

Challenges and Developments in Preserving and Publishing of Large Audio/Video Data

Some projects and perspectives from the Lund University Humanities Lab

JOHAN FRID & JENS LARSSON

HUMANISTLABORATORIET



LUND
UNIVERSITY

an autonomous department
a multi-user facility for

technology/method
training
interdisciplinary interface

HUMANISTLABORATORIET



SWE-CLARIN

Vi ger dig verktyg för framtidens forskning

Common Language Resources and
Technology Infrastructure



CLARIN
K CENTRE 

PDF: <https://hdl.handle.net/11372/DOC-13>

multimodal and sensor-based methods

Sensor-based technology



eye-tracking



keystroke logging



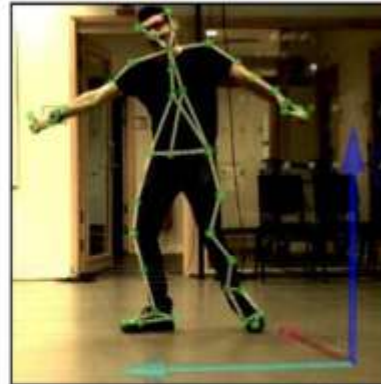
EEG



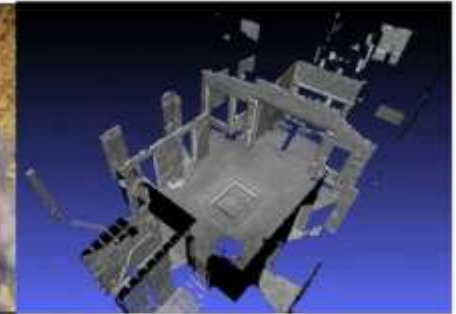
bio pac



articulography



motion tracking



3D scanning

© Marianne Gullberg, 2018



COLLECT
DIGITISE

DOCUMENT
STORE
SHARE

EXPLORE
ENRICH
COLLABORATE

Our archive

LAT (MPI, Nijmegen)

Currently: > 5TB

One large (> 80%) 'customer':

RWAAI (Languages and Cultures in Southeast Asia & India)

Continuous data provider

Many 'small'

Lacola/Diacl (Historical/Typological data, ~250G)

Swedia (Swedish dialects in Sweden and Finland, ~180G)

Data comes in 'spurts'

IMDI Browser

IMDI-Browser

- Semang
 - General information
 - Languages
 - Batek Deq
 - Jahai
 - Lexicon
 - Media
 - Conversation
 - Elicitation
 - Narratives
 - Other activities
 - Places
 - Ritual
 - Species
 - Subsistence activities
 - Blowpipe
 - Blowpipe hunting
 - Blowpipehunting1
 - Blowpipehunting2
 - Blowpipehunting3
 - Blowpipe parts
 - Darts
 - Ipoh poison
 - Quiver
 - Fishing
 - Food preparation
 - Fruit

© Marianne Gullberg, 2018

e.g., audio, video, text, GIS/GPS, census data



Language, cognition
and landscape
(Burenhult)



00:09:10.000 00:09:11.000 00:09:12.000 00:09:13.000 00:09:14.000 00:09:15.000 00:09:16.000 00:09:17.000 00:09:18.000 00:09:19.000 00:09:20.000

with: 0:11
gross: 0:12
branch: 0:13

ton, koo-wr	ban7 loh lak-wor Jahaj	ton, koo-wr 7et 70n	ton
DEM to make stone line+FO	DEM EMP ID-to make stone line+FO+ Jah	DEM REL = 70 FAM to look D	DEM
there, a row of stones	now that's a Jahai stone row	there, which you see there	there

How are people using the data
aka
what has been done with it?

Diacl: lexeme example

Lexeme – Details – kirves

Eurasia / [Finnish](#)

Lexeme

Transcription	kirves
Script	
IPA	
Meaning	axe
Meaning Note	
Grammatical Data	
Note	

[Back](#)

Sources

Source	Taina Cronhamn [Native speaker Taina Cronhamn]
---------------	---

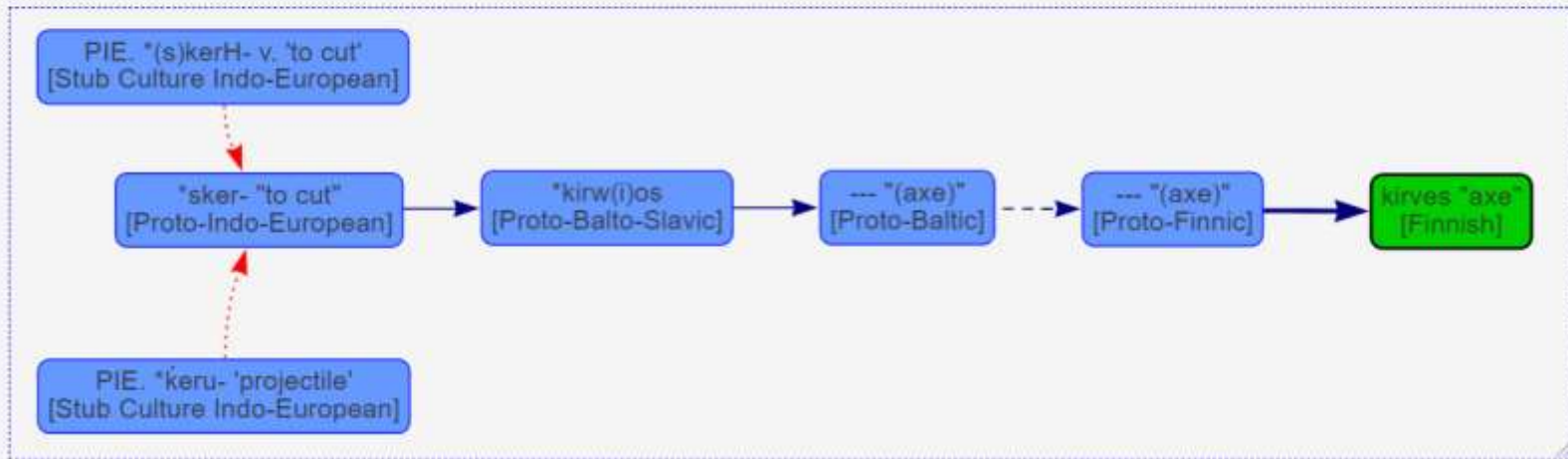
Diacl: ancestry diagram

— Ancestry Diagram

The element in question is marked in green. When the etymological tree is displayed in hierarchical style, the nodes can be repositioned manually only within their level. (Note that the node configuration is not saved in any way.)

Arrange nodes hierarchically.

The graph as a whole can be zoomed in on by scrolling and can be enlarged by dragging the right-lower corner. Left-click on a node or edge to show additional information and links. In some browsers, the graph may be savable as an image when right-clicking on it.



Diacl: geographical context

This map shows the etymological links in a geographic context.



Sweden

Dialect project

Recorded around 2000

Audio only

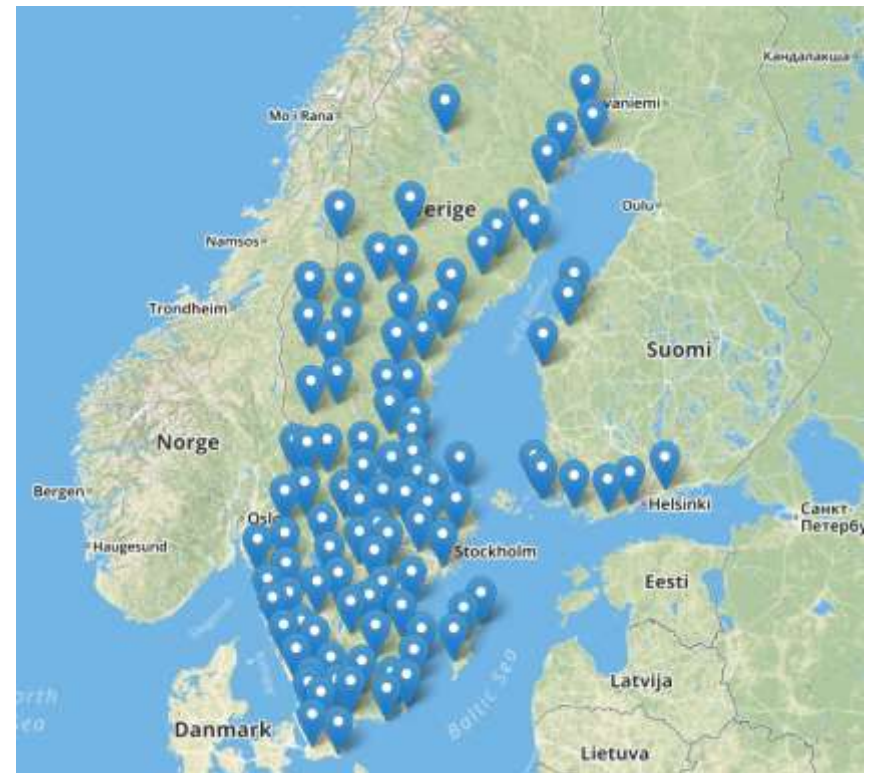
100 villages in Sweden
and Finland

12 speakers/village

Word lists

Interviews

Dialect 'trips'



Augmented data: an archive issue?

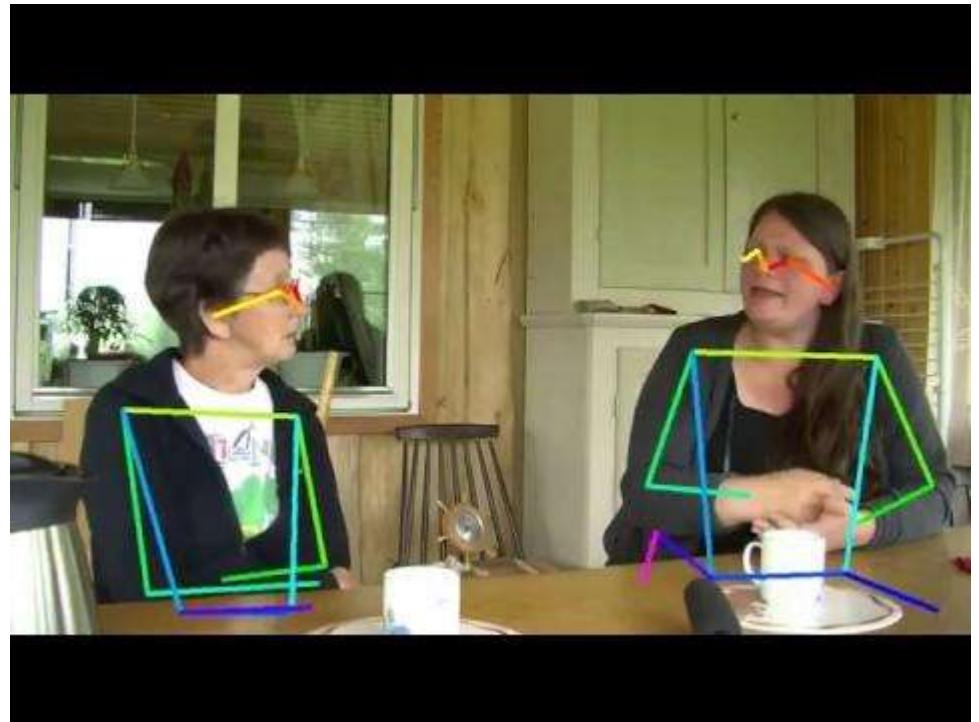
Video of two consultants conversing in Elfdalian, using gestures

any type of 'spatial locationing' in videos and images

pose estimation, geospatial mapping, handwriting recognition etc

adding new layers may be very useful (particularly for annotation), but may duplicate lots of data

linkage to audio, video, annotations and object recognition



Eye-tracked frog stories: rescuing old data

(shaky) video of eye tracking while looking at and describing a picture

gaze behaviour in relation to spoken description

do people look at what they talk about?

famous task, 'frog stories', known to induce 'cultural' differences



Issues

we have the file with the eye tracking data with coordinates etc but these are related to the screen not the video

picture might be retrieved, but not how it was displayed

1. x-y scaling
2. the actual monitor is nowhere to be found

nowadays object detection is easy enough, so we can detect the red circle and probably the screen too

but still need to realign with audio

all this creates new files and new archiving, new PIDs?

layered approach, versioning for video? Git?