"Strategies for developing machine translation for minority languages"

==

5th SALTMIL Workshop on Minority Languages
on Tuesday May 23rd 2006 (morning)
Magazzini del Cotone Conference Centre, Genoa, Italy

Organised in conjunction with LREC 2006: Fifth International Conference on
Language Resources and Evaluation, Genoa, Italy, 24-26 May 2006

Background
=========

This workshop continues the series of LREC workshops organized by SALTMIL (SALTMIL is the ISCA Special Interest Group for Speech And Language Technology for Minority Languages: http://isl.ntf.uni-lj.si/SALTMIL/):

The minority or "less resourced" languages of the world are under increasing pressure from the major languages (especially English), and many of them lack full political recognition. Some minority languages have been well researched linguistically, but most have not, and the vast majority do not yet possess basic speech and language resources (such as text and speech corpora, lexicons, POS taggers, etc) which would enable the commercial development of products.

The workshop aims to share information on tools and best practice, so that isolated researchers will not need to start from nothing. An important aspect will be the forming of personal contacts, which can minimise duplication of effort. There will be a balance between presentations of existing language resources, and more general presentations designed to give background information needed by all researchers present.

Format
======

The workshop will begin with the following presentations from invited speakers:

- Delyth Prys (University of Wales, Bangor): "The BLARK matrix and its relation to the language resources situation for the Celtic languages."
- Hermann Ney (Rheinisch-Westfälische Technische Hochschule, Aachen, Germany): "Statistical Machine Translation with and without a bilingual training corpus"
- Mikel Forcada (Universitat d'Alacant, Spain): "Open source machine translation: an opportunity for minor languages"
- Lori Levin (Carnegie Mellon University, USA): "Omnivorous MT: Using whatever resources are available."
- Anna Sågvall Hein (University of Uppsala, Sweden): "Approaching new languages in machine translation."

These talks will then be followed by a poster session featuring contributed papers.

Papers
======

Papers are invited that describe research and development in the following areas:

- The BLARK (Basic Language Resource Kit) matrix at ELDA, and how it relates to minority languages.
- The advantages and disadvantages of different corpus-based strategies for developing MT, with reference to a) speed of development, and b) level of researcher expertise required.
- What open-source or free language resources are available for developing MT?
- Existing resources for minority languages, with particular emphasis on software tools that have been found useful.

All contributed papers will be presented in poster format. All contributions will be printed in the workshop proceedings (CD). They will also be published on the SALTMIL website.

Timetable:
========

- Paper submission deadline: Feb 17, 2006
- Notification of acceptance: March 10, 2006
- Final version of paper: April 10, 2006
- Workshop: May 23, 2006 (morning)

Submissions:
===========

Abstracts should be in English, and up to 4 pages long. The submission format is PDF.

Papers will be reviewed by members of the organising committee. The reviews are not anonymous.

Accepted papers may be up to 6 pages long. The final papers should be in the format specified for the proceedings by the LREC organisers.

Each submission should include: title; author(s); affiliation(s); and contact author's e-mail address, postal address, telephone and fax numbers.

Abstracts should be sent via e-mail to *Briony Williams* at b.williams @ bangor.ac.ukk. The deadline for submission is **February 17th**.

Organising committee
================

- Briony Williams (University of Wales, Bangor, UK: b.williams @ bangor.ac.uk)
- Kepa Sarasola (University of the Basque Country: ksarasola @ si.ehu.es)
- Bojan Petek (University of Ljubljana, Slovenia: bojan.petek @ uni-lj.si)
- Julie Berndsen (University College Dublin, Ireland: julie.berndsen @ ucd.ie)
- Atelach Alemu Argaw (University of Stockholm, Sweden: atelach @ dsv.su.se)