

Human language technologies find international communality

Over the past few years the South African government and the Committee of Ministers of the Dutch Language Union have often expressed their interest in cooperation and exchanging of expertise in the field of human language technologies (HLT) in a multilingual context. This expression of interest recently culminated in a workshop, HLT collaboration between South Africa and the Low Countries (Netherlands and Flanders), held in Cape Town on 24-25 November 2008. Organised by the local National HLT Network (NHN) and the Dutch Language Union ([DLU](#)), it was attended by some 30 local and international delegates.

The NHN is an initiative of the Meraka Institute at the CSIR, which fosters local and international collaboration between HLT researchers, while DLU is an international treaty organisation that promotes joint efforts by the governments of the Netherlands, Flanders and Suriname to strengthen the position of the Dutch language. The event was sponsored jointly by the DLU and South Africa's National Research Foundation, with additional support by South Africa's Department of Science and Technology.



One of the special aims of the event was to focus on the potential of HLT to promote multilingualism as well as discussing mechanisms that would allow HLT researchers from the Low Countries and South Africa to collaborate and plan. The programme focused on issues of common interest for the DLU and the NHN, such as multilingual and multicultural contexts, new methods and technologies for languages with limited resources and technology transfer.

Although 22 million people worldwide consider Dutch as their mother language, the market for HLT for Dutch is still deemed to be too limited for major investments by industry. This is the context in which cross-border cooperation between governments, industry and academia from the Netherlands and Flanders has been established.

In South Africa, different imperatives drive priorities in the HLT domain. Various HLT research groups investigate how HLT can be adapted and applied to benefit a developing country, and pursue basic and applied research relevant to the local context. HLT is seen as an enabling technology for information empowerment by providing access through speech-driven technology. It also supports language diversity in an affordable and equitable fashion. Through various initiatives — one such example is that of the Meraka Institute's Lwazi project — industry and government are assisted to make services and documents available in the 11 official South African languages. A special emphasis is placed on supporting language diversity and encouraging the sharing of methods and data as open resources.

Professor Gerhard van Huyssteen, research group leader for HLT at the Meraka Institute, is excited about the outcome of the November event. "Due to the interaction, researchers have come up with some exciting projects to work on, which will contribute significantly to the advancement of technologies, as well as to end-users' experience of these technologies." A framework document for collaboration has been formulated in which mechanisms for research collaboration and human capital development, specifically through reciprocal

teaching and exchange of researchers, have been set out. Additional training opportunities include the development of intensive study courses for Master's and PhD studies as well as internships. "This programme will not only be of benefit to scholars from all three countries, but will also stimulate innovation and create opportunities in emerging markets," Van Huyssteen said.