Report on the Special joint LTC-FLaReNet session
« Getting Less-Resourced Languages On-Board ! »
LTC'09 Conference
Poznan (Poland), 6-8 November 2009

A special session on the way to develop Language Technologies for Less-Resourced Languages was organized by FLaReNet within the LTC'09 conference in Poznan (Poland). It had been decided within the FLaReNet Steering Committee to organize a joint satellite workshop in connection with the Language & Technology Conference (LTC'09) in Poznan (6-8 November 2009). The choice was finally to organize it as a special session during the conference, on November 6.

The final conclusions extracted from the report of this event (written by J. Mariani (LIMSI-CNRS & IMMI) and K. Choukri (ELDA-ELRA) were the following:

  - Generally speaking, a strong political will (more than only lip-service) to consider the language dimension and enough funds are necessary.
  - This must go with the awareness that Language Technologies and Language Resources are important.
  - There should be specialists in the processing of that language, reaching a critical mass, and young researchers should be trained.
  - An infrastructure must exist, including:

  - a writing system/a transcription code/an agreed orthography,
  - Language Resources (sufficient in quantity and quality),
  - tools (especially language independent (based on statistical training) ones, if possible as Open Source),
  - metadata, annotation schemes, standards,
  - development platforms,
  - evaluation means (adapted to the language specificities (such as for Machine Translation of morphologically-rich languages)).

  - The effort should be devoted in the long-term, resulting in a necessary strong foundation.
  - Dialects variants and sociolinguistics should also be taken into account.
  - Addressing only the short-term development of a specific product or service for that language (as a kind of simple toy), should be avoided, whereas demonstrating applications based on a strong foundation should be favoured.

- When a majority language also exists, both should be studied together, and it would save time and efforts to consider a family of languages all together.
- Bootstrapping approaches facilitate the coverage of a language.
- Cooperation among countries or programs would greatly help by providing the less advanced ones with examples and Best Practices, such as the definition of a commonly agreed basic set of Language Resources which have already been proven necessary to correctly produce the corresponding technologies for a given language, and the identification of gaps and roadmaps should be aimed at.
- Master keywords should be Interoperability and Sustainability.