LREC conference 2004

Peter Witteriburg
MPI, N megen

The Language Resource and Evaluation Conference (Lisbon, May 2004) is certainly one of the most relevant conferences for all matters regarding linguistics and language engineering resources. A number of issues were discussed that were of direct relevance for the work of the MPI team and also for the technological

aspects in the DOBES programme. Of course, we can only report a few selected topics:

- The ISO TC37/SC4 initiative is about to come up with frst standards. The nature of the Data Category Registry that is seen as a repository for all relevant linguistic concepts to increase semantic interoperability is widely agreed. The IMDI metadata concepts for example have already been inserted. The Lexicon Markup Framework is currently being worked out as a fexible lexicon format. Its design is infuenced by the requirements of the DOBES programme with all its difering lexicons. The MPI team is busy building one of the frst lexicon tools based on LMF.

- In the area of metadata description, IMDI and OLAC are well-known and widely recognized by the community. Increasingly more people realize the importance of describing their resources with the help of metadata to make them visible to others.

- It seems that the number of video annotation tools is decreasing. Some major developments in the last few years such as the ATLAS project from NIST and LDC have been stopped. ELAN seems to be one of the few that will be continued. Yet it is not clear why several projects have been stopped but video handling is still not simple and requires many resources.

- A very interesting panel organized by Brian McWhinney (Childes CMU) showed that there are two major ideas concerning how to achieve fexible collaborative environments: (a) peer-to-peer frameworks that allow collaboration via instant messaging, i.e. collaborative commentary or annotation is done in real-time; and (b) collaborative commentary on web-based presentations that are not done in real-time. The MPI team was one of the frst demonstrating a peer-to-peer collaborative annotation framework based on an ELAN extension.

The MPI team gave 13 presentations at the conference and conference workshops on topics such as metadata, multimodal annotation, the relevance of XML for archiving and distributed archiving as intended in DELAMAN. See http://www.lrec-conf.org/lrec2004/