

humaine

WP 5:Data and Databases

WP5: The Context

◆ **Databases in the HUMAINE Context**

- ➔ Databases should represent the type and range of emotional colouring that characterise naturally occurring emotional states

◆ **Challenges**

- ➔ address the issue of acquiring naturalistic data
- ➔ address the need for multimodal data & range of modalities
- ➔ address the labelling issues

◆ **Basis for Development**

- ➔ relevant expertise in the network
 - on collection of naturalistic data & multimodal data
 - on labelling

WP5: What we have done

◆ **Made available existing databases**

- ➔ via list of relevant databases on portal
 - ongoing project
 - WP1 will work on dedicated Plone data type for database list entries

- ➔ release of 2 datasets to HUMAINE community
 - AIBO database (CEICES agreement)
 - SAL (Sensitive Artificial Listener) database

WP5: What we have done

- ◆ **Established core principles of exemplar**
 - ➔ development of appropriate techniques NOT delivery of large fully labelled databases based on status quo
 - ➔ focus on induction techniques & labelling systems for induced and naturalistic data
 - ➔ a library of recordings and labelling schemes, not one unitary database
 - ➔ move from relatively manageable data to more complex data (from supportive to provocative)

WP5: What we have done

◆ Established core elements of exemplar

- ➔ a library of recordings that can underpin meaningful research into emotion as it appears in action and interaction in everyday life
- ➔ will cover signs in visual, audio, and physiological modalities
- ➔ will provide data across a range of cultures
- ➔ will provide descriptive schemes for labelling emotion expressed and signs of emotion
- ➔ these schemes will be applied to a subset of the data to show good practice

WP5: What we have done

◆ Taken core decisions on nature of data

Two types of data - induced data and naturalistic data

➔ Induction techniques

- Recall Technique piloted by the Tel Aviv team (emotion in monologue),
- SAL (Sensitive Artificial Listener) Technique piloted by the QUB team (emotion in dialogue)
- Driving Simulator Technique piloted by QUB (emotion in action)

➔ Naturalistic data: three bodies of data

- the Belfast Naturalistic Database
- the EMoTV database
- the Reality TV collection of recordings

WP5: What we have done

◆ Taken core decisions on labelling approach

➔ Emotion content

- verbal categorical, broad dimensional, appraisal related labelling

➔ Signs of emotion

- low order, automatically generated
- high order, perceptually determined

➔ Context labels

- expression of emotion is dependent on context at multiple levels
- appropriate labels need to be developed

WP5: What we have done

Allocation of tasks to partners

DATA COLLECTION

Leading teams: QUB, LIMSI, TAU, TWENTE, ICCS,

Others contributing: UNIGE, FAU-Erlangen, TCD, Augsburg (?)

DATA LABELLING

Leading teams: QUB & CNRS + UNIGE (WG3.1)

Others contributing: FAU Erlangen, TAU, Augsburg (?)

+ input from KCL on Ethical Considerations

WP5: Work in progress

Data Collection: Development of Induction Techniques

- ◆ **The Recall Technique**
- ◆ **The SAL concept (Sensitive Artificial Listener)**
 - User interacts or 'converses' with a computer simulation consisting of four characters
 - The SAL presents four characters equipped with a set of characteristic responses encouraging the user to respond in differing emotional states
 - Each character tries to pull the user towards his/her own emotional state
 - The SAL has no intelligence, only prespecified stock responses



WP5: Work in progress

Data collection: Refinement/extension of induction techniques

- ◆ SAL data
 - ➔ QUB has refined SAL technique & collected & distributed data
 - ➔ QUB & TWENTE initial collaboration on WoZ development of SAL
 - ➔ ICCS has translated SAL into Greek & done preliminary testing,
 - ➔ TAU has preliminary version of SAL in Hebrew
- ◆ RECALL Technique
 - ➔ TAU is refining for audiovisual use

WP5: Work in progress

Data Collection: Naturalistic Data

- ◆ Television data
 - ➔ QUB and CNRS collections of naturalistic television data
 - ➔ QUB & CNRS exchange of data
 - ➔ Reality TV data acquired
- ◆ Driving Simulator data (emotion in action)
 - ➔ Long term mood induction techniques developed & ready to test
 - ➔ Driving scenarios piloted

WP5: Work in progress

- ◆ Data Collection: other relevant data
 - UNIGE acted data
 - Augsburg physiological data
 - Erlangen databases
 - TAU computer game induction techniques
 - QUB adventure data



WP5: Work in progress

Labelling

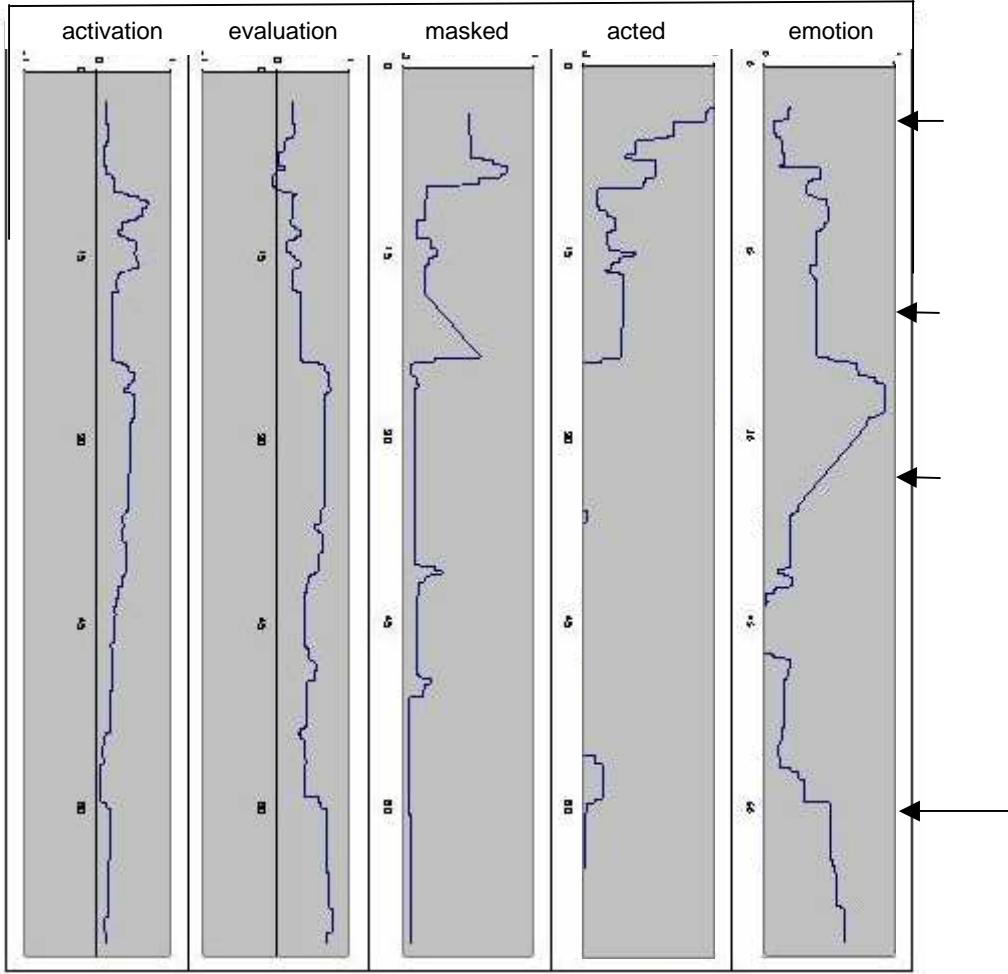
- ◆ Emotion labelling (R Cowie & C Cox leading with input from WP3)
- ◆ Context labelling (I Sneddon leading)
- ◆ Gesture/posture labelling (L Devillers, J-C Martin leading)
- ◆ High order speech labelling (E Douglas-Cowie)
- ◆ Low order labelling (WP4)

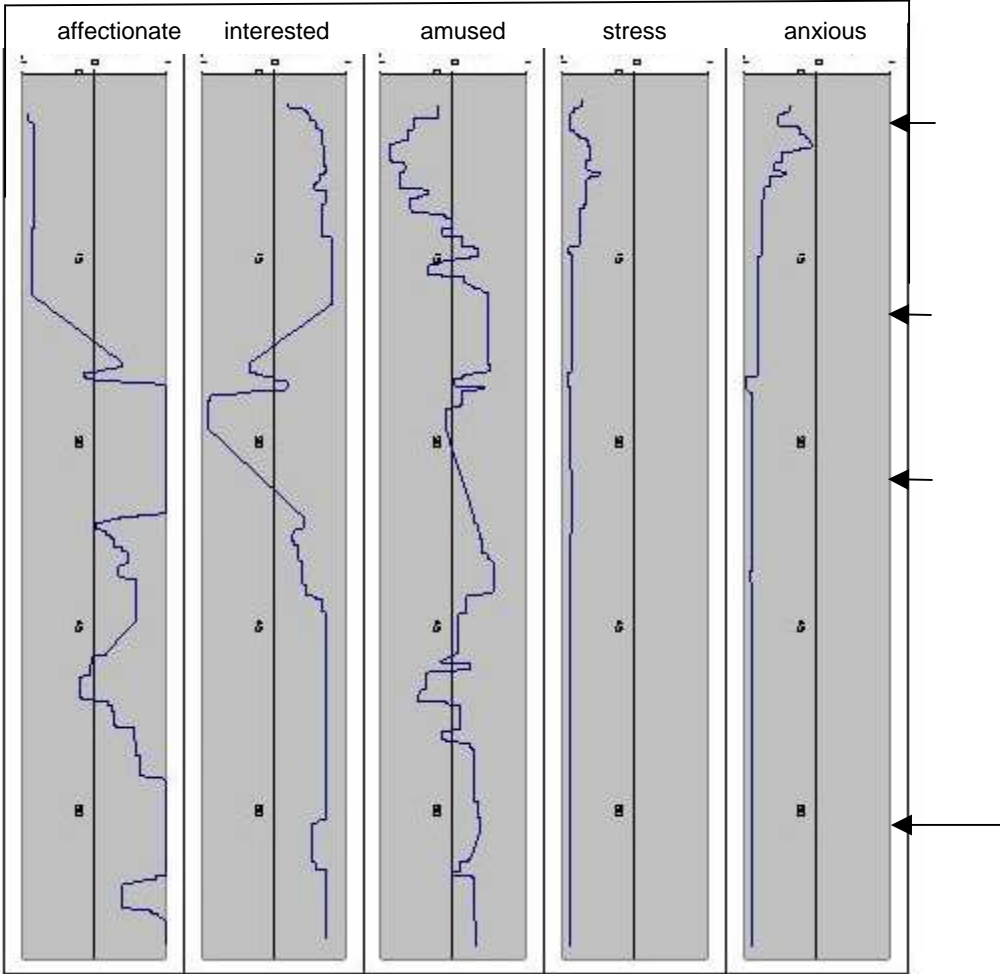
WP5: Work in progress

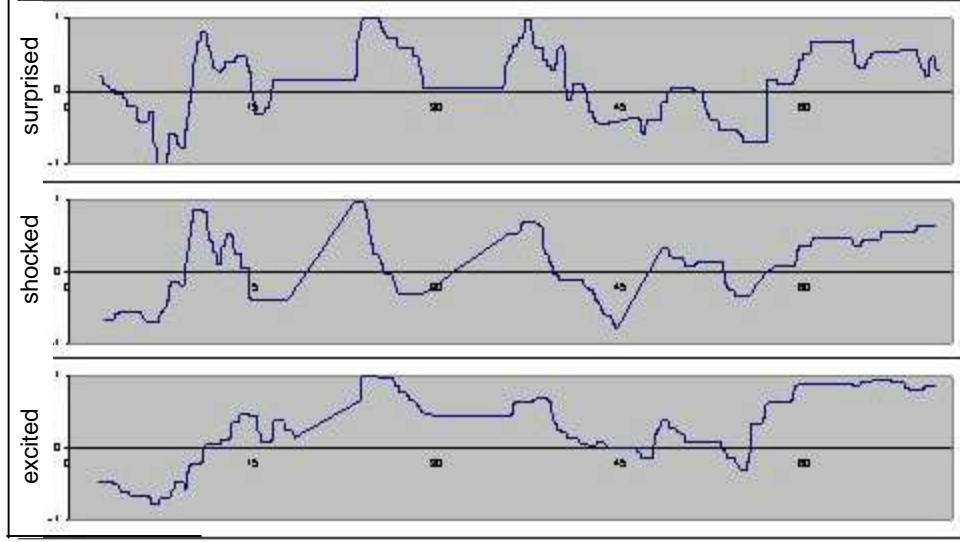
Emotion labelling

- ◆ Multi-level approach – to include broad dimensional, verbal categorical, appraisal related labelling
- ◆ Tools developed
 - ➔ Actrace
 - ➔ MaskTrace
 - ➔ ETrace+
 - ➔ AppTrace
 - ➔ Feeltrace

*Demonstrate **Actrace**, **ETrace** and **AppTrace** on 12d*







WP5: Work in progress

◆ Context labelling/Gesture labelling

- ➔ Agent characteristics (age, gender, race)
- ➔ Recording-context (recording style, acoustic quality, video quality)
- ➔ Intended audience (kin, colleagues, public)
- ➔ Overall communicative-goal (to claim, to sway, to share a feeling, etc).
- ➔ Social setting (none, passive other, interactant, group)
- ➔ Spatial focus (physical focus, imagined focus, none)
- ➔ Physical constraint (unrestricted, posture constrained, hands constrained)
- ➔ Social constraint (pressure to expressiveness, neutral, pressure to formality)

Jean-Claude Martin and Laurence Devillers will present gesture labelling at WP5 session tomorrow

WP5: Focus for tomorrow

- ◆ Timetable and specific tasks in data collection
 - ➔ Need to discuss physiological data collection
- ◆ Sharing of new tools for labelling
- ◆ Decisions on individuals leading different aspects of labelling