

Florian Schiel

# CLARIN: Prosodic/Gesture Corpora

Existing Resources,  
Best Practice and Formats



## CLARIN : European Infrastructure Initiative for Speech and Language Resources

Focus on linguistic resources and users in humanities

National Centers and cross-national infrastructures

Examples:

- Repositories
- Virtual Language Observatory
- Federated Content Search



# Context PAGE: Prosodic and Gestural Entrainment

- (1) Existing re-usable Speech Corpora?
- (2) Existing re-usable Gesture Corpora?
- (3) Existing best practice, inter-operable formats, infrastructures?

# (1) Existing re-usable speech corpora

Looking for:

- Spontaneous/semi-spontaneous speech
- Two speakers interacting
- Orthographic transcription
- Prosodic annotation
- Availability

<http://www.bas.uni-muenchen.de/forschung/Bas/BasFormatseng.html>

# (1) Existing re-usable speech corpora

**Problem: very few prosodically annotated dialogue corpora**

- **Verbmobil1/2** (German, English, Japanese): 2400 speakers  
two speakers, domain: business meeting scheduling  
VM1 : GToBI in 907 dialogue turns (PRB)  
VM2 : GToBI in 3584 dialogue turns (LBP/LBG),  
Syntactic-Prosodic word tagging in 12358 turns (PRO)

**available**

<http://www.bas.uni-muenchen.de/forschung/Bas/BasFormatseng.html>

- **Verbmobil1/2** - Example

PRB:	5824	-1	TON: L+H*; FUN: PA
PRB:	9895	-1	BRE: B3; TON: L-L%
PRB:	32381	-1	TON: L+H*; FUN: EK
PRB:	43983	-1	BRE: B3; TON: L-H%
PRB:	58397	-1	TON: H*; FUN: PA
PRB:	65185	-1	BRE: B9; TON: L-
PRB:	69565	-1	TON: H*; FUN: NA
PRB:	81779	-1	BRE: B3; TON: !H-H%
PRB:	93596	-1	TON: L+H*; FUN: PA
PRB:	102053	-1	BRE: B2; TON: H-
PRB:	104956	-1	TON: L+H*; FUN: PA
PRB:	112280	-1	TON: !H*; FUN: NA
PRB:	118112	-1	TON: !H*; FUN: PA
PRB:	122959	-1	BRE: B3; TON: L-L%



# (1) Existing re-usable speech corpora

- **Smartkom** (German): 450 speakers

human-machine (WOZ), domain: information retrieval  
segmentation of prosodic events/pauses 6257 turns (USP)

[PAUSE\_PHRASE] [PAUSE\_WORD] [PAUSE\_SYLL] [LENGTH\_SYLL] [EMPHASIS]  
[STRONG\_EMPH] [CLEAR\_ART] [HYPER\_ART] [LAUGHTER]

**available**

<http://www.bas.uni-muenchen.de/forschung/Bas/BasFormatseng.html#USP>

# (1) Other speech corpora with prosodic annotation

- only read speech
- GToBI or similar
- **BAS PD2, BAS BITS, BAS SI1000P**

(see <http://www.bas.uni-muenchen.de/Bas>)



## (2) Existing gesture corpora in CLARIN

Looking for:

- Speech + video (+ motion capture)
- Spontaneous/semi-spontaneous speech
- Two or more actors interacting
- Orthographic transcription
- Gesture annotation

<http://www.bas.uni-muenchen.de/forschung/Bas/BasFormatseng.html#GES>

## (2) Existing gesture corpora in CLARIN

General problem: very specific designs depending on specific research questions

- **Smartkom** (German): 450 actors

human - machine, mostly index gestures on interaction surface, intentional-morphological tagging and segmentation (GES)

available

example

<http://www.bas.uni-muenchen.de/forschung/Bas/BasFormatseng.html#GES>

## (2) Existing gesture corpora

- **SaGA (Speech and Gesture Alignment)**: 50 actors  
human-human (1 actor recorded), descriptive gestures,  
morphological tagging, alignment to linguistic transcript,  
5000 gestures, 50000 words

available ETA 2014

*example*

## (2) Existing gesture corpora in CLARIN

- **MoCap (Motion Capture Corpus)** : 20 actors  
human-human (1 actor recorded), descriptive  
quantitative gestures, morphological tagging  
20 dialogues, 3.5h recording time

available ETA 2014

*example*

## (2) Existing gesture corpora in CLARIN

*Other CLARIN resources (**not available**)*

- **'Gesture Project'** MPI Nijmegen (American English)  
*story telling elicited by cartoon, monologue*  
*annotation: unknown*
- **'Co-Speech Gestures'** MPI Nijmegen (Turkish, German)  
*543 annotated videos, monologue*  
*annotation: unknown*

## (2) Existing gesture corpora in CLARIN

*Other CLARIN resources (**available**)*

- 'Caused Motion Event Representation Project'  
MPI Nijmegen (Turkish)  
*elicited motion gestures produced by kids, monologue*  
*annotation: unknown*  
**available**
- 'Picture Pointing Experiment' MPI Nijmegen (Dutch)  
**tangram elicited** gestures, 26 dialogues, 52 actors  
*annotation: unknown, often only 1 actor visible*  
**available** **example**



(1 + 2 ) in context of entrainment:

- possibly usage of Verbmobil1/2
- probably no suitable corpora for gestures

### (3) Best practice, interoperable formats

Recordings, formats, metadata, legal advice:

- **CLARIN-D User Guide**

<http://de.clarin.eu/en/language-resources/userguide.html>

meta data standards, Persistent Identifier (PID), long term storage, indexing in meta-infrastructures, tools

- **BAS White Book: Speech Corpora**

<http://www.bas.uni-muenchen.de/forschung/BITS/TP1/Cookbook/>

hands-on guidelines for recording, tools



## (3) Best practise, interoperable formats

### Annotating video/audio

- **FiESTA** (in progress)

Peter Menke, John McCrae, Philipp Cimiano: Releasing multimodal data as Linguistic Linked Open Data: An experience report. University of Bielefeld, in print.

data model for multimodal data streams,  
annotations, structure, XML file format,  
multiple time lines/spatial scales

## Summary

- very few suitable corpora for prosody
- no suitable corpora for gesture
- CLARIN best practice, BAS guidelines
- CLARIN repositories for long-term storage, metadata/PID support