CoCo-FLaRE

Reinforcing International Collaboration on Language Resources and Evaluation

A joint COCOSDA – FLaReNet Workshop

Saturday, May 26th, 2012

CALL for PARTICIPATION

Language Resources and Evaluation are gathering more and more momentum, with the increasing success of the LREC conference and of the LRE journal, and with new program initiatives flourishing in Europe (FLaReNet, META-SHARE, CLARIN, etc.), as well as in other parts of the world (TDIL, NHN, Language GRID, ALAGIN, BOLT, SILT etc.).

COCOSDA, the Coordinating Committee on Speech Databases and Speech I/O Systems Assessment, paved the way with its pioneering creation in 1991 to serve the international speech processing community.

Since then, COCOSDA continues to gather the international speech community under its informal umbrella, while the European initiatives extend their scope to language in general, through ELRA and LREC, and Oriental COCOSDA fosters spoken language resources with an increasing participation of Asian countries in its annual conferences.

This 2012 workshop will be the occasion to share views on ways to reinforce international collaboration in the field of Language Resources and Evaluation at a time when data sharing is coming to the forefront of science, in many different areas such as biology, astrophysics, humanities, etc.

The workshop will comprise two parts: a morning session specifically devoted to spoken language resources and evaluation, organized by COCOSDA, and an afternoon session enlarging the scope to speech & language in general. A final discussion will aim at proposing new measures to increase and improve international collaboration and coordination, in order to foster Language Resources and Evaluation.

The workshop is open to all interested parties at no charge, but registration is mandatory. Those wishing to contribute topics at the workshop are invited to contact the organizers by March 15th.

Organizers:
- Nicoletta Calzolari (glottolo@ilc.cnr.it) (Co-Chair)
- Dafydd Gibbon (gibbon@uni-bielefeld.de) (Co-Chair)
- Nick Campbell (nick@tcd.ie)
- Khalid Choukri (choukri@elda.org)
- Joseph Mariani (Joseph.Mariani@limsi.fr)

Morning session

"COCOSDA: New developments in multimodal resources and assessment"

This session will consist of regional and topical reports, with discussion, by COCOSDA members and invited presenters.

A central push forward in multimodal speech research has come from the increasing amount of work on the languages of smaller speech communities, resulting in large amounts of new data from typologically very different languages with very different sound inventories, prosody, grammar, and speech dialogue conventions.

Tentative Schedule:

Introduction (D. Gibbon)

Part 1: Topical reports
with a special emphasis on resources and evaluation for :

- speech interaction with robotic and virtual agents,

- types of multimodal systems,

- types of mobile systems,

- under-resourced languages.

Coffee break

Part 2: Regional reports
Presentations will summarize regional developments in Africa, the Americas, Asia, Europe, and Oceania in current major topic areas.

Part 3: Information exchange and discussion

Lunch Break

Afternoon session

"International collaboration and coordination on Language Resources and Evaluation"

This session will be devoted to the various aspects of international collaboration on Language Resources and Evaluation such as:
- Charting the field
- Ensuring the best possible language coverage
- Fostering language resources
- Sharing language resources
- Identifying and tracking the resources
- Recognizing the merits of LR producers and LT evaluators
- Networking the community, etc.

It will rely on the activities conducted in this area in various part of the world, including the recent conclusions and recommendations of the FLaReNet project.

Tentative Schedule:

Introduction (N. Calzolari)

Part 1: Information and Instruments
- Charting the field
  LRE Map and Language Matrices, FLaReNet International Wiki survey and National Contact Points, META-NET Language White Papers and Tables, ...
- Ensuring the best possible language coverage
  Highlights of National Programs on LR, Endangered and local languages, ...
- Sharing Resources
  META-SHARE, Language Grid, NSF policy, the Language Library, ...

Coffee break

Part 2: Actions
- Fostering language resources
  The FLaReNet Strategic Language Resource Agenda
- Identifying and tracking resources
  The new ISLRN initiative
- Recognizing the merits of LR producers and LT evaluators
  Towards a Language Resource Impact Factor?
- Networking the community
  The LRE Forum

Part 3: General discussion: *The way forward*