slh
^^^       section label file (s1 = SAMPA labels; h = aspiration labels)
101xyz00.wav
^^^       section signal file (16-bit/16kHz)
101xyz00.kom
^^^       section commentary file
101xyz.wav
^^^       complete signal file of speaker xyz
101xyz.slh
^^^       complete label file of speaker xyz
101xyz.tur
^^^       complete label file of speaker xyz with turn transition marking
101xyz.kdb
^^^       data bank of speaker xyz
101_dual.wav
^^^^     both speakers of dialogue 101 on two channels
101_mono.wav
^^^^     both speakers of dialogue 101 on one channel
101.trl
^^^      transliteration file
```

4 File formats

4.1 Signal files

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

4.2 Text files (orthographic and label)

Text files: 7-bit ASCII files, DOS format (CR/LF at line end)

4.2.1 Transliteration files (*.trl)

Orthographic transliteration was carried out on the two-channel recording of a dialogue. Initially the dialogue was segmented into consecutive dialogue sections of approximately 30 sec, bounded by silence on both channels. The speech sections from both speakers in these chunks were then transliterated and numbered consecutively throughout the dialogue, in such a way that the section from the speaker who started speaking first in each chunk was given the next consecutive number first. All the transliterated sections for each dialogue were entered into one transliteration file. Each section has an identification comprising three letters (speaker) and three digits (section number, beginning at 000). The text of each section is indented 8 spaces and an empty line separates each

section from the next. Global commentaries are placed after ; at the beginning of a line. Transliteration follows DUDEN before the spelling reform, and applies the T_EX conventions to represent umlaut and "ß": "a = ä, "A = Ä, "o = ö, "O = Ö, "u = ü, "U = Ü, "s = ß. Word-initial small letters are kept sentence-initially. For further conventions, particularly those regarding non-lexical material and disfluencies, see Kohler et al., AIPUK 29 (1995).

Example

```
;Dialog: L01
CJE000: also , +/mein/+ meine Folge hatte gar keinen <:<#> Namen:> .
        +/das/+ die ging einfach los . <A> die bei dir hie"s sie
        richtig so ? nee , bei <:<#Kopfh> mir/-:> <P> m . <P>
        <Ger"ausch> ja , <P> ja . <:<#Kopfh> <A> <Schlucken> <A>:> ja
        Valle . <P> m , mit diesem Typ . <A> genau , die hat ihm das
        gesteckt , da"s das Kind nicht von ihm ist , und da hat er sie
        die Treppe runtergeschubst , ne ? <A> ..

TEV001: <Schmatzen> <A> <:<#> ach so:> <A> . <A> ja , bei mir wird 's
        richtig <"ahm> <A> +/aufge=/+ <P> d=/+ also , gezeigt , ja .
        <A> ja , und dann <A> , ja , <%> wie fing 's dann an , dann
        fing 's an hier mit , <A> na , wie hei"st sie noch ? <Lachen>
        Walze , ne ? <A> Walze +/mit dem/+ <P> mit dem/+ <P> mit dem
        anderen da im Bett . <A> ..

TEV002: ach so , das hab' ich nicht gesehen <A> . <Lachen> <A> aber ej
        sag' mal , ich denk' , das ist ja Phonetik und nich<Z>t
        Psychologie . <P> <Ger"ausch> <P> und was ist das denn ? na ja
        , <A> gut . <P> also/- <A> <Ger"ausch> <Schmatzen> <A> ach ,
        das hab' ich nicht gesehen . <P> ..
        .
        .
        .
```

4.2.2 Label files (*.s1h)

Section label files have the following structure:

- a. file name
- b. 7-bit orthographic representation which is a copy of the corresponding section in the transliteration file, bounded by 'oend' at the beginning of the next line
- c. canonical transcription of the section, bounded by 'kend' at the beginning of the next line
- d. segmental transcriptions (modified SAMPA) and prosodic annotations (PROLAB) of the section, constructed from the labels placed by the segmenter, and bounded by 'hend' at the beginning of the next line
Cf. Kohler et al. (1995) for segmental labelling; Peters and Kohler (2004) for PROLAB labelling, summarized in English in section 4.2.2.1 below; Kohler, Kleber, and Peters (2005) for the analysis of prosodically labelled data.
- e. label line with time stamp (sample number), label, starting time in sec

The complete label file l0?xyz.s1h for a speaker xyz has the same structure as the individual section files of that speaker. The orthographic, canonical, variant and label line parts each contain the linear sequence of the corresponding parts in all the section files of speaker xyz. In the orthographic part, only the identification of the first section is kept.

Example of a section label file

l01cje00.slh

```
CJE000: also , +/mein/+ meine Folge hatte gar keinen <:<#> Namen:> .
+/das/+ die ging einfach los . <A> die bei dir hie"s sie
richtig so ? nee , bei <:<#Kopfh> mir/-:> <P> m . <P>
<Ger"ausch> ja , <P> ja . <:<#Kopfh> <A> <Schlucken> <A>:> ja
Valle . <P> m , mit diesem Typ . <A> genau , die hat ihm das
gesteckt , da"s das Kind nicht von ihm ist , und da hat er sie
die Treppe runtergeschubst , ne ? <A>
```

oend

```
Q a l z o:+ , m a I n+ /+ m a I n @+ f 'O l g @ h a t @+ g 'a: 6
k 'a I n @ n :k n 'a: m @ n . d a s+ /+ d i:+ g 'I N
Q 'a I n #f "a x l 'o: s . h: d i:+ b a I+ d i: 6+ h 'i: s z i:+
r 'I C t I C z o:+ ? n 'e: , b a I+ :k m i: 6+ /- p: m . p:
g: j 'a: , p: j 'a: . :k h: :k w: :k h: j 'a: v 'a l @ .
p: m , m I t+ d i: z @ m+ t 'y: p . h: g @ n 'a U , d i:+
h a t+ Q i: m+ d a s+ g @ S t 'E k t , d a s+ d a s+ k 'I n t
n I C t+ f O n+ Q i: m+ Q I s t+ , Q U n t+ d a:+ h a t+
Q e: 6+ z i:+ d i:+ t r 'E p @ r 'U n t 6 #g @ S "U p s t , n @+
? h:
```

kend

```
c: &0 %Q a %l z o:+ , &2( m ' ' a I n+ /+ &0. &PG/ &3(
m ' ' a I n @+ &0 f 'O %l g @ &0 h a t -h @+ &0. &2^ g -h 'a: 6
&0 k -h 'a I n @- %n &0. &1^% :k %n 'a: m @- %n-m . &2; &=PGn
c: &0 d a s+ /+ &0 d -h i:+ &2^ g -h 'I N &%0
Q- -q 'a I n #f "a x &0. &2) l 'o: s . &2. &PGn c: h: &0
d i:+ &0 b a I+ &0 d i: 6+ &2^ h 'i: s &0 z -s i:+ &0
r 'I C t -h I C &0 z o:+ ? c: &0 n 'e: , &0 b a I+ &1. &2^
:k m ' ' i: 6+ /- &0. &PG/ p: &2- ' ' m . &0. &PGn c: p: g:
&2)% j 'a: , &1. &PGn p: &2) j 'a: . &1. &PGn c: :k h:
:k w: :k h: &2^ j 'a: &1. &2( v 'a l @ . &2. &PGn c: p:
&2- ' ' m , &0. &PGn &HP1 &0 %m I t+ &0 %d -h i: z @- m+ &2^
t -h 'y: p . &1. &PGn c: &HP1 h: &0 %g -h @ n 'a U , &0
d i:+ &1^ h ' ' a t -h+ &0 Q- i: m+ &0 d a s+ &2>
g -h @ S t -h 'E k %t , &0 %d -h a s+ &0 d -a s+ &1. &2^
k -h 'I n t &0 n I C t-+ &0 f O n+ &1. &|2^ Q- ' ' i: m+ &0
Q- I s t-+ , &1; &=PGn &PC &HP1 &0 Q- U n t-+ &0 %d-n a:+
&1^ h ' ' a t -h+ &0 Q- e: 6+ &0 z i:+ &0 d i:+ &%1. &2)%
t -h r 'E p @ &0. &%2^ r 'U n t 6 &%0. &%2^% #g @ S ' ' "U p s t ,
&2. &1]% n ' ' @+ ? &, &PGn c: h:
```

hend

```
187 #c: 0.0116250
187 #&0 0.0116250
187 ##%Q 0.0116250
579 $a 0.0361250
887 $%1 0.0553750
1158 $z 0.0723125
1983 $o:+ 0.1238750
2456 #, 0.1534375
2456 #&2( 0.1534375
2456 ##m 0.1534375
3401 $' ' 0.2125000
3401 $aI 0.2125000
5024 $n+ 0.3139375
```

```
.
.
.
```

4.2.2.1 Principles of PROLAB and list of prosodic labels

A *General principles for segmental and prosodic labels*

- (1) All segmental and prosodic labels (including syntactic punctuation) are placed on one tier. But see section 4.2.3 for further prosodic markers on a separate commentary 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), **#l:** (laugh), **#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 section B(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 4.2.2).
- (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 C.
- (4) 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 **%**, eg. **&%2^** or **&2^%** or **&%2^%**.

B *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
 - &PG/** after truncations and false starts.
 - 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. See section 4.2.5.
- (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." 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. **##f '\$O6 \$t \$f '\$a: \$r \$@ \$n fortfahren**
- 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 wunder'bar.**

(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 peak (maximum before the stressed vowel)
&2^ MEDIAL peak (maximum within the stressed vowel)
&2(LATE peak (maximum late within, 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(Sonntag ##&2. ##&PGn
"on Saturday or Sunday" either – or alternatives

&0 am &2^ Samstag &0 oder &0. &2) Sonntag &2. #&PGn
 "on Saturday or Sunday" = "at the weekend"
&2^ tschüß &1., #&PGn "bye"

&2^ ##t \$-h \$E6 \$m \$'i: \$n \$&0. \$&2^ \$#v \$' '\$'Y \$n \$\$ \$&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 valley (minimum before the stressed vowel)
&2[LATE valley (minimum within the stressed vowel),
 similarly for **&1** and **&3**

CONTINUATION
&, low rising pitch
&? high rising pitch, ending in the speaker's high voice range

Examples

&0 er &0 ist &0 in &2] Rom &? #&PGn "he is in Rome?" asking for confirmation
&0 er &0 ist &0 in &3[Rom &? #&PGn "he is in Rome?" surprise question
&0 er &0 ist &0 in &2[Rom &, #&PGn "he is in Rome." friendly non- categorical statement

- FLAT pattern
 SYNCHRONIZATION
&2- &1- &3-

CONTINUATION
&0. &1. &2.

Examples

&2^ heute &0 ist &1. &2(schönes &0. &2- warmes &0. &2) Sommerwetter
&2. #&PGn
 "today we have got beautiful warm summer weather"
&0 er &0 hat &0 einen &2(Brief &2. &1- geschrieben &2. #&PGn
 "he has written a letter"

(5) FORCE ACCENT

In addition to the PROLAB conventions applied to the APPOINTMENT-MAKING corpus, a label **>** is used to mark a FORCE ACCENT in the LINDENSTRASSE corpus, put after the accent marker **&2** or **&3**, either on its own for a non-tonal accent, or followed by an intonation marker for a tonal overlay of a force accent, e.g. **&2>** or **&2>(.**

Example

&0 das &0 ist &2>^ beschissen#&2. #&PGn "That's bloody awful."
&0 wie &0 er &0 sie &0 die &2^ Treppe &2. runter\$&3> \$#k\$' 'ickt #&2. #&PGn
 "as he kicks her down the stairs"

(6) 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 DOWNSTEP continues across the prosodic phrase boundary the latter is marked as **&=PGn** .

(7) 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.

(8) RATE and REGISTER

In the labelling of the APPOINTMENT-MAKING SCENARIO, deviations from an impressionistically judged mean speech rate of a speaker were labelled as
&RP speeding up
&RM slowing down;
deviations from an impressionistically judged mean pitch level of a speaker were labelled as register changes
&LR change to a lower pitch level
&HR change to a higher pitch level
These changes within a speaker's delivery are in general associated with prosodic boundaries and therefore marked at the beginning of prosodic phrases in which they take place.
In the labelling of the LINDENSTRASSE SCENARIO, RATE and REGISTER labelling were removed from the label files *.s1h and incorporated in the commentary files *.kom, see section 4.2.3.

C *Ordering of prosodic labels*

Punctuation marks, labels for CONTINUATION, for PHRASE BOUNDARIES, for sentence beginning, for HIGH PRE-HEAD, RATE, REGISTER, 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 PG/]	PHRASE BOUNDARY
#c:	sentence beginning
#&[HP1 HP2 RM RP LR HR]	HIGH PRE-HEAD (RATE, REGISTER in APPOINTMENT-MAKING SCENARIO); the ordering among these labels is free
#& [1 2 3] > [(^) []]	ACCENT, FORCE ACCENT, SYNCHRONIZATION of PITCH PATTERN
##	word boundary
\$\$&[0. 1. 2.]	CONTINUATION part of PITCH PATTERN inside compound
\$\$ [1 2 3] > [(^) []]	ACCENT, FORCE ACCENT, SYNCHRONIZATION of PITCH PATTERN inside compound
\$#	boundary of compound element
\$' '	inserted lexical stress before vowel to receive ACCENT

4.2.3 Commentary files (*.kom)

In the labelling of the LINDENSTRASSE corpus, an additional formalized commentary level is introduced to note labelling problems as well as additional phonetic phenomena which are not captured by the PROLAB notation. The label **&PC** in the *.s1h file links the main label and commentary levels. The following features are represented: register (RG), speech rate (RT), range (RN), loudness (VO), voice quality (VQ), imitation of another voice, dialect etc. (IM), and problem section (PS). The notations have the following structure:

PHENOMENON_B_ATTRIBUTE

where B marks the beginning of the PHENOMENON RG, RT, RN, VO, VQ, IM, or PS. The attribute is plus or minus for RG, RT, RN, and VO. It is creak, breathy, whispered, etc. for VQ, and a free commentary for IM and PS. The attributes of RG, RT, RN, VO and VQ may be further specified by supplementary comments. The complex label is set at the prosodic phrase boundary after which the phenomenon occurs. The end is marked by PHENOMENON_E at the prosodic boundary after the phenomenon. RG and RT labels at the commentary level replace the HR/LR (high/low register) and RP/RM (rate plus/rate minus) labels, which are used in the *.s1h files of the APPOINTMENT-MAKING corpus. The format follows the label files.

Example l06mpi39.kom

```
oend
kend
hend
      1 VO_B_minus      0.0000000
    92754 VO_E          5.7970625
   328843 RG_B_minus   20.5526250
   346379 RG_E         21.6486250
   346379 VO_B_plus    21.6486250
   368935 VO_E         23.0583750
   368935 VO_B_minus   23.0583750
   375599 VO_E         23.4748750
```

A *.s1h file was supplemented by a *.kom file if, and only if, such additional phonetic information needed to be provided in the labelling of the speech signals.

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 index (includes section number in dialogues)
- 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

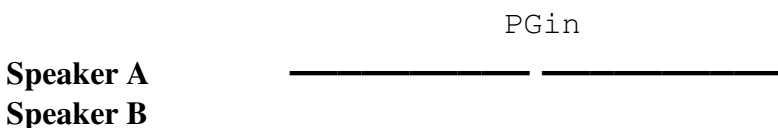
Elements of an entry are separated by tabs. Cf. Pätzold (1997).

4.2.5 Complete label files with turn transition markers (*.tur)

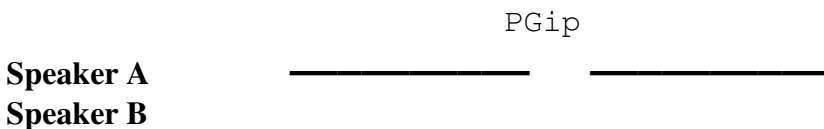
In the prosodic labelling of *.slh files, only one type of boundary category marker is used (&PGn). For each speaker, an additional segmental and prosodic label file of a complete dialogue is created (*.tur), in which a subclassification of phrase boundary marking is introduced, providing information about different boundary types and about the synchronizing of turns between the two speakers in a dialogue. The following cases are differentiated.

(1) Turn-internal prosodic boundaries without overlaying talk

In the case of a turn-internal prosodic boundary without overlaying talk or a pause between the phrases the label PGn is changed to PGIN.



A turn-internal prosodic boundary with pause or breathing between the phrases is labelled PGip.



(2) Turn-internal prosodic boundaries with overlaying talk

When phrases are overlayed by utterances from the dialogue partner the label PGio is used, except for overlays by backchanneling like *uhm* or *m*.



When more than one phrase is overlayed by simultaneous talk, the end of every phrase with overlay is labelled PGio.



(3) Turn-internal prosodic boundaries before simultaneous start of both speakers

When both speakers start talking simultaneously, the label PGib is used. *Simultaneous* refers to a delay shorter than 100 ms between the starting points of the utterances of speaker A und B.

PGib



(4) Turn-final prosodic boundaries without overlay or pause

Prosodic boundaries are classified as *turn-final* when speakers end their utterances at the boundary and the dialogue partner starts talking. Such prosodic boundaries between adjacent turns are labelled PGtn when there is no overlay or pause of more than 100 ms between the speakers' turns.

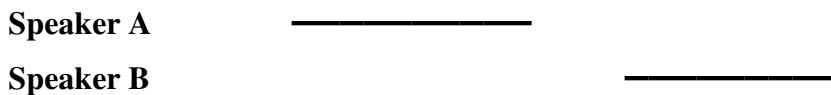
PGtn



(5) Turn-final prosodic boundaries with pause

Turn-final boundaries with a pause of more than 100 ms between the turns are labelled PGtp.

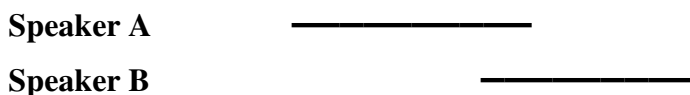
PGtp



(6) Turn-final prosodic boundaries with overlay

Turn-final boundaries with an overlay of more than 100 ms of the turns are labelled PGto.

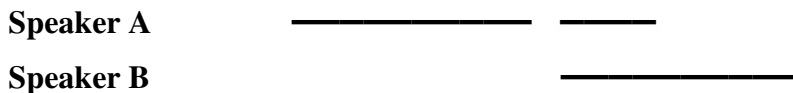
PGto



(7) Turn-final prosodic boundaries after simultaneous starts

Turn-final boundaries are labelled PGto when there is a delay shorter than 100 ms between the starting points of the utterances of speakers A and B, and speaker A stops speaking while the incomer continues.

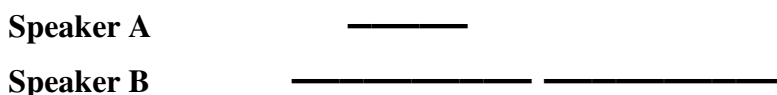
PGib PGto



(8) Turn-external prosodic boundaries in overlaying talk

Turn-external utterances are stretches of talk that overlay continuous speech for a short time. The label PGeo is used, when a speaker produces a short stretch of speech overlaying the other speaker's utterance and stops speaking before the other speaker.

PGeo



(9) Prosodic boundaries after backchanneling

Prosodic boundaries after backchanneling like *uhm* or *m* are labelled PGer.

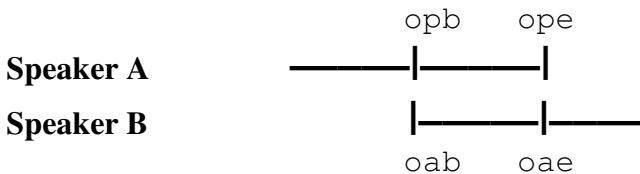


(10) Ambiguous cases

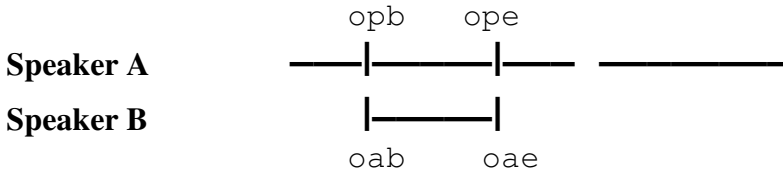
Ambiguous cases are labelled PGx.

(11) Delimitation of overlap

The starting and end points of overlaying stretches of speech are labelled in both channels. The following illustration shows a turn transition with overlap. The starting point of the overlay is labelled *opb* (*overlap passive begin*) for the turn holder and *oab* (*overlap active begin*) for the incomer. The endpoint of the overlay is labelled *ope* (*overlap passive end*) for the turn holder and *oae* (*overlap active end*) for the incomer.



The following illustration shows the labelling of overlaying speech *inside* the turn of speaker A.



(12) Delimitation of pauses between turns

The following illustration shows a turn transition with a pause. The starting point of the pause between the turn is labelled *pb* (*pause begin*) the end of the pause *pe* (*pause end*). The end of the pause is determined by the point of time the incomer starts speaking. Breathing is treated in the same way as silent pauses.



This turn labelling offers the following possibilities for analysis:

- separating turn-internal and turn-final prosodic boundaries
- distinguishing different types of turn transitions between dialogue partners with a pause, with direct continuation, with overlap
- automatic computation of pause and transition durations

- separating reception signals from ordinary turns, thus allowing the exclusion or the specific analysis of phrase boundaries after this type of utterance.

Cf. Peters (2006).

5 Speaker information

The table below provides information about the speakers. The columns contain the following:

- Dialogue
- Time of recording (month/year)
- Dialogue partners acquainted (k = yes; uk = no)
- Speaker abbreviation
- Gender (f= female; m = male)
- Age at time of recording
- German spoken (classification made by speakers themselves):
SH - Schleswig-Holstein
- Profession
- Duration of recording (two channels, labelled separately)

101	01/2000	k	CJE	f	28	SH	student	09:37
			TEV	f	23	SH	student	
102	02/2000	k	AHA	f	22	SH	highschool pupil	13:02
			TLU	f	20	SH	highschool pupil	
103	01/2000	k	APE	f	27	SH	student	10:03
			KPA	f	32	SH	designer	
104	01/2000	k	KTH	f	23	SH	student	09:23
			CHE	f	23	SH	student	
105	11/2000	k	UHO	m	33	SH	technician	15:27
			JKO	m	31	SH	tradesman	
106	10/2000	k	MPI	m	35	SH	ship builder	11:26
			TRA	m	31	SH	technician	
								68:58 Total time

6 References

- IPDS (1997). xassp User's Manual (Advanced Speech Signal Processor under the X Window System. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, vol. 32 (Institut für Phonetik und digitale Sprachverarbeitung, Kiel), 31-115.
- Kohler, K. J., Pätzold, M., and Simpson, A. (1995). From scenario to segment: the controlled elicitation, transliteration, segmentation and labelling of spontaneous speech. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, vol. 29 (Institut für Phonetik und digitale Sprachverarbeitung, Kiel).
- Kohler, K. J., Pätzold, M., and Simpson, A. (1997). From the acoustic data collection to a labelled speech data bank of spoken Standard German. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, vol. 32 (Institut für Phonetik und digitale Sprachverarbeitung, Kiel), 1-29.
- Kohler, K. J., Kleber, F., and Peters, B. (eds) (2005). Prosodic Structures in German Spontaneous Speech. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, vol. 35a (Institut für Phonetik und digitale Sprachverarbeitung, Kiel)
URL: <http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html>
- Pätzold, M. (1997). KielDat - Data bank utilities for the Kiel Corpus. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, vol. 32 (Institut für Phonetik und digitale Sprachverarbeitung, Kiel), 117-126.
- Peters, B. (2001). 'Video Task' oder 'Daily Soap Szenario' - Ein neues Verfahren zur kontrollierten Elizitation von Spontansprache (" 'VideoTask' or 'Daily Soap Szenario' - A new method for the controlled elicitation of spontaneous speech"). *Manuscript*.
URL: http://www.ipds.uni-kiel.de/pub_exx/bp2001_1/Linda21.html.
- Peters, B. (2006). Form und Funktion prosodischer Grenzen im Gespräch. PhD Diss., University of Kiel, 2006.
URL: <http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html>
- Peters, B., and Kohler, K. J. (2004). Trainingsmaterialien zur prosodischen Etikettierung mit dem Kieler Intonationsmodell KIM ("Training Material for Prosodic Labelling with the Kiel Intonation Model KIM"). *Manuscript*.
URL: <http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html>

Queries and comments

In case there are problems reading the DVD, the user should first try another DVD drive.
Any queries or comments can be directed to:

Dr. Michel Scheffers
Christian-Albrechts-Universität zu Kiel
Institut für Phonetik und digitale Sprachverarbeitung
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-DVD #1, however the Editors are not responsible for any damage arising out of the use of this DVD-ROM or the data contained on it.

Copyright notice

The data on the DVD-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 DVD-ROM or any part thereof may not be reproduced in any form without the written permission from

The Editors of Kiel-DVD #1
Christian-Albrechts-Universität zu Kiel
Institut für Phonetik und digitale Sprachverarbeitung
D - 24098 Kiel
Germany
© August 2006