Anvil

A Generic Annotation Tool for Multimodal Dialogue

Michael Kipp
University of the Saarland & DFKI
Germany

kipp@dfki.de
Annotation of Video and Spoken Language
Motivation

- Find new **gestures** and their semantic and temporal relation to **speech**
- Analysis of empirical data on video files
- **Notepad** for parallel events (video synchronized):
  - speech
  - head movements
  - gestures
- Parallel information layers = **Tracks**
- Notes for one video = **Annotation**
Annotation

Tracks

<table>
<thead>
<tr>
<th>words</th>
<th>Track Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welcome</td>
<td>to</td>
</tr>
<tr>
<td>Anvil</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>head movements</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>gestures</th>
</tr>
</thead>
</table>

Elements within one track *must not* overlap!
Annotation

Tracks

- words
- head movements
- gestures

Track Elements

- token = "Welcome"
- gesture class = deictic
- repetition = false
- certainty = 4

- gesture class: ValueSet {deictic, emblem, iconic}
- repetition: Boolean
- certainty: Number(1,4)
Track Hierarchies

primary and secondary tracks

words

sentences

gesture phases

gestures

gesture cluster

Welcome to Anvil
• Track hierarchies express a very specific kind of relation
• How to code general interdependencies between elements in the same or different tracks?
• Put logical links into a new type of attribute
Links

words
Welcome to Anvil

token = "Anvil"

category = "deictic"
lexical affiliate = {}
Reciprocal Links

words
- Welcome
- to
- Anvil

token = „Anvil“
backlink = {
}

category = „deictic“
lexical affiliate = {
}

head movements

gestures

lexical affiliate: ReciprocalLink(backlink)
Specification

• Specify
  – Tracks  \textit{name + type + reference track}
  – For each track: Attributes  \textit{name + type}
    • String, Boolean, Number(x, y)
    • ValueSet \{set of user-defined labels\}
    • MultiLink, ReciprocalLink(A)

• This is the \textit{annotation scheme}  
  \textit{= specification file (XML)}
Project Tool

• Anvil only good for **one annotation** at a time
• A *project* is a **set of annotations** with equal specification (same scheme)
• What can you do with projects?
  – **overview** (coder, date, progress)
  – **search & find** across annotations
  – **export** found units for analysis
Project Tool

- search engine
- annotation files
- annotation info:
  - coder
  - last modification time
  - summary/notes
  - each track’s size
Import/Export

- Anvil is not suited for speech transcription or statistical analysis
- **Import** PRAAT tier to Anvil track
- **Import** XWAVES speech transcription
- **Export** single track of an annotation as ASCII text
- **Export** single track of a project to text tables
- **Import/Export** by working with Anvil XML annotation files (e.g. using XSL)
Extending Anvil using Plug-ins

- File reader/writers/converters (other formats!)
- Different views on the annotation: tables, matrices, trees, 3D graphics, etc.
- Analysis modules
Gesture Analysis Plug-in

- single gesture
- list of lexical affiliates
- handedness
- relative timing

- emblem.attention (34)
  - avg. num of phases: 2.7058823
  - num of single strokes: 10
  - num of multi-strokes: 20
  - num of indep. holds: 1
  - RH = 11
  - LH = 13
  - 2H = 2
  - average init/span [sSpeech-sDest] timing (1): -0.31 sec ± 0.32 sec
  - average direct/indir [sSpeech-sDest] timing (13): 0.05 sec ± 0.26 sec

Distribution of connected argument elements:
- toumin (19)
- claim: 9 (47.4%)
- claim-preparation: 4 (21.1%)
- data: 3 (15.0%)
- qualifier-certain: 1 (5.3%)
Technical Data

- Developed on PentiumIII PC 500MHz 256MB
- Platform-independent (Java + JMF)
- XML-based
- Supports QuickTime and AVI
- Small (1.2MB) and easy to install
- Documented 43 page manual + papers + demo files

FREEWARE
Circulation of Anvil

• Papers + presentations at
  – Agents 2001 (Workshop)
  – Oralité et Gestualité (Orage) 2001
  – EuroSpeech 2001
• Part of ISLE (formerly EAGLES) survey
• Anvil mailing list: 80 entries
  
  
  (HCI, Ling., Phonetics, CA, Psycholing., Ethology, Animation, Tools/Corpora - Research & Teaching)
• Model function for NITE, LDC, Observer and TalkBank project
Summary

• Anvil is a free video annotation tool for
  – coding single annotations
  – browsing multiple annotations
• Clean separation of tool and coding scheme
• Leaves special tasks to other programs
  *speech transcription, statistical analysis*
• Extensible via plug-in architecture

www.dfki.de/~kipp/anvil