Vertical ShowCase 'Web Application WebMAUS'

Phonetic alignment, conversion and visualisation in a single workflow

Thomas Kisler

Dieter Van Uytvanck
Vertical ShowCase 'Web Application WebMAUS'

Phonetic alignment, conversion and visualisation in a single workflow

Thomas Kisler

Dieter Van Uytvanck
Elements of the CLARIN-D Infrastructure

(Web) Applications
- TCF Visualiz.
- TLA
- WebLicht
- FCS
- WebAnno
- VLO
- Annex
- WebMaus
- PWS

Protocols and Specifications
- OAI-PMH
- PID
- AAI
- SRU/CQL
- ISOCat

Dataformats
- CMDI
- TEI
- TCF

Nine CLARIN-D Data Repositories
Elements of the CLARIN-D Infrastructure

(Web) Applications
- TCF Visualiz.
- TLA
- WebLicht
- FCS
- WebAnno
- VLO
- Annex
- WebMaus
- PWS

Protocols and Specifications
- OAI-PMH
- PID
- AAI
- SRU/CQL
- ISOCat

Dataformats
- CMDI
- TEI
- TCF

Nine CLARIN-D Data Repositories
please call Stella and bring these things from the store: small bag of fresh snow peas, thick slabs of brie, and maybe a small bottle of chile sauce.
Segmentation & Labeling

- Time Consuming
- Error Prone
- Phoneticians / trained personnel
- Interlabeller Agreement
- Tedious Task
WebMAUS – S&L Problems

Segmentation & Labeling

- Time Consuming
- Error Prone
- Phoneticians / trained personnel
- Interlabeller Agreement
- Tedious Task

Solution => Automation
WebMAUS – System Overview

Balloon
Text-to-Phoneme Conversion

Web Services & Interfaces

MAUS
Segmentation & Labeling
WebMAUS – System Overview

- Balloon (Text-to-Phoneme Conversion)
- MAUS (Segmentation & Labeling)
- Web Services & Interfaces

Start
WebMAUS – System Overview

Balloon
Text-to-Phoneme Conversion

MAUS
Segmentation & Labeling

Web Services & Interfaces

Start
Balloon

Text-to-Phoneme Conversion

- **Input:** Plain text: “please”

- **Output:** SAMPA transcribed phonemes (including stress) aka canonical pronunciation: “pli:z”
WebMAUS – System Overview

- Balloon
  - Text-to-Phoneme Conversion

- MAUS
  - Segmentation & Labeling

- Web Services & Interfaces

Start
MAUS - Munich Automatic Segmentation System

- Canonical pronunciation
- Pronunciation hypothesis space (graph)
- Graph with probabilities
  - Decoding with Viterbi
  - Segmentation & Labeling

MAUS Segmentation & Labeling
- Canonical pronunciation

  - Pronunciation hypothesis space (graph)

  - Graph with probabilities
    - Decoding with Viterbi

  - Segmentation & Labeling
WebMAUS – System Overview

Balloon
Text-to-Phoneme Conversion

Web Services & Interfaces

MAUS Segmentation & Labeling

Start

Intervals [16]
\( x_{\text{min}} = 5.520000 \)
\( x_{\text{max}} = 5.600000 \)
• RESTful web services
• Demonstration for web services
• Wide user acceptance
Vertical ShowCase 'Web Application WebMAUS' Thomas Kisler

Phonetic alignment, conversion and visualisation in a single workflow Dieter Van Uytvanck
Elements of the CLARIN-D Infrastructure

(Web) Applications
- TCF Visualiz.
- TLA
- WebLicht
- FCS
- WebAnno
- VLO
- Annex
- WebMaus
- PWS

Protocols and Specifications
- OAI-PMH
- PID
- AAI
- SRU/CQL
- ISOCat

Dataformats
- CMDI
- TEI
- TCF

Nine CLARIN-D Data Repositories
Elements of the CLARIN-D Infrastructure

(Web) Applications
- TCF Visualiz.
- TLA
- WebLicht
- FCS
- WebAnno
- VLO
- Annex
- WebMaus
- PWS

Protocols and Specifications
- OAI-PMH
- PID
- AAI
- SRU/CQL
- ISOCat

Dataformats
- CMDI
- TEI
- TCF

Nine CLARIN-D Data Repositories
Steps

• User uploads
  • Text: (orthographic) transcription
  • Sound fragment: WAV file
• Federated Login to WebLicht
• Creation of web service workflow:
  • phonetic analysis + alignment (BAS)
  • conversion Praat > ELAN (HZSK)
• Visualisation with Annex (MPI), TextGrid viewer (BAS)
• Post-processing with Praat, ELAN, EXMARaLDA, …
Important aspects

- **Useful** for researcher:
  - saves **time**
  - high **quality** annotations > more detailed analysis
  - **freedom** to use preferred tools

- **Easy** to use
  - simple login procedure
  - click-and-go

- **Fully distributed:**
  - select modules à la carte
  - but transparent to user