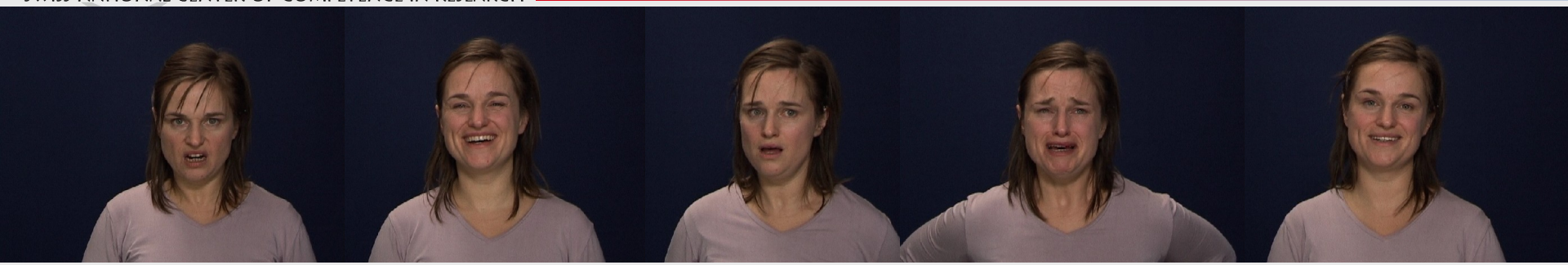
Affective Sciences
SWISS NATIONAL CENTER OF COMPETENCE IN RESEARCH

GEMEP



Multimodal data management and annotation
Introducing the GEneva Multimodal Emotional Portrayals

Martijn Goudbeek

Geneva Emotion Research Group / Swiss Center for Affective Sciences
Riederalp, September 2, 2008

Outline

- The GEMEP project
- Short introduction of the corpus
- Annotation
  - Ratings (judgment studies)
  - Acoustic and phonetic annotation
  - FACS coding
  - Gesture annotation
- Current practices and outlook
- Results

People involved

- Project leaders
  - Klaus Scherer & Tanja Bänziger
- The GEMEP group:
  - Nele Dael, body posture and gesture
  - Eva Krumhuber, face and dynamics
  - Marc Mehu, face and social interaction
  - Marcello Mortillaro, multimodal interaction
  - Martijn Goudbeek, speech and data management
  - Lucas Tamarit, technical support
- Internal and external collaborators
  - Antonio Camurri, Jeffrey Cohn, Thomas Ethofer, Donald Glowinski, Jean Philippe Goldman, Didier Grandjean, Michael Kipp, David Sander, Hannes Pirker, Valentijn Visch, Gualtiero Volpe

Design

- 10 actors
- 4 simultaneous recordings

(+ audio)

- In interaction (Stanislawski)
- 18 emotions
- 3 verbal contents:
  - Né kal ibam soud molén!
    (I don't believe it!)
  - Koun sé mina lod bélam ?
    (Do you really think that?)
  - AAA

valence

Arousal	positive	negative
high	elation amusement pride	hot anger (rage) panic fear despair
low	(sensual) pleasure relief interest	cold anger (irritation) anxiety (worry) sadness (depression)

Additional states:
shame, surprise, disgust
admiration, contempt, tenderness

Some examples

Some examples of the GEMEP portrayals

Rating studies

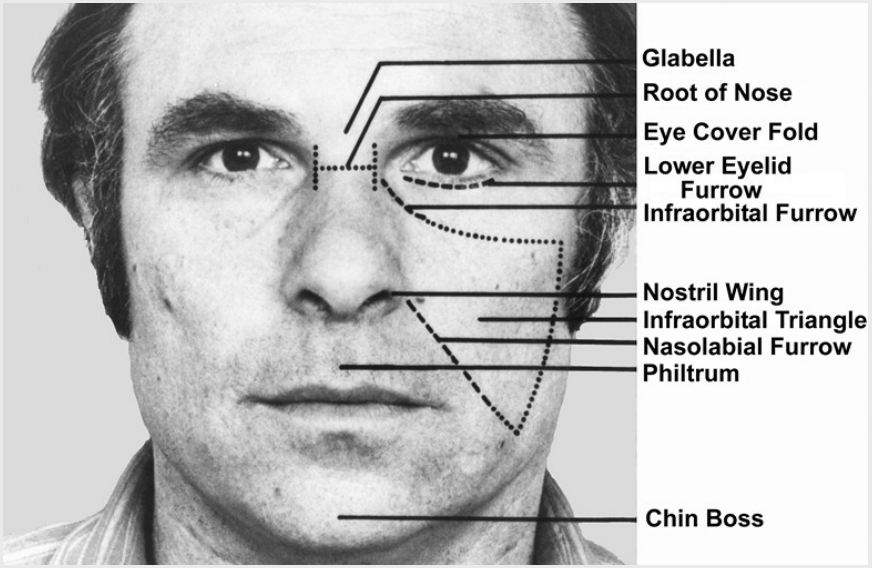
- First task: **selection** of subsets / core set out of large amount of content (25'000 portrayals)
- 1'260 portrayals (standard content) selected for rating studies
- Following this, more detailed ratings on a subset of the best (150) portrayal
- Ratings of audio/video, audio-only, video-only
  - *recognition* of expressive intentions, believability
  - *accuracy* of communication (cat. forced choice)
  - dimensional ratings (valence, arousal, naturalness)

Annotation and coding

- The 150 coreset portrayals are being annotated and coded in terms of:
  - Facial expressions
  - Acoustic parameters (1260)
  - Phonetic description and segmentation (1260)
  - Body and gesture
- Will be used as:
  - stimulus material in recognition / neuroimaging experiments
  - as a database for the community (e.g., automatic facial, vocal and bodily annotation)

FACS Coding

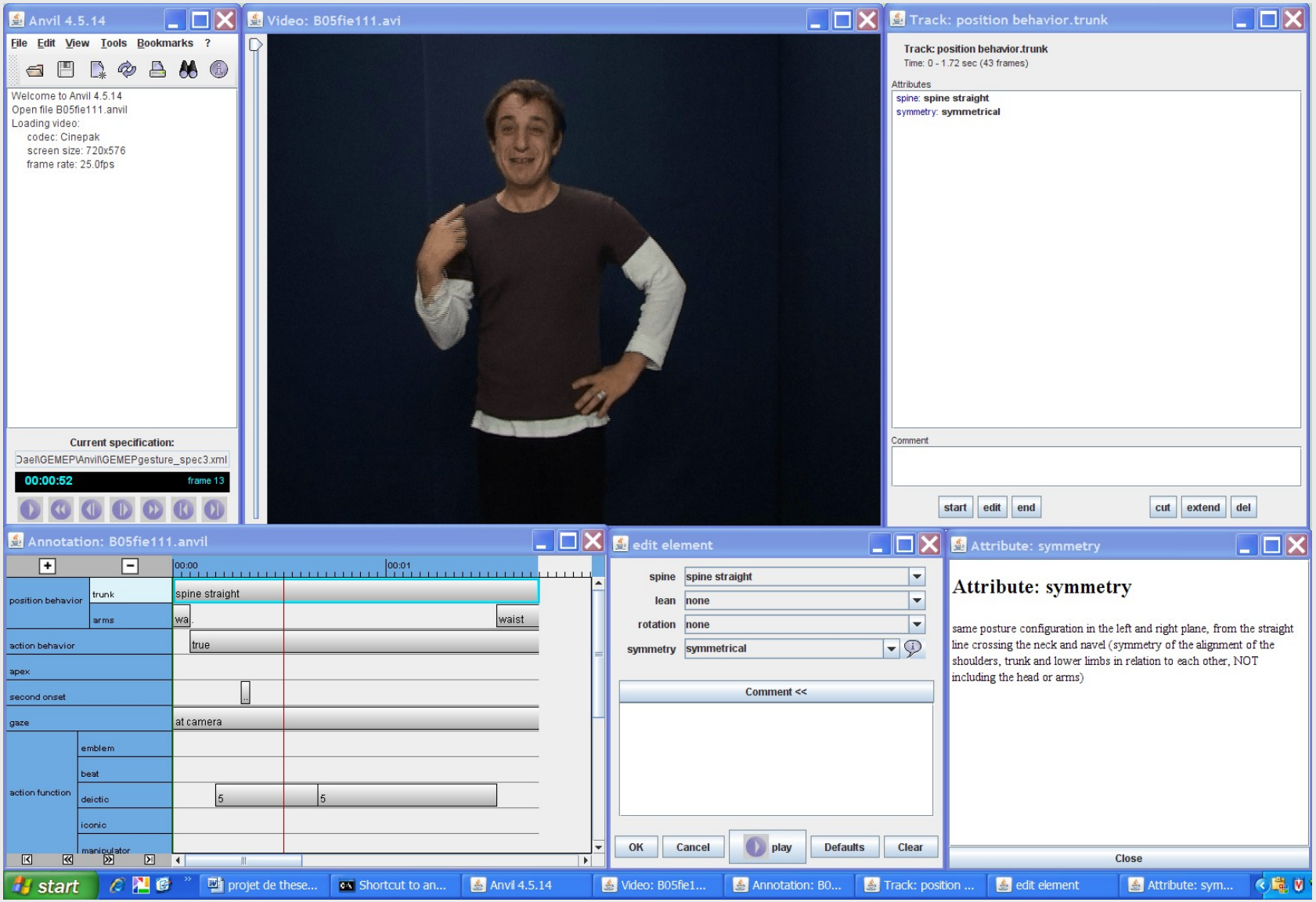
- Facial Action Coding System (Ekman & Friesen, 1978)
- Manual frame by frame coding of AU and intensity
- Multiple FACS coders: reliability measures
- Text based output



- Plus: automatic annotation

Graphical user interface for manual coding (Anvil)

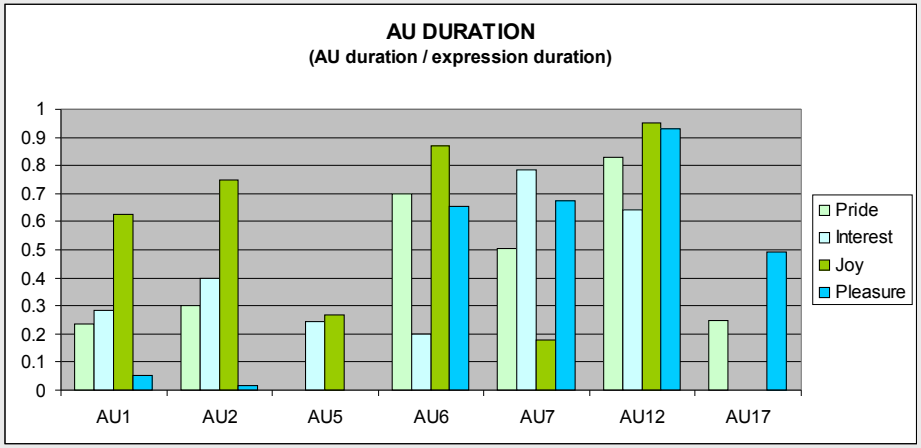
Affective Sciences
SWISS NATIONAL CENTER OF COMPETENCE IN RESEARCH



FACS coding: first results for positive emotions

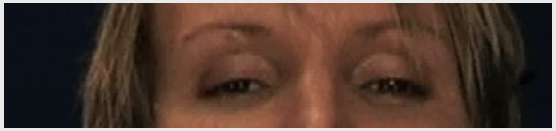
INTEREST –JOY – PRIDE – PLEASURE

These emotions are not clearly associated with emotion specific facial configurations.



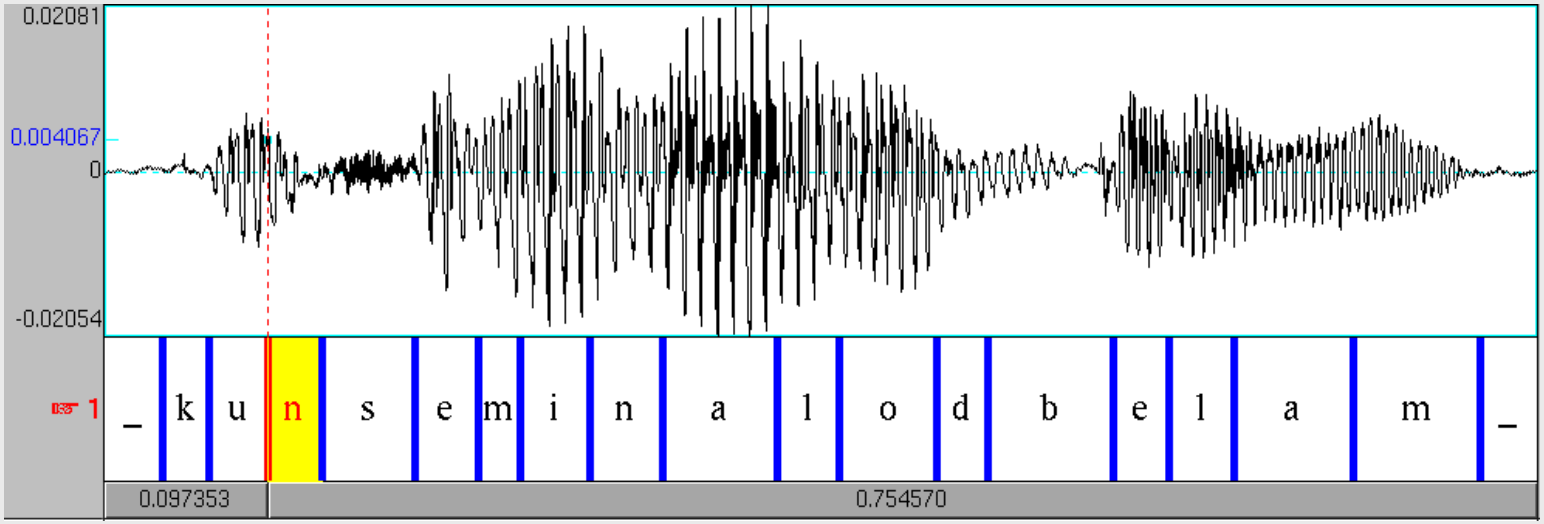
In these cases, facial expressions are more easily differentiated according to the underlying appraisal checks.

NOT SUDDEN ⟶ SUDDEN



Phonetic annotation

- Manual annotation of 1260 utterances:
  - Phonetic/syllabic segmentation of the corpus
  - Voice quality annotation (breathy, creaky, harsh)
  - Additional information
- Output: PRAAT/SFS annotation files, text



Acoustic annotation

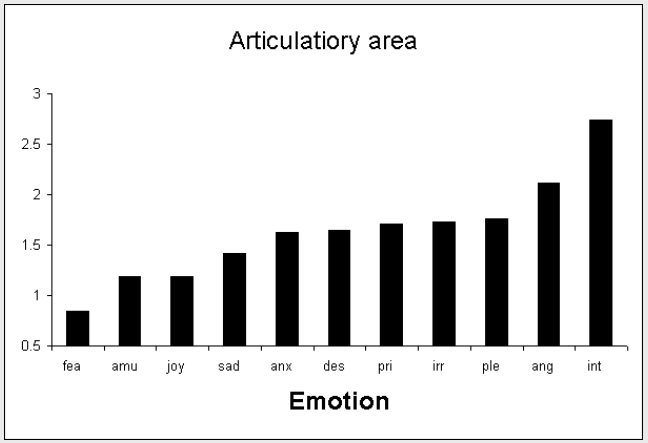
- Habitual acoustic measures (pitch, intensity, duration, spectral measures)
- Formant tracking
- Prosodic analysis
- Fine grained acoustic analysis at the phoneme level

- Output: Praat datafiles, text

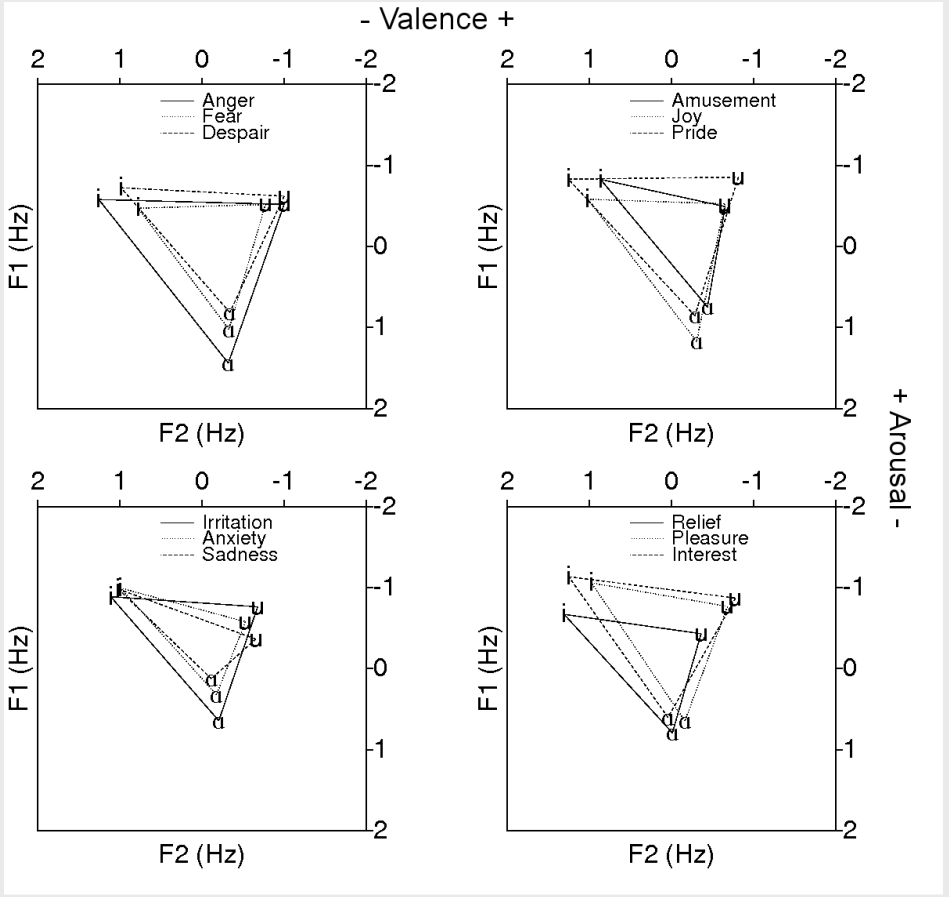
Acoustic analysis first results: articulation and potency

Emotions and vocalic area

We plotted the first and second formant of the vowels i, a, and u of the utterances present in the GEMEP corpus.



In line with the CPM, emotions differ in their size of their vocalic triangle: emotions with a high in poteny (anger, interest) have a larger vocalic triangle



Gesture annotation

- In-house developed Gesture coding scheme (Dael)
- Manual frame by frame coding of posture and movement
- Eclectic combination of gesture (deictics, emblems) and and posture (Wallbott & Scherer, 1994)
- XML / Text based output
- Plus: automatic annotation (see Dael/Glowinski)

Gesture coding: first results for anger and pride

Marked difference in trunk lean for emotion: forward leaning only in hot anger, backward leaning in pride



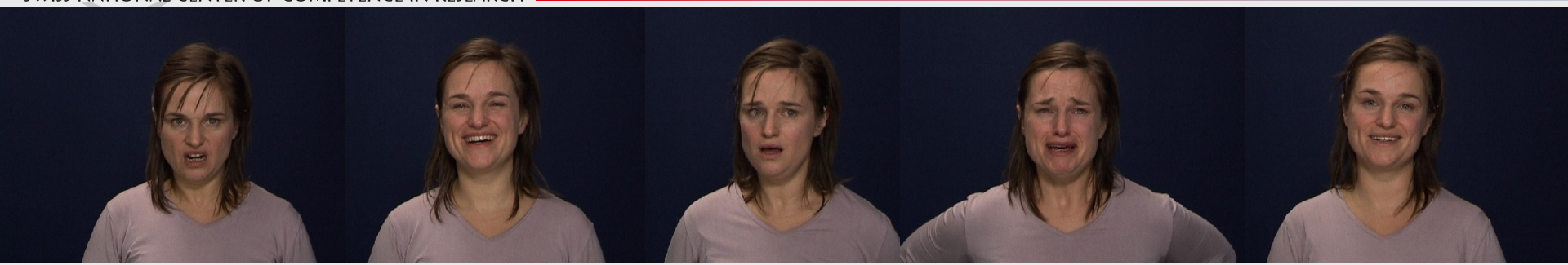

In sum: the content of our multimodal database

- Video and audio portrayals (1260/150) of 18 emotions portrayed at four intensity level and three utterances
- Rating data
- Phonetic annotation
- Acoustic annotation
- FACS coding (manual and automatic)
- Gesture coding (manual and automatic)

- How are we currently managing and organizing this data?

I wish, I wish, I wish...

- Future plans:
- Integrating the database application that enables:
  - Searching and selecting/downloading
  - Contributing new annotations
  - Viewing portrayals/annotations (multiple viewers)
  - Selection at the frame level ?
  - Exploratory analysis?
  - Displaying annotation data in graphical format



Thank you for your attention