

[Home](#) | [What's New](#) | [Theme](#) | [Evaluation Campaign](#) | [Important Dates](#) | [Submission](#) | [Downloads/data](#) | [Program](#) | [Proceedings](#) | [Author Index](#) | [Photo Gallery](#) | [Registration](#) | [Venue Information](#) | [Organizers](#) | [Contact](#) | [References](#)



International Workshop on Spoken Language Translation

Home

- [What's new](#)
- [Theme](#)

Important dates

Evaluation campaign

- [Schedule/Data Status](#)
- [Translation directions](#)
- [Data Tracks/Restrictions](#)
- [Data](#)
- [Downloads](#)
- [Evaluation specification](#)
- [Test Data](#)
- [System description paper](#)
- [Evaluation Servers](#)

Submissions

Workshop

- [Registration](#)
- [Program](#)
- [Author Index](#)
- [Proceedings](#)
- [Venue](#)
- [Information](#)
- [Photo Gallery](#)
- [Organizers](#)

Contact

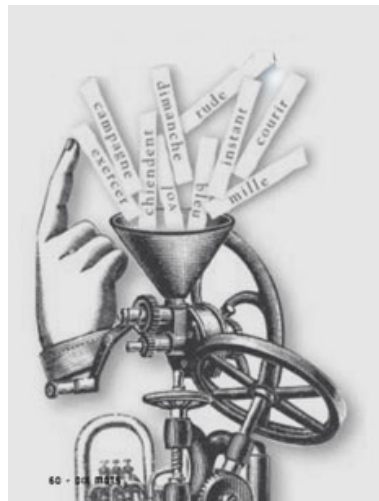
References

Welcome

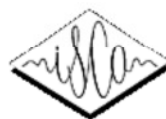
the International Workshop on Spoken Language Translation 2005 will be held:

Date: **October 24th - 25th**

Venue: **Pittsburgh, PA**



Sponsored and supported by the following organizations:





International Workshop on Spoken Language Translation

Home

- [What's new](#)
- [Theme](#)

Important dates

Evaluation campaign

- [Schedule/Data Status](#)
- [Translation directions](#)
- [Data Tracks/Restrictions](#)
- [Data](#)
- [Downloads](#)
- [Evaluation specification](#)
- [Test Data](#)
- [System description paper](#)
- [Evaluation Servers](#)

Submissions

Workshop

- [Registration](#)
- [Program](#)
- [Author Index](#)
- [Proceedings](#)
- [Venue](#)
- [Information](#)
- [Photo Gallery](#)
- [Organizers](#)

Contact

References

Theme:

The Consortium for Speech Translation Advanced Research (C-STAR) projects has attempted to solve issues in speech translation. Speech translation systems are designed as a system combined machine translation (MT) system with automatic speech recognition (ASR) system. Researches on text translation with large vocabulary such as newspaper text and HTML documents by using statistical approaches are being actively investigated. Text translation is still difficult and speech translation is more complicated than text translation. The difficulty in speech translation in terms of verbal information is caused by processing spoken language and ASR output. To accomplish speech translation, we have to solve problems in MT for spoken language and handle ASR output. In this workshop, we will focus on handling spontaneous aspects in spoken language and speech recognition output including multiple recognition hypotheses.

Speech recognition results are not always perfect. Speech translation performance is degraded by recognition errors. When we translate speech so that maintains original information in source speech as much as possible by handling recognition errors, we have concrete questions that we hope to see addressed include, but are not limited to the following issues.

- How can ASR output be translated more accurately even if recognition errors exist?

- How much MT performance could be enhanced by considering multiple hypotheses?

- Which hypothesis can contribute for MT performance?

- How to select best hypothesis in recognition results which can be translated more accurately?

[Home](#) | [What's New](#) | [Theme](#) | [Evaluation Campaign](#) | [Important Dates](#) | [Submission](#) | [Downloads/data](#) | [Program](#) | [Proceedings](#) | [Author Index](#) | [Photo Gallery](#) | [Registration](#) | [Venue Information](#) | [Organizers](#) | [Contact](#) | [References](#)



International Workshop on Spoken Language Translation

Home

- [What's new](#)
- [Theme](#)

Important dates

Evaluation campaign

- [Schedule/Data Status](#)
- [Translation directions](#)
- [Data Tracks/Restrictions](#)
- [Data Downloads](#)
- [Evaluation specification](#)
- [Test Data](#)
- [System description paper](#)
- [Evaluation Servers](#)

Submissions

Workshop

- [Registration](#)
- [Program](#)
- [Author Index](#)
- [Proceedings](#)
- [Venue](#)
- [Information](#)
- [Photo Gallery](#)
- [Organizers](#)

Contact

References

Organizers:

Alex Waibel (CMU, USA / UKA, Germany; Chair)

Chiori Hori (CMU, USA, Co-Chair)

Stephan Vogel (CMU, USA, Program Chair)

Tanja Schultz (CMU, USA, Program Co-Chair)

Christian Boitet (CLIPS, France)

Gianni Lazzari (ITC-irst, Italy)

Youngjik Lee (ETRI, Korea)

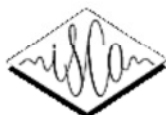
Satoshi Nakamura (ATR, Japan)

Chengqing Zong (CAS, China)

Local Arrangements:

Céline Carraux (CMU, USA)

Sponsored and supported by the following organizations:



Copyright(c) 2005 interACT All rights reserved.

[Home](#) | [What's New](#) | [Theme](#) | [Evaluation Campaign](#) | [Important Dates](#) | [Submission](#) | [Downloads/data](#) | [Program](#) | [Proceedings](#) | [Author Index](#) | [Photo Gallery](#) | [Registration](#) | [Venue Information](#) | [Organizers](#) | [Contact](#) | [References](#)



International Workshop on Spoken Language Translation

Home

- [What's new](#)
- [Theme](#)

Important dates

Evaluation campaign

- [Schedule/Data Status](#)
- [Translation directions](#)
- [Data Tracks/Restrictions](#)
- [Data Downloads](#)
- [Evaluation specification](#)
- [Test Data](#)
- [System description paper](#)
- [Evaluation Servers](#)

Submissions

Workshop

- [Registration](#)
- [Program](#)
- [Author Index](#)
- [Proceedings](#)
- [Venue Information](#)
- [Photo Gallery](#)
- [Organizers](#)

Contact

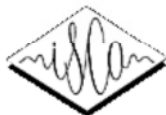
References

References

[IWSLT 04 - International Workshop on Spoken Language Translation 2004](#)

[C-Star 3 - Consortium for Speech Translation Advanced Research](#)

Sponsored and supported by the following organizations:



[Home](#) | [What's New](#) | [Theme](#) | [Evaluation Campaign](#) | [Important Dates](#) | [Submission](#) | [Downloads/data](#) | [Program](#) | [Proceedings](#) | [Author Index](#) | [Photo Gallery](#) | [Registration](#) | [Venue Information](#) | [Organizers](#) | [Contact](#) | [References](#)



International Workshop on Spoken Language Translation

Home

- [What's new](#)
- [Theme](#)

Important dates

Evaluation campaign

- [Schedule/Data Status](#)
- [Translation directions](#)
- [Data Tracks/Restrictions](#)
- [Data Downloads](#)
- [Evaluation specification](#)
- [Test Data](#)
- [System description paper](#)

Submissions

Workshop

- [Registration](#)
- [Program](#)
- [Author Index](#)
- [Proceedings](#)
- [Venue Information](#)
- [Photo Gallery](#)
- [Organizers](#)

Contact

References

Monday 24th Oct

Breakfast

Workshop Opening

Welcome talk

Alex Waibel (CMU/UKA)

Overview of the IWSLT05 Evaluation Campaign

Chiori Hori and Matthias Eck
(CMU)

Session 1: Statistical MT system -Part I-

Chair: *Stephan Vogel*

The ITC-irst SMT System for IWSLT-2005

10:00-10:25

Boxing Chen, Roldano Cattoni, Nicola Bertoldi, Mauro Cettolo, Marcello Federico (ITC-irst)

NUT-NTT Statistical Machine Translation System for IWSLT 2005

10:25-10:50

Kazuteru Ohashi, Kazuhide Yamamoto, Kuniko Saito and Masaaki Nagata
(Nagaoaka University of Technology & NTT)

10:50-11:05

Coffee Break

Tuning a phrase-based statistical translation system for the IWSLT 2005 Chinese to English and Arabic to English tasks

11:05-11:30

Marta R. Costa-jussa and Jose A. R. Fonollosa
(TALP)

11:30-11:55

The TALP Ngram-based SMT System for IWSLT'05

Josep M. Crego, Adria de Gispert and Jose B. Marino (TALP)

IBM Statistical Machine Translation for Spoken Languages

11:55-12:20

Young-Suk Lee (IBM)

12:20-13:20

Lunch

Session 2: Statistical MT system-Part II-

Chair: *Yutaka Sasaki*

The CASIA Phrase-Based Machine Translation System

13:20-13:45

Wei Pang, Zhendong Yang, Zhenbiao Chen, Wei Wei, Bo Xu and Chengqing Zong (NLPR)

Edinburgh System Description for the 2005 IWSLT Speech Translation Evaluation

13:45-14:10

Philipp Koehn, Amittai Axelrod, Alexandra Birch Mayne, Chris Callison-Burch, Miles Osborne and David Talbot (Edinburgh University)

The NTT Statistical Machine Translation System for IWSLT2005

14:10-14:35

Hajime Tsukada, Taro Watanabe, Jun Suzuki, Hideto Kazawa and Hideki Isozaki (NTT)

The MIT-LL/AFRL MT System

14:35-15:00

Wade Shen, Brian Delaney and Tim Anderson (MIT and Airforce)

15:00-15:15

Coffee Break

Session 3: MT for ASR output

Chair: *Marcello Federico*

The RWTH Phrase-based Statistical Machine Translation System

15:15-15:50

Richard Zens, Oliver Bender, Sasa Hasan, Shahram Khadivi, Evgeny Matusov, Jia Xu, Yuqi Zhang and Hermann Ney (RWTH)

A Decoding Algorithm for Word Lattice Translation in Speech Translation

15:50-16:25

Ruiqiang Zhang, Genichiro Kikui, Hirofumi Yamamoto and Wai-Kit Lo (ATR)

**Using Multiple Recognition Hypotheses to
Improve Speech Translation**

*Ruiqiang Zhang, Genichiro Kikui and Hirofumi
Yamamoto (ATR)*

**The CMU Statistical Machine Translation
System for IWSLT2005**

16:25-17:00

*Sanjika Hewavitharana, Bing Zhao, Almut Silja
Hildebrand, Matthias
Eck, Chiori Hori, Stephan Vogel and Alex Waibel
(CMU)*

17:00-17:15

Coffee break

Session 4: Special session

chair: *Alex Waibel*

**Ongoing and future projects for speech
translation**

17:15-18:15

*Joe Olive (DARPA)
Mari Maeda (DARPA)
Gianni Lazzari (ITC-irst)
Satoshi Nakamura (ATR)*

Banquet

18:45

(Bus transfer from the workshop venue to the
banquet venue)

Tuesday 25th Oct

9:00-9:30

Breakfast

Session 5: Invited Talk

9:30-10:05

Jaime Carbonell (CMU)

Session 6: Example-based, Rule-based and
Hybrid MT

Chair: *Christian Boitet*

**A Pattern-Based Machine Translation
System - Yakushite Net MT Engine**

10:05-10:30

Miki Sasaki and Toshiki Murata (OKI)

10:30-10:55

**Example-based Machine Translation
Pursuing Fully Structural NLP**

*Sadao Kurohashi, Toshiaki Nakazawa and
Kauffmann Alexis, Daisuke Kawahara
(University of Tokyo)*

10:55-11:10 **Coffee Break**

11:10-11:35 **ALEPH: an EBMT system based on the
preservation of proportional analogies
between sentences across languages**

Yves LEPAGE and Etienne DENOVAL (ATR)

11:35-12:00 **Nobody is Perfect: ATR's Hybrid Approach
to Spoken Language Translation**

*Michael Paul, Takao Doi, Youngsook Hwang,
Kenji Imamura, Hideo Okuma and Eiichiro
Sumita (ATR)*

12:00-13:30 **Lunch**

Session 7: Syntax base MT
Chair: *Chengqing Zong*

13:30-13:55 **ISI's 2005 Statistical Machine Translation
Entries**

Steve DeNeefe and Kevin Knight (ISI)

13:55-14:20 **Microsoft Research Treelet Translation
System: IWSLT Evaluation**

Arul Menezes and Chris Quirk (Microsoft)

Session 8: Technical Paper
Chair: *Jun Park*

13:50-14:15 **Evaluating Machine Translation Output with
Automatic Sentence Segmentation**

*Evgeny Matusov, Gregor Leusch, Oliver Bender
and Hermann Ney (RWTH)*

14:15-14:40 **Machine Translation Evaluation Inside
QARLA**

*Jesus Gimenez, Enrique Amigo and Chiori Hori
(UPC, UNED & CMU)*

14:40-15:00 **Coffee Break**

15:00-15:25 **Rapid Development of an Afrikaans-English
Speech-to-Speech Translator**

Herman A. Engelbrecht and Tanja Schultz
(University of Stellenbosch, South Africa &
CMU)

**Integrated Chinese Word Segmentation in
Statistical Machine Translation**

15:25-15:50

*Jia Xu and Evgeny Matusov and Richard Zens
and Hermann Ney (RWTH)*

**Ngram-based versus Phrase-based
Statistical Machine Translation**

15:50-16:15

*Josep M. Crego, Marta R. Costa-jussa, Jose B.
Mario and Jose A. R. Fonollosa*
(TALP)

**Low Cost Portability for Statistical Machine
Translation based on N-gram Frequency
and TF-IDF**

16:15-16:40

Matthias Eck, Stephan Vogel and Alex Waibel
(CMU)

Closing remarks

16:40-17:00

Alex Waibel (CMU/UKA)

Copyright(c) 2005 interACT All rights reserved.