
ANVIL:
A Multi-Level Video Coding
and Analysis Tool
Michael Kipp
DFKI, Germany
http://www.anvil-software.de
© R.L. Segal

Overview

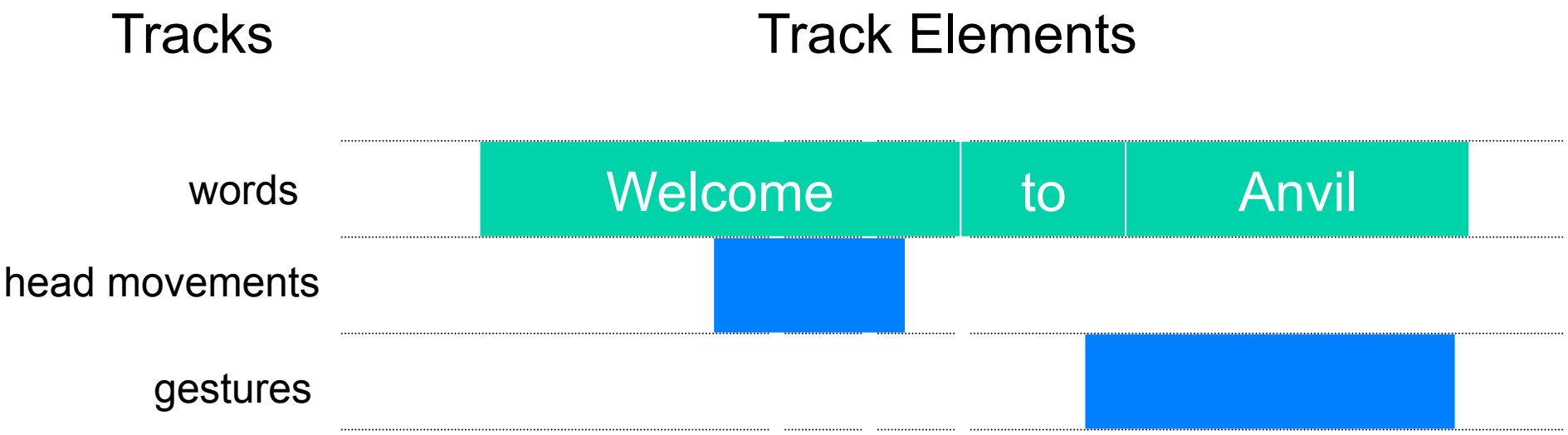
- Introduction to ANVIL
  - who knows it?
  - who has used it?
- Interesting (advanced) features, including
  - Analysis possibilities
  - Working with Motion Capture
- Today:
  - Lots of demos
  - Please ask questions any time...

Introduction to ANVIL

What is ANVIL?

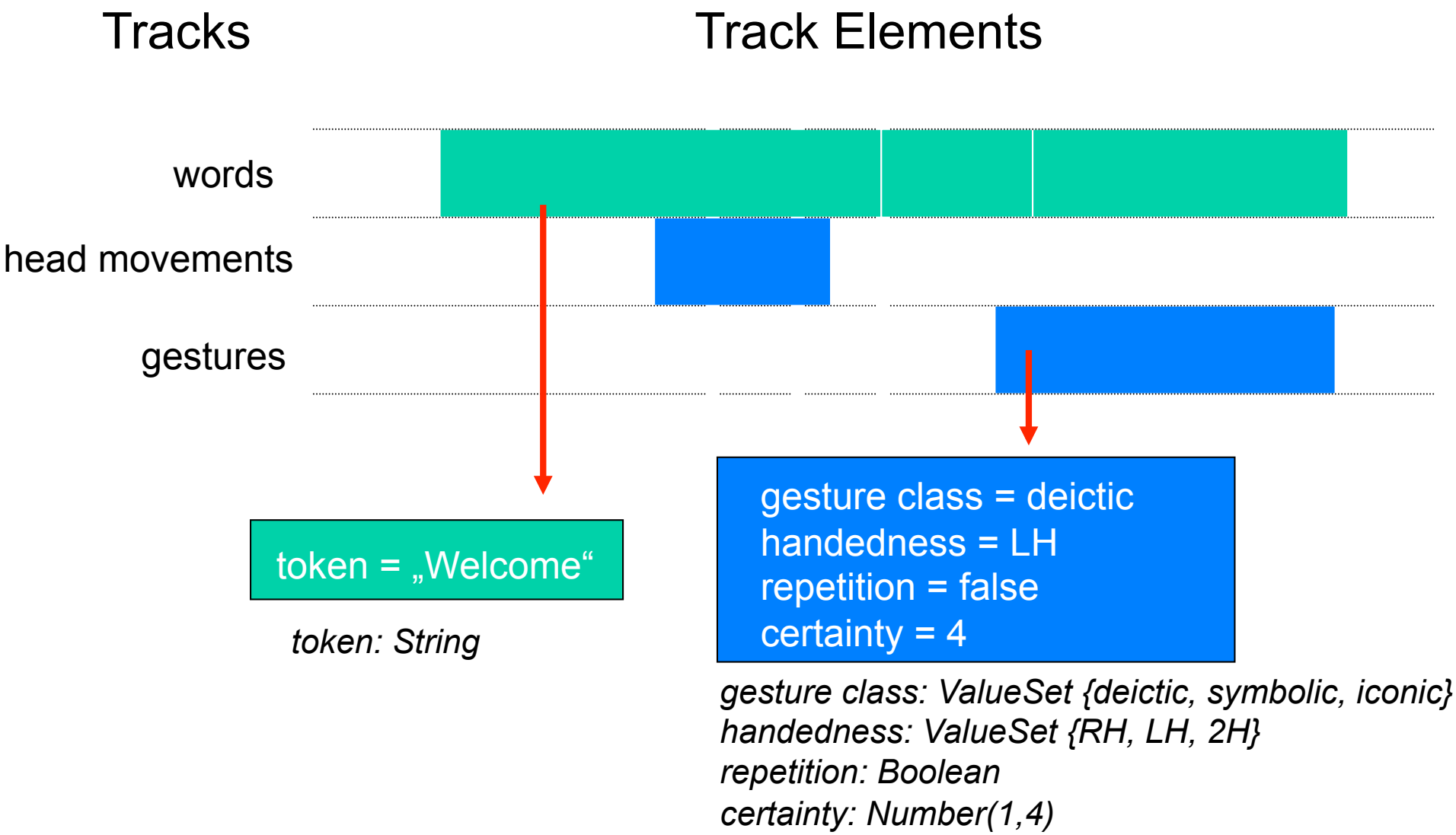
- **Manual annotation** of digital videos
  - „Notepad“ for events in time
  - See temporal relationships between events
- Examples
  - gesture and speech
  - human-computer interaction
  - parent and child interaction
  - dance analysis
  - movie analysis
  - sign language analysis
  - oceanography
  - …
- **Java**, version 5 beta 16
- Originally: **Gesture research** (incl. avatars)
- > 2000 research licenses in 25+ countries. Standard tool in European projects (e.g. HUMAINE), corpus collection initiatives (e.g. LDC, GEMEP, MUMIN, OTIM), research labs (MIT, LIMSI, NII…), individuals

Annotation Elements



Elements within one track cannot overlap!

Elements are Complex



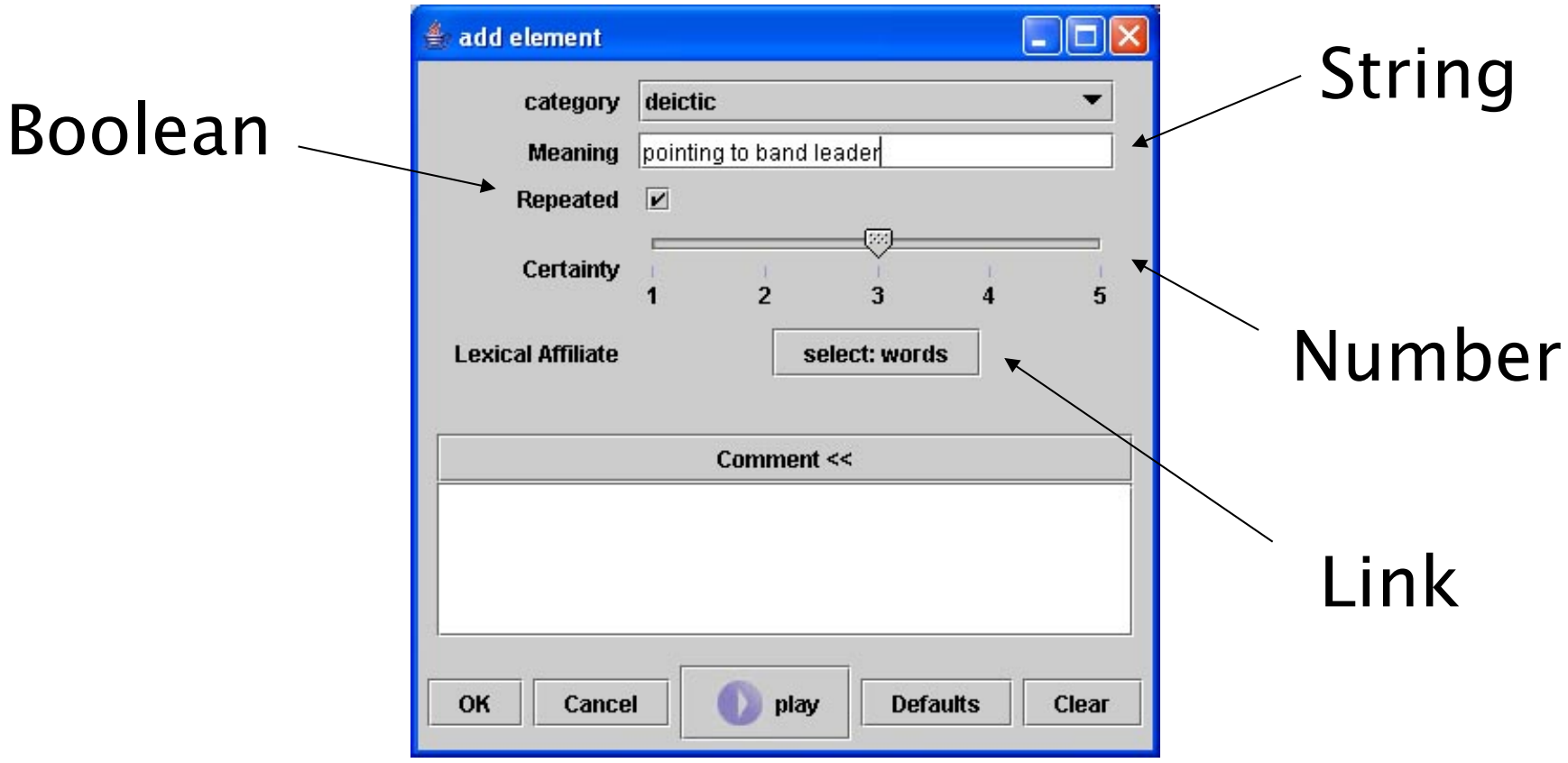
Why multiple, typed attributes?

- Multiple attributes
  - encode many aspects of one „event“
    alternative: use more tracks
- Types
  - efficient input: type-specific interfaces
    - number => number slider
    - set of tokens => menu
    - yes/no => check box
    - alphanumeric string => text field
  - fewer errors
    - input restricted (options, check box etc.)

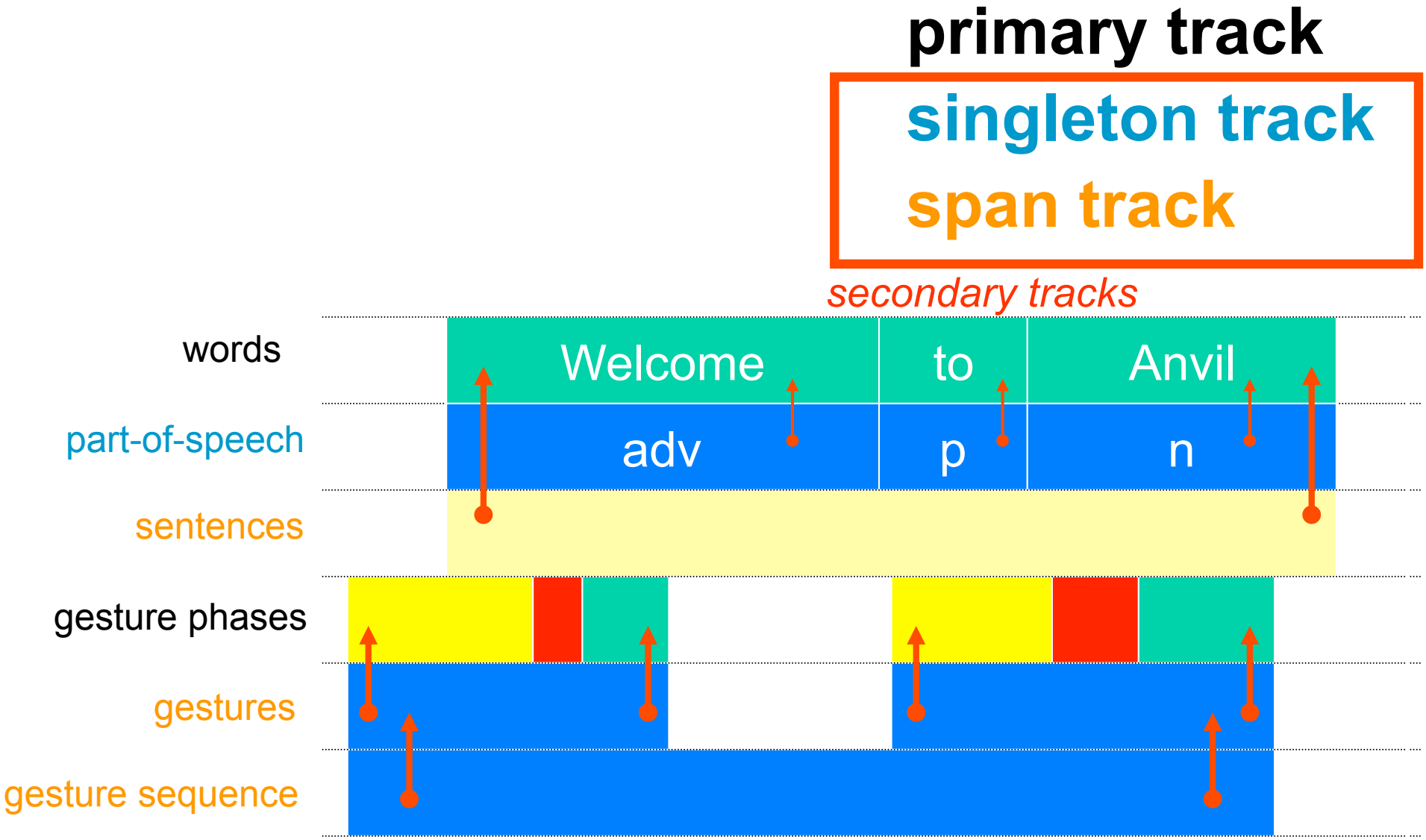
Type-specific Interface

```
<attribute name=„type“>
    <value-el>preparation</value-el>
    <value-el>stroke</value-el>
    <value-el>hold</value-el>
    <value-el>retraction</value-el>
</attribute>
```

Type-specific Interface



Track Types


primary track
singleton track
span track
secondary tracks
words
part-of-speech
sentences
gesture phases
gestures
gesture sequence
Welcome
to
Anvil
adv
p
n

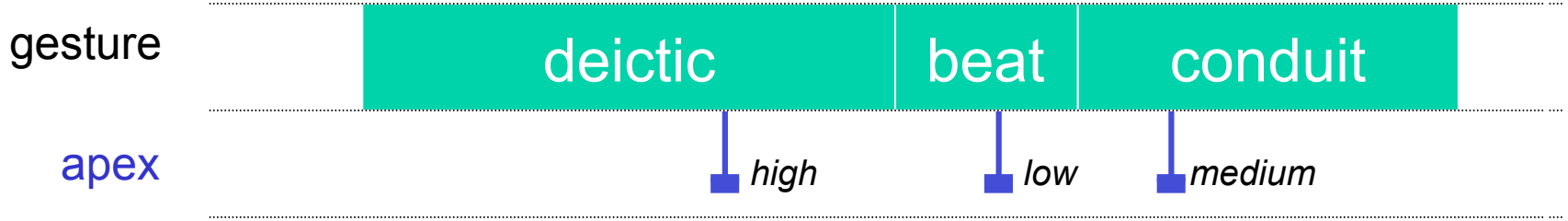
NEW: primarypoint

primarypoint track

primary track

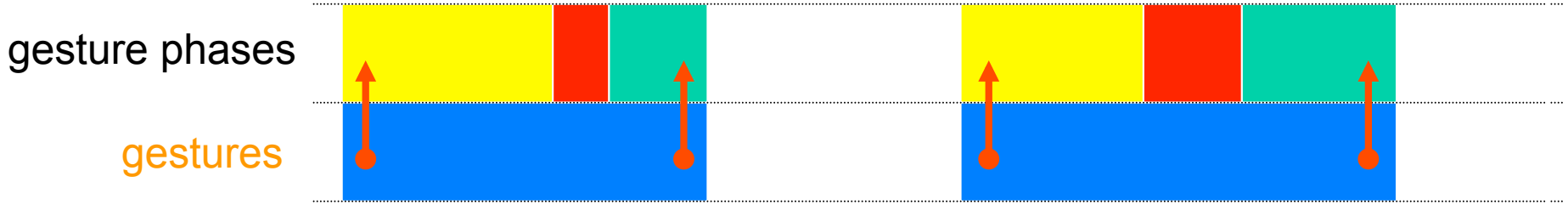
singleton track

span track

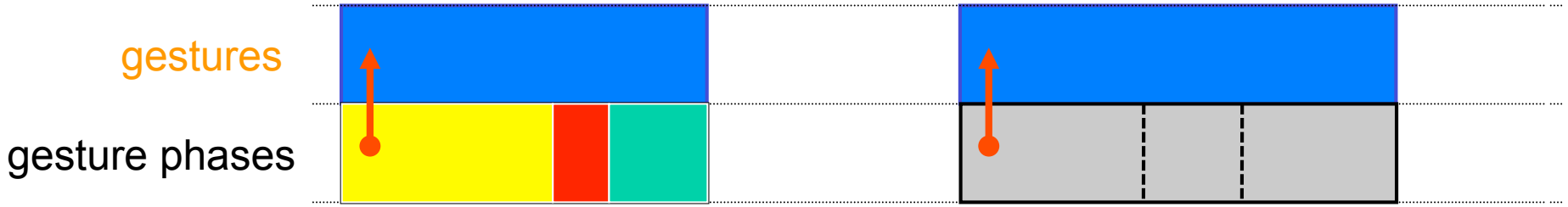


NEW: subdivision

Bottom-Up Coding with „span“ track



Top-Down Coding with „subdivision“ track

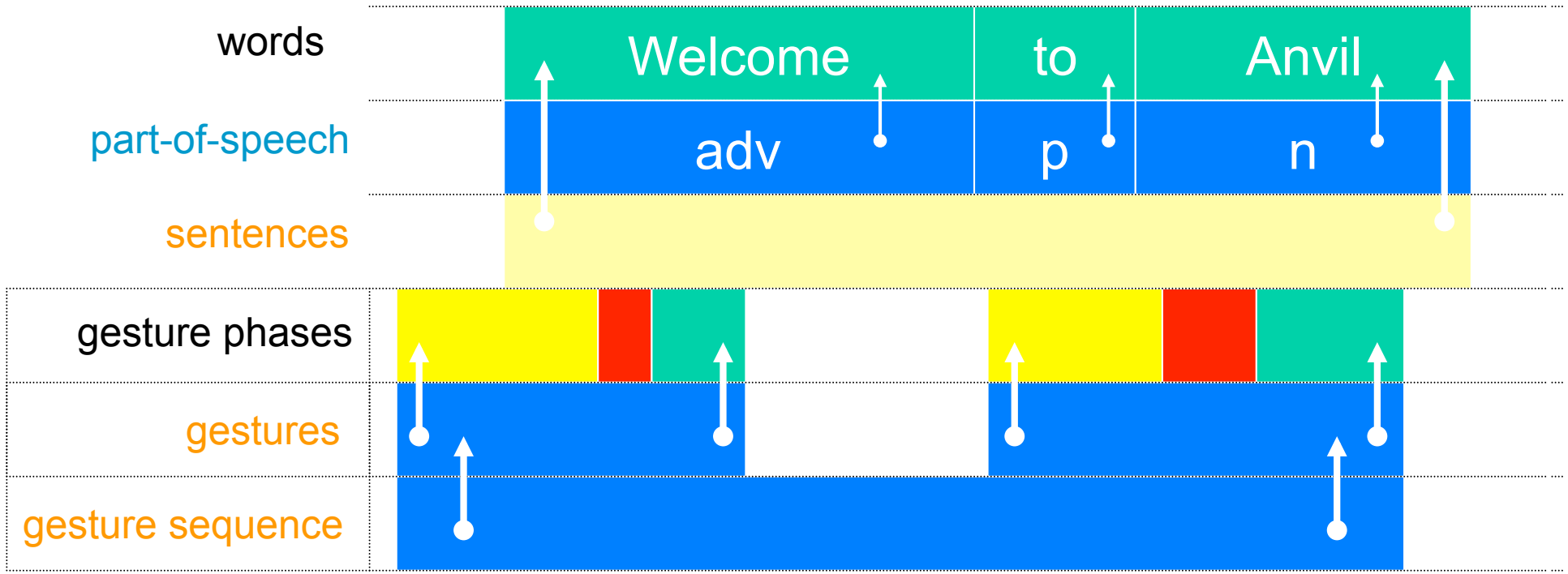


Summary: Track Types

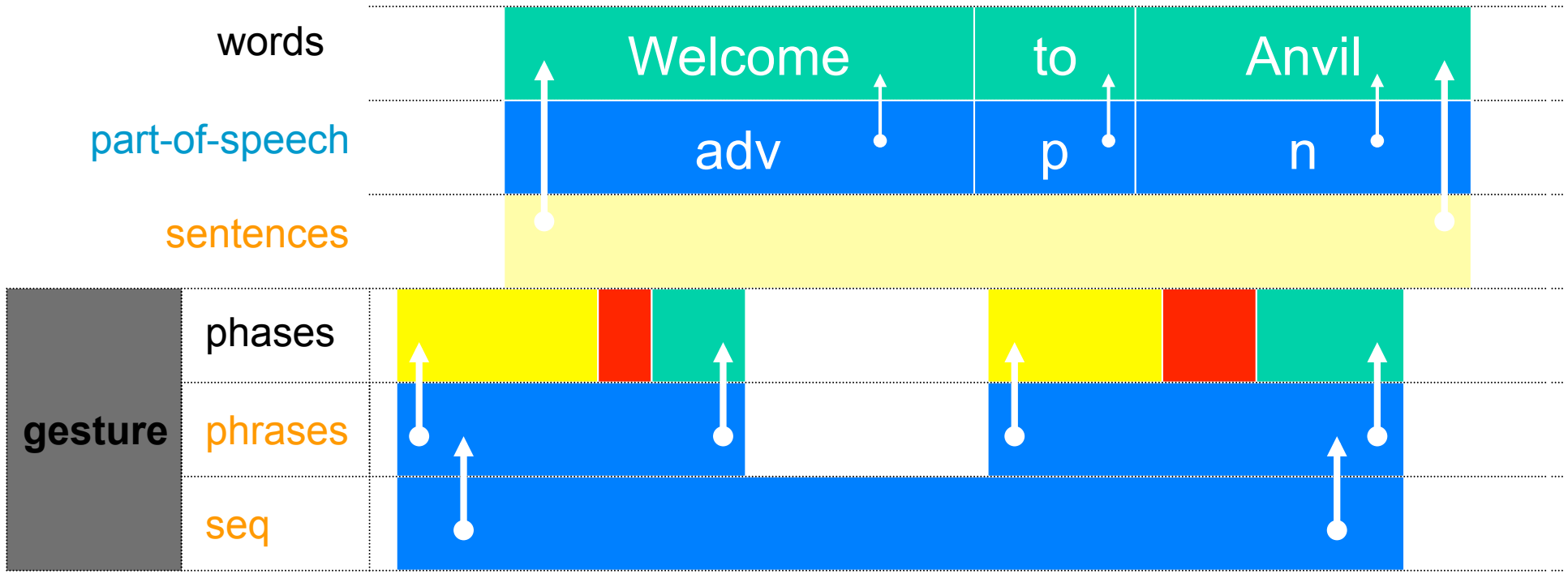
- Base types:
  - primary
  - primarypoint
- Secondary types:
  - singleton
  - span
  - subdivision

➔ DEMO (JL, tracks)

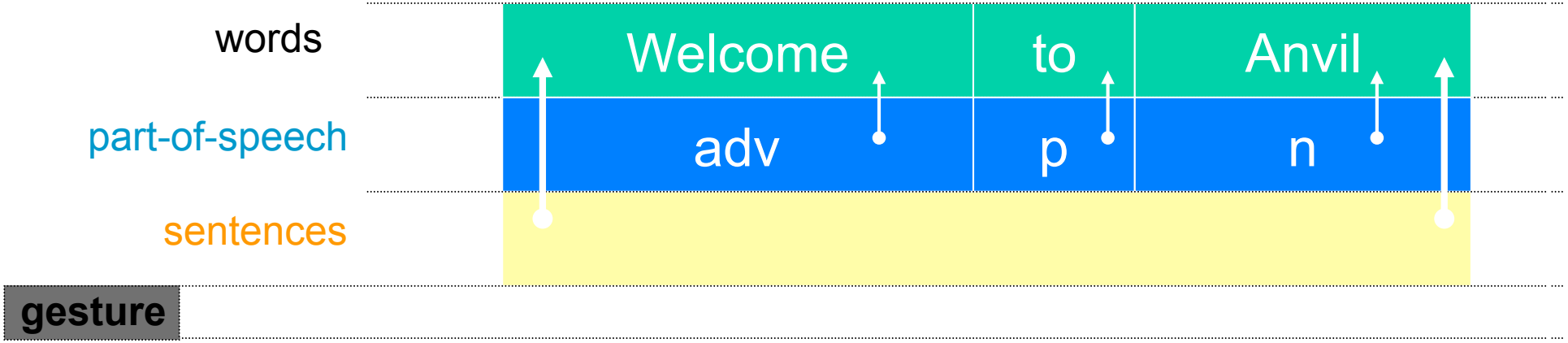
Groups


words
part-of-speech
sentences
gesture phases
gestures
gesture sequence
Welcome
to
Anvil
adv
p
n

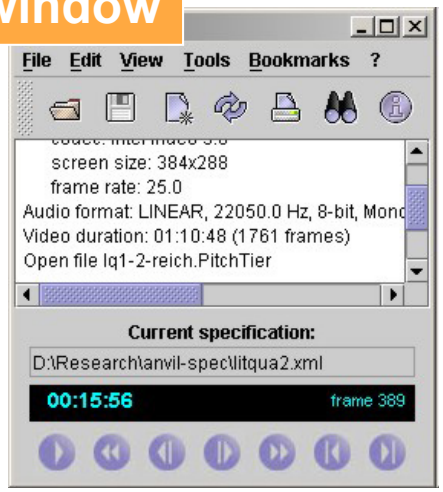
Groups

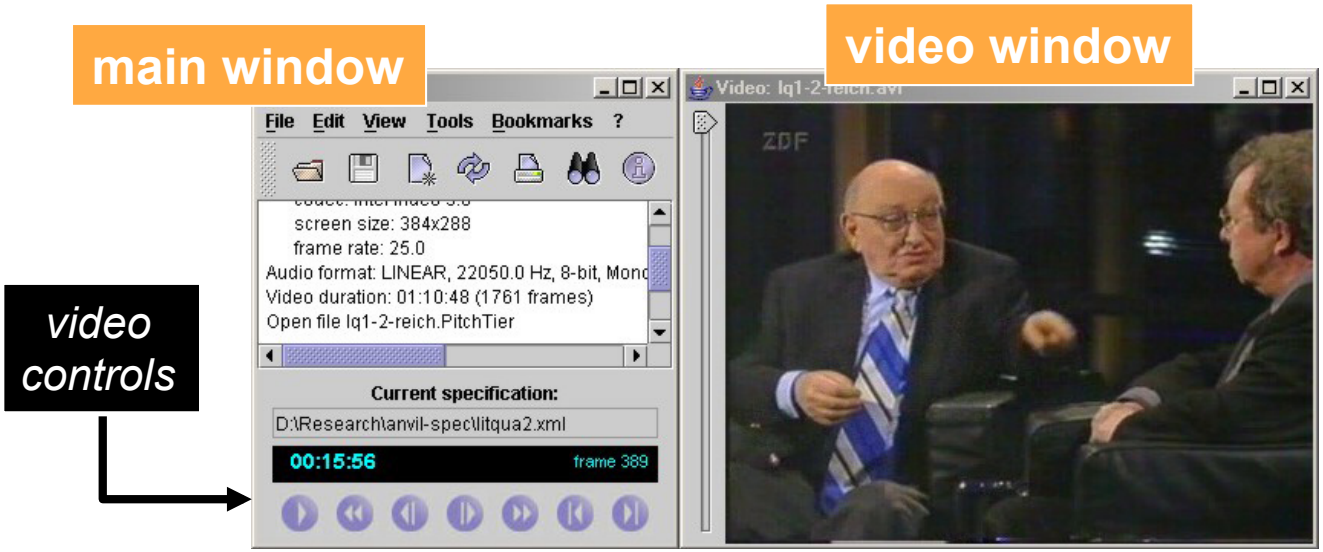

words
part-of-speech
sentences
Welcome
to
Anvil
adv
p
n
gesture
phases
phrases
seq

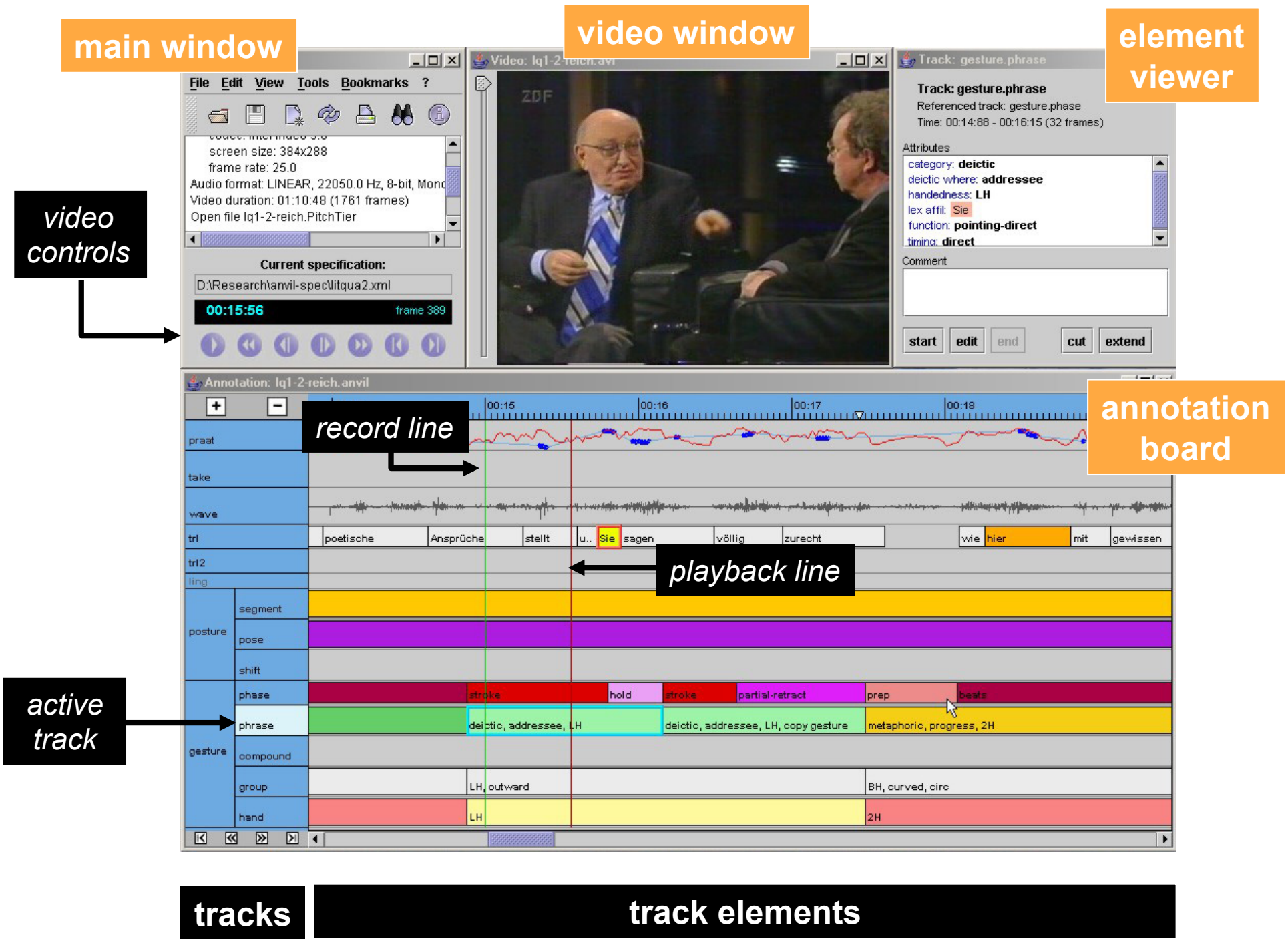
Groups


words
part-of-speech
sentences
gesture
Welcome
to
Anvil
adv
p
n

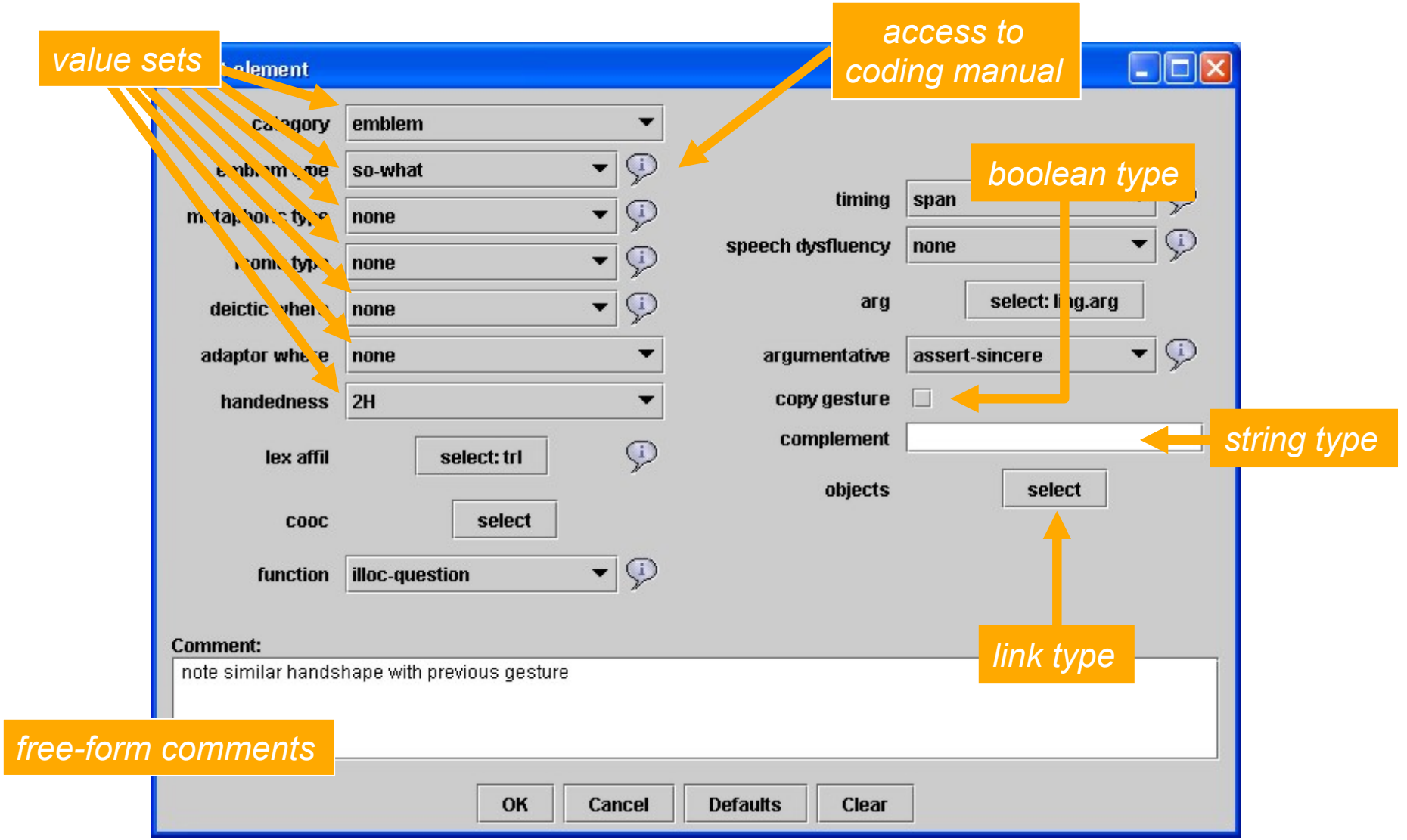
main window


File Edit View Tools Bookmarks ?
screen size: 384x288
frame rate: 25.0
Audio format: LINEAR, 22050.0 Hz, 8-bit, Mono
Video duration: 01:10:48 (1761 frames)
Open file lq1-2-reich.PitchTier
Current specification:
D:\Research\anvil-spec\litqua2.xml
00:15:56
frame 389


main window
video window
video controls
File Edit View Tools Bookmarks ?
screen size: 384x288
frame rate: 25.0
Audio format: LINEAR, 22050.0 Hz, 8-bit, Mono
Video duration: 01:10:48 (1761 frames)
Open file lq1-2-reich.PitchTier
Current specification:
D:\Research\anvil-spec\litqua2.xml
00:15:56
frame 389
ZDF


main window
video window
element viewer
video controls
File Edit View Tools Bookmarks ?
screen size: 384x288
frame rate: 25.0
Audio format: LINEAR, 22050.0 Hz, 8-bit, Mono
Video duration: 01:10:48 (1761 frames)
Open file lq1-2-reich.PitchTier
Current specification:
D:\Research\anvil-spec\litqua2.xml
00:15:56
frame 389
Video: lq1-2-reich.avi
ZDF
Track: gesture.phrase
Track: gesture.phrase
Referenced track: gesture.phase
Time: 00:14:88 - 00:16:15 (32 frames)
Attributes
category: deictic
deictic where: addressee
handedness: LH
lex affil: Sie
function: pointing-direct
timing: direct
Comment
start
edit
end
cut
extend
Annotation: lq1-2-reich.anvil
annotation board
record line
playback line
00:15
00:16
00:17
00:18
praat
take
wave
trl
trl2
ling
poetische
Ansprüche
stellt
u..
Sie
sagen
völlig
zurecht
wie
hier
mit
gewissen
posture
segment
pose
shift
gesture
phase
phrase
compound
group
hand
stroke
hold
stroke
partial-retract
prep
beats
deictic, addressee, LH
deictic, addressee, LH, copy gesture
metaphoric, progress, 2H
LH, outward
BH, curved, circ
LH
2H
active track
tracks
track elements

New Annotation Element

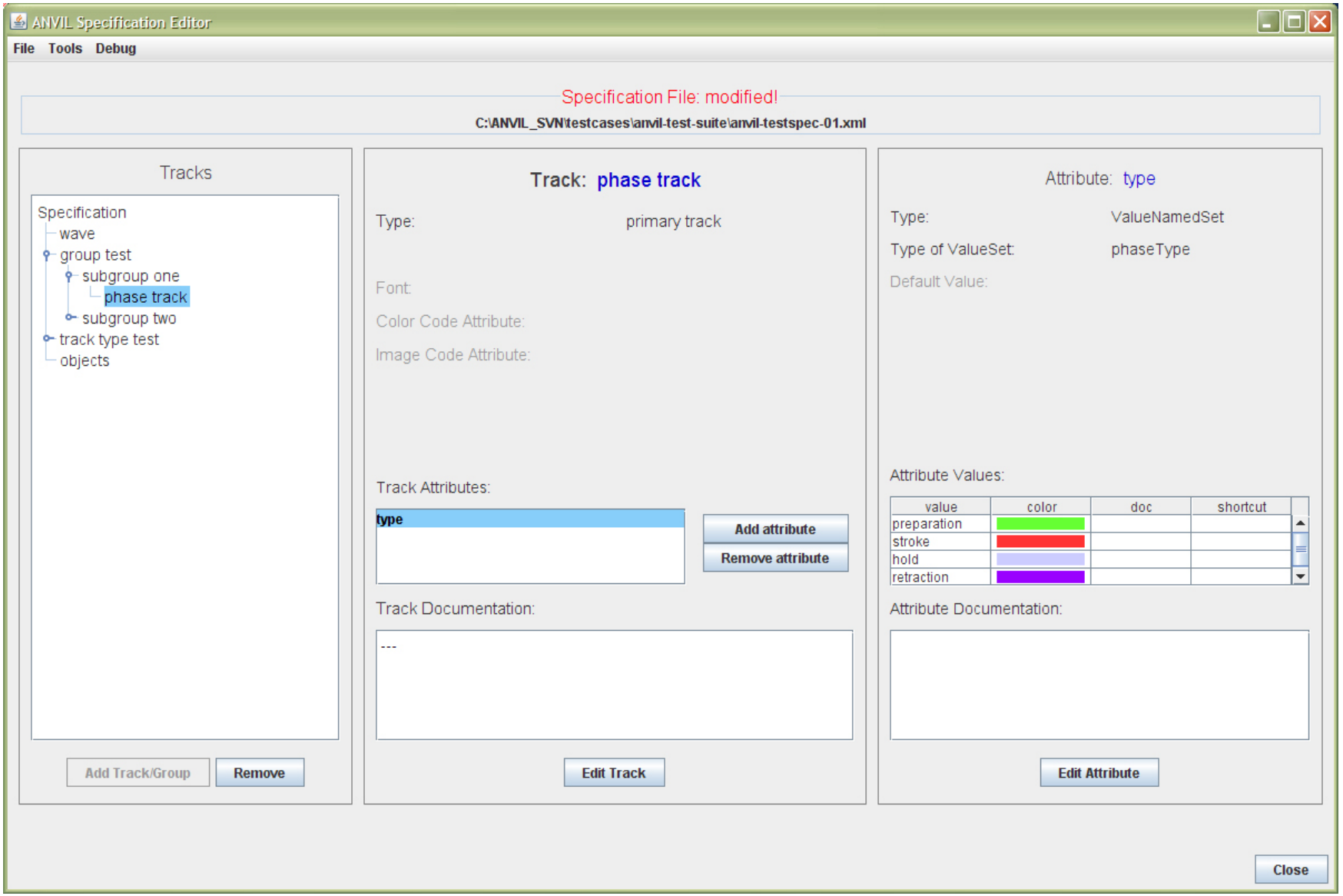


➔ DEMO (prof)

Coding Scheme = Specification

- Specify
  - tracks *name + type + ref*
  - groups *name + contained tracks/groups*
  - each track: attributes *name + type + values*
- Bad news: XML file!
- Good news: Graphical editor in Anvil 5

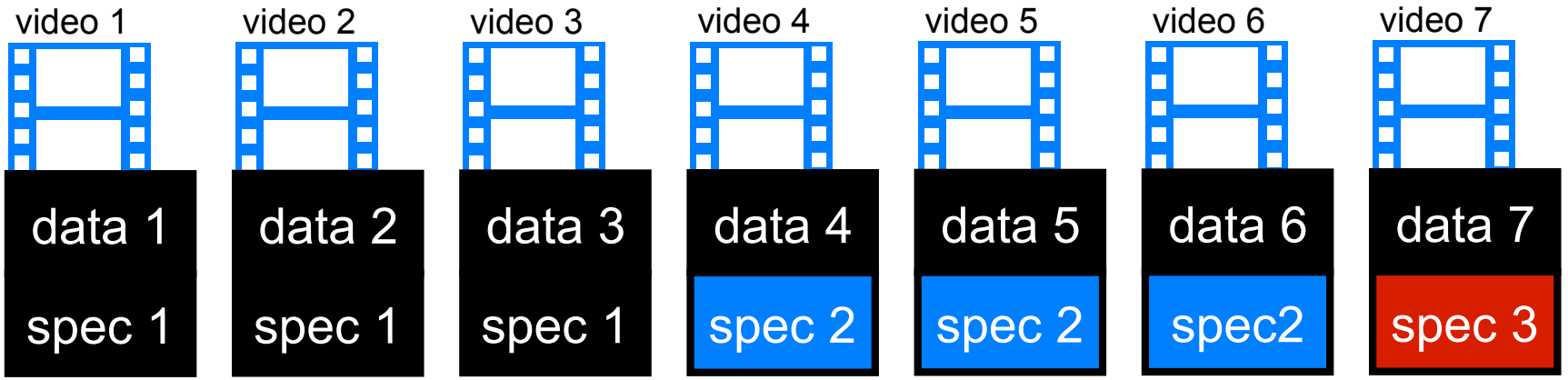
Specification Editor

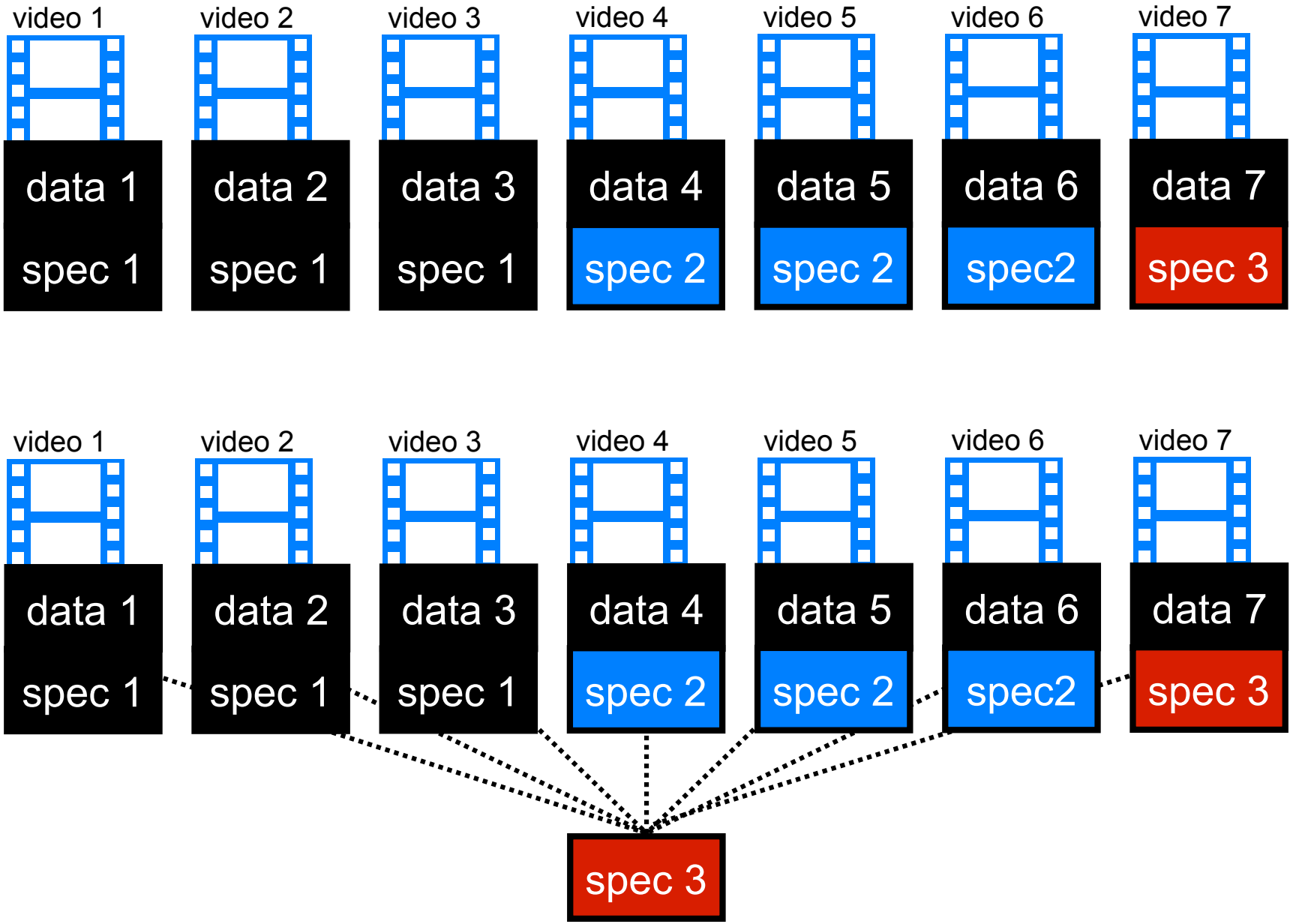


ANVIL 5

Annotation Data vs. Coding Scheme

- Store in one file
  - pro: all data in one place
  - pro: consistency between scheme and data guaranteed
  - con: modify scheme → modify all files!
- Store in separate files
  - pro: modify scheme → modify 1 file
  - pro: trivial to check whether N files have equal scheme


video 1
video 2
video 3
video 4
video 5
video 6
video 7
data 1
spec 1
data 2
spec 1
data 3
spec 1
data 4
spec 2
data 5
spec 2
data 6
spec2
data 7
spec 3

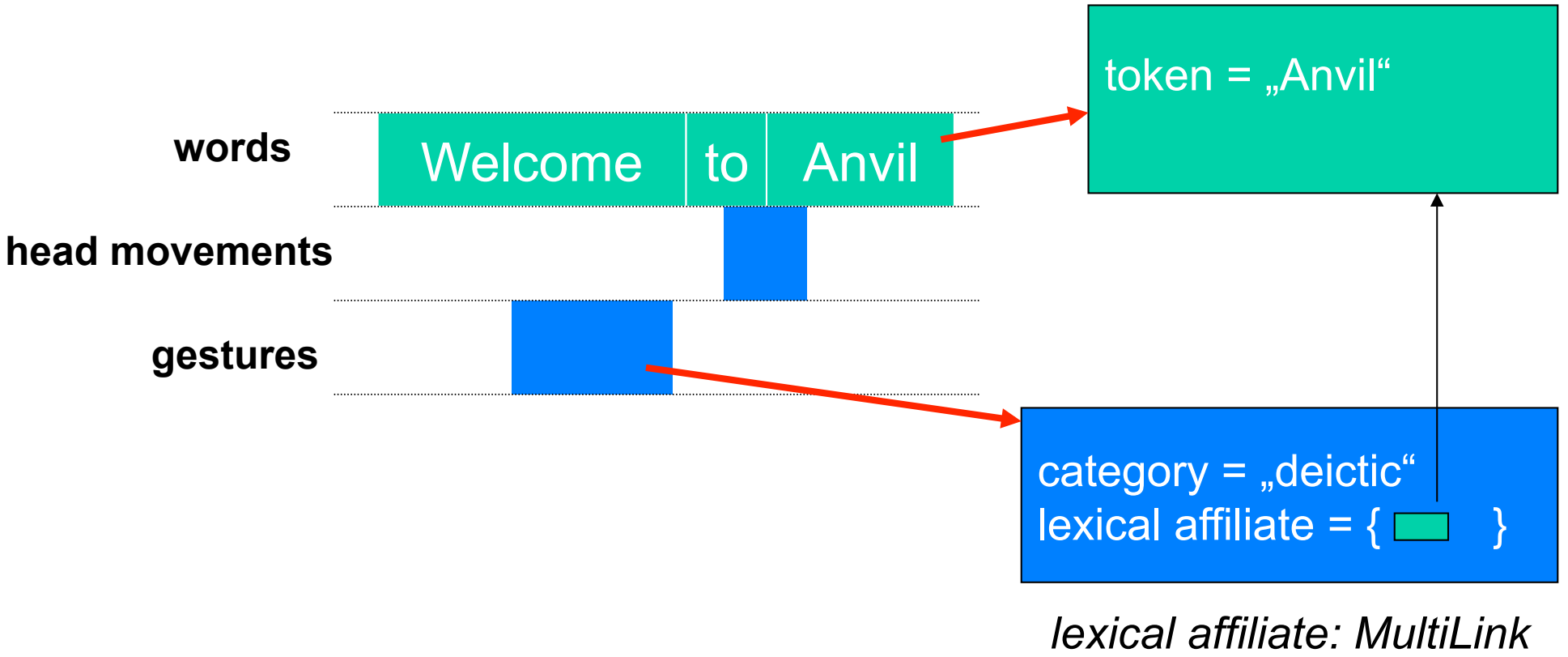

video 1
video 2
video 3
video 4
video 5
video 6
video 7
data 1
data 2
data 3
data 4
data 5
data 6
data 7
spec 1
spec 1
spec 1
spec 2
spec 2
spec2
spec 3
video 1
video 2
video 3
video 4
video 5
video 6
video 7
data 1
data 2
data 3
data 4
data 5
data 6
data 7
spec 1
spec 1
spec 1
spec 2
spec 2
spec2
spec 3
spec 3

Interesting Features

Cross-level Links

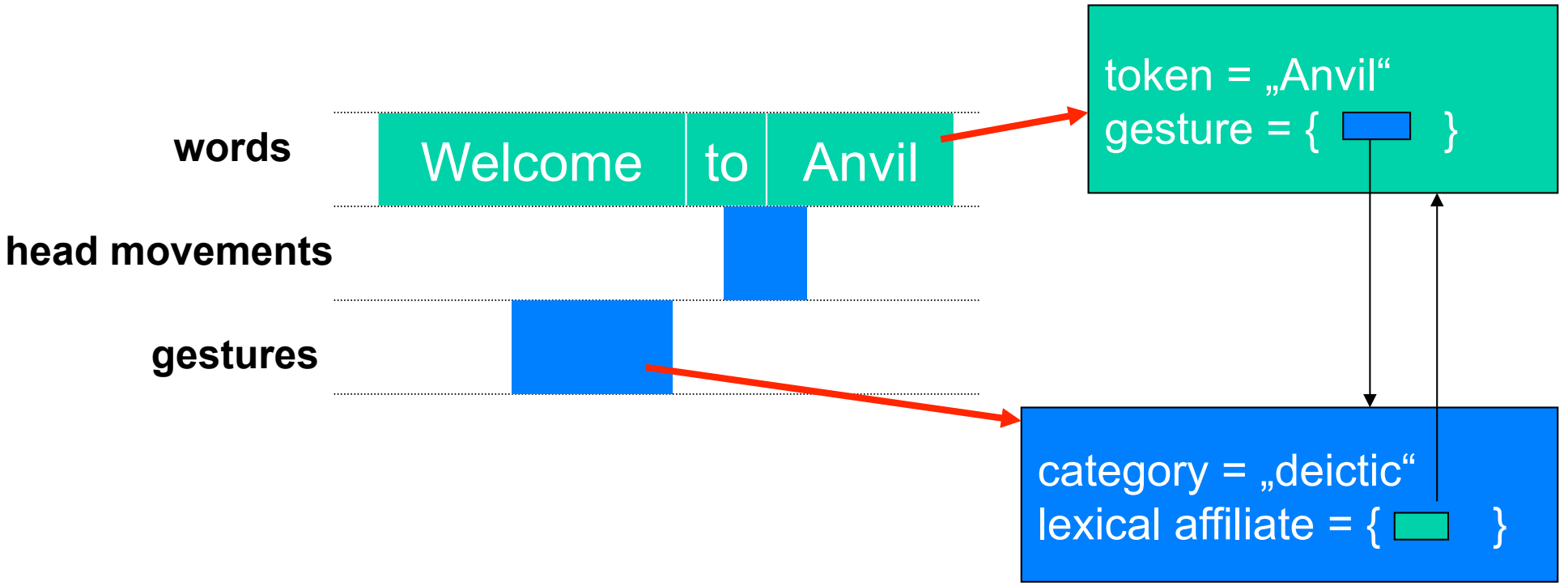
- problem
  - code relations between elements in one or across tracks
- examples
  - gesture and word(s) have same meaning
  - two gestures refer to same object/person/location
- solution
  - allow putting logical links into an attribute

Links



lexical affiliate: MultiLink

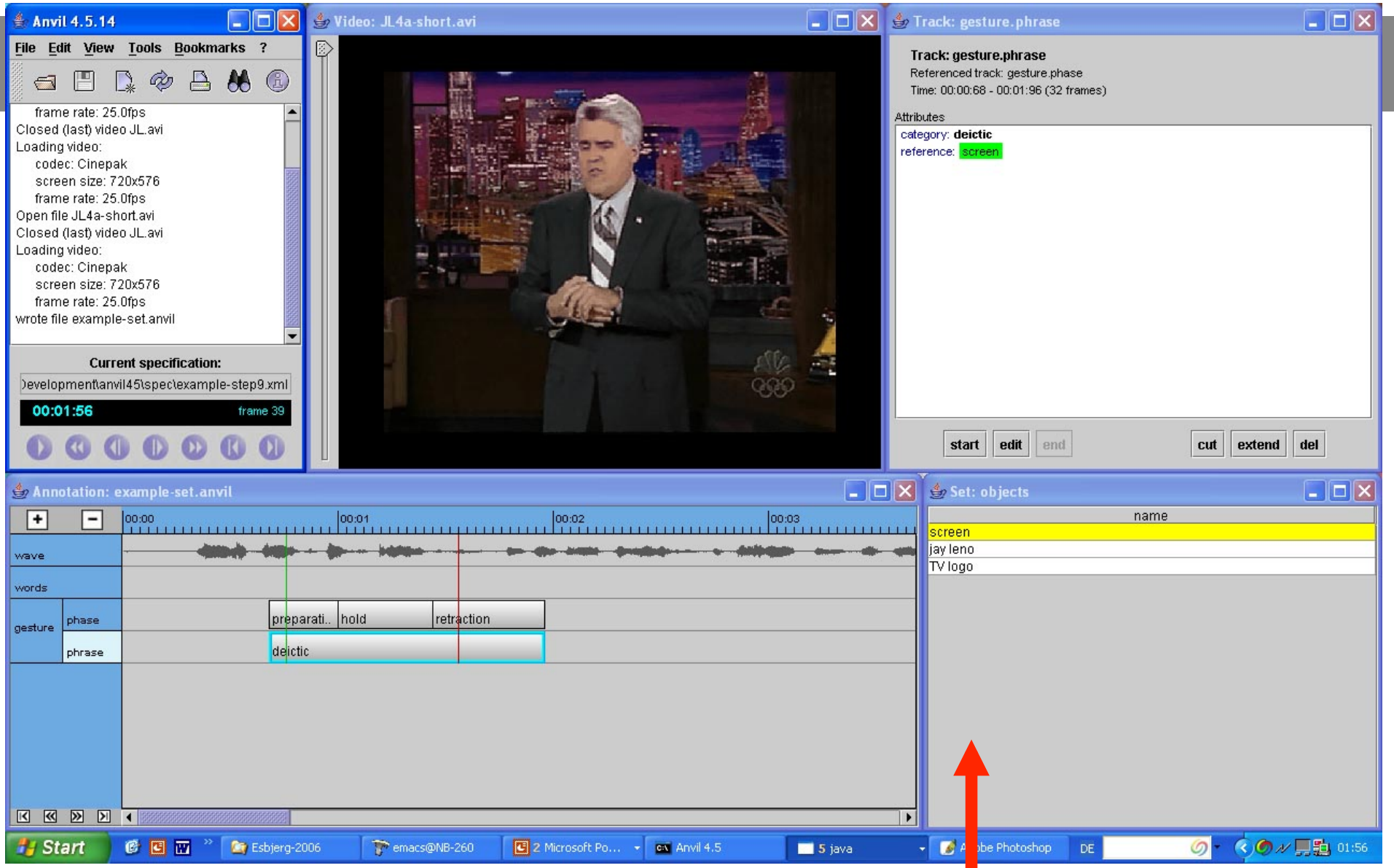
Reciprocal Links



lexical affiliate: ReciprocalLink(gesture)

Non-Temporal Elements

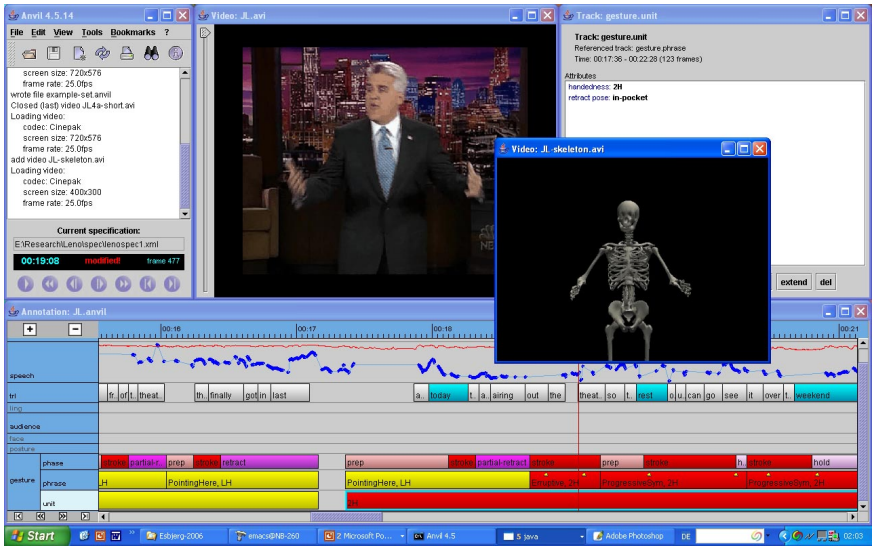
- Some things do not change for whole video
  - identity of person
  - writing on blackboard
  - object on table …
- Refer to these items with link
  - special container for „non-temporal elements“
  - called a SET



a „Set“ containing non-temporal elements

Multiple Videos

- Multiple videos displayed in synchrony
  - multiple camera set-up
    (e.g. one for screen / face / overall ..)
- Video sync (offset) possible **ANVIL 5**
  - no graphical interface
- Video that is first opened is „master“
  - Spatial annotation only
    on master



Spatio-Temporal Annotation

- What does an encoded element refer to… spatially?
  ➔ whole frame(s)
- Restrict to screen region/point:
  - mark up face region
  - mark up a distance (e.g. hand to hand)
  - mark up movement path (gesture)
- Options
  - store single location (constant time)
  - store moving location (timestamps)
  - region may change size (morphing)
- Where store it?
  - object-centric
  - interval-centric

Current State

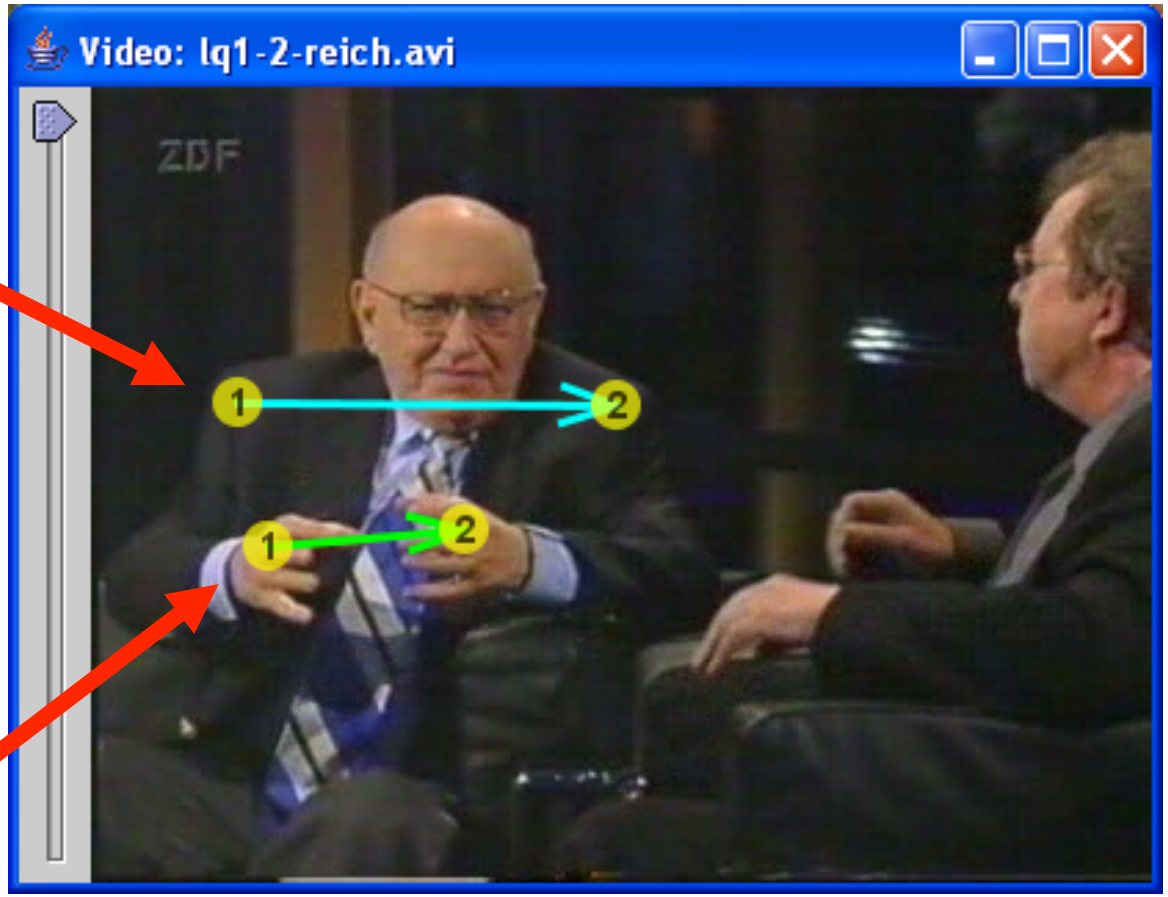
- What
  - single points
- Where
  - attribute in element (= interval)
- How
  - draw directly on video frame
- What you get
  - attribute *myPath* = $< (t_1, x_1, y_1) \ldots (t_n, x_n, y_n) >$
- Feature „under construction“

Example: Hand-to-hand Distance

measure
shoulder width = s
to normalize value

normalized:
d* = d / s = 0.4

hand-to-hand distance
= d



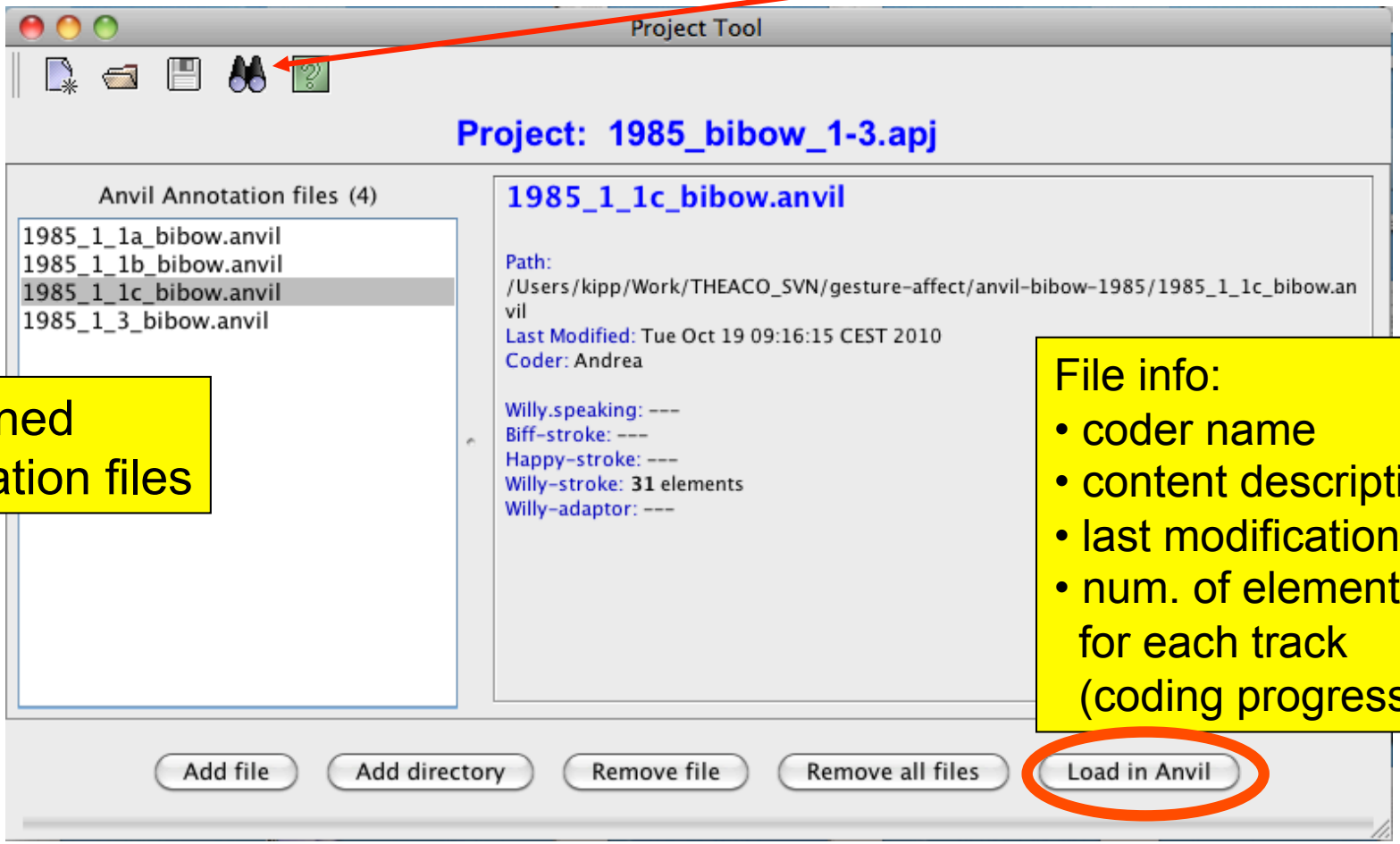
➔ DEMO, MR

Projects

- So far: annotations for single video/session
- Usually: corpus = many videos+annotations
- Definition:
  - project = annotation files with same specification
- Project Tool
  - overview: metadata, progress
  - search
  - analysis over corpus
  - export for external analysis

Project Tool

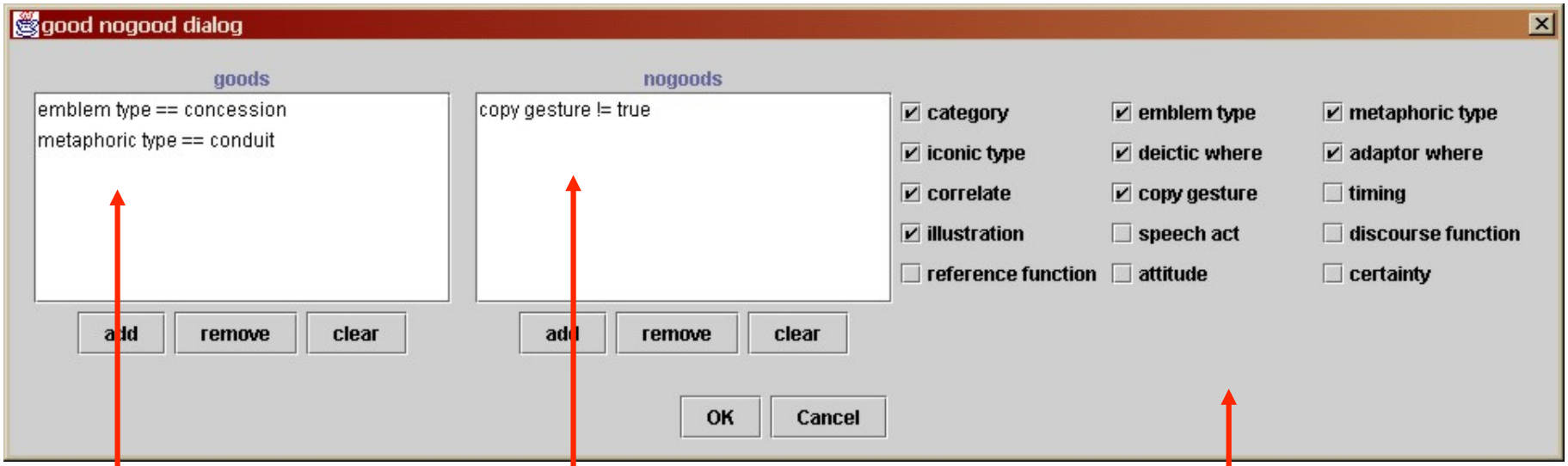
search engine



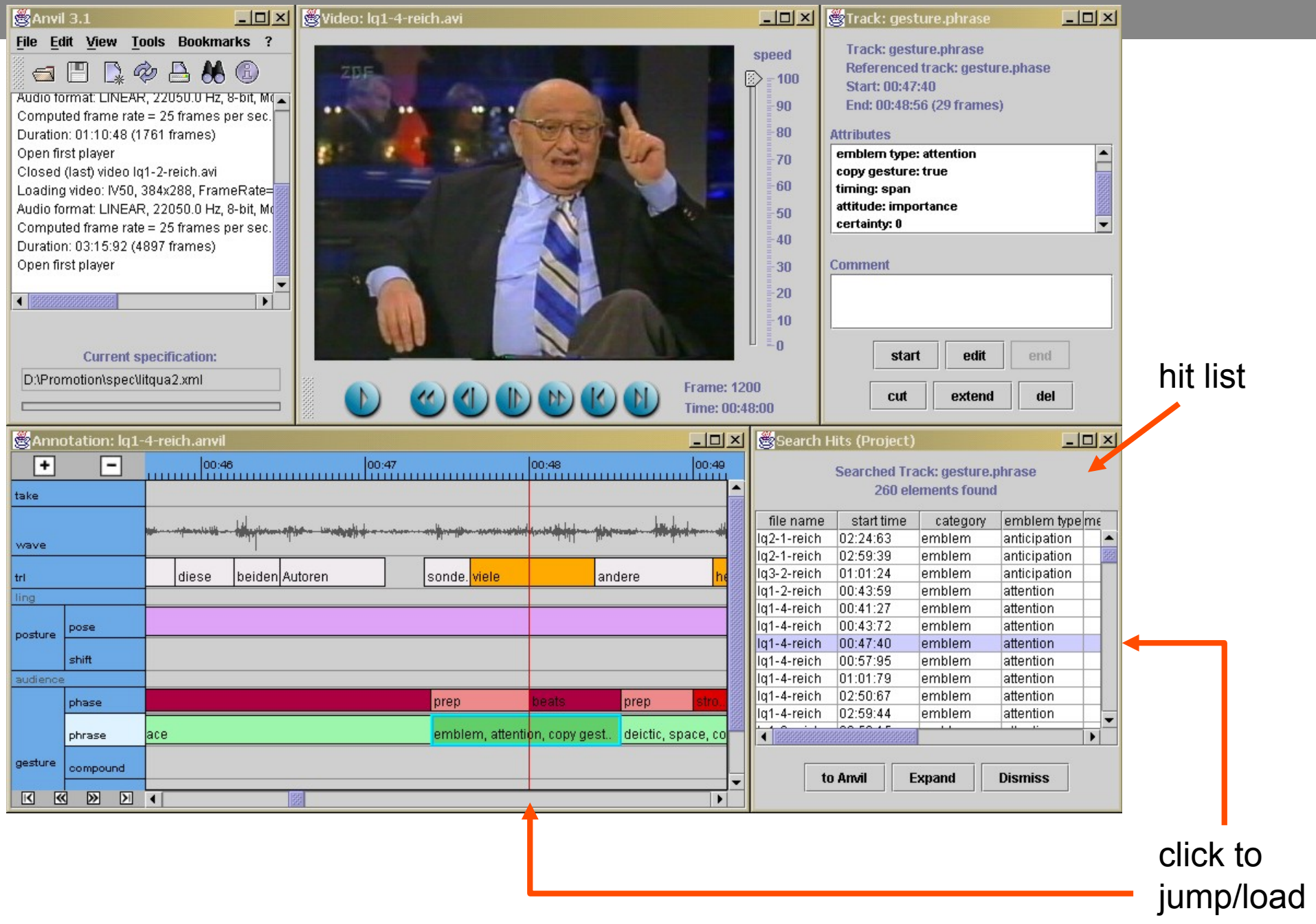
contained annotation files

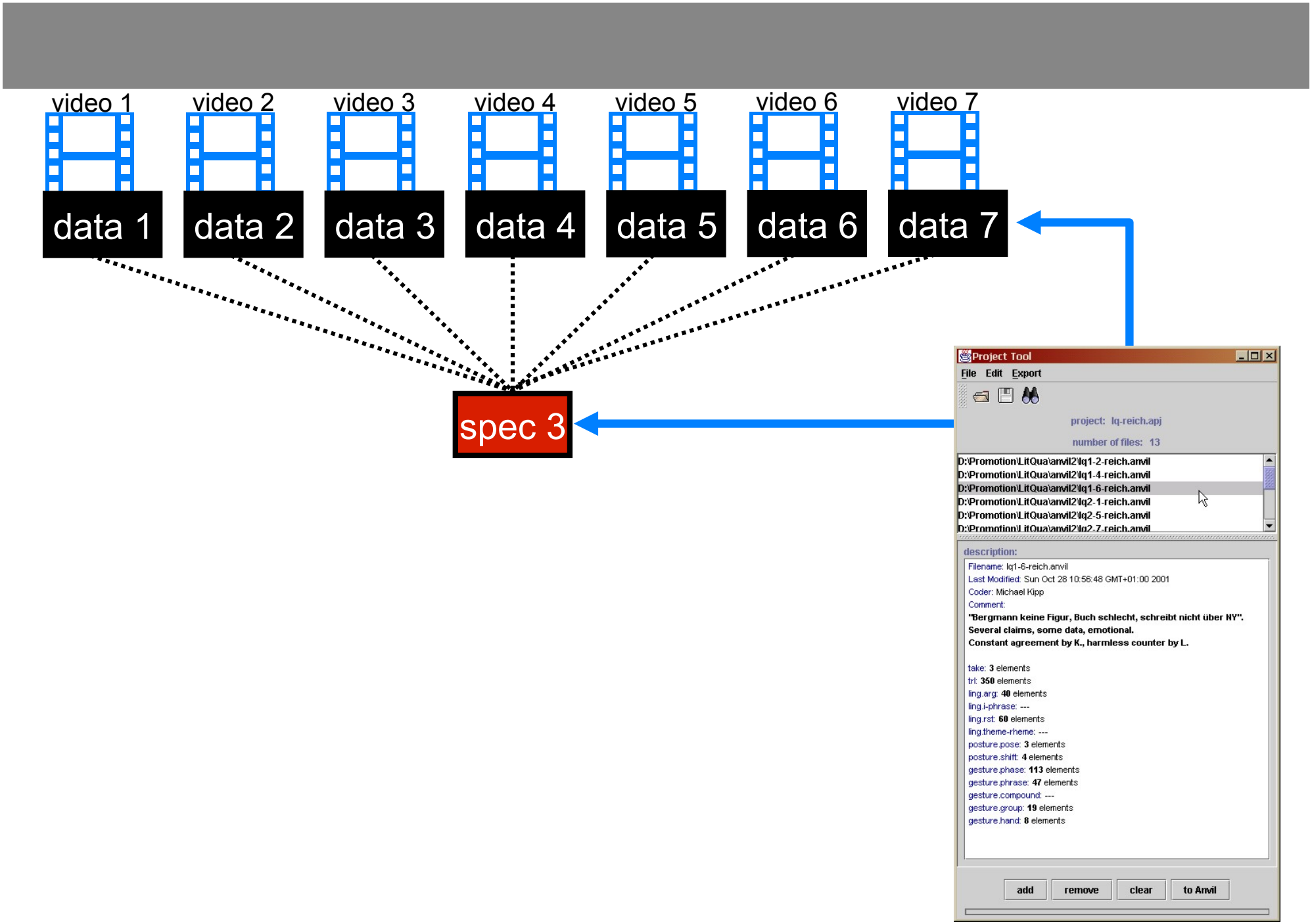
File info:
- coder name
- content description
- last modification time
- num. of elements for each track (coding progress !)

Search



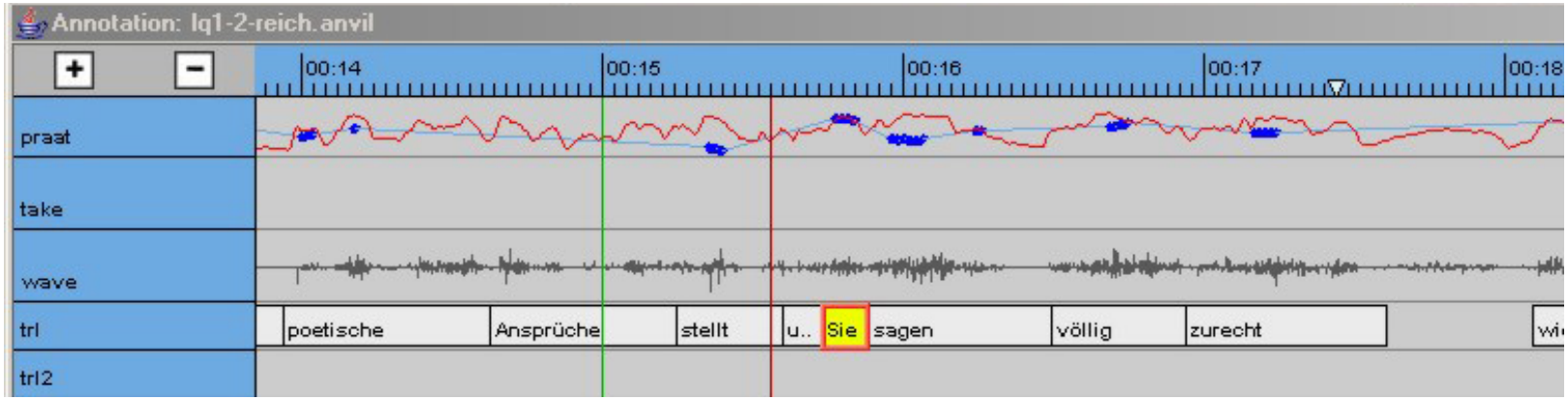
include & exclude view options


Anvil 3.1
File Edit View Tools Bookmarks ?
Computed frame rate = 25 frames per sec.
Duration: 01:10:48 (1761 frames)
Open first player
Closed (last) video lq1-2-reich.avi
Loading video: IV50, 384x288, FrameRate=
Computed frame rate = 25 frames per sec.
Duration: 03:15:92 (4897 frames)
Open first player
Current specification:
D:\Promotion\spec\litqua2.xml
Video: lq1-4-reich.avi
speed
100
90
80
70
60
50
40
30
20
10
0
Frame: 1200
Time: 00:48:00
Track: gesture.phrase
Track: gesture.phrase
Referenced track: gesture.phrase
Start: 00:47:40
End: 00:48:56 (29 frames)
Attributes
emblem type: attention
copy gesture: true
timing: span
attitude: importance
certainty: 0
Comment
start
edit
end
cut
extend
del
Annotation: lq1-4-reich.anvil
00:46
00:47
00:48
00:49
take
wave
trl
diese
beiden
Autoren
sonde..
viele
andere
ling
posture
pose
shift
audience
phase
prep
beats
prep
phrase
ace
emblem, attention, copy gest..
deictic, space, co
gesture
compound
Search Hits (Project)
Searched Track: gesture.phrase
260 elements found
file name
start time
category
emblem type
lq2-1-reich 02:24:63 emblem anticipation
lq2-1-reich 02:59:39 emblem anticipation
lq3-2-reich 01:01:24 emblem anticipation
lq1-2-reich 00:43:59 emblem attention
lq1-4-reich 00:41:27 emblem attention
lq1-4-reich 00:43:72 emblem attention
lq1-4-reich 00:47:40 emblem attention
lq1-4-reich 00:57:95 emblem attention
lq1-4-reich 01:01:79 emblem attention
lq1-4-reich 02:50:67 emblem attention
lq1-4-reich 02:59:44 emblem attention
to Anvil
Expand
Dismiss
hit list
click to
jump/load


video 1
video 2
video 3
video 4
video 5
video 6
video 7
data 1
data 2
data 3
data 4
data 5
data 6
data 7
spec 3
Project Tool
File Edit Export
project: lq-reich.apj
number of files: 13
D:\Promotion\LitQua\anvil2\lq1-2-reich.anvil
D:\Promotion\LitQua\anvil2\lq1-4-reich.anvil
D:\Promotion\LitQua\anvil2\lq1-6-reich.anvil
D:\Promotion\LitQua\anvil2\lq2-1-reich.anvil
D:\Promotion\LitQua\anvil2\lq2-5-reich.anvil
D:\Promotion\LitQua\anvil2\lq2-7-reich.anvil
description:
Filename: lq1-6-reich.anvil
Last Modified: Sun Oct 28 10:56:48 GMT+01:00 2001
Coder: Michael Kipp
Comment:
"Bergmann keine Figur, Buch schlecht, schreibt nicht über NY".
Several claims, some data, emotional.
Constant agreement by K., harmless counter by L.
take: 3 elements
trl: 350 elements
ling.arg: 40 elements
ling.i-phrase: ---
ling.rst: 60 elements
ling.theme-rheme: ---
posture.pose: 3 elements
posture.shift: 4 elements
gesture.phase: 113 elements
gesture.phrase: 47 elements
gesture.compound: ---
gesture.group: 19 elements
gesture.hand: 8 elements
add
remove
clear
to Anvil

Import from PRAAT

- Import PRAAT interval data
  - words, syllables, phonemes…
- Import pitch + intensity
  - data files are linked up

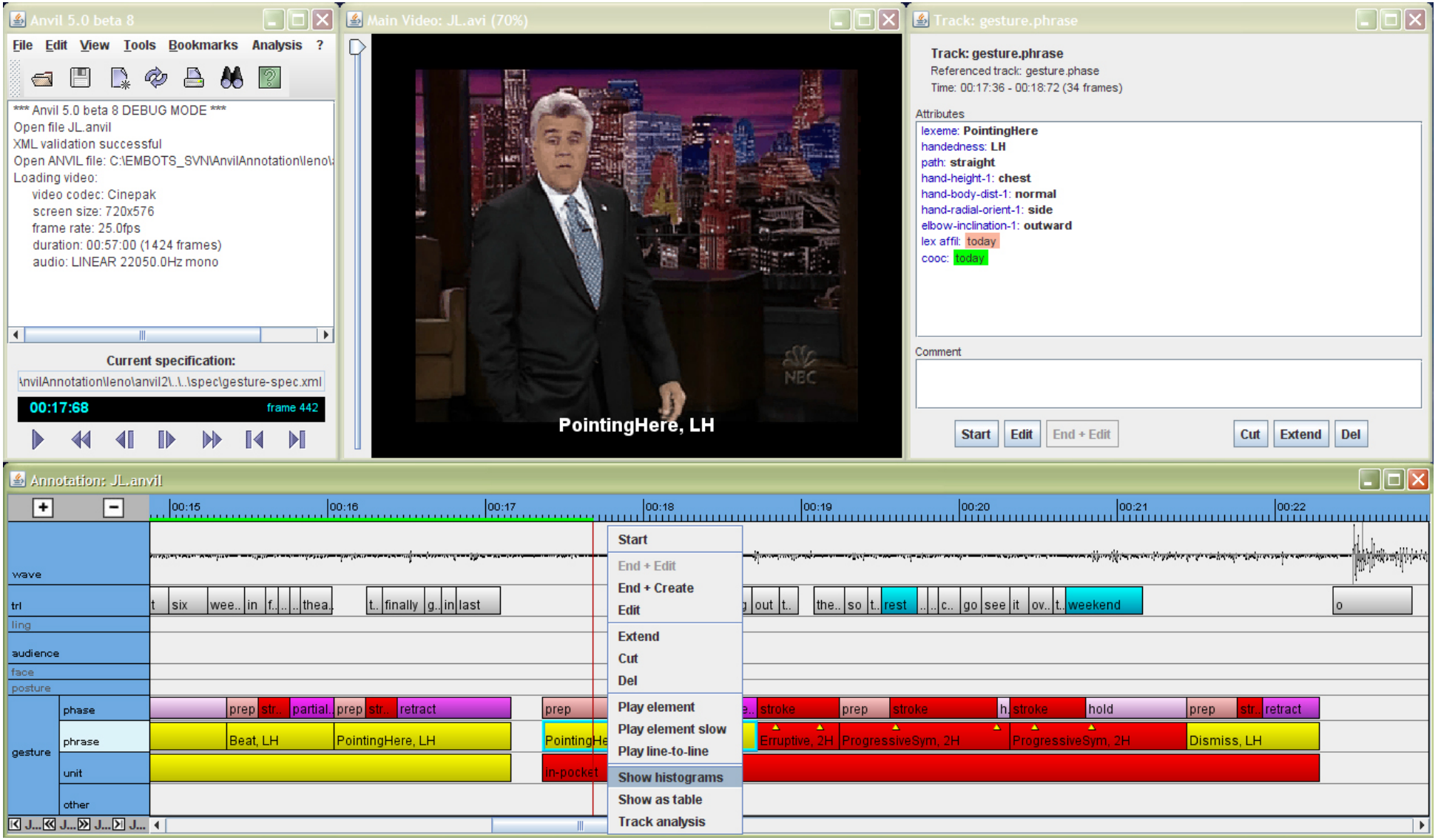


Analysis

Analysis

- Work on single annotations **and** projects
- Features
  - Histograms
  - Agreement
  - Transition Diagrams
  - Association
  - Export

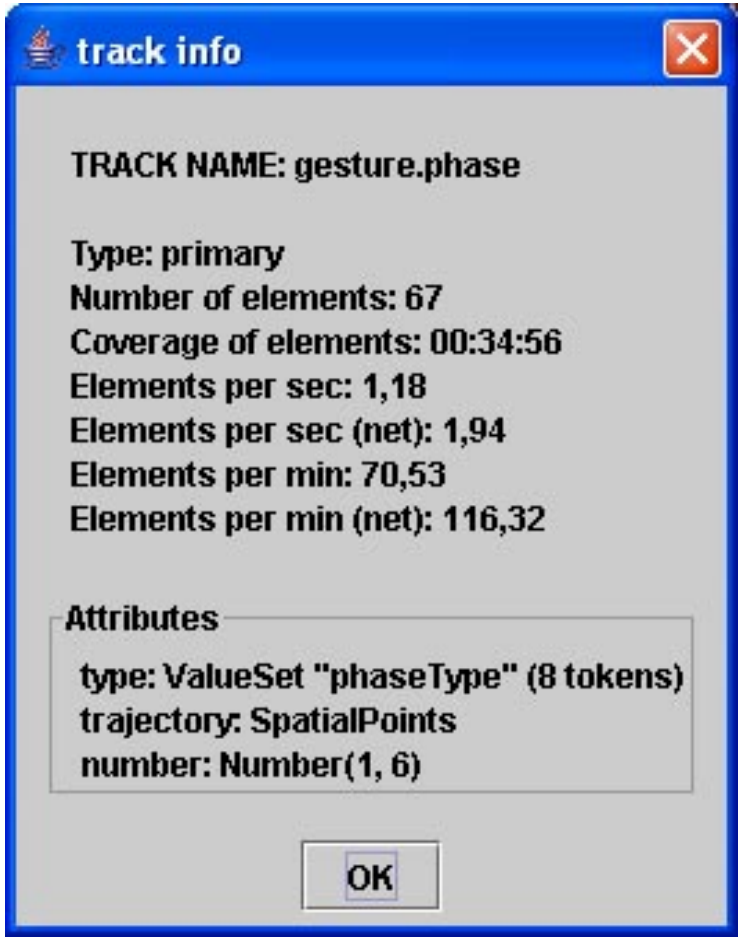
Analysis in main menu



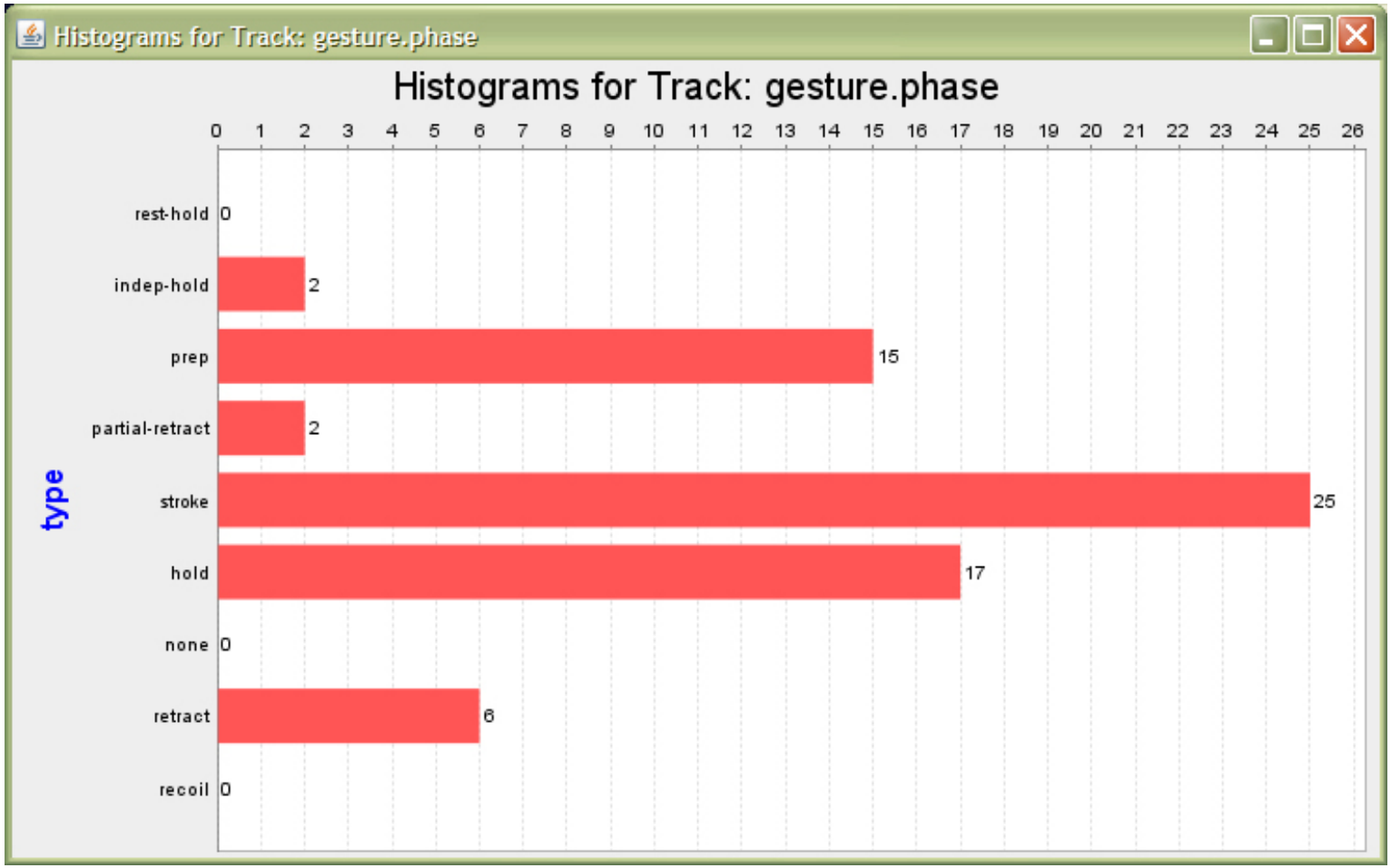
Analysis in context menu

Track Analysis

- descriptive stats
  - num. of elements
  - coverage (time)
  - elements / sec
  - elements / min
- „activity“
  - gesture rate
  - smile rate ...



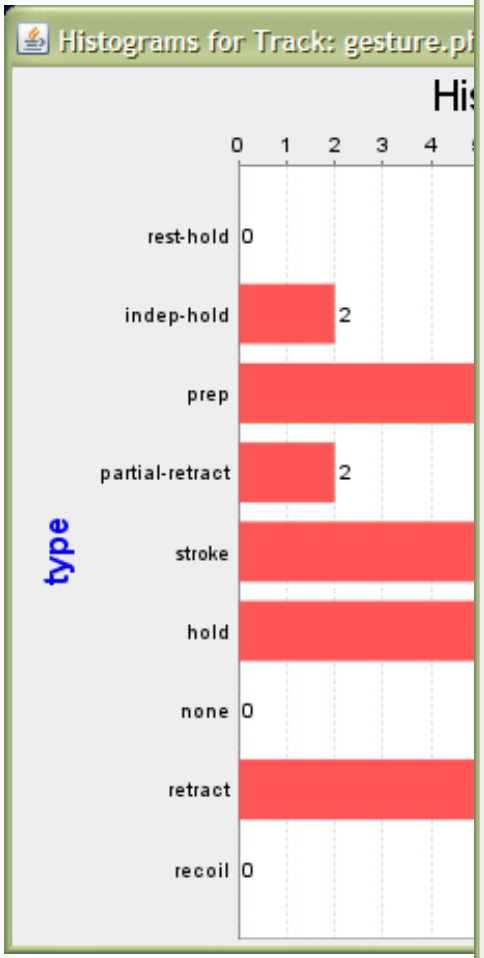
Histogram

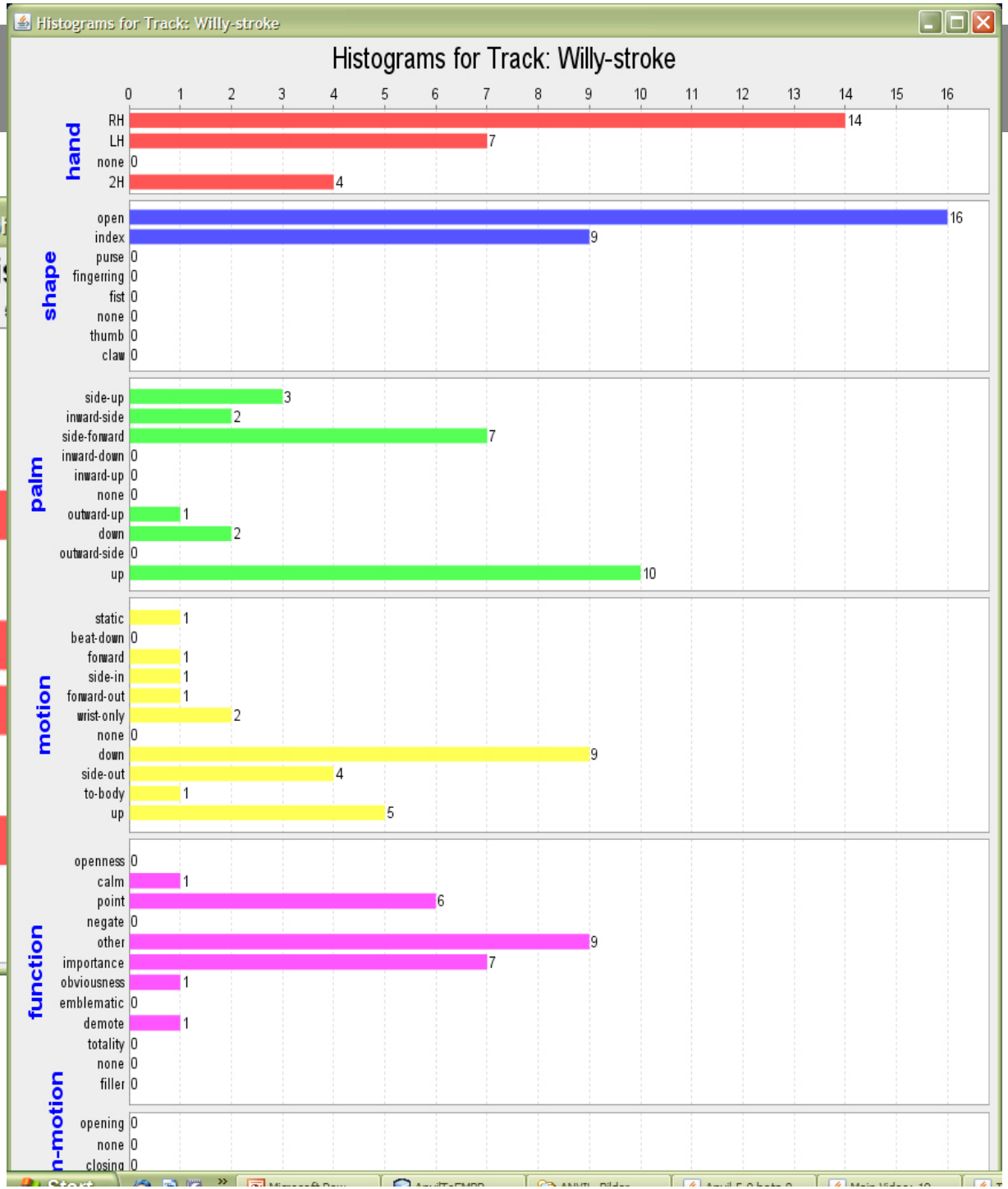


Distribution of labels in a single attribute („controlled vocab“)

ANVIL 5

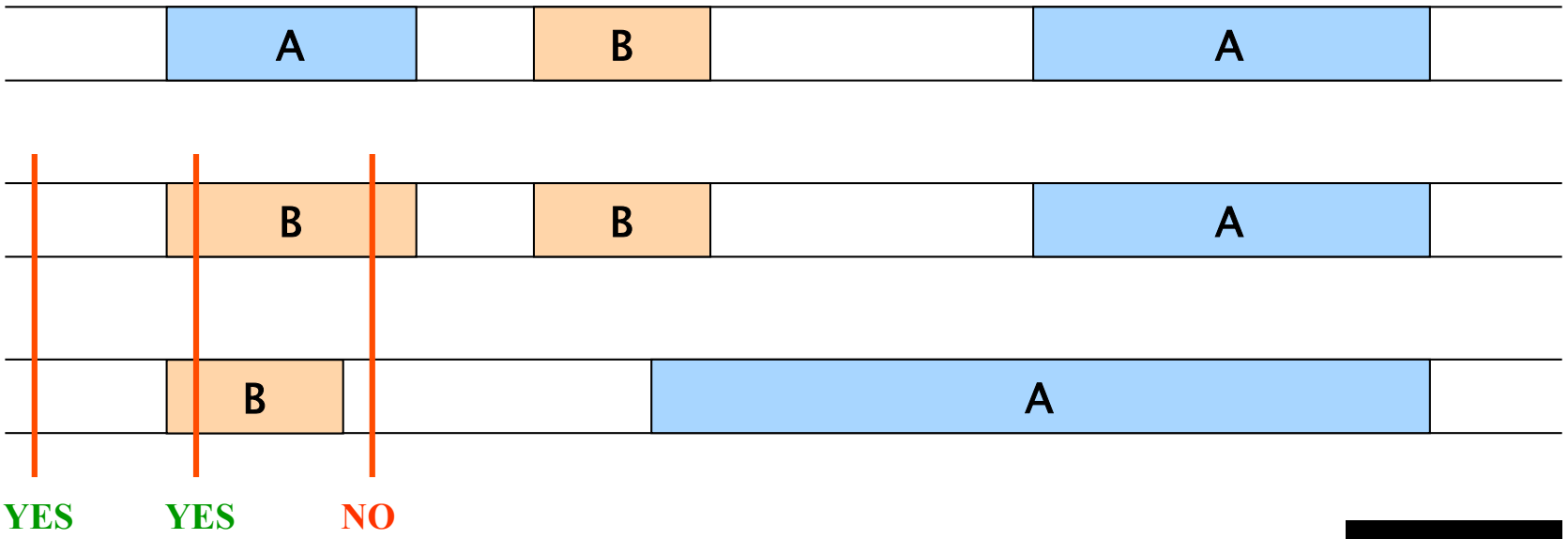
Histogram


Histograms for Track: gesture.pf
type
rest-hold 0
indep-hold 2
prep
partial-retract 2
stroke
hold
none 0
retract
recoil 0


Histograms for Track: Willy-stroke
Histograms for Track: Willy-stroke
hand
RH 14
LH 7
none 0
2H 4
shape
open 16
index 9
purse 0
fingering 0
fist 0
none 0
thumb 0
claw 0
palm
side-up 3
inward-side 2
side-forward 7
inward-down 0
inward-up 0
none 0
outward-up 1
down 2
outward-side 0
up 10
motion
static 1
beat-down 0
forward 1
side-in 1
forward-out 1
wrist-only 2
none 0
down 9
side-out 4
to-body 1
up 5
function
openness 0
calm 1
point 6
negate 0
other 9
importance 7
obviousness 1
emblematic 0
demote 1
totality 0
none 0
filler 0
n-motion
opening 0
none 0
closing 0

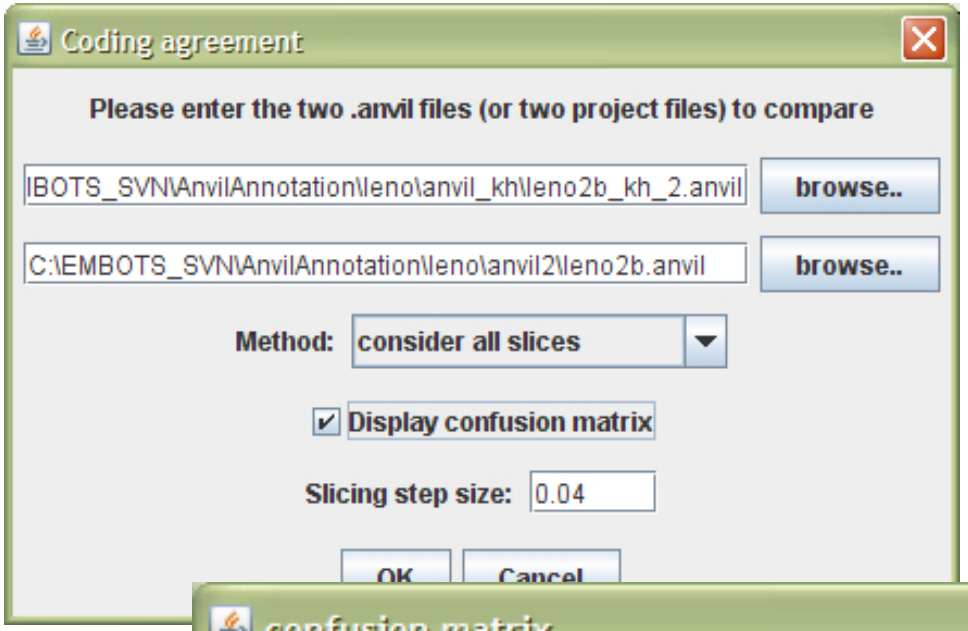
Agreement

- Measure agreement between coders
- Needs confusion matrix
- Which elements to compare?
- Solution: slice + compare



ANVIL 5

Agreement Results



confusion matrix

	V-O-I-D	hold	indep-hold	partial-retract	prep	retract	stroke
V-O-I-D	1.875	109	0	0	20	12	49
hold	83	484	7	12	7	10	17
indep-hold	0	14	2	0	3	4	5
partial-retract	0	0	0	32	0	0	0
prep	37	15	9	0	274	1	78
retract	32	2	0	0	0	201	16
stroke	7	38	3	10	120	17	529

Save..

ANVIL 5

Results

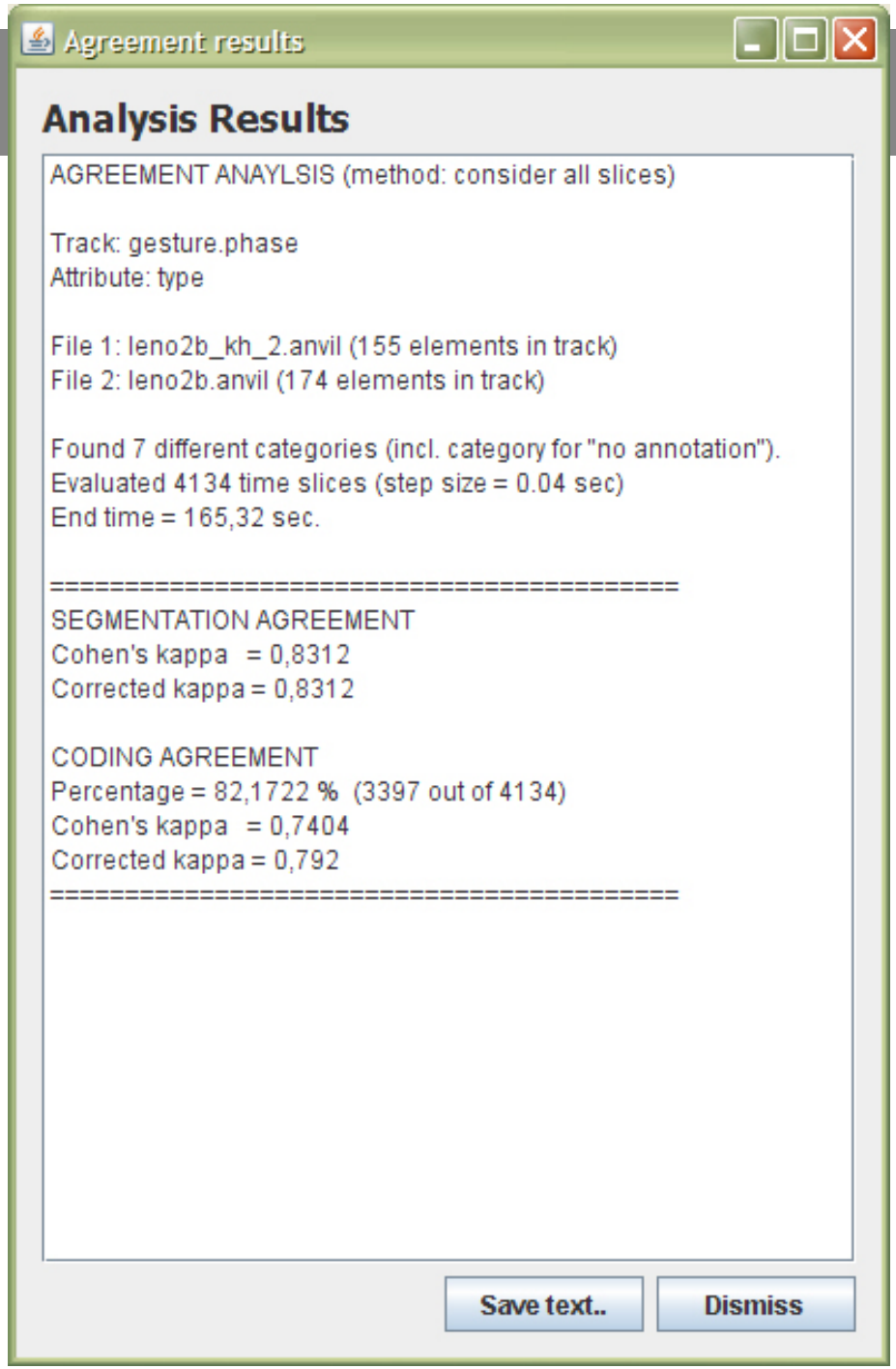
- Cohen's kappa
  [Cohen 1960]

$$\frac{P_0 - P_e}{1 - P_e}$$

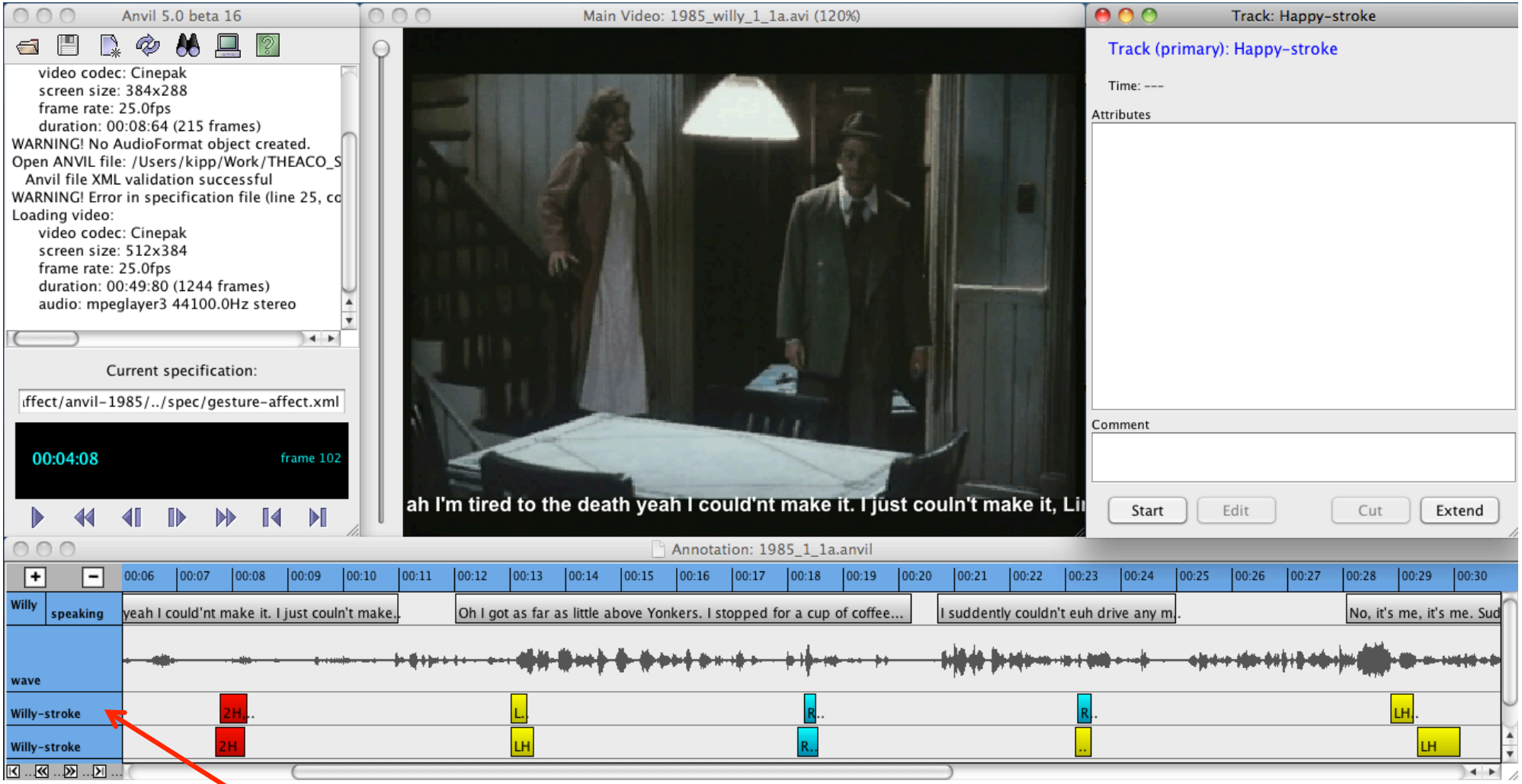
- Corrected kappa
  [Brennan, Prediger 1981]

$$\frac{P_0 - (1/z)}{1 - (1/z)}$$

- Segmentation only: treats all values as 1
- Also: compare whole projects



Beta feature: View multiple tracks

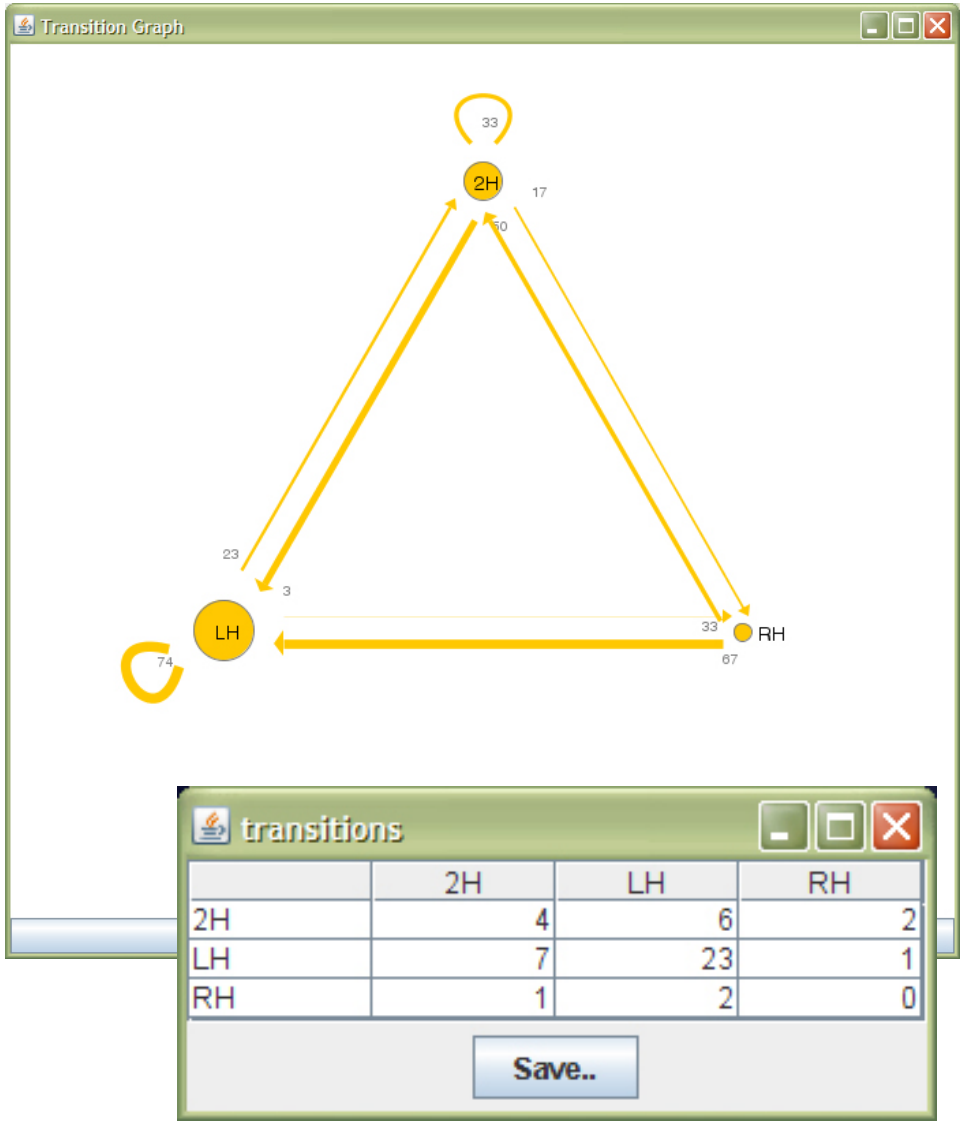


Annotations by two different coders

Right-click here: import track...

Transition Diagrams

- Models transitions between states
  - e.g., changing hands (LH, RH, 2H)
- Compute bigram (Markov) model from data
- Transition probablity from state A to B given as percentage on edge



➔ DEMO, JL

ANVIL 5

Association

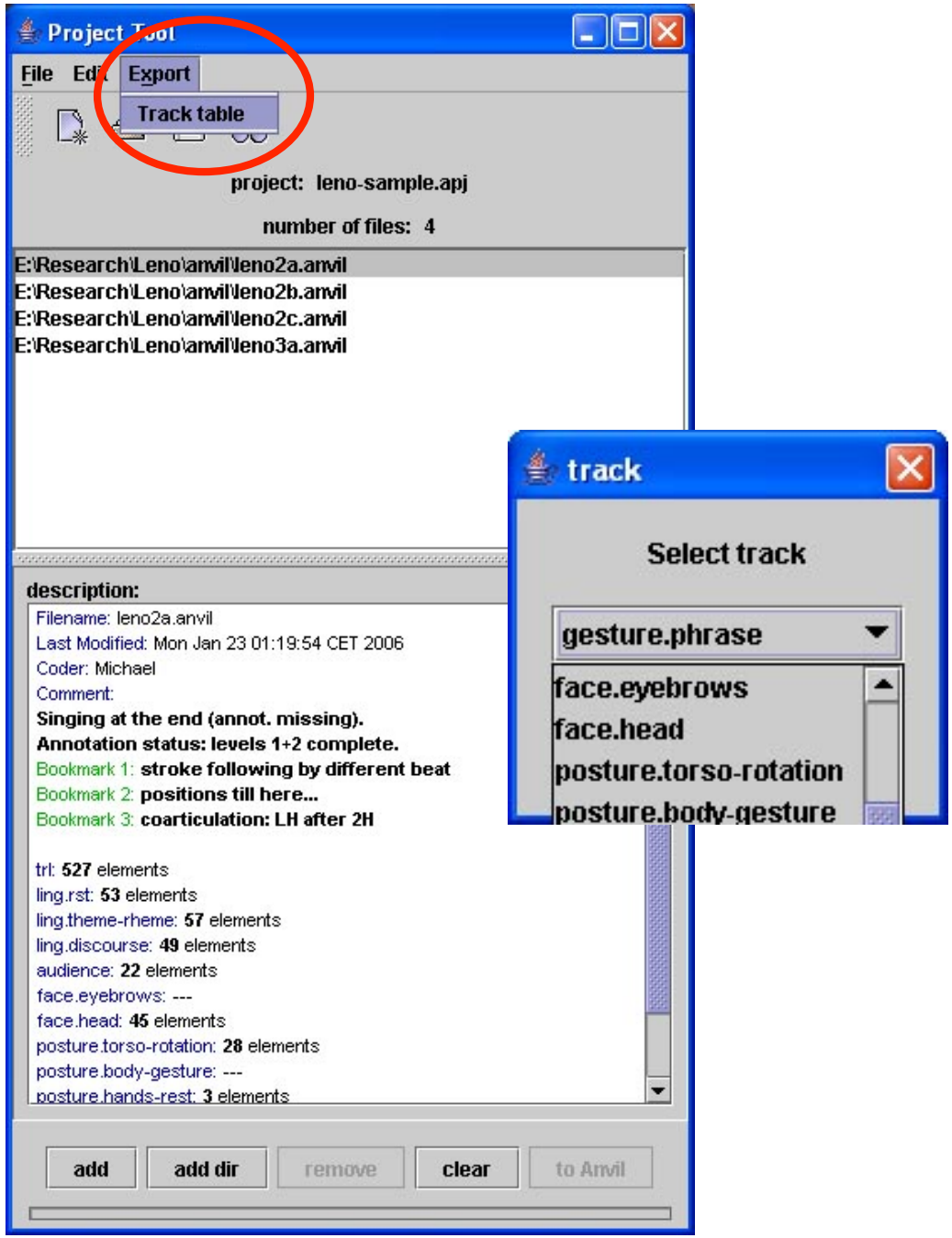
- Compare association between events on two tracks
- Example: gesture + emotion
- Step 1: decide which entities to compare
  ➔ SQL query
- Step 2: build contingency table
- chi square analysis
- compare tables for precise interpretation

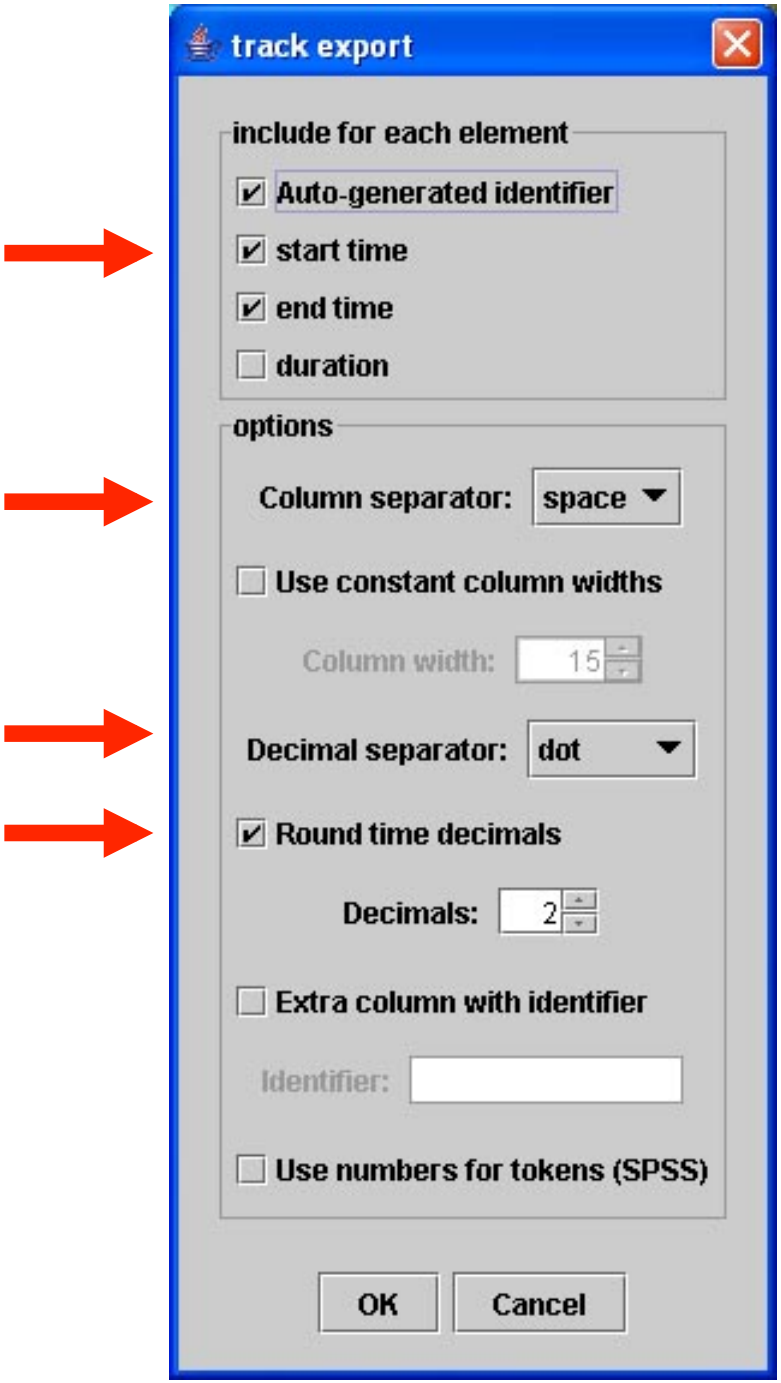
	LH	RH	2H
Happy	12	4	1
Angry	5	2	20

	LH	RH	2H
Happy	6.6	2.3	8.1
Angry	10.4	3.7	12.9

Export / Analysis

- Single track → text table
- Many tracks → text table
- Import to Excel, Statistica, SPSS
- Export formats:
  - element-by-attribute matrix
  - framenumber-by-attribute matrix
  - WEKA (machine learning software)
- Alternatives:
  - use XML stylesheets
  - Anvil plug-ins (Java)


Project Tool
File
Edit
Export
Track table
project: leno-sample.apj
number of files: 4
E:\Research\Leno\anvil\leno2a.anvil
E:\Research\Leno\anvil\leno2b.anvil
E:\Research\Leno\anvil\leno2c.anvil
E:\Research\Leno\anvil\leno3a.anvil
description:
Filename: leno2a.anvil
Last Modified: Mon Jan 23 01:19:54 CET 2006
Coder: Michael
Comment:
Singing at the end (annot. missing).
Annotation status: levels 1+2 complete.
Bookmark 1: stroke following by different beat
Bookmark 2: positions till here...
Bookmark 3: coarticulation: LH after 2H
trl: 527 elements
ling.rst: 53 elements
ling.theme-rheme: 57 elements
ling.discourse: 49 elements
audience: 22 elements
face.eyebrows: ---
face.head: 45 elements
posture.torso-rotation: 28 elements
posture.body-gesture: ---
posture.hands-rest: 3 elements
add
add dir
remove
clear
to Anvil
track
Select track
gesture.phrase
face.eyebrows
face.head
posture.torso-rotation
posture.body-gesture


track export
include for each element
Auto-generated identifier
start time
end time
duration
options
Column separator:
space
Use constant column widths
Column width:
15
Decimal separator:
dot
Round time decimals
Decimals:
2
Extra column with identifier
Identifier:
Use numbers for tokens (SPSS)
OK
Cancel

Element-by-attribute matrix

ID	start	end	lexeme	handedness	path
f0e0	0.76	2.32	Cup	LH	straight
f0e1	5.76	7.52	Frame	2H	straight
f0e2	13.08	15.36	Wipe	LH	straight
f0e3	15.36	16.04	Beat	LH	straight
f0e4	16.04	17.16	PointingHere	LH	straight
f0e5	17.36	18.72	PointingHere	LH	straight
f0e6	18.72	19.24	Erruptive	2H	curved
f0e7	19.24	20.32	ProgressiveSym	2H	curved
...					
...					
...					

WEKA Export

- WEKA
  - Decision Trees
  - SVM Classifiers
  - Rule Learning
  - …
- Format: ARFF

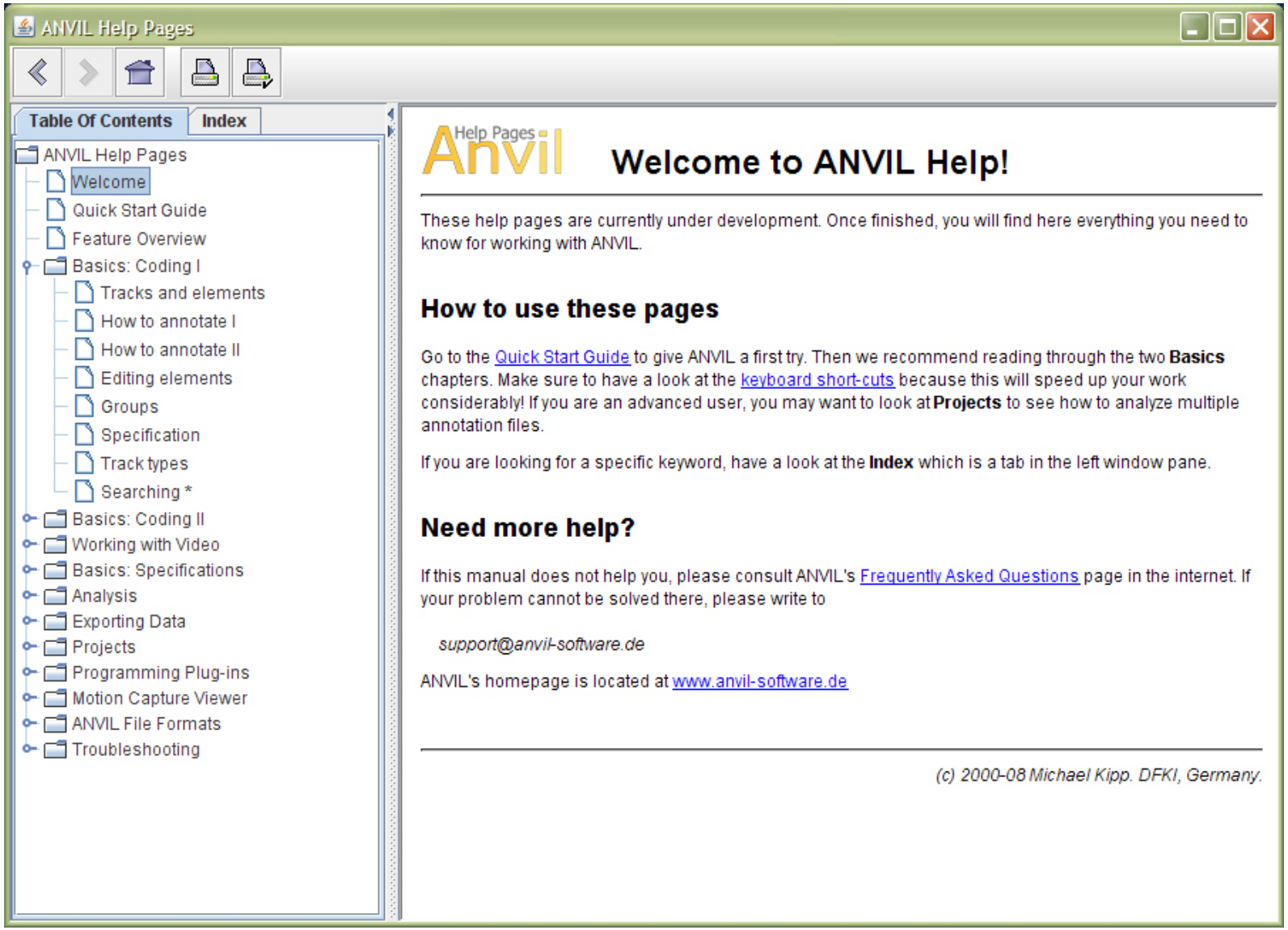
```
@relation gesture.phrase

@attribute handedness { none, LH, RH, 2H }
@attribute path { none, straight, curved }
@attribute handshape { none, open-flat-relax,
                open-flat-relax-thumb, open-flat-tense, ... }

@data
2H, straight, open-flat-relax
LH, straight, open-flat-relax
LH, curved, open-flat-relax
LH, straight, open-flat-tense
…
```

More ANVIL 5 Features

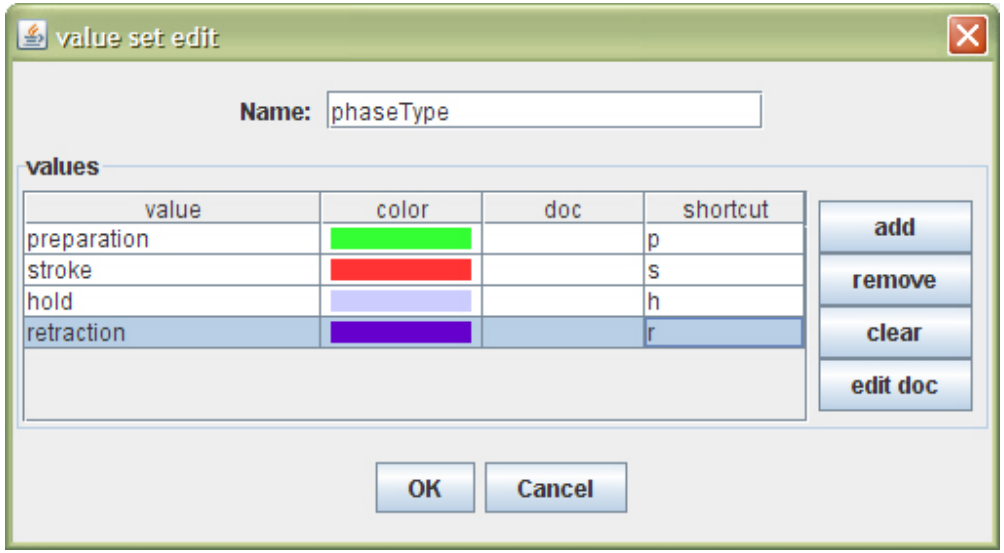
Online Manual Inside



ANVIL 5

Keyboard Shortcuts

- Drastic speed-up in coding!
- Shortcuts for placing start / end (F1 + F2) (on a Mac: 1 + 2)
- Shortcuts for categories
  - track-specific!
  - e.g. p=preparation, s=stroke, r=retraction



➔ DEMO, JL

ANVIL 5

Database Integration

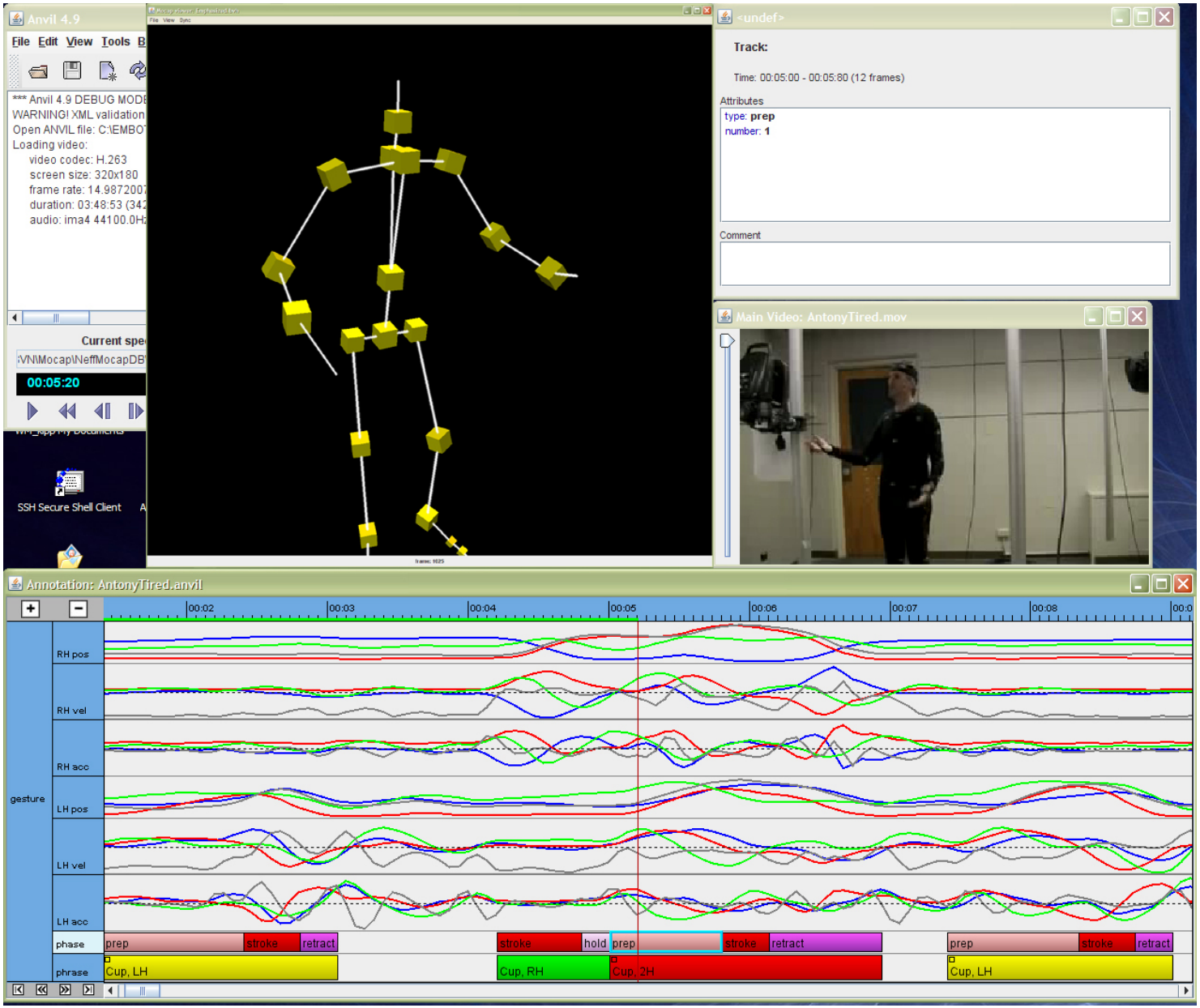
- Annotation mapped to temporary SQL DB
  - Track = table
  - Annotation element = row
  - Attribute = column
- SQL queries for track elements (by attribute)
  - [mytrack , (att1 = 2H OR att1 = LH) AND anotherAtt <> null]
- SQL query for temporal relation:
  - R[overlaps, .8] [firstTrack, hand = 2H] [otherTrack, hand <> null]

temporal relation *restricts track A* *restricts track B*

Mocap Viewer

- Accepts data in BVH and ASF/AMC
  - ASF: skeleton morphology
  - AMC: motion data
- Java3D
- Synchronization with video possible
- Allows full 3D view on person
- 3D traces (curves) of joints
  - position (x / y / z)
  - velocity (x / y / z / total)
  - acceleration (x / y / z / total)

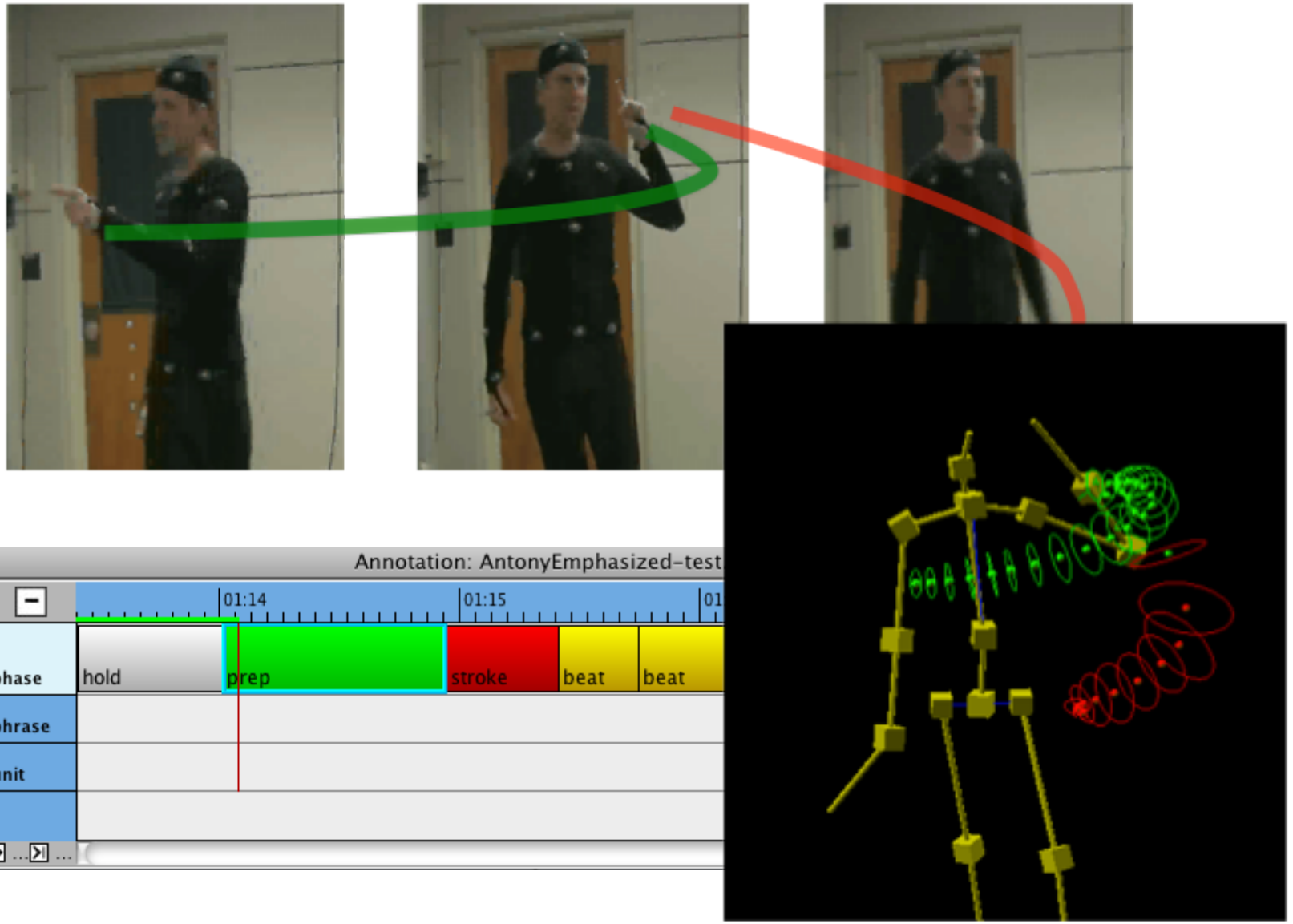
Mocap Viewer



→ DEMO

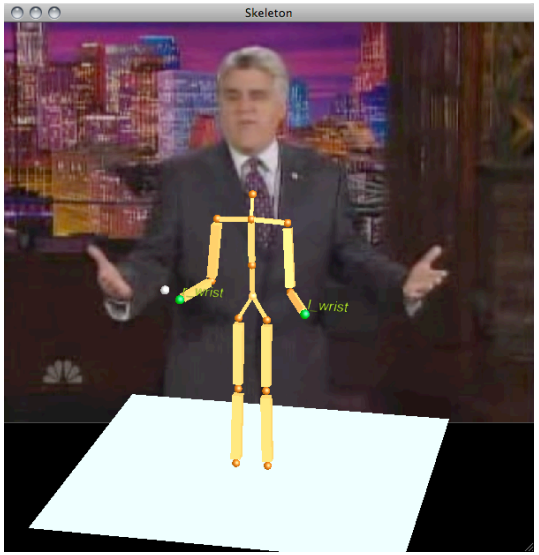
ANVIL 5

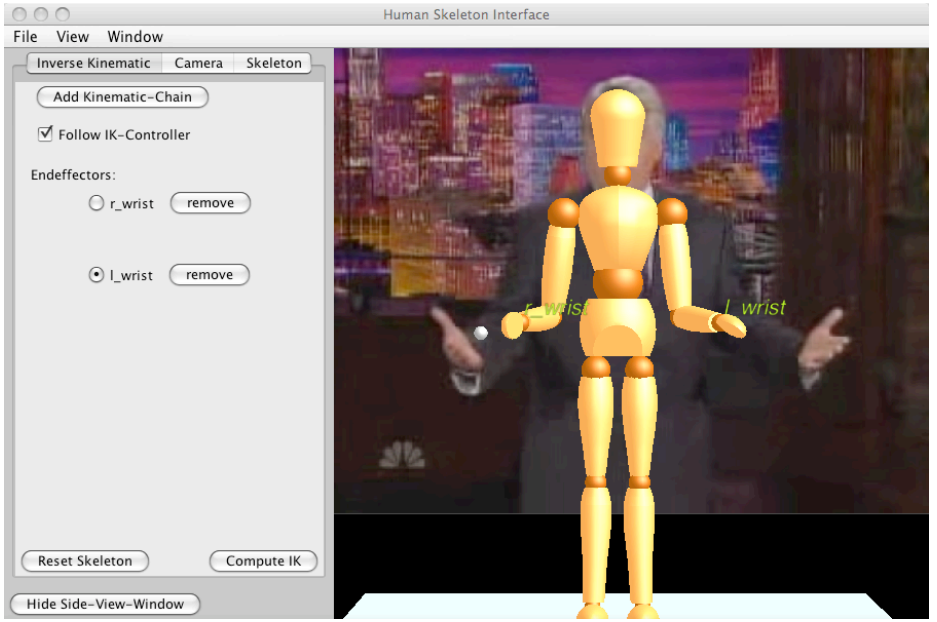
Motion capture



Skeleton-based annotation

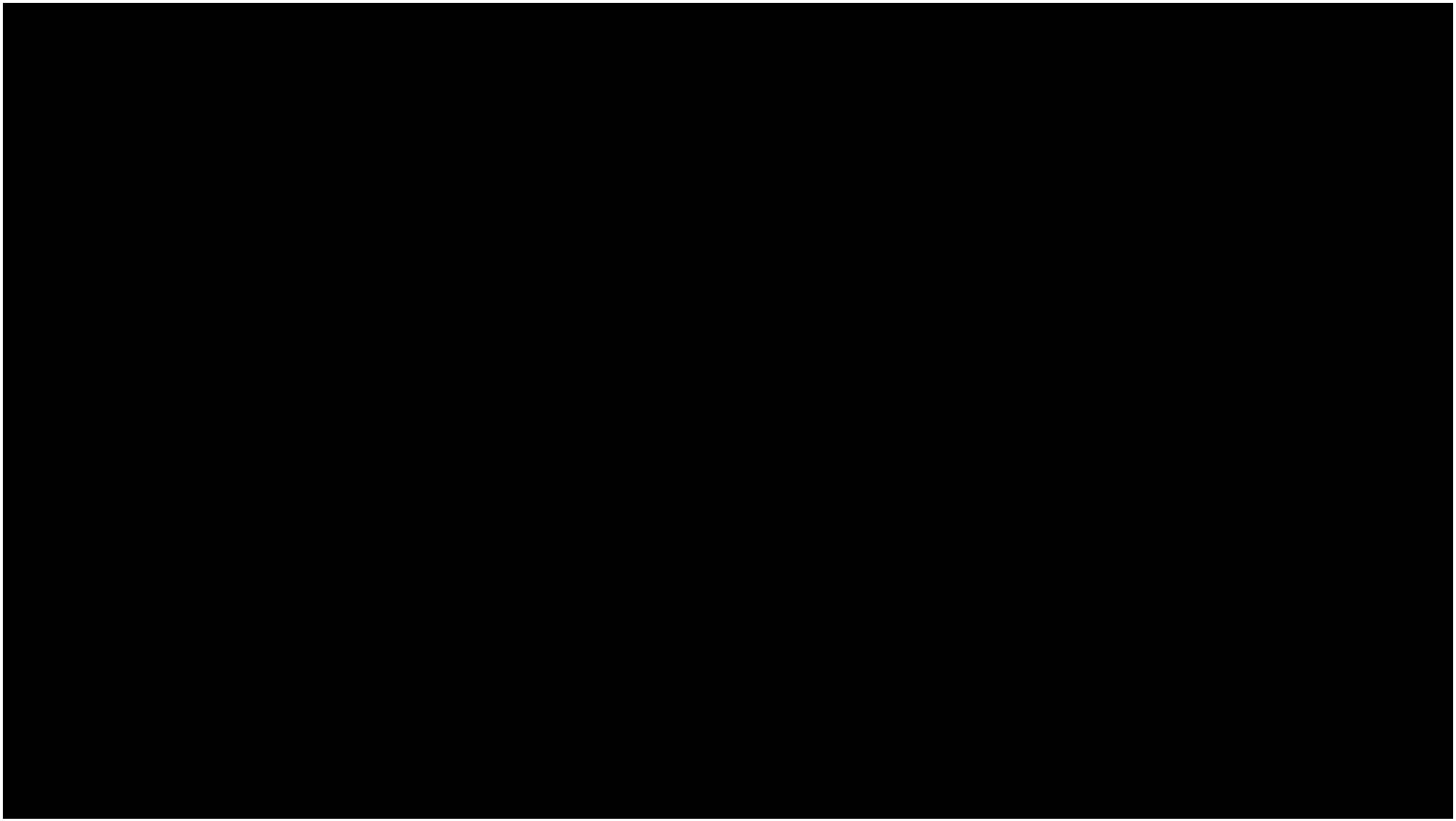
- Idea
  - add 3D key poses
  - Interpolate in-betweens
- Potentials
  - animation
  - motion templates
  - motion analysis





ANVIL 5+ ?

Gantool



User participation

Anvilforum

Search... Search
Advanced search

Board index

FAQ Register Login

It is currently Fri Nov 19, 2010 8:27 am

View unanswered posts • View active topics

WELCOME	TOPICS	POSTS	LAST POST

GENERAL	TOPICS	POSTS	LAST POST
ANVIL News New ANVIL versions, new features, new plans...	3	3	by Michael Fri Jul 23, 2010 10:06 am
Everybody's News Announce your annotation projects, publications and corpora...	1	1	by jean-claude.martin Thu Sep 23, 2010 3:54 pm
Feature wishlist If you have urgent features, killer features, nice-to-have features... here's the place to leave them.	1	1	by Michael Wed Jul 21, 2010 5:58 pm

TECHNICAL ISSUES	TOPICS	POSTS	LAST POST
Installation Problem due to operating systen, error messages, work-arounds...	5	12	by Bart Thu Sep 23, 2010 12:15 pm
Annotation/Coding Annotating, creating specifications, adjusting the annotation board, problems with speed ...	3	11	by sgilbert92 Thu Nov 04, 2010 12:20 pm
Video Video conversion, finding codecs, difference between operating systems, video length, video playback etc.	4	16	by jessyj Thu Nov 11, 2010 1:28 pm
Analysis Computing coder agreement, histograms, transition diagrams ...	4	4	by paul Wed Oct 20, 2010 3:08 pm
Anything else... If none of the above categories apply, click here...	3	3	by **Quan** Wed Oct 20, 2010 4:10 pm

RESEARCH ISSUES	TOPICS	POSTS	LAST POST
Application in research Questions related to how to make a coding scheme for a particular research area, how to conduct practical coding or perform analyses.	0	0	No posts
Related tools Tools related to ANVIL	0	0	No posts

ANVIL and ELAN

- ANVIL: user interface + many features
- ELAN: stable, many videos formats supported
- Use best of both:
  - import ELAN files to ANVIL
- Why?
  - kappa statistics on multiple files
  - search over multiple files
  - export for WEKA
  - add 2D annotations
  - mocap annotation
- Future: ELAN exporter ☺

Future Work

- Semi-automatic coding:
  exploit motion capture
  - determine handedness automatically
    (simple rule, 83% correct, 269 items)
  - detect movement phases (preparation, stroke...)
- Sign language transcription
  - sentence-oriented transcription (Exmaralda-like)
  - connection to animation
- Interoperability
  - lessons from import/export
  - towards a unified European framework
    (ELAN + Exmaralda + ANVIL + ...)


Thank you
for your attention
kipp@dfki.de
www.anvil-software.de
http://www.anvil-software.de
© R.L. Segal