In June 2002, a workshop following the Third International Language Resources and Evaluation Conference LREC 2002 was held in Las Palmas de Gran Canaria, Spain.

Motivation and Aims

There are more than 6000 languages in the world, yet only a small number possess the resources required for implementation of Human Language Technologies (HLT). This imbalance in technical resources available to languages of the world is likely to result in a significant linguistic divide that further exacerbates global social and economic inequities unless decisive action is taken relatively soon. One potential means of ameliorating this imbalance in technology resources is through encouraging research in the portability of human language technology for multilingual application.

Portability issues in HLT are important in the structuring and acquisition of local language resources. The primary objective of the workshop is to bring together participants from academia and industry to discuss and disseminate the current state of the art in multilingual research and development in the context of cross-language HLT transfer. Major challenges for HLT-portability research will also be discussed. The workshop will focus on the following topics and languages:

- Linguistic corpora and portability (new models, language maps, novel systems for creating and managing multilingual data).

- Automatic Speech Recognition (generic design in acoustic modelling, task portability, cross-language portability):

- Acoustic modelling (monolingual and multilingual modelling, cross-language transfer, finite state automata,decision trees and data-driven methods).
- Dictionary development (word-definition issues, automatic dictionary acquisition).
- Language modelling (the Internet as a linguistic resource, language modelling in spoken language processing).

- Natural Language Processing (cross-language transfer of HLT):

- Parsing.
- Translation.

The proposed workshop is intended to continue the series of [SALTMIL]([ISCA SIG]) LREC workshops related to the integration of local and global languages ("Language Resources for European Minority Languages" (LREC'98)) as well as the workshop on "Developing Language Resources for Minority Languages: Re-usability and Strategic Priorities" (LREC'00)).

Workshop Agenda

14:30   *Registration, and preparation of posters*

Oral Session: Portability Issues in HLT

14:50	Workshop Welcome and Introduction Bojan Petek
14:55	Multilingual Time Maps: Portable Phoneme-Based Models for Speech Julie Carson-Berndsen
15:20	Units for Automatic Language Independent Speech Processing

Jan Černocký

15:45	Some Issues in Speech Recognizer Portability Leonid
16:10	Seven Dimensions of Portability for Language Documentation and Description Steven Bird and Gary Simons

16:35   *Break*

Oral Session: HLT and the Coverage of Languages

17:00	Challenges and Opportunities in Portability of Human Language Technologies Bojan Petek
17:25	The Atlantis Observatory: Resources Available on the Internet to Speakers and Learners of Minority Languages Salvador Climent, Miquel Strubell, Marta Torres, and Glyn Williams

17:50	Towards the definition of a basic toolkit for HLT	Kepa Sarasola

18:15 *Poster Session*

Ubiquitous Multilingual Corpus Management	Daryl Shibley	in Computational Fieldwork
A Theory of Portability	Hyo-Kyung Lee	
A Requirement Analysis for the Set of Human	Fredrik Olsson	Language Technology Tasks
Taking Advantage of Spanish Speech Resources to	José B. Padrell	Improve Catalan Acoustic HMMs
José	B. Mariño	
Portability Issues of Text Alignment Techniques		

António Ribeiro, Gabriel Lopes, and João Mexia

SPE Based Selection of Context Dependent Units for Speech Recognition	Matěj Rodina, Bojan Petek, and Tom Brøndsted	
VIPTerm: The Virtual Terminology Information	Fixed System	Point for the Dutch Language. A Supranational Project

20:00 *End*

Workshop Organizing and Program Committee

Julie Carson-Berndsen	University College Dublin	Dublin, Ireland
Steven Greenberg	International Computer Science Institute	Berkeley, USA
Bojan Petek	University of Ljubljana	Ljubljana, Slovenia
Kepa Sarasola	University of the Basque Country	Donostia, Basque Country

Submission

Papers are invited that describe research and development in the area   of Human Language Technology portability. All contributed papers will   be presented in poster format. Each submission should include: title;   author(s); affiliation(s); and contact author's e-mail address, postal   address, telephone and fax numbers. Abstracts (maximum 500 words,   plain-text format) should be sent via email to: <u>Bojan</u>

Petek

All contributed papers will be printed in the workshop proceedings.

Important Dates

Deadline for workshop abstract submission	8th March 2002
Notification of acceptance	15th March 2002
Final version of the paper for workshop proceedings	8th April 2002
Workshop	1st June 2002

Workshop Registration Fees

The registration fees for the workshop are:

- If you are   not attending LREC: 70 EURO
- If you are   attending LREC: 45 Euro

 These fees cover the following: a copy of the proceedings of the attended   workshop, coffee-breaks and refreshments.
 Participation in the workshop is limited by the venue. Requests for   participation will be processed on first come first served basis.   Registration will be handled by the LREC Secretariat.

LREC02. SALTMIL workshop. Portability Issues in Human Language Technologies (HLT)

4 / 4