

Transcription Tools, Transcription Conventions, and the TEI Guidelines for Transcriptions of Speech

Thomas Schmidt, University of Hamburg

Andreas Witt, University of Tübingen

This talk



Three worlds:

- **Multi-layer transcription tools**
- **Transcription conventions of conversation analysis (and similar fields)**
- **The Text Encoding Initiative**

How to bring them together?

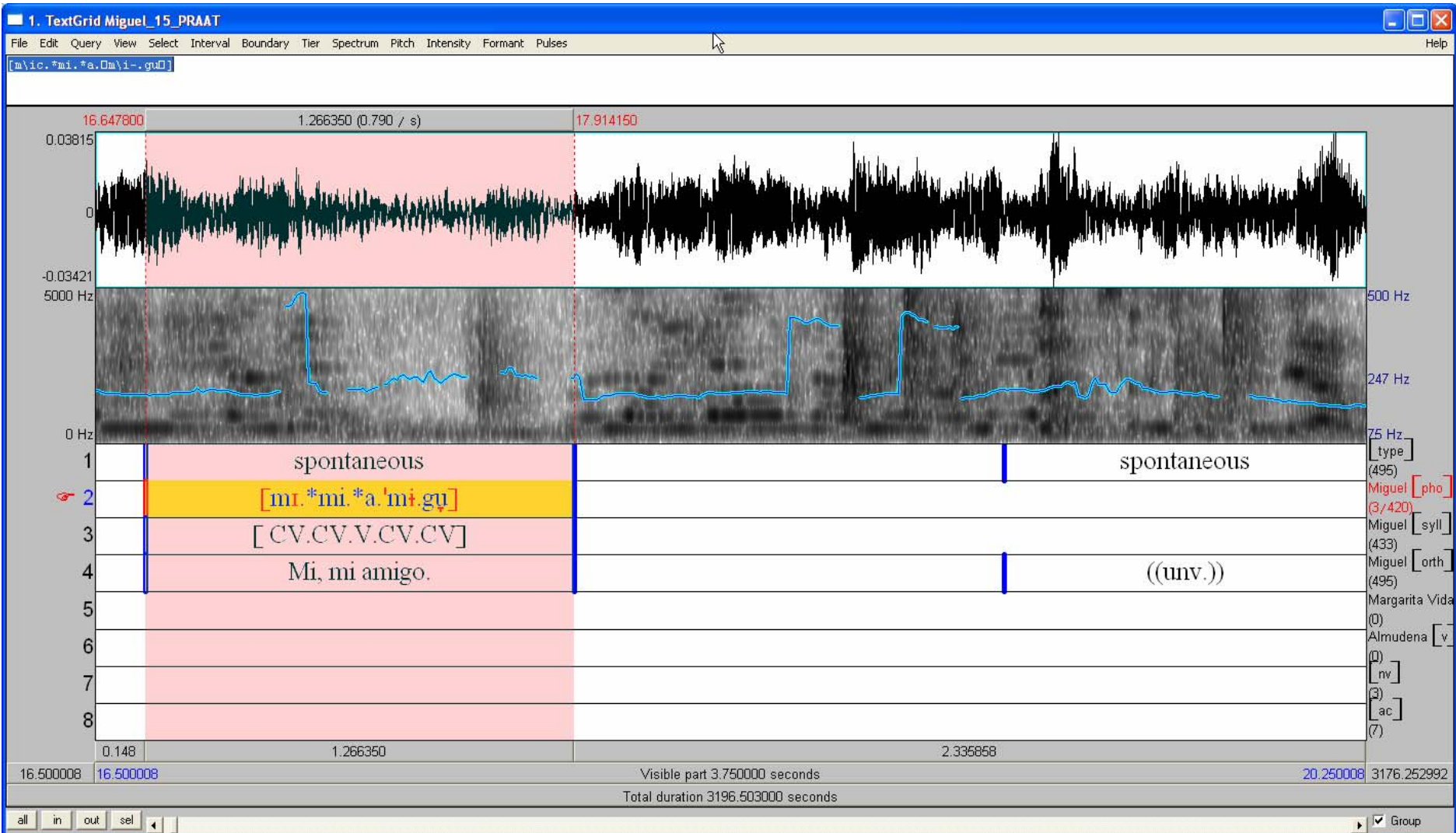
World 1: Multi Layer Transcription Tools

- ANVIL, Praat, EXMARaLDA, ELAN, TASX-Annotator, MacVisSTA
- Developed for
 - Study of multimodal behaviour
 - Phonetic/Phonological analysis
 - Study of multi party conversation
 - Transcription + Glossing + Translation (field linguistics, endangered Languages)
- Also used in
 - Dialectology
 - Language acquisition studies
 - Multi-layer annotation of written texts

The screenshot displays the ANVIL software interface, which is used for video analysis and gesture annotation. It consists of several main windows:

- Anvil 3.6 (Top Left):** The main control window. It shows the current video file being loaded: "IV50, 384x288, FrameRate=25". It also displays the video frame rate (25.0), audio format (LINEAR, 22050.0 Hz, 8-bit, Mono), and duration (03:43:92). The current specification is "D:\Research\anvil-spec\litqua2.xml". The current time is 03:08:76, and the frame number is 4719. Playback controls are visible at the bottom.
- Video: lq1-7-reich.avi (Top Middle):** A video player window showing a man in a suit and glasses, identified as Michael Reich, with his arms raised in a gesture. The ZDF logo is visible in the top left corner of the video frame.
- Track: gesture.phrase (Top Right):** A window showing the current track. It lists the track name, the referenced track, and the time range (03:07:63 - 03:10:27, 66 frames). It also displays attributes: category: iconic, iconic type: smash, handedness: 2H, cooc: Rivas, function: emblematic, and timing: direct. A comment field contains "compare with lq1-8 at 0:28". Control buttons for start, edit, end, cut, extend, and del are at the bottom.
- Annotation: lq1-7-reich.anvil (Bottom):** A timeline window showing the annotated video. It includes a waveform (wave), Praat spectrogram (praat), and a transcript (trl) with the text: "übersetzen", "ich spüre den poetischen Stil", "den ich bei Rivas", "na nur die Bemühung". Below the transcript, there are tracks for posture (pose, shift), phase (prep, beats, prep, hold, stroke, hold, prep, stroke, hold), phrase (metaphoric, heart, 2H; iconic smash, 2H; emblem, so-what, 2H), and gesture (compound). The "Rivas" segment in the transcript and the corresponding "iconic smash, 2H" and "emblem, so-what, 2H" annotations are highlighted in green.

ANVIL: Michael Kipp, DFKI Saarbrücken



Praat: Paul Boersma, University of Amsterdam

Elan - r03_v20_s5.eaf

File Edit Annotation Tier Type Search View Options Window Help

Grid Text Subtitles Controls

Gloss2

| Nr | Annotation | Begin Time | End Time | Duration |
|----|---|--------------|--------------|--------------|
| 1 | he's speaking into his 1/2 bush knife | 00:00:31.708 | 00:00:32.708 | 00:00:01.000 |
| 2 | I am saying that he's talking into a half bushknife | 00:00:37.141 | 00:00:38.501 | 00:00:01.360 |
| 3 | everything we are saying is going inside | 00:00:39.336 | 00:00:40.536 | 00:00:01.200 |
| 4 | See this fellow: he seems to me like a little frog | 00:00:45.426 | 00:00:47.106 | 00:00:01.680 |
| 5 | it's collecting all these things? | 00:00:53.400 | 00:00:54.460 | 00:00:01.060 |
| 6 | it's recording | 00:00:55.521 | 00:00:56.056 | 00:00:00.535 |
| 7 | You have sung them, sing them again | 00:00:59.412 | 00:01:00.942 | 00:00:01.530 |
| 8 | You sing those two parts of the MYââ mbwaa | 00:01:01.120 | 00:01:02.462 | 00:00:01.342 |
| 9 | You (P) ask him (Kp) | 00:01:02.602 | 00:01:03.982 | 00:00:01.380 |
| 10 | that thing, you sung those 2 parts of Myââ, sing them again | 00:01:05.591 | 00:01:07.931 | 00:00:02.340 |
| 11 | the next one | 00:01:32.024 | 00:01:32.694 | 00:00:00.670 |

00:01:01.115 Selection: 00:01:01.115 - 00:01:02.182 1067

00 00:01:00.000 00:01:01.000 00:01:02.000 00:01:03.000 00:01:04.000 00:01:05.000 00:01:06.000 00:01:07.000 00:01:08.000 00:01:09.000

Gloss1

Gloss2 You have sung them, sing t You sing those two part You (P) ask him (Kp) that thing, you sung those 2 parts of Myââ,

Kp mm :êê?

Pius Myââ mbwaa dê a ngî Myââ kn:ââ lô y:i a nyi cha

K mu cha ngî dé, mède ngî d Myââ mbwaa dê... u kwo péé, nyi u kwo pé nyi u kwo p mu tpile, mu Myââ u mbwaa dê mu cha n

Ricky

man outside vie = K?

child 1

loudspeaker

ELAN: Han Sloetjes, MPI Nijmegen

EXMARaLDA Partitur-Editor 1.4 [T:\VTP-Z2\DATEN\BEISPIELE\LYON\Pear Story\BasicTransPear.xml]

File Edit View Tier Event Timeline Format Segmentation Help

| | 0 [00.00] | 1 [01.90] | 2 [01.10] | 3 [03.21] | 4 [04.10] | 5 [09.20] |
|---------|-------------------------------|--------------------------------|---------------------|----------------------------------|-------------------|-----------|
| | | <i>louder</i> | | | | |
| X [v] | So it starts out with: A | roo | ster crows | . ((1,3s)) ((takes breath 0,5s)) | And the | |
| X [nv1] | <i>rHA on rKN, lHA on ISH</i> | <i>rHA up and to the right</i> | <i>rHA stays up</i> | | <i>rHA back c</i> | |
| X [nv2] | | | <i>HE nods once</i> | | | |
| X [nv3] | | <i>emphasizes the crow</i> | | | | |
| Y [v] | | | | | | |

Done.

Audio/Video panel

..DATEN\BEISPIELE\LYON\Pear Story\pear.mov

Start: sync

Position: sync

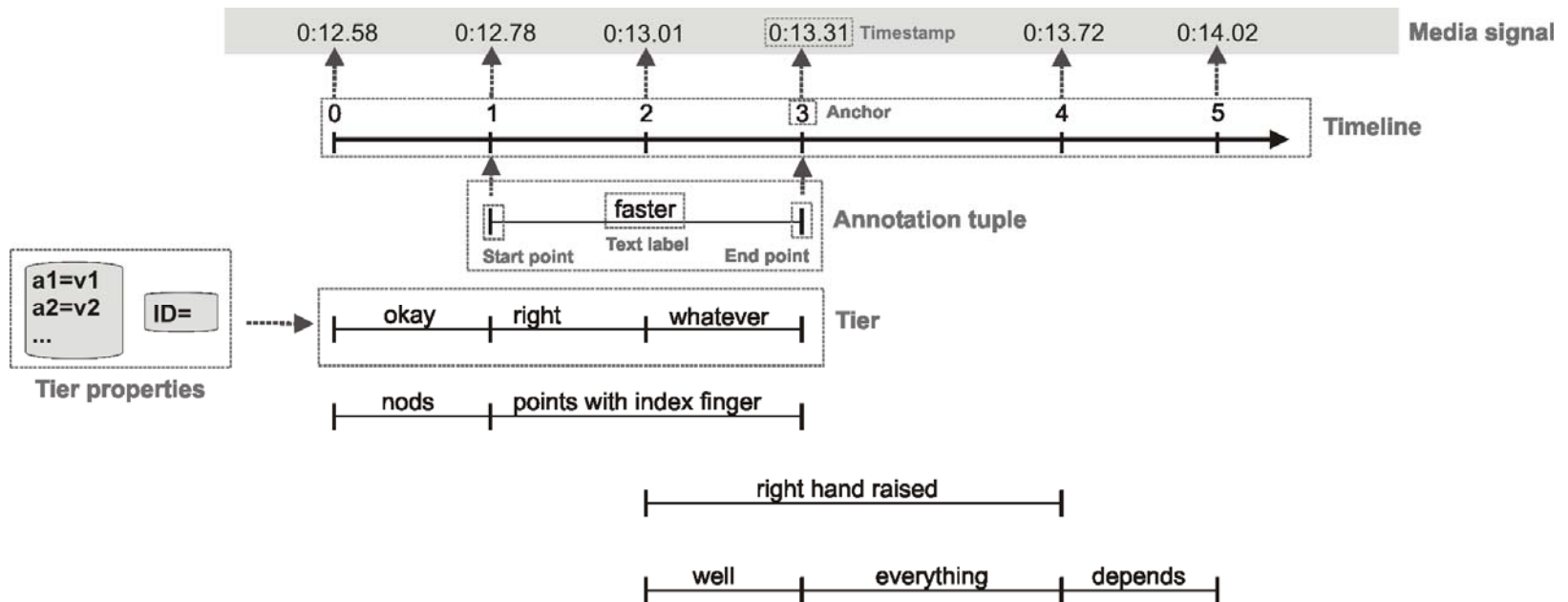
Stop: sync

0.0 3.2 5.1 43.8 43.8 sec

Playback halted

EXMARaLDA: Thomas Schmidt, University of Hamburg

Model: Single timeline, multiple tiers



- Annotation tuples: text label + timeline reference
 - Timeline: fully ordered, reference to a recording
 - Tiers: collections of annotations of a specific category, a specific speaker, annotations in a tier do not overlap
- Annotation Graph Framework (Bird/Lieberman 2001)

Data formats

| | 0 [0] | 1 [1.3] | 2 [2.6] | 3 [4] |
|----------|---------------------|---------------|---------|-------|
| Max [v] | You always | interrupt me. | | |
| Max [nv] | <i>gesticulates</i> | | | |
| Tom [v] | | No I | don't. | |

```
<basic-transcription>
  <timeline>
    <timepoint id="T0" time="0.0"/>
    <timepoint id="T1" time="1.33"/>
    <timepoint id="T2" time="2.66"/>
    <timepoint id="T3" time="4.0"/>
  </timeline>
  <tier id="TIE0" speaker="MAX" category="v">
    <event start="T0" end="T1">You always </event>
    <event start="T1" end="T2">interrupt me. </event>
  </tier>
  <tier id="TIE1" speaker="MAX" category="nv">
    <event start="T0" end="T2">gesticulates</event>
  </tier>
  <tier id="TIE2" speaker="TOM" category="v">
    <event start="T1" end="T2">No I</event>
    <event start="T2" end="T3">don't. </event>
  </tier>
</basic-transcription>
```

EXMARaLDA Basic Transcription:

- (Flat) hierarchy of events in tiers
- Use of ID and IDREFS to encode temporal relations
- No additional markup, no „deep“ semantics

Data formats

| | 0 [0] | 1 [1.3] | 2 [2.6] | 3 [4] |
|-----------------|---------------------|---------------|---------|-------|
| Max [v] | You always | interrupt me. | | |
| Max [nv] | <i>gesticulates</i> | | | |
| Tom [v] | | No I | don't. | |

```

<basic-transcription>
  <timeline>
    <timeline> <ANNOTATION_DOCUMENT>
      <timep <TIME_ORDER>
      <timep <TIME_SLOT TIME_SLOT_ID="T0" TIME_VALUE="0"/>
      <timep <TIME_SLOT TIME_SLOT_ID="T1" TIME_VALUE="1333"/>
      <timep <TIME_SLOT TIME_SLOT_ID="T2" TIME_VALUE="2667"/>
      <timep <TIME_SLOT TIME_SLOT_ID="T3" TIME_VALUE="4000"/>
    </timelin </TIME_ORDER>
    <tier id=" <TIER PARTICIPANT="MAX" LINGUISTIC_TYPE_REF="v">
      <annotation>
        <ALIGNABLE_ANNOTATION TIME_SLOT_REF1="T0" TIME_SLOT_REF2="T1">
          <ANNOTATION_VALUE>You always </ANNOTATION_VALUE>
        </ALIGNABLE_ANNOTATION>
      </annotation>
    <annotation>
        <ALIGNABLE_ANNOTATION TIME_SLOT_REF1="T1" TIME_SLOT_REF2="T2">
          <ANNOTATION_VALUE>interrupt me. </ANNOTATION_VALUE>
        </ALIGNABLE_ANNOTATION>
      </annotation>
    </annotation>
    </TIER>
    <annotation>
        <TIER PARTICIPANT="MAX" LINGUISTIC_TYPE_REF="nv">
          <ANNOTATION>
            <ALIGNABLE_ANNOTATION TIME_SLOT_REF1="T0" TIME_SLOT_REF2="T2">
              <ANNOTATION_VALUE>gesticulates</ANNOTATION_VALUE>
            </ALIGNABLE_ANNOTATION>
          </ANNOTATION>
        </annotation>
      </annotation>
    </TIER>
  </basic-trans
  
```

- EXMARaLDA
- ELAN

Data formats

| | 0 [0] | 1 [1.3] | 2 [2.6] | 3 [4] |
|-----------------|---------------------|---------------|---------|-------|
| Max [v] | You always | interrupt me. | | |
| Max [nv] | <i>gesticulates</i> | | | |
| Tom [v] | | No I | don't. | |

```

<basic-transcription>
  <timeline>
    <timeorder> <TIME_ORDER> file type = "ooTextFile"
    <timeobject> <TIME_SLOT TIME> object class = "TextGrid"
    <timeorder> <TIME_SLOT TIME> xmin = 0.0
    <timeorder> <TIME_SLOT TIME> xmax = 4.0
  </timelin
  <tier id="1" <TIER PARTICIPANT> tiers? <exists>
    <event <ALIGNABLE_AN item []:
      <event <ANNOTATION item [1]:
        </tier> </ALIGNABLE_AI class = "IntervalTier"
        <tier id="2" <ANNOTATION name = "Max [v]"
          <event <ALIGNABLE_AI xmin = 0.0
            <ANNOTATION xmax = 4.0
              </tier> </ALIGNABLE_AI intervals: size = 3
                <event <ALIGNABLE_AI intervals [1]:
                  <ANNOTATION xmin = 0.0
                    </tier> </ALIGNABLE_AI xmax = 1.33
                      <TIER> text = "You always "
                        <event <ALIGNABLE_AI intervals [2]:
                          <ANNOTATION xmin = 1.33
                            </tier> </ALIGNABLE_AI xmax = 2.66
                              <TIER> text = "interrupt me. "
                                </basic-trans
                                  </ALIGNABLE_AI
                                    </ANNOTATION>
                                      </TIER>
                                        item [2]:
                                          class = "IntervalTier"
  
```

- EXMARaLDA
- ELAN
- Praat

Data formats

- Schmidt, Loehr et al. (2008): An exchange format for multimodal annotations.
 - XML format
 - based on Atlas Interchange Format
 - allowing lossless exchange of common denominator information between seven tools with STMT data models
 - import/export routines implemented in the tools

World 2: CA transcription systems

- HIAT (Ehlich/Rehbein 1976, Rehbein 2004)
- DIDA (Klein/Schütte 2001)
- GAT (Selting et al. 1998, 2008)
- DT (DuBois 1992)
- CA transcription
 - Original variant (Sacks/Schegloff/Jefferson 1978)
 - CA-Bank/CHAT variant (MacWhinney 2000)
 - Danish variant (Wagner/MacWhinney 2008)

World 2: CA transcription systems

- Conversation Analysis, Discourse Analysis, Functional Pragmatics, Interactional Linguistics
- Used in the study of social interaction, with linguistic and/or sociological research interest
- Paper and pencil tradition
- Text editors, word processors, specialized applications (syncWriter, HIAT-DOS), STMT editors
- Many commonalities, but no widely accepted agreement on data processing standards

Jefferson transcription

1. CPO: Is that o[-ka:y.]
2. Caller: [Fine.] =yes.
3. [°that's fine.°]
4. CPO: [˘Brilliant] okay,
5. Caller: °.Hh° (0.2) u:m (0.1) >I'm sorry
6. I'm a little bit< emo:~tional
7. to-d[ay~ .hih]
8. CPO: [Tch Oh::] go:sh I'm so:rry,
9. Caller: ~I've got a little four year old grandson,~
10. [huh]
11. CPO: [Yea]h:,
12. (0.3)
13. Caller: ~My son w(h)as s(h)ixtee:n~ (0.5) er fif-teen when

HIAT transcription

DS [sup]

faster

DS [v]

Okay. D'accord d'accord.

DS [en]

Okay. Agreed, agreed.

DS [nv]

right hand raised

FB [v]

Alors ça dépend ((cough)) un petit peu.

FB [en]

That depends then, a little bit.

FB [pho]

[étipø:]

Common ground

- Speaker assignment
- Temporal structure
 - sequence
 - overlap
- Basic building blocks
 - words
 - pauses
 - non-verbal, para-verbal, etc. entities
 - prosodic features: stress, lengthening, etc.
 - uncertain transcription, unintelligible passages etc.

Surface variation

| | |
|----------------------|----------------|
| he makes • a pause | (HIAT) |
| he makes (.) a pause | (GAT) |
| he makes * a pause | (DIDA) |
| he makes .. a pause | (DT) |
| he makes # a pause | (CHAT) |
| he makes / a pause | (Czech system) |

Model variation

- Multi-word entities („sentence equivalents“)
 - utterance (HIAT)
 - intonation phrase (GAT)
 - intonation unit (DT)
 - turn construction unit (CA)
 - none (DIDA)
- Accent, Stress
 - according to intonation contours (GAT)
 - according to deviations from the standard (HIAT)

World 3: TEI

- TEI Guidelines, Chapter 8 “Transcriptions of Speech”
 - transcriptions = “texts of a special kind”
 - elements unique to spoken language transcription

<u>, <pause>, <vocal>, <kinesic>, <incident>, <shift>

- Other relevant parts of the guidelines
 - Chapter 2 “TEI Header”
 - Chapter 13 “Names, Dates, People, and Places”
 - Chapter 16 “Linking, Segmentation, and Alignment”
 - Chapter 17 “Simple Analytic Mechanisms”

<person>, <when>, <anchor>, <w>, <seg>

World 3: TEI

I don't have utterances as units in my descriptive system, and indeed many transcription systems don't have rigorously defined utterances. [...] I will personally **not accept TEI** if it requires me to have an utterance under [this] definition [...], because that is far too rigorous for me and it doesn't represent the world, as far as I'm concerned.

[Sinclair 1995]

The TEI guidelines for 'Transcriptions of Speech' [...] offer access to a very broad range of representational techniques drawn from other aspects of the TEI specification. The TEI report sketches or alludes to a correspondingly wide range of possible issues in speech annotation. All of these seem to be encompassed within [the AG] framework, but it does not seem appropriate to speculate at much greater length about this, given that this portion of the TEI guidelines **does not seem to have been used** in any published transcriptions to date.

[Bird/Lieberman 2001]

Summary: Three Worlds

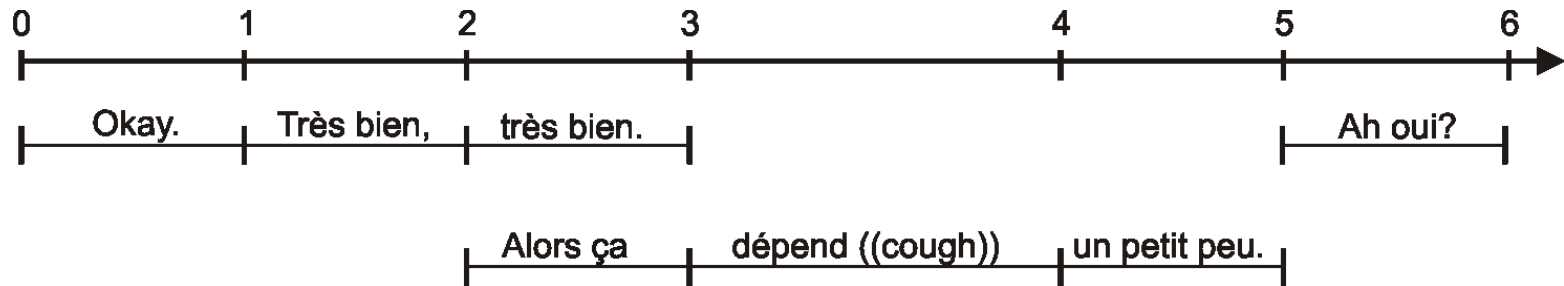
- Transcription tools
 - good acceptance
 - good interoperability
 - no OHCO paradigm
 - no „deep“ semantics
- Transcription conventions
 - (almost) no (inter)operability
 - surface variation, model variation
 - „deep“ semantics
- TEI guidelines
 - low acceptance
 - potential to improve interoperability, harmonize surface variation, respect model variation
 - OHCO paradigm
 - „deep“ semantics



Summary: Three Worlds

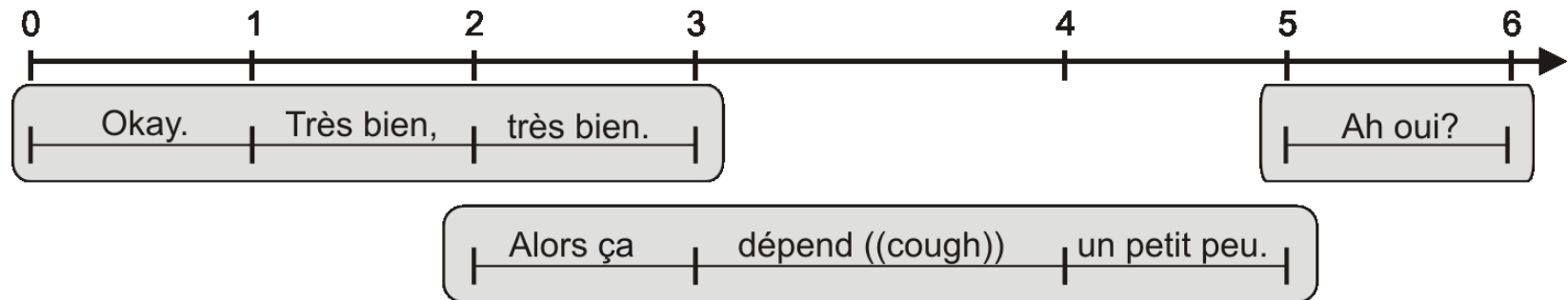
- TEI as the missing link
 - Transform STMT to a TEI conformant representation
 - Transform text representations to a TEI conformant representation
- Open questions
 - How to get from STMT to OHCO?
 - How to get from idiosyncratic text representations to standardized markup?
 - How to operationalize existing TEI guidelines for that purpose?

STMT to OHCO transformation



- Segment chain = any temporally connected chain of annotations within one tier

STMT to OHCO transformation



- Segment chain = any temporally connected chain of annotations within one tier
- Assumption: all other hierarchical structure beneath the level of segment chains
- Correspondence: segment chain \leftrightarrow $\langle u \rangle$

STMT to OHCO transformation

```
<STMT>
  <tier speaker="A">
    <annotation start="T0" end="T1">Okay.</annotation>
    <annotation start="T1" end="T2">Très bien, </annotation>
    <annotation start="T2" end="T3">très bien.</annotation>
    <annotation start="T5" end="T6">Ah oui.</annotation>
  </tier>
  <tier speaker="B">
    <annotation start="T2" end="T3">Alors ca </annotation>
    <annotation start="T3" end="T4">dépend ((cough)) </annotation>
    <annotation start="T4" end="T5">un petit peu.</annotation>
  </tier>
</STMT>
```

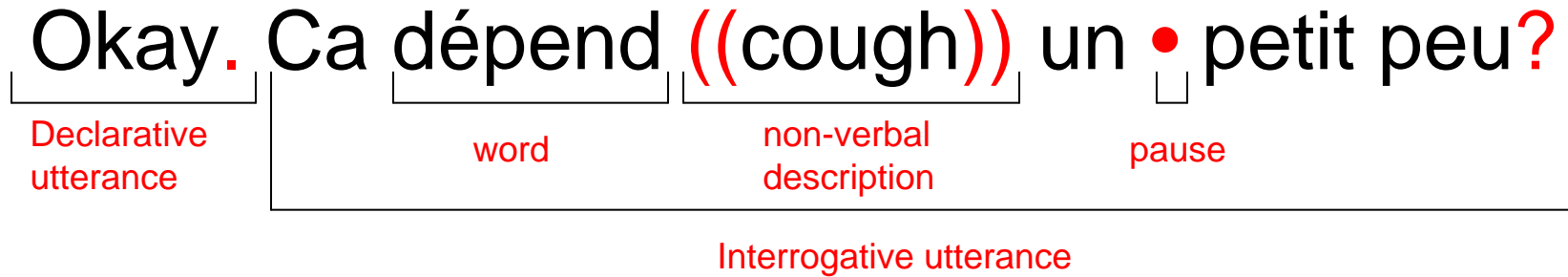
```
<OHCO>
  <u who="A" start="T0" end="T3">
    Okay.
    <anchor synch="T1"/>
    Très bien,
    <anchor synch="T2"/>
    très bien.
  </u>
  <u who="B" start="T2" end="T5">
    Alors ca
    <anchor synch="T3"/>
    dépend ((cough))
    <anchor synch="T4"/>
    un petit peu.
  </u>
  <u who="A" start="T5" end="T6">
    Ah oui.
  </u>
</OHCO>
```

Text to XML transformation

Okay. Ca dépend ((cough)) un • petit peu?

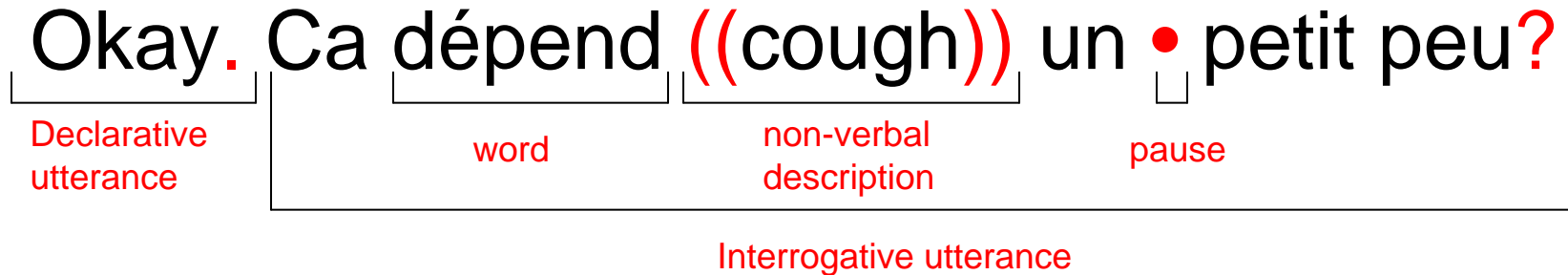
- This *is* text + markup!

Text to XML transformation



- This *is* text + markup!

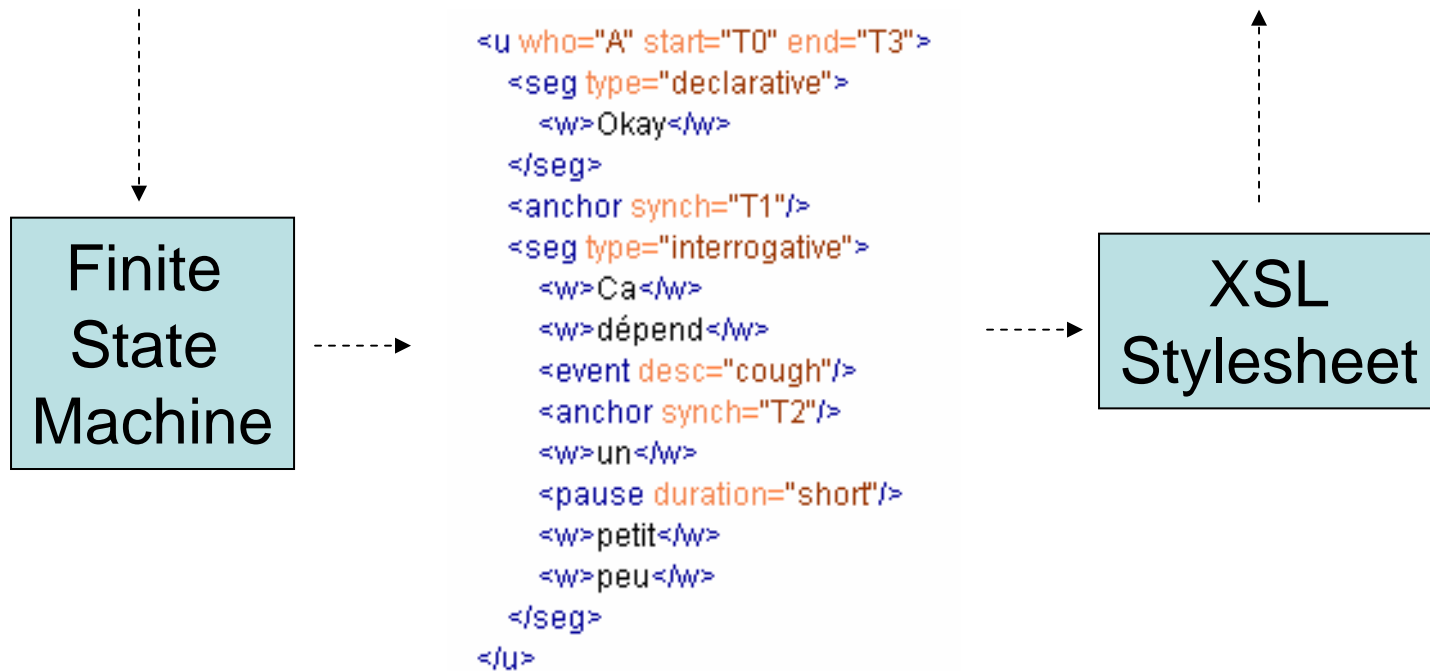
Text to XML transformation



- This *is* text + markup!
- Use of punctuation, boundary markers etc. is 100% predictable (→ conventions)
- (Simple) formal grammar for each transcription system for translation of implicit markup to XML markup

Text to XML transformation

Okay. Ca dépend ((cough)) un • petit peu?



XML to XML transformation

```
<u who="A" start="T0" end="T3">  
  <seg type="declarative">  
    <w>Okay</w>  
  </seg>  
  <anchor synch="T1"/>  
  <seg type="interrogative">  
    <w>Ca</w>  
    <w>dépend</w>  
    <event desc="cough"/>  
    <anchor synch="T2"/>  
    <w>un</w>  
    <pause duration="short"/>  
    <w>petit</w>  
    <w>peu</w>  
  </seg>  
</u>
```

XSL
Stylesheet

Finite
State
Machine

```
<u who="A" start="T0" end="T3">  
  Okay.  
  <anchor synch="T1"/>  
  Ca dépend ((cough))  
  <anchor synch="T2"/>  
  un • petit peu?  
</u>
```

Three worlds

World 3: TEI & OHCO

World 2: Transcription Systems

```
<u who="A" start="T0" end="T3">  
  <seg type="declarative">  
    <w>Okay</w>  
  </seg>  
  <anchor synch="T1"/>  
  <seg type="interrogative">  
    <w>Ca</w>  
    <w>dépend</w>  
    <event desc="cough"/>  
    <anchor synch="T2"/>  
    <w>un</w>  
    <pause duration="short"/>  
    <w>petit</w>  
    <w>peu</w>  
  </seg>  
</u>
```

XSL
Stylesheet

```
<u who="A" start="T0" end="T3">  
  Okay.  
  <anchor synch="T1"/>  
  Ca dépend ((cough))  
  <anchor synch="T2"/>  
  un • petit peu?  
</u>
```

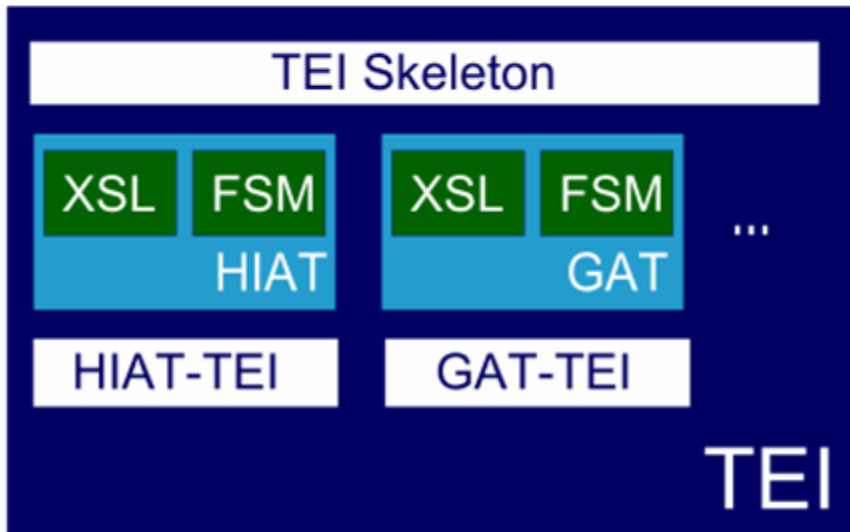
Finite
State
Machine

World 1: Transcription Tools





(Generic) STMT data models



Generic OHCO data model

Specific OHCO data model

What is wrong with the TEI guidelines?

- What is missing?
 - nothing, really...
- What needs to be changed?
 - <u>: name and definition
- What needs to be clarified?
 - one way of encoding temporal structure
 - define temporal „integrity“

Implementation 1

- EXMARaLDA
 - Proof of concept implementation: Thomas Schmidt (2005). Time-based data models and the TEI guidelines for transcriptions of speech. Working papers in Multilingualism. Hamburg.
 - TEI-Versions of all HIAT-Corpora at the Research Centre on Multilingualism
 - Interpreted Doctor-Patient-Communication
 - Turkish-German bilingualism
 - Simultaneous Interpreting Portuguese/German
 - Semicommunication in Scandinavia

Implementation 2

- FOLKER
 - FOLK = Forschungs- und Lehrkorpus Gesprochenes Deutsch
 - Under construction at IDS Mannheim
 - Based on the GAT transcription system

Folker [T:\TP-Z2\GAT\Demokorpus\HART_ABER_FAIR\HART_ABER_FAIR_02_APRIL_2008...]

Datei Bearbeiten Ansicht Transkription Hilfe

[00:08.63 00:08.63 00:13.36]

00:00 00:04 00:08 00:12 00:16 00:20 00:24 00:28 00:32 00:36 00:40 00:44 00:48 00:52 00:56 01:00

00:00 00:04 00:08 00:12 00:16 00:20 00:24 00:28 00:32 00:36 00:40 00:44 00:48 00:52 00:56 01:00

Ereignisse Partitur Beiträge

| Start | Ende | Spr... | Transkriptionstext | ... |
|-----------|-----------|--------|--|-----|
| 00:00.... | 00:03.... | FP | herr pofalla wissen sie was im moment n liter diesel kostet | ✓ |
| 00:03.... | 00:04.... | RP | ja | ✓ |
| 00:04.... | 00:04.... | FP | wieviel | ✓ |
| 00:04.... | 00:08.... | RP | der liter diesel äh kostet derzeit um die eins dreißig | ✓ |
| 00:08.... | 00:13.... | FP | eins dreiunddreißig sinds so ist der bundesdurchschnitt sie ham viell... | ✓ |
| 00:13.... | 00:15.0 | FP | was macht das mit ihrem budget | ✓ |
| 00:15.0 | 00:22.3 | RP | ja meine frau fährt äh äh diesel und äh insofern hat sich die dieselpri... | ✓ |
| 00:22.3 | 00:23.... | NN | (--) | ✓ |
| 00:23.... | 00:27.... | FP | wie das frag ich lieber frau kastenhuber wie wirkt sich die dieselprei... | ✓ |
| 00:27.... | 00:30.... | FK | also bei uns in bayern wars heut früh eins sechsunddreißig ((atmet ... | ✓ |
| 00:30.... | 00:34.... | FK | ähm und ich werd einfach weiterhin nicht zum friseur gehen die näd... | ✓ |

Transkription T:\TP-Z2\GAT\Demokorpus\HART_ABER_FAIR\HART_ABER_FAIR_02_APRIL_2008_exmaralda4.xml geöffnet.

Folker [T:\VTP-Z2\GAT\Demokorpus\HART_ABER_FAIR\HART_ABER_FAIR_02_APRIL_2008...

Datei Bearbeiten Ansicht Transkription Hilfe

[00:08.63 00:16.44 00:15.0]

00:00 00:04 00:08 00:12 00:16 00:20 00:24 00:28 00:32 00:36 00:40 00:44 00:48 00:52 00:56 01:00

00:00. 00:03. 00:04. 00:04. FP tig 00:04.29 00:04.95 FP wieviel ✓ 00:08. RP 00:04.95 00:08.63 RP der liter diesel äh kostet derzeit um die eins dreißig ✓ 00:13. NN 00:08.63 00:15.0 FP eins dreiunddreißig sinds so ist der bundesdurchschnitt sie ham vielleicht ne günstige tankstelle in berlin was macht das mit ihrem budget ✓ 00:22. FK 00:15.0 00:22.3 RP ja meine frau fährt äh äh diesel und äh insofern hat sich die dieselpreisentwicklung natürlich auch auf äh uns ausgewirkt ✓ 00:27. EK 00:30. X 00:30. HB eins dreiunddreißig sinds so ist der bundesdurchschnitt sie ham vielleicht ne günstige tankstelle in berlin 00:13.36 was macht das mit ihrem budget

Playback stopped.

Summary

- Making TEI work for spoken language transcription
- TEI as a mediator between tools and transcription conventions
- Two simple ideas for transformation
 - Segment chains: layers → hierarchy
 - Parsing: Implicit markup → explicit markup
- Two working implementations