

"Strategies for developing machine translation for minority languages"

[PDF file of full Proceedings](#) (5.7 megabytes)

Speakers' presentations

- [Welcome](#) : Briony Williams (University of Wales, Bangor, UK)
- Stephen Krauwer (University of Utrecht, Netherlands) " [ENABLER, BLARK, what's next?](#) "
- Delyth Prys (University of Wales, Bangor, UK) " [The BLARK matrix and its relation to the language resources situation for the Celtic languages](#) "
- Lori Levin (Carnegie-Mellon University, USA) " [Omnivorous MT: Using whatever resources are available](#) "
- Mikel Forcada (University of Alacant, Spain) " [Open source machine translation: an opportunity for minor languages](#) "
- Anna Sígvald Hein (University of Uppsala, Sweden) " [Approaching new languages in machine translation](#) "
- Daniel Jacob (Director, Ge'ez Frontier Foundation) " [Unicode Development for Under-Resourced Languages](#) "
- Maja Popovic and Hermann Ney (RWTH Aachen University, Germany) " [Statistical Machine Translation with and without a bilingual training corpus](#) "

Posters

- Dafydd Jones and Andreas Eisele, "Phrase-based statistical machine translation between Welsh and English".
- Carme Armentano-Oller and Mikel Forcada, "Open-source machine translation between small languages: Catalan and Aranese Occitan".
- Kevin Scannell, "Machine translation for closely related language pairs".
- Sara Morrissey and Andy Way, "Lost in Translation: The Problems of Using Mainstream MT Evaluation Metrics for Sign Language Translation".
- John Wogan, Brian Ó Raghallaigh, Áine Ní Bhriain, Eric Zoerner, Harald Berthelsen, Ailbhe Ní Chasaide and Christer Gobl, "Developing a Spoken Corpus and a Synthesiser for Irish".
- Svetla Koeva, Svetlozara Lesseva and Maria Todorova, "Bulgarian sense tagged corpus".
- Jorge González, Antonio L. Lagarda, José R. Navarro, Laura Eliodoro, Adriá Giménez, Francisco Casacuberta, Joan M. de Val and Ferran Fabregat, "SisHiTra: A Spanish-to-Catalan hybrid machine translation system".
- Adriá de Gispert and José B. Mariño, "Catalan-English Statistical Machine Translation without Parallel Corpus: Bridging through Spanish".

- Arantza Casillas, Arantza Díaz de Illarraza, Jon Igartua and Kepa Sarasola, "A Spanish-Basque weather forecast corpus for probabilistic speech translation".
- Saba Amsalu and Sisay Fissaha Adafre, "Machine Translation for Amharic: Where we are".
- Mirjam Sepesy Maučec and Zdravko Kačič, "Statistical machine translation using the IJS-ELAN Corpus".
- Pavel Zheltov, "An Attribute-Sample Database System for describing Chuvash affixes".
- Bostjan Dvorák, Petr Homola, Vladislav Kubon, "Exploiting Similarity in the MT into a Minority Language".
- Tod Allman, "A Natural Language Generator for Minority Languages".
- Arantza Casillas, Arantza Díaz de Illarraza, Jon Igartua and Kepa Sarasola, "A Spanish-Basque compiled corpus and its structure".
- Oliver Streiter, Mathias Stuflesser and Qiu Lan Weng, "Models of Cooperation for the Development of NLP Resources: A Comparison of BLARK and XNLRDF".