

## Data Resources, Coding Schemes and Coding Tools for Natural Interactivity

Presented by Laila Dybkjær and Niels Ole Bernsen NISLab, Odense, Denmark

COCOSDA meeting, Denver, 16 September 200



## **This presentation**

- Background: the ISLE project
- Approach
- NIMM data resources
- NIMM annotation schemes
- NIMM annotation tools
- Conclusions and ongoing work



# The ISLE project

- International Standards for Language Engineering (ISLE) 1.1.2000 – 31.12.2002
- Successor to Eagles I and II
- Three working groups: lexicons, evaluation, NIMM
- NIMM focuses on data resources, coding schemes, annotation tools and meta-data
- US NIMM loosely connected with EU NIMM
- isle.nis.sdu.dk

# Euro an ISLE Ning

DA meeting, Denver, 16 September 200



#### Approach

- Selection criteria
- Common description template
- Interaction with, and verification by, developers whenever possible

- Information on creation, purposes, use, user, needs, best practice, standardisation, etc.
- For data resources and coding schemes division into facial expression (plus speech) and gesture (plus speech)



## **Data resources (1)**

- Information on 64 data resources world-wide
- Survey of market and user needs done by ELRA
- 28 filled questionnaires collected at Dagstuhl workshop

 Resources primarily found via web and via proceedings



## Data resources (2)

36 facial and 28 gesture resources:

- Dynamic facial resources
  - with audio
  - without audio
- Static facial resources
- Lesser known/used facial resources
- Gesture resources
- Lesser know/used gesture resources



## **Data resources (3)**

- Static images:
  - image analysis, specialised databases, special conditions (light, size, face orientation. ...)
- Dynamic video (mostly with audio):
  - e.g. lip movements during speech, audiovisual speech recognition, correlation between speech and gesture



# Needs and purposes (1)

- The collected data resources reflect among others the following needs and purposes:
- Automatic analysis and recognition of facial expressions, including lip movements
- Audio-visual speech recognition
- Study of emotions, communicative facial expressions, phonetics, multimodal behaviour, etc.



# **Needs and purposes (2)**

- Creation of synthetic graphical interface characters, including, e.g., talking heads
- Automatic person identification
- Training of speech, gesture and emotion recognisers
- Multimodal system specification and development



#### Market study

- Questionnaire sent to more than 150 people
- 25 answers received
- Questions on
  - types of data resources needed/used/offered
  - kinds of tasks for which data resources are well-suited
  - areas in which data resources are being used



#### **Types of data resources**

- Audio most popular (84%)
- Video (52%)
- Image (28%)
- Advantage if a resource is annotated
- Many users produce their own resources



#### Tasks resources are used for

- Six general categories in questionnaire with several possibilities subsumed
- Authentication: Speech verification
- Recognition: Speech recognition, face recognition
- Analysis: Speech/lips correlation
- Synthesis: Multimedia development
- Control: Voice control
- Other: Information retrieval





## **Coding schemes**

Information on 21 coding schemes

- 7 facial
- 14 gesture

Nearly all schemes meant for video markup

- A couple of them can also be used for static image markup
- Schemes for single modality and modality combination markup



#### **Purposes and practices**

Many coding schemes tailored to particular task and only used by creator or creator's site

Few coding schemes fairly general and frequently used



#### **Facial coding schemes**

- MPEG-4 is considered a standard
- FACS used by many (not good for lip movements)
- ToonFace used by some (2D caricature)



#### **Gesture coding schemes**

- Very often combined with other modalities
- Only sign language focuses on gesture alone
- No real standardisation for gesture markup HamNoSys most frequently used among schemes looked at
- Many gesture schemes created to support application development – used by one or a few people
- Picture supported by answers to resource questionnaires at Dagstuhl workshop



## **Coding tools**

- 12 coding tool projects reviewed
- All available tools tested with focus on functionality and interface
- Most tools run on Windows, but java-based tools are cross-platform
- XML is very frequently used for data representation
- Some tools built for a very specific purpose while other tools aim at more general use



COCOSDA meeting, Denver, 16 September 200



#### Conclusions

- Non-exhaustive information collection
- Focus on accessible data resources and coding schemes
- People mostly create own data and schemes
- Often tailored to specific purpose, e.g. focus on a particular kind of interaction, modality combination or lighting condition
- No strong guidance by best practice and standards and no particular focus on sharing data and schemes
- On the other hand, ELRA and LDC exist



# **Ongoing work**

- Guidelines for the creation of NIMM data
  resources
- Guidelines for the creation of NIMM annotation schemes
- We are collecting information on ongoing standardisation activities
- We would greatly appreciate any information you may have on such activities
- We would also greatly appreciate input on draft versions of our work
- Please contact: laila@nis.sdu.dk



## **Standardisation efforts (1)**

- TC37/SC4: ISO committee dedicated to the representation and management of language resources which are the results of manual annotation or automatic processes
- OLAC (Open Language Archives Community): international partnership of institutions and individuals creating a worldwide virtual library of language resources by:
  - developing consensus on best current practice for the digital archiving of language resources
  - developing a network of interoperating repositories and services for housing and accessing such resources.



## **Standardisation efforts (2)**

- DCMI (Dublin Core Metadata Initiative): open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. Aims to promote widespread acceptance of metadata standards and practices.
- SPEX: founded in 1987 with the objective to develop and provide software, tools and databases for companies and institutes engaged in speech technology. Main activities at present are the creation, annotation and validation of spoken language resources.



## We need your help

- ISLE cannot update and extend the collected information
- If you have a data resource, coding scheme or coding tool you are willing to share and want others to know about

GO TO isle.nis.sdu.dk AND FILL IN INFORMATION ABOUT YOUR CORPUS, CODING SCHEME, OR TOOL



# **THANK YOU**

COCOSDA meeting, Denver, 16 September 200