



October 20-21, 2008

International Workshop on Spoken Language Translation

Evaluation Campaign on Spoken Language Translation

- What's New
- Theme
- Evaluation Campaign
- Important Dates
- Downloads
- Resources
- Submission
- Evaluation Server
- Registration
- Accommodation
- Program
- Keynote Speech
- Proceedings**
- Author Index
- Bibliography
- Venue
- Gallery
- Organizers
- Contact
- References

Co-located Events

- AMTA2008
(Oct 22-25)
- Metrics MATR
(Oct 25)
- EMNLP
(Oct 25-27)



Copyright(c) 2007-2008 All rights reserved.





October 20-21, 2008

International Workshop on Spoken Language Translation

Evaluation Campaign on Spoken Language Translation

[What's New](#) - [Theme](#) - [Evaluation Campaign](#) - [Important Dates](#) - [Downloads](#) - [Resources](#) - [Submission](#) - [Evaluation Server](#) - [Registration](#) - [Accommodation](#) - [Program](#) - [Keynote Speech](#) - [Proceedings](#) - [Author Index](#) - [Bibliography](#) - [Venue](#) - [Gallery](#) - [Organizers](#) - [Contact](#) - [References](#) -

Theme

Last Update: 06/17/2009 08:25:17

The International Workshop on Spoken Language Translation (IWSLT) is a [yearly](#), open evaluation campaign for spoken language translation followed by a scientific workshop, in which both system descriptions and scientific papers are presented. IWSLT's evaluations are not competition-oriented, but their goal is to foster cooperative work and scientific exchange. In this respect, IWSLT proposes challenging research tasks and an open experimental infrastructure for the scientific community working on spoken and written language translation.

For developing state-of-the-art machine translation systems, IWSLT will provide links to relevant open software tools as well as a limited amount (20k sentence pairs for each translation direction) of linguistic resources. In addition, a list of links to permitted public resources is available on the [resources page](#). The participants of IWSLT 2008 are encouraged to look for additional public resources and to share their availability by sending the respective information to the [evaluation campaign chair](#). All declared resources will be permitted to be used for the training of the systems. Each participant in the evaluation campaign is requested to submit a paper describing the MT system, the utilized resources, and results using the provided test data. Contrastive run submissions using only the bilingual resources provided by IWSLT as well as investigations in the contribution of each utilized resource are highly appreciated. Moreover, all participants are requested to present their papers at the workshop.

Evaluation Campaign

The focus of this year's evaluation campaign will be the translation of spontaneous speech recorded in a real situation. Foreign travelers were provided with a state-of-the-art speech-to-speech translation hand-held device and were asked to carry-out specific tourism-related tasks (e.g., buying entrance tickets) using the device to communicate with local staff. Speech data was collected for 50 English and 50 Chinese travelers at 5 different locations, each carrying out 3-4 tasks. For the **Challenge Task**, IWSLT participants will have to translate the Chinese/English output of the automatic speech recognizers (lattice, N/1BEST) into English/Chinese, respectively.

Another innovative aspect of this year's edition will concern the feasibility of pivot-language-based approaches. In the **Pivot Task**, participants will be provided with read-speech recordings (lattice, N/1BEST) of Chinese utterances from the travel domain and have to apply *Chinese-English* and *English-Spanish* systems to produce the Spanish output. The results will be compared to the direct translation between *Chinese-Spanish*.

Like in previous IWSLT events, a standard **BTEC task**, i.e. the translation of read-speech recordings (lattice, N/1BEST) and correct recognition results (text) of frequently used utterances in the travel domain, will be provided for *Arabic-English* and *Chinese-English*.

Scientific Paper

In addition to the evaluation campaign, the IWSLT 2008 workshop also invites scientific paper submissions related to spoken language technologies. Possible topics include, but are not limited to:

- Text and speech translation systems
- Search algorithms for MT
- Phrase alignment methods for MT
- Re-ordering models for MT
- Semantic models for MT
- Syntax-based MT
- Pivot-language-based MT
- MT evaluation
- Integration of ASR and MT
- Open source software for MT
- Language resources for MT
- Task adaptation and portability in MT



October 20-21, 2008

International Workshop on Spoken Language Translation

Evaluation Campaign on Spoken Language Translation

[What's New](#) - [Theme](#) - [Evaluation Campaign](#) - [Important Dates](#) - [Downloads](#) - [Resources](#) - [Submission](#) - [Evaluation Server](#) - [Registration](#) - [Accommodation](#) - [Program](#) - [Keynote Speech](#) - [Proceedings](#) - [Author Index](#) - [Bibliography](#) - [Venue](#) - [Gallery](#) - [Organizers](#) - [Contact](#) - [References](#) -

Evaluation Campaign

Last Update: 06/17/2009 08:22:56

The evaluation campaign is carried out using BTEC (*Basic Travel Expression Corpus*), a multilingual speech corpus containing tourism-related sentences similar to those that are usually found in phrasebooks for tourists going abroad. Details about the supplied corpus, the data set conditions for each track, the guidelines on how to submit one's translation results, and the evaluation specifications used in this workshop are given below.

[\[Corpus Specifications\]](#)

[\[Translation Input Conditions\]](#)

[\[Evaluation Specifications\]](#)

Corpus Specifications

Training Corpus:

- data format:
 - each line consists of three fields divided by the character '\'
 - sentence consisting of words divided by single spaces
 - format: <SENTENCE_ID>\01<MT_TRAINING_SENTENCE>**
 - Field_1: sentence ID
 - Field_2: paraphrase ID
 - Field_3: MT training sentence
- example:
 - TRAIN_00001\01\This is the first training sentence.
 - TRAIN_00002\01\This is the second training sentence.
- Arabic-English (AE)
- Chinese-English (CE)
- Chinese-Spanish (CS)
- Chinese-(English)-Spanish (CES)
- English-Chinese (EC)
 - 20K sentences randomly selected from the BTEC corpus
 - coding: UTF-8
 - word segmentations according to ASR output segmentation
 - text is case-sensitive and includes punctuations

[TOP](#)

Develop Corpus:

- ASR output (lattice, NBEST, 1BEST), correct recognition result transcripts (text), reference translations of previous IWSLT test data sets
- data format:
 - **1-BEST**
 - each line consists of three fields divided by the character '\'
 - sentence consisting of words divided by single spaces
format: <SENTENCE_ID>\01\<MT_INPUT_SENTENCE>
 - Field_1: sentence ID
 - Field_2: paraphrase ID
 - Field_3: best recognition hypothesis
 - example (input):
DEV_001\01\best ASR hypothesis for 1st input
DEV_002\01\best ASR hypothesis for 2nd input
...
 - **N-BEST**
 - each line consists of three fields divided by the character '\'
 - sentence consisting of words divided by single spaces
DEV_001\01\best ASR hypothesis for 1st input
DEV_001\02\2nd-best ASR hypothesis for the 1st input
...
DEV_001\20\20th-best ASR hypothesis for the 1st input
DEV_002\01\best ASR hypothesis for the 2nd input
...
 - **word lattices** → [HTK Standard Lattice Format \(SLF\)](#)
 - **reference translations**
 - each line consists of three fields divided by the character '\'
 - sentence consisting of words divided by single spaces
format: <SENTENCE_ID>\<PARAPHRASE_ID>\<REFERENCE>
 - Field_1: sentence ID
 - Field_2: paraphrase ID
 - Field_3: reference translation
 - example:
DEV_001\01\1st reference translation for 1st input
DEV_001\02\2nd reference translation for 1st input
...
DEV_002\01\1st reference translation for 2nd input
DEV_002\02\2nd reference translation for 2nd input
...
- Arabic-English
 - CSTAR03 testset: 506 sentences, 16 reference translations
 - IWSLT04 testset: 500 sentences, 16 reference translations
 - IWSLT05 testset: 506 sentences, 16 reference translations
 - IWSLT06 devset: 489 sentences, 7 reference translations
 - IWSLT06 testset: 500 sentences, 7 reference translations
 - IWSLT07 testset: 489 sentences, 6 reference translations
- Chinese-English
 - CSTAR03 testset: 506 sentences, 16 reference translations
 - IWSLT04 testset: 500 sentences, 16 reference translations
 - IWSLT05 testset: 506 sentences, 16 reference translations

- IWSLT06 devset: 489 sentences, 7 reference translations
- IWSLT06 testset: 500 sentences, 7 reference translations
- IWSLT07 testset: 489 sentences, 6 reference translations
- IWSLT08 devset: 250 sentences of the field-experiment data (challenge task)
- Chinese-Spanish
- Chinese-(English)-Spanish
 - IWSLT05 testset: 506 sentences, 16 reference translations
- English-Chinese
 - IWSLT05 testset: 506 sentences, 16 reference translations
 - IWSLT08 devset: 250 sentences of the field-experiment data (challenge task)

[TOP](#)

Test Corpus:

- Challenge Task
 - Chinese-English
 - English-Chinese
 - 500 sentences of the field-experiment data
 - coding: → see Develop Corpus
 - data format: → see Develop Corpus
- BTEC Task
 - Arabic-English
 - Chinese-English
 - Chinese-Spanish
 - 500 unseen sentences of the BTEC evaluation corpus
 - coding: → see Develop Corpus
 - data format: → see Develop Corpus
- PIVOT Task
 - Chinese-(English)-Spanish
 - 500 unseen sentences of the BTEC evaluation corpus
 - coding: → see Develop Corpus
 - data format: → see Develop Corpus

[TOP](#)

Translation Input Conditions

Spontaneous Speech

- Challenge Task
 - Chinese-English
 - English-Chinese

→ ASR output (word lattice, N-best, 1-best) of ASR engines provided by IWSLT organizers

Read Speech

- BTEC Task
 - Arabic-English
 - Chinese-English
 - Chinese-Spanish
- PIVOT Task
 - Chinese-(English)-Spanish

→ ASR output (word lattice, N-best, 1-best) of ASR engines provided by IWSLT organizers

Correct Recognition Results

- *Challenge Task*
 - *Chinese-English*
 - *English-Chinese*
- *BTEC Task*
 - *Arabic-English*
 - *Chinese-English*
 - *Chinese-Spanish*
- *PIVOT Task*
 - *Chinese-(English)-Spanish*

→ text input

[TOP](#)

Evaluation

Subjective Evaluation:

- *Metrics:*
 - **ranking**
→ all primary run submissions
 - **fluency/adequacy**
→ top-scoring (according to average of BLEU and METEOR scores) primary run submission + up to 3 additional primary runs selected by organizers (according to level of innovation of translation approach)
- *Evaluators:*
 - 3 graders per translation

Automatic Evaluation:

- *Metrics:*
 - **BLEU**
 - **METEOR**

→ up to 7 reference translations
→ all run submissions
- *Evaluation Specifications:*
 - *Official:*
 - case sensitive
 - with punctuation marks tokenized
 - *Additional:*
 - case insensitive (lower-case only)
 - no punctuation marks
- *Data Processing Prior to Evaluation:*
 - *English MT Output:*
 - simple tokenization of punctuations (see 'tools/ppEnglish.case+punc.pl' script)
 - *Spanish MT Output:*
 - simple tokenization of punctuations (see 'tools/ppSpanish.case+punc.pl' script)
 - *Chinese MT Output:*
 - segmentation into characters (see 'tools/splitUTF8Characters' script)

[TOP](#)



Organizers

Last Update: 06/17/2009 08:24:31

Organizers

- Marcello Federico (FBK, Italy)
- Satoshi Nakamura (NICT/ATR, Japan)
- Alex Waibel (CMU, USA / UKA, Germany)

Chairs

- Tanja Schultz (CMU, USA / UKA, Germany; **Workshop**)
- Michael Paul (NICT/ATR, Japan; **Evaluation Campaign**)
- Marcello Federico (FBK, Italy; **Technical Paper**)

Program Committee

- Nicola Bertoldi (FBK, Italy)
- Laurent Besacier (LIG, France)
- Francisco Casacuberta (ITI-UPV, Spain)
- Diamantino Caseiro (INESC, Portugal)
- Boxing Chen (I2R, Singapore)
- Katrin Kirchhoff (UW, USA)
- Philipp Koehn (UEDIN, UK)
- Philippe Langlais (RALI, Canada)
- Yves Lepage (GREYC, France)
- José B. Mariño (TALP-UPC, Spain)
- Hermann Ney (RWTH, Germany)
- Liu Qun (ICT, China)
- Eiichiro Sumita (NICT/ATR, Japan)
- Stephan Vogel (CMU, USA / UKA, Germany)
- Shen Wade (MIT-LL, USA)
- Taro Watanabe (NTT, Japan)
- Andy Way (DCU, Ireland)
- Dekai Wu (HKUST, Hong Kong)
- Richard Zens (GOOGLE, USA)
- Chengqing Zong (CASIA, China)

Local Arrangements

- Lisa Mauti (CMU, USA)

Supporting Organizations

 **ATR** Advanced Telecommunication Research Institute International

 **NICT** National Institute of Information and Communications Technology



October 20-21, 2008

International Workshop on Spoken Language Translation

Evaluation Campaign on Spoken Language Translation

[What's New](#) - [Theme](#) - [Evaluation Campaign](#) - [Important Dates](#) - [Downloads](#) - [Resources](#) - [Submission](#) - [Evaluation Server](#) - [Registration](#) - [Accommodation](#) - [Program](#) - [Keynote Speech](#) - [Proceedings](#) - [Author Index](#) - [Bibliography](#) - [Venue](#) - [Gallery](#) - [Organizers](#) - [Contact](#) - [References](#) -

Keynote Speech

Last Update: 06/17/2009 08:24:24

October 20, 2008

Language Technology in Humanitarian Aid and Disaster Response

09:10-10:10

Scott HOURIN (U.S. Marine Corps Forces, USA)

Every year, the U.S. military expends significant amounts of money and manpower providing humanitarian aid and disaster response (HA/DR) assistance to the developing world. While interpreters are usually available for programs such as the Medical Civic Action Program (MEDCAP), there rarely are enough to provide the level of doctor ? patient interaction those of us in the developed world have come to expect. This is just one example where machine translation, while not necessarily at a maturity level to supplant an interpreter, could facilitate more efficient and effective employment of a corps of interpreters. The greatest hurdle to development of less common language technologies, however, is economic. There is little to no commercial benefit to developing translation systems for the majority of the languages spoken in the developing world. Add to that the complexity and diversity of these languages and the hurdle becomes monumental. While the military has great interest (and funding) in developing some of these capabilities, it can benefit greatly from the full focus of the linguistic and technical communities to figure out how to best tackle the problem of linguistic diversity.

Scott Hourin has worked with machine translation for the U.S. military since 2003. As a U.S. Marine Corps Officer he facilitated some of the first operational assessments of machine translation technologies. After college at the University of Richmond, he began his career as an infantry Marine in 1994 then moved small mountains of men and materiel in the logistics field before re-joining the civilian world in 1998. After four and a half years as a civilian he was mobilized for Operation Iraqi Freedom in 2003 and spent two years on active duty. It was during this second period on active duty in the Marine Corps that he began working with machine translation, among other technologies, in his role in the Innovation Technology staff section at I Marine Expeditionary Force. After leaving active duty a second time 2005 he continues much of the same work in his current role as a Project Lead with the U.S. Marine Corps Forces, Pacific Experimentation Center.



October 20-21, 2008

International Workshop on Spoken Language Translation

Evaluation Campaign on Spoken Language Translation

[What's New](#) - [Theme](#) - [Evaluation Campaign](#) - [Important Dates](#) - [Downloads](#) - [Resources](#) - [Submission](#) - [Evaluation Server](#) - [Registration](#) - [Accommodation](#) - [Program](#) - [Keynote Speech](#) - [Proceedings](#) - [Author Index](#) - [Bibliography](#) - [Venue](#) - [Gallery](#) - [Organizers](#) - [Contact](#) - [References](#) -

Program

Last Update: 06/17/2009 08:24:42

October 20, 2008

Workshop Opening

9:00	9:10	Welcome Remarks
------	------	-----------------

Invited Talk

09:10	10:10	Language Technology in Humanitarian Aid and Disaster Response <i>Scott HOURIN</i> (U.S. Marine Corps Forces, USA)
-------	-------	--

coffee ☕ break

Evaluation Campaign : Overview

10:30	11:00	Overview of the IWSLT 2008 Evaluation Campaign <i>Michael PAUL</i>
-------	-------	---

Evaluation Campaign : Oral I

11:00	11:20	The TCH Machine Translation System for IWSLT 2008 <i>Haifeng WANG, Hua WU, Xiaoguang HU, Zhanyi LIU, Jianfeng LI, Dengjun REN, Zhengyu NIU</i>
-------	-------	---

11:20	11:40	FBK @ IWSLT-2008 <i>Nicola BERTOLDI, Roldano CATTONI, Marcello FEDERICO, Madalina BARBAIANI</i>
-------	-------	--

11:40	12:00	The NICT/ATR Speech Translation System for IWSLT 2008 <i>Masao UTIYAMA, Andrew FINCH, Hideo OKUMA, Michael PAUL, Hailong CAO, Hirofumi YAMAMOTO, Keiji YASUDA, Eiichiro SUMITA</i>
-------	-------	---

12:00	12:20	I²R Multi-Pass Machine Translation System for IWSLT 2008 <i>Boxing CHEN, Deyi XIONG, Min ZHANG, Aiti AW, Haizhou LI</i>
-------	-------	--

lunch 🍽️ break

Evaluation Campaign : Poster I

14:00	15:00	Exploiting Alignment Techniques in MaTrEx: the DCU Machine Translation System for IWSLT08 <i>Yanjun MA, John TINSLEY, Hany HASSAN, Jinhua DU, Andy WAY</i>
-------	-------	---

14:00	15:00	The GREYC Machine Translation System for the IWSLT 2008 Evaluation Campaign <i>Yves LEPAGE, Adrien LARDILLEUX, Julien GOSME, Jean-Luc MANGUIN</i>
-------	-------	--

14:00	15:00	The ICT System Description for IWSLT 2008 <i>Yang LIU, Zhongjun HE, Haitao MI, Yun HUANG, Yang FENG, Wenbin JIANG, Yajuan LU, Qun LIU</i>
14:00	15:00	The LIG Arabic/English Speech Translation System at IWSLT08 <i>Laurent BESACIER, Atef Ben YOUSSEF, Herve BLANCHON</i>
Technical Paper : Oral I		
15:00	15:25	Improving Statistical Machine Translation by Paraphrasing the Training Data <i>Francis BOND, Eric NICHOLS, Darren Scott APPLING, Michael PAUL</i>
15:25	15:50	Investigations on Large-Scale Lightly-Supervised Training for Statistical Machine Translation <i>Holger SCHWENK</i>
15:50	16:15	Analysing Soft Syntax Features and Heuristics for Hierarchical Phrase Based Machine Translation <i>David VILAR, Daniel STEIN, Hermann NEY</i>
16:15	16:40	Improvements in Dynamic Programming Beam Search for Phrase-based Statistical Machine Translation <i>Richard ZENS, Hermann NEY</i>
Demo Reception (with Coffee Break)		
16:40	18:00	(to be announced)

October 21, 2008

Discussion Panel

09:20	10:20	(to be announced)
-------	-------	-------------------

coffee ☕ break

Technical Paper : Oral II

10:40	11:05	Phrase-Based Statistical Machine Translation with Pivot Languages <i>Nicola BERTOLDI, Madalina BARBAIANI, Marcello FEDERICO, Roldano CATTONI</i>
11:05	11:30	Rapid Development of an English/Farsi Speech-to-Speech Translation System <i>Chia-Lin KAO, Shirin SALEEM, Rohit PRASAD, Fred CHOI, Prem NATARAJAN, David STALLARD, Kriste KRSTOVSKI, Matin KAMALI</i>
11:30	11:55	Evaluating Productivity Gains of Hybrid ASR-MT Systems for Translation Dictation <i>Alain DÉSILETS, Marta STOJANOVIC, Jean-Françoise LAPOINTE, Rick ROSE, Aarthi REDDY</i>
11:55	12:20	Simultaneous German-English Lecture Translation <i>Muntsin KOLSS, Matthias WÖLFEL, Florian KRAFT, Jan NIEHUES, Matthias PAULIK, Alex WAIBEL</i>

lunch  break

Evaluation Campaign : Poster II

14:00	15:20	The LIUM Arabic/English Statistical Machine Translation System for IWSLT 2008 <i>Holger SCHWENK, Yannick ESTÈVE, Sadaf Abdul RAUF</i>
14:00	15:20	NTT Statistical Machine Translation System for IWSLT 2008 <i>Katsuhito SUDO, Taro WATANABE, Jun SUZUKI, Hajime TSUKADA, Hideki ISOZAKI</i>
14:00	15:20	POSTECH Machine Translation System for IWSLT 2008 Evaluation Campaign <i>Jonghoon LEE, Gary Geunbae LEE</i>
14:00	15:20	The QMUL System Description for IWSLT 2008 <i>Simon CARTER, Christof MONZ, Sirvan YAHYAEI</i>
14:00	15:20	The TALP & I ² R SMT Systems for IWSLT 2008 <i>Maxim KHALILOV, Marta R. COSTA-JUSSÀ, Carlos A. HENRÍQUEZ, José A. R. FONOLLOSA, Adolfo HERNÁNDEZ, José B. MARIÑO, Rafael E. BANCHS, Chen BOXING, Min ZHANG, Aiti AW, Haizhou LI</i>
14:00	15:20	Statistical Machine Translation without Long Parallel Sentences for Training Data <i>Jin'ichi MURAKAMI, Masato TOKUHISA, Satoru IKEHARA</i>
14:00	15:20	The TUBITAK-UEKAE Statistical Machine Translation System for IWSLT 2008 <i>Coskun MERMER, Hamza KAYA, Ömer Farukhan GÜNES, Mehmet Ugur DOGAN</i>

coffee  break

Evaluation Campaign : Oral II

15:40	16:00	The MIT-LL/AFRL IWSLT-2008 MT System <i>Wade SHEN, Brian DELANEY, Timothy ANDERSON, Raymond SLYH</i>
16:00	16:20	The CMU Syntax-Augmented Machine Translation System: SAMT on Hadoop with N-best Alignments <i>Andreas ZOLLMANN, Ashish VENUGOPAL, Stephan VOGEL</i>
16:20	16:40	The RWTH Machine Translation System for IWSLT 2008 <i>David VILAR, Daniel STEIN, Yuqi ZHANG, Evgeny MATUSOV, Arne MAUSER, Oliver BENDER, Saab MANSOUR, Hermann NEY</i>
16:40	17:00	The CASIA Statistical Machine Translation System for IWSLT 2008 <i>Yanqing HE, Jiajun ZHANG, Maoxi LI, Licheng FANG, Yufeng CHEN, Yu ZHOU, Chengqing ZONG</i>

Banquet

17:30		Leave for Banquet Dinner
18:30	20:30	Banquet Dinner at the Turtle Bay Resort