

Barcelona, June 19th – June 23rd

Evaluation Workshop and TC-STAR Review Meeting

Barcelona, June 19-23

Demonstrations:

UKA:

- **Crossing the Linguistical Divide in a Digital Society**

IBM:

- **IBM Tales: International News Monitoring Solution**

UPC, RWTH, ELDA, IBM:

- **TC-STAR UIMA demonstrator with partner engines**

ITC-irst:

- **Translation from the web**
- **ASR and MT chain – EPPS domain**

LIMSI/UPC:

- **The LIMSI/UPC Speech-to-Speech Translation Demonstration**

RWTH:

- **TC-STAR Evaluation 2006: Transcription and Translation of Parliamentary Speeches**

Barcelona, June 19th – June 23rd

Crossing the Linguistical Divide in a Digital Society

UKA, ISL

Overview

We will present simultaneous translation of lectures and speeches in real-time from English into Spanish and German. Therefore the audio and video of the lecturer will be recorded. The audio will be translated in real-time and presented to the audience in different ways: as subtitles on a canvas, spoken through a targeted audio device, and through translation goggles; which are heads-up display goggles that are worn to see the simultaneous translation as subtitles while we listen to someone speak across the language divide.

With growing international information structures and affordable international travel, the dissemination of knowledge in this globalized world would be increasingly possible – if the language barrier among diverse people can be overcome. Lectures are an effective method of dissemination. Personal talks are preferable, because they allow the speaker to tailor his presentation to specific audiences, and in return allow the listeners to get access to the information relevant to them through interaction with the speaker. Also, personal communication fosters the cooperative creation of ideas and generates collaborations and ties between distant units, e.g. scientific laboratories or companies. At the same time it is desirable to allow the participants in those talks and lectures to express themselves in their native language. Because, no matter how intensively one studies a foreign language one will always be more expressive, more fluent, and more precise in one's native tongue. For this problem human translators are currently the only solution. But they are very expensive, often prohibitively expensive, so that many lectures just cannot happen because of the language barrier. However, the use of modern machine translation techniques can potentially provide affordable translation services to a wide audience, making it possible to overcome the language barrier for almost everyone.

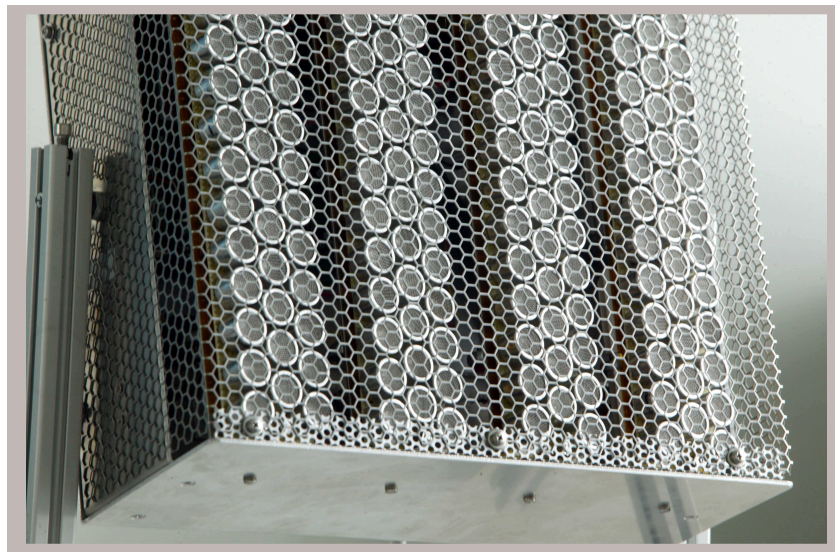


The Lecture Translation System in action.

Barcelona, June 19th – June 23rd

Output devices

The targeted audio device is a beam-steered loudspeaker, consisting of several small ultrasound loudspeakers. It outputs audio in a beam with a width of about 1-2 meters starting from the device and going e.g. until the back of the room. Persons sitting in the beam are able to hear everything, persons outside of the beam not. For the future you can think of having several such targeted audio devices, one for each language and chairs in the lecture room marked as e.g. Spanish or German. The targeted audio device was developed by DaimlerChrysler within the CHIL project.



The targeted audio device.



Translation goggles.

Barcelona, June 19th – June 23rd

IBM TALES: International News Monitoring Solution

IBM

TALES enables an English speaker to search foreign language news content. It currently supports Spanish, Chinese and Arabic content and ingests both Broadcast TV and news web sites. TALES is a 24x7 solution for monitoring 4 simultaneous TV broadcast channels in Arabic, Chinese, and Spanish and for ingesting/updating tens of web sites on a daily basis. It provides

1. a search interface to allow an English speaker to search the ingested foreign language content.
2. In addition to searching and browsing archived news (TV and web sites), TALES allows live monitoring of all ingested TV channels with about four and half minute delay.

TALES uses UIMA components for Chinese and Arabic STT and UIMA components for Spanish, Chinese, and Arabic to English statistical machine translation.

The screenshot displays the IBM TALES Main Result Page in Microsoft Internet Explorer. The browser's address bar shows the URL <http://kmc1.1b08.watson.ibm.com:9080/tales/>. The search bar contains the text "evacuation hurricane". A callout box labeled "English Query" points to the search bar. Below the search bar, there are several search results. One result is highlighted with a callout box labeled "Merged Arabic and Chinese search results". This result includes a video player showing a news anchor. A callout box labeled "Arabic/Chinese/Spanish Video in pop-up player" points to the video player. Another callout box labeled "Translated Speech" points to the video player. A callout box labeled "English Translation" points to the video player. The page also features a sidebar with navigation options like "Headlines", "Usage Help", "Search History", "Bookmarks", "Media Type", "Language", "Sort By", and "Media Player". The date "07/17/2005 08:00:07 ET" is visible at the bottom right.

Barcelona, June 19th – June 23rd

TC-STAR UIMA Demonstrator

UPC, RWTH, ELDA, IBM

Overview

In order to demonstrate continued progress of the TC-STAR infrastructure development as well as progress of the speech and language technologies (ASR, SLT and TTS) we present an end-to-end showcase based on TC-STAR UIMA and the latest TC-Star components from different partners. The UIMA framework is an emerging standard for text and media processing, and is becoming a key infrastructure for speech analyses systems (e.g. DARPA's GALE program). UIMA is derived from an IBM research initiative, and has been open-sourced late 2005.

Apart of the infrastructure which takes care of data delivery to and from the major processing engines (ASR, SLT and TTS), the TC-STAR Type System supports data interoperability between components by using input and output data formats based on common definitions standardized among project partners.

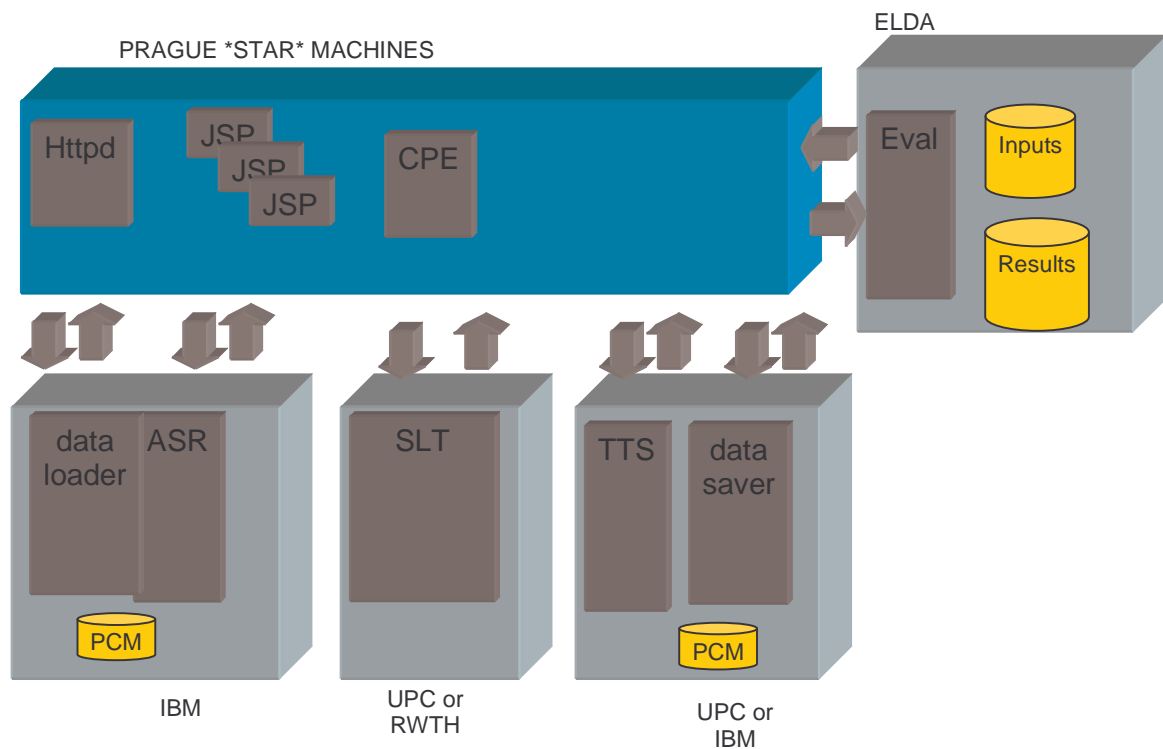


Figure 1: Topology of the inter-partner end-to-end showcase. Data is loaded from the ELDA site, processed through text and media processing engines running at different machines at participating sites, and uploaded back to the ELDA site and evaluated. The processing workflow is controlled from a Web console and UIMA' Collection Processing Engines that runs at IBM Prague.

Barcelona, June 19th – June 23rd

The End-to-End Evaluation Scenario

From a web interface, which controls the end-to-end experiments, we first select the ASR, SLT, TTS component from different partners (called: profile) and then the set of audio files to be processed. This setup is to mimic real deployments where the evaluation can be done distributed as well as at a single location.

TC-STAR UIMA sends the audio files sequentially to the technology components and uploads the translated and synthesized speeches to the web site. The result web page displays the recognized and translated text as well as icons to allow listening to the input and output audio. The web page also displays the result of the evaluation. The processing is done distributed at partner's sites to allow processing of the best, "non-realtime" evaluation engines from the second evaluation period. Details of the architecture will be explained during the presentation.

Available Profiles

Language direction English to Spanish:

- IBM English ASR (UPC English-Spanish SLT (UPC Spanish TTS (ELDA Evaluation
- IBM English ASR (UPC English-Spanish SLT (IBM Spanish TTS (ELDA Evaluation
- IBM English ASR (RWTH English-Spanish SLT (UPC Spanish TTS (ELDA Evaluation

Language direction Spanish to English:

- With IBM only components

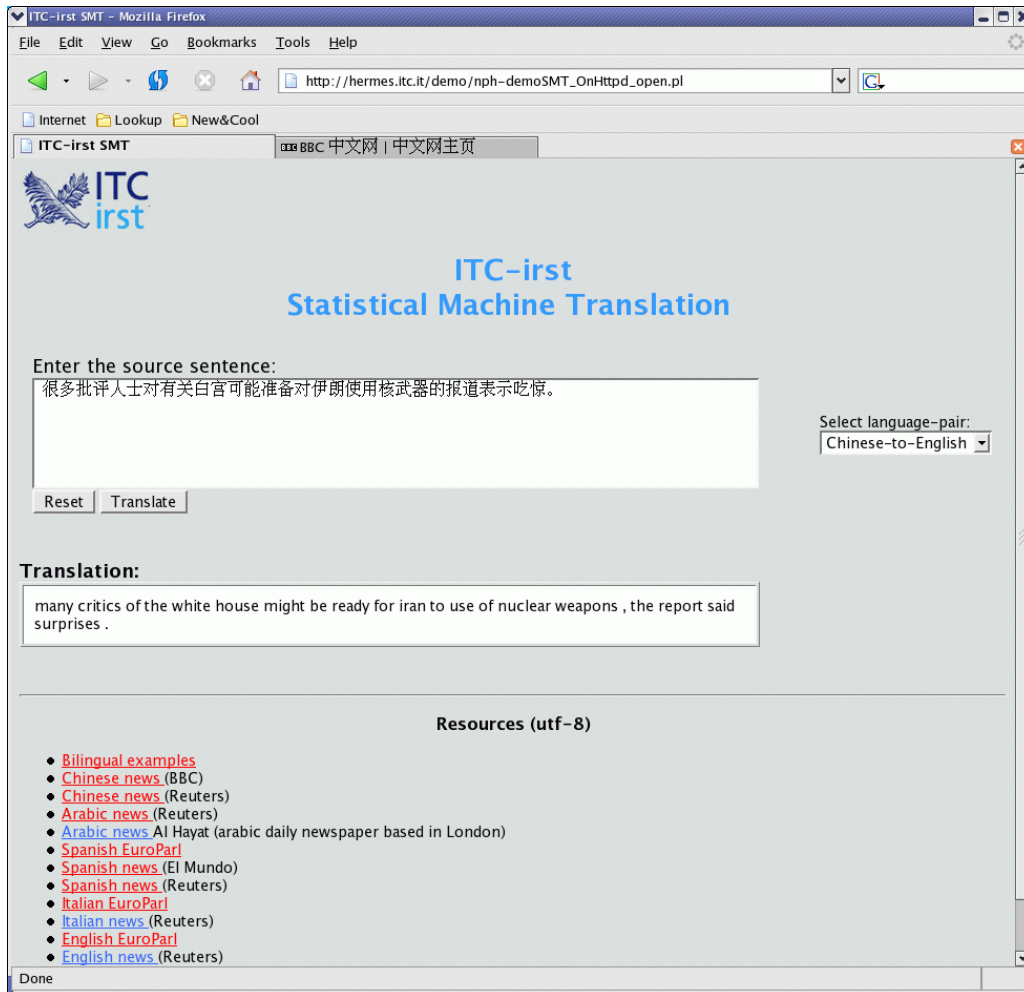
Barcelona, June 19th – June 23rd

TRANSLATION FROM THE WEB

ITC-irst

Overview

We present a demonstrator of a multi-lingual phrase-based Statistical Machine Translation system accessible by means of a Web page. The user can issue translation requests from Chinese or Spanish into English and from English into Italian. The same phrase-based statistical technology is employed to realize all the supported language-pairs: what changes is basically the training data utilized to build the models. The client-server software architecture of the demonstrator allows new language-pairs to be easily added into the system. The Web-based interface includes links to some sites that provide information in the supported language sources, as for example news agencies.



Barcelona, June 19th – June 23rd

ASR and MT Chain – EPPS Domain

ITC-irst

Overview

This off-line demo consists of a video showing about 30 minutes of an European Parliament Plenary Session (EPPS), original audio channel (only politicians speaking their own language, no interpreters), recorded September 7th, 2005. The corresponding audio stream is part of the TC-STAR eval06 set, both English and Spanish, and has been manually divided into chunks according to language only. Chunks labelled as English and Spanish that are part of the eval06 set have been processed with the corresponding ASR system, whose output is fed to the MT system. Chunks were divided automatically in segments having a maximum length of 30 seconds. Both ASR and MT technologies have been fully developed at ITC-irst.

Aligned with the video there are three windows, showing respectively:

- ASR output, word by word;
- ASR output, segment by segment;
- MT output, segment by segment, aligned with b).

The demo is built on RealPlayer software. ASR and MT output were processed to become RealText files, aligned with the video, so it is possible to play the video and watch, listen, and read the corresponding outputs.

The screenshot shows a RealPlayer window with a video of a woman speaking at a podium. The video is overlaid with three text windows:

- Automatic Speech Recognition**:
croata en nombre de la - Presidencia . - por eso quiero
aprovechar su presencia en - este foro y - la del Comisario
Frattini para insistir
- ASR segment by segment**:
gracias Presidente reconozco que la lucha contra el terrorismo no soy neutral apoyo completamente
el enfoque y las prioridades que hoy nos ha expresado el ministro croata en nombre de la
Presidencia por eso quiero aprovechar su presencia en este foro y la del Comisario Frattini para
insistir en la necesidad de actuar de que tomemos la iniciativa politica luchar contra el
terrorismo es algo mas que perseguir y detener a los criminales y consolar a las victimas
- Machine Translation**:
Thank you President I recognise that the fight against terrorism not neutral am I fully support
the approach and the priorities that today we have expressed the Minister Croatian on behalf of
the Presidency I would therefore like to take advantage of its presence in this forum and the
Commissioner Frattini to insist on the need for action for us to take the political initiative
fighting terrorism is more than to pursue and to stop the criminals and console the victims

The RealPlayer interface shows the video is paused, with a progress bar at 20:22,6 / 30:17,0 and a volume control at 150 kbps.

Barcelona, June 19th – June 23rd

The LIMSI/UPC Speech-to-Speech Translation Demonstration

LIMSI, UPC

Overview

Fully automatic offline Speech-to-Speech Translation of European Parliament Plenary Sessions is demonstrated for English and Spanish. The sample video is from a plenary session recorded on 20 July 2004. The audio data is automatically segmented and transcribed using the LIMSI automatic speech recognition (ASR) system and the output is fed to a statistical Machine Translation (MT) system (UPC or LIMSI). The text produced by MT is then synthesized by the UPC Text-to-Speech (TTS) system. The translation alignment from MT and the word time codes provided by the ASR system are used to synchronize the speech output to the original speech signal. The interface allows the user to select the translation direction, the MT system, and the audio signal corresponding to the original speech, the TTS output or the human (simultaneous) translation.

The screenshot shows the 'EPPS demo - (C) LIMSI-CNRS 2005' window. The title bar reads 'LIMSI / UPC Automatic Speech to Speech Translation'. The interface includes a header with logos for ASR and LIMSI, and a globe icon. Below the header, there are controls for 'Audio language' (with flags for UK, EU, Spain, and TTS) and 'Translation' (with flags for UK and UPC). The main area contains a list of two audio clips from the European Parliament plenary session on 07/20/2004. A video window on the right shows a man speaking at a podium. Below the video, there are volume and cursor controls. The bottom section displays the translated text in two languages: English and Spanish. The English text is: 'Mister Watson now Thank you Mister President -- Mr. President I'd like to add my words to those of my colleagues Hans Gert Poettering and Martin Schultz in thanking all the contestants in this competition for the Presidency of our House I think the result that we've is an honourable one -- I'm proud that the political family I represent gave the European Parliament a'. The Spanish text is: 'Señor Watson Gracias Señor Presidente, Señor Presidente Ah, quisiera añadir mis palabras a aquellos de mis colegas a Hans Gert Poettering y Martin Schulz dando las gracias a todos los participantes en esta competencia a la Presidencia de nuestra Asamblea, creo que el resultado que hemos constatado honorable, estoy completamente orgulloso de que la familia política a la que represento diera al Parlamento Europeo una primera'.

Barcelona, June 19th – June 23rd

TC-STAR Evaluation 2006: Transcription and Translation of Parliamentary Speeches

RWTH

Overview

An off-line video presents the results of the RWTH transcription and translation system used in the TC-STAR Evaluation 2006. Examples of the Spanish ASR and Spanish to English SLT evaluation output for the European Parliament Plenary Session (EPPS) task are included. The demo provides Spanish speeches (audio and video tracks) held by Members of the European Parliament and the corresponding punctuated, case-sensitive, text output from the RWTH ASR and SLT engines.



RWTH AACHEN
UNIVERSITY



European Parliament

la prosperidad para toda la humanidad .

y lo que se puede decir es que básicamente esa cumbre lo que ha conseguido es frenar los intentos de desmantelamiento de las Naciones Unidas y

rights the peace and prosperity for the whole of humanity .

and what we can say is that basically this summit what it has achieved is to stop the attempts of dismantling of the United Nations and