Spoken language translation technologies attempt to cross the language barriers between people with different native languages who each want to engage in conversation by using their mother-tongue. The importance of these technologies is increasing because there are many more opportunities for cross-language communication in face-to-face and telephone conversation, especially in the domain of tourism.

Novel technologies have been proposed to tackle the problems in spoken language translation research. A number of institutes are developing huge bilingual or multilingual spoken language corpora. MT technologies based on machine learning, such as statistical MT and example-based MT, are being applied to the translation of spoken language by using these corpora. Some of the characteristics of spoken language seem suitable for the application of machine-learning-based MT in comparison with written language. However, there is still no concrete standard methodology for comparing the translation qualities of spoken language translation systems.

One of the prominent research activities in spoken language translation is the work being conducted by the Consortium for Speech Translation Advanced Research (C-STAR III), which is an international partnership of research laboratories engaged in automatic translation of spoken language. Current members include ATR (Japan), CAS (China), CLIPS (France), CMU (USA), ETRI (Korea), ITC-irst (Italy), and UKA (Germany). One of C-STAR's ongoing projects is the joint development of a speech corpus that handles a common task in multiple languages. The creation of such a corpus will not only enable translation among multiple languages but will also facilitate exchange and discussion of research results among member labs. As a first result of this activity, a Japanese-English speech corpus comprising tourism-related sentences, originally compiled by ATR, has been translated into the native language of C-STAR members.

In this workshop, an "evaluation campaign" of spoken language translation technologies will be held by using the multilingual speech corpus containing the tourism-related sentences developed by ATR and C-STAR members. Two types of submissions are invited: 1) participants in the evaluation campaign of spoken language translation technologies; and 2) technical papers on related issues. An overview of the evaluation campaign is as follows:

**Main Theme: Evaluation of spoken language translation systems**

**Corpus used for the evaluation campaign:**

- Basic Travel Expression Corpus (BTEC)
- Languages: Chinese-English, Japanese-English
- Domain: tourism-related sentences
Tracks of the Evaluation Campaign:

- Translation Directions:
  - Chinese to English
  - Japanese to English

- Resources Used:
  - Supplied corpus only (C-to-E, J-to-E)
  - Supplied corpus + additional linguistic resources available from LDC (C-to-E)
  - Unrestricted (C-to-E, J-to-E)

Evaluation Methodology of Translated Results

- Subjective Evaluation
- Automatic Evaluation (BLEU, NIST, WER, etc.)

The workshop also invites technical papers related to spoken language translation. Possible topics for the session include, but are not limited to:

- MT Evaluation Measures
- MT Algorithms
- Word / Phrase Alignment
- Multilingual Lexicon / Translation Rule Extraction
- Multilingual Parsing
The objective of this workshop is NOT to organize a competition in order to rank current state-of-the-art machine translation systems. We would rather like to provide a framework for the validation of existing evaluation methodologies concerning their applicability to the evaluation of spoken language translation technologies and open new directions on how to improve current methods. In order to achieve this goal and to support future evaluation research efforts, we plan to release the supplied corpus and obtained translation results (provided that each participant agrees) after the workshop.

The evaluation campaign is carried out using a multilingual speech corpus. It contains tourism-related sentences similar to those that are usually found in phrasebooks for tourists going abroad. Details about this Basic Travel Expression Corpus (BTEC), the different 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

**Supplied Corpus:**

- **[C-to-E]**
  - 20K sentences randomly selected from the BTEC corpus
  - coding:
    - Chinese: EUC-china
    - English: ISO-8859-1
  - word segmentation for Chinese
  - tokenizer applied to English text

- **[J-to-E]**
  - 20K sentences randomly selected from the BTEC corpus
  - coding:
    - Japanese: EUC-japan
    - English: ISO-8859-1
  - word segmentation for Japanese (using CHASEN)
  - tokenizer applied to English text

- data format:
each line consists of two fields divided by the character `\'
  sentence consisting of words divided by single spaces

format: <SENTENCE_ID>\<MT_TRAINING_SENTENCE>
  Field_1: sentence ID
  Field_2: MT training sentence

example:
  00001\this is the first training sentence
  00002\this is the second training sentence

Sample Corpus:

- 50 sentences randomly selected from the supplied corpus for [C-to-E] and [J-to-E], respectively
- sample corpus will be sent two weeks prior to the training corpus release

Test Corpus:

- 500 sentences from the BTEC corpus reserved for evaluation purposes
- coding:
  - Japanese: EUC-japan
  - Chinese: EUC-china
- data format:
  - each line consists of two fields divided by the character `\'
  - sentence consisting of words divided by single spaces
    Note: word segmentation carried out automatically

format: <SENTENCE_ID>\<MT_INPUT_SENTENCE>
  Field_1: sentence ID (as given in the Test Corpus)
  Field_2: MT input sentence

example:
  001\this is the first input sentence
  002\this is the second input sentence

Evaluation Corpus:

- up to 16 man-made English reference translations of the Test Corpus (the original corpus sentences were given to five native speakers of American English who provided up to three paraphrased translations each)
- coding: ISO-8859-1

Data Set Conditions

Small Data Track: (C-to-E, J-to-E)

- the training data of the MT systems is limited to the supplied corpus only

Additional Data Track: (C-to-E)

- The Additional Data condition limits the use of bilingual resources. There are no restrictions on monolingual resources. Besides the supplied corpus, the following bilingual resources available from the LDC are permitted:
  - LDC2000T46 Hong Kong News Parallel Text
  - LDC2000T47 Hong Kong Laws Parallel Text
there are no limitations on the linguistic resources used to train the MT systems.

**Evaluation Specifications**

**Subjective Evaluation:**

- Human assessments of translation quality with respect to the "fluency" and "adequacy" of the translation (similar to the evaluation guidelines used in projects by NIST) is carried out by native speakers of American English using a browser-based evaluation tool.
- "Fluency" indicates how the evaluation segment sounds to a native speaker of English. The evaluator grades the level of English used in the translation using one of the following phrases:
  - "Flawless English"
  - "Good English"
  - "Non-native English"
  - "Disfluent English"
  - "Incomprehensible"
- The "adequacy" assessment is carried out after the fluency judgement is done. The evaluator is presented with the "gold standard" translation and has to judge how much of the information from the original translation is expressed in the translation by selecting one of the following grades:
  - "All of the information"
  - "Most of the information"
  - "Much of the information"
  - "Little information"
  - "None of it"

In order to minimize grading inconsistencies between evaluators due to contextual misinterpretations of the translations, the situation in which the sentence is uttered (corpus annotations like "sightseeing" or "restaurant") is provided for the adequacy judgment.

- Evaluator Assignment:
  - six evaluators per track
  - the test sentences are divided into two, whereby three evaluators are assigned to the first and the remaining three to the second subset
  - each evaluator grades all MT system outputs of the respective subset

Therefore, each translation of a single MT system will be evaluated by three judges.

**Automatic Evaluation:**

- N-gram co-occurrence scoring:
  - BLEU/NIST (mteval-v11a.pl)
WER/PER (word error rate, position-independent WER)
GTM (general text matcher, v1.2)

Evaluation Parameter:
- case insensitive (lower-case only)
- no punctuation marks (remove '.',' ','?','!' '"')
  (periods that are parts of a word should not be removed, e.g., abbreviations, like "mr.", "a.m.", remain as they occur in the corpus data)
- no word compounds (substitute hyphens '-' with blank space)
- spelling-out of numerals

Evaluation Procedure:
- Usage of MT system output file (plain text) as it is. Participants are responsible to provide translation output in agreement with above mentioned MT output constraints.
- A freely available tagger (http://www.cs.jhu.edu/~brill/RBT1_14.tar.Z) will be applied to the MT output and reference translations, respectively.
- Each automatic scoring software is applied to the respective data files.

Run Submission Specifications

Run Submission:
- Access evaluation server (URL will be announced before the test data release).
- Multiple run submissions for each track permitted (one file per each submission).
  - automatic evaluation applied to ALL submissions
  - human assessment only for one run of each track (the participants can select/mark the run that should be evaluated by humans at submission time)
- Submission format of MT output file:
  - plain text file
  - one translation per line
  - each translation consists of a sequence of words separated by a single blank space
  - order of translation has to agree with those of the test set (input data of the MT systems)
  - if MT system fails to output a translation you have to add an empty line in your output files

Note: The submission format differs from the initial format we announced on our website. The change was necessary due to the browser-based interface.

Evaluation Procedure:

Evaluation Server:
- An easy-to-use browser-based interface to upload run submissions.
  - UserId and PassCode are required to access the run submission page
  - specify the type of your MT system
  - select the language and track of your run submission
  - upload your MT output file (see "submission format" below)
  - press "Submit"
- Evaluation scripts will be applied automatically after you have pressed "Submit" and the evaluation results will be sent back to you by email. In case of an error, an error message (instead of the evaluation results) is sent back to you by email. Possible errors are:
  - sentence count mismatch between MT output and reference translations (please correct your MT output results and submit again)
  - non-Ascii characters included in your MT output file (please correct your MT
output results and submit again)
- unsuccessful termination of evaluation tool (we checked it carefully, but ...)

- Data sets provided:
  - development set (CSTAR03)
  - test set (IWSLT04) [from: 08/09 (0:01 JST) until: 08/12 (23:59 JST)]

  **Note:** The automatic feedback of the evaluation results via email will be disabled for the submission of the test set (IWSLT04) results.

The evaluation server will be taken off-line at 08/12 (23:59 JST) in order to prepare the evaluation data for the subjective evaluation. However, we plan to put in on-line again after the data preparations are finished (this time with the email feedback enabled).

**Evaluation Results:**
- For the development set (CSTAR03), the automatic scoring results of each MT system will be sent to the respective participant by email shortly after the run submission. Each participant will receive:
  - a summary of the overall test set scores
  - sentence-wise scoring results (for all evaluation metrics)

- For the test set (IWSLT04), the automatic scoring results of each MT system will be sent to the respective participant by email after the run submission deadline.

- The subjective evaluation results will be sent to the respective participant by email around September 10, 2004. Each participant will receive:
  - the ranking list and evaluation scores of all MT systems (anonymous reference to MT systems using code names)
  - the code name of the participant’s MT system
  - detailed scoring results for each translation (fluency, adequacy)

The reference translation data set, and the MT system output results (given the participants consensus) will be made available to the participants after the workshop. These resources can be used as a benchmark for future research on MT systems and MT evaluation techniques.
Organizers

- **Seiichi Yamamoto** (ATR, Japan; Chair)
- **Christian Boitet** (CLIPS, France)
- **Gianni Lazzari** (ITC-irst, Italy)
- **Youngjik Lee** (ETRI, Korea)
- **Alex Waibel** (CMU, USA / UKA, Germany)
- **Chengqing Zong** (CAS, China)

Program Committee

- **Marcello Federico** (ITC-irst, Italy; Co-chair)
- **Hiromi Nakaiwa** (NTT, Japan; Co-chair)
- **Herve Blanchon** (CLIPS, France)
- **Key-Sun Choi** (KAIST, Korea)
- **Casacuberta Francisco** (ITI, Spain)
- **Sadao Kurohashi** (Univ. of Tokyo, Japan)
- **Hermann Ney** (RWTH, Germany)
- **Franz Josef Och** (Google, USA)
- **Seung-Shin Oh** (ETRI, Korea)
- **Michael Paul** (ATR, Japan)
- **Keh-Yih Su** (Behavior Design, Taiwan)
- **Stephan Vogel** (CMU, USA)
- **Dekai Wu** (HKUST, Hong Kong)
- **Bo Xu** (CAS, China)
- **Chengqing Zong** (CAS, China)

Evaluation Committee [tentative]

- **Jun'ichi Tsujii** (Univ of Tokyo, Japan; Chair)
- **Yasuhiro Akiba** (ATR, Japan)
- **Marcello Federico** (ITC-irst, Italy)
- **Noriko Kando** (NII, Japan)
- **Hiromi Nakaiwa** (NTT, Japan)
- **Michael Paul** (ATR, Japan)
- (representatives of organizations participating in the Evaluation Campaign)

Local Arrangements

- **Eiichiro Sumita** (ATR, Japan)
- **Michael Paul** (ATR, Japan)

Supporting Organizations
## Invited Talk Abstracts

### September 30, 2004

**The Statistical Approach to Spoken Language Translation**

**13:30-14:30 Prof. Hermann Ney** (Rheinisch Westfälische Technische Hochschule)

During the last few years, the statistical approach has found widespread use in machine translation of both written and spoken language. In many comparative evaluations, the statistical approach was found to be competitive or superior to the existing conventional approaches. Like other natural language processing tasks, machine translation requires four major components:

1. an error measure for the decision rule that is used to generate the target sentence from the source sentence;
2. a set of probability models that replace the true but unknown probability distributions in the decision rule,
3. a training criterion that is used to learn the unknown model parameters from training data;
4. an efficient implementation of the decision rule, which is referred to as generation or, like in speech recognition, as search or decoding.

We will consider each of these four components in more detail and review the attempts that have been made to improve the state of the art. In addition, we will address the problem of recognition-translation integration which is specific of spoken language translation.

### October 1, 2004

**How long will we be able to ignore linguistic knowledge and their formalisms?**

**13:30-14:30 Prof. Jun'ichi Tsujii** (Department of Computer Science, University of Tokyo)

The paradigms of MT proposed so far have their own attractions such as SBMT being good for rapid development of MT systems, EBMT for non-compositional translation, etc. However, it is becoming increasingly clear that proper theories of language are also crucial for quality of NLP systems. In this talk, we will argue that grammar in proper linguistic formalisms can improve performances of systems based on ill-conceived grammar, and that it is the time for another paradigm shift in NLP in general and MT in particular.

Our experience in parsing has show a parser that uses linguistically sound formalisms with substantial knowledge of lexical items can not only supersede the performance of
parsers based on arbitrary forms of grammar but also improve adaptability towards specific domain and widen the scope of applicability in actual NLP application systems. Good grammar formalisms also provide better bases for statistical language models. Since MT have to deal with diverse aspects of language, we need to avoid the naïve distinction of different MT paradigms and start to pursue possible integration of good ideas in different paradigms.
Toward the Evaluation of Speech Translation

The evaluation of conversational-speech translation systems rises many technical issues. For the sake of stimulating the discussion, some general problems and proposals are briefly introduced, which will be integrated with the presentations given by the invited panelists.

1. Speech translation requires carefully considering the goal of the task itself. While, e.g., broadcast news translation can be treated similarly to written text translation, different ideas of translation could be considered for conversational speech. For this task, humans professional translators typically refer to three "interpreting modalities": simultaneous, consecutive and liason. Simply speaking, all modalities require the human interpreter to listen to a given amount of speech, to recount what has been said, to listen again, and so on. Probably, the less ambitious scenario for automatic SLT might be the one of simultaneous interpreting, which typically requires the human to translate at very short intervals, e.g. few seconds, or even in real-time. Besides being physically very demanding, simultaneous interpreters, due to the strict time constraints, are less able to exploit their linguistic and domain knowledge. Both reasons make users accept less fluent and almost close to literal translations.

2. Given that speech translation relies on automatic speech recognition (ASR), the task should be tailored to the affordable ASR accuracy. In the past, interlingua-based systems have been applied to resemble the way a liason interpreter works, e.g. at a meeting or appointment. In particular, the interpreter is assumed to be familiar with the subject under discussion and uses psychological skills to facilitate communication. While the mediator metaphor seemed appropriate, especially in the presence of noisy input, interlingua approaches have shown little ability to cope with poor speech recognition performance, and to work significantly worse than purely data-driven translation models. Nevertheless, any plan for speech translation evaluation should take into account progress in the area of speech recognition and scale up difficulty of the considered tasks accordingly.

3. Human and automatic evaluation should take into account important differences
between written and spoken language. Practically, how should input sentences containing disfluencies and syntactic errors be treated? What kind of human translations should be taken as target references? The simultaneous interpreting scenario would suggest to put more emphasis on adequacy rather than fluency. Moreover, appropriate reference translations could be obtained by transcribing human interpreters working in realistic conditions.
## Program

### September 30, 2004

#### Workshop Opening

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:50</td>
<td>Address of Welcome</td>
</tr>
<tr>
<td></td>
<td>Seiichi YAMAMOTO</td>
</tr>
<tr>
<td>9:00</td>
<td>Overview of the IWSLT04 Evaluation Campaign</td>
</tr>
<tr>
<td></td>
<td>Yasuhiro AKIBA, Marcello FEDERICO, Noriko KANDO, Hiromi NAKAIWA, Michael PAUL, Jun'ichi TSUJII</td>
</tr>
<tr>
<td></td>
<td>Coffee break</td>
</tr>
</tbody>
</table>

#### Evaluation Campaign: "Statistical MT" (chair: Chengqing ZONG)

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>&quot;The ISL Statistical Translation System for Spoken Language Translation&quot;</td>
</tr>
<tr>
<td></td>
<td>Stephan VOGEL, Sanjika HEAVVITHARANA, Muntsin KOLSS and Alex WAIBEL</td>
</tr>
<tr>
<td>10:25</td>
<td>&quot;IBM Spoken Language Translation System Evaluation&quot;</td>
</tr>
<tr>
<td></td>
<td>Young-Suk LEE and Salim ROUKOS</td>
</tr>
<tr>
<td>10:50</td>
<td>&quot;The ISI/USC MT System&quot;</td>
</tr>
<tr>
<td></td>
<td>Emil ETTELAIE, Kevin KNIGHT, Daniel MARCU, Dragos Stefan MUNTEANU, Franz J. OCH, Ignacio THAYER and Quamrul TIPU</td>
</tr>
<tr>
<td>11:15</td>
<td>&quot;The ITC-irst Statistical Machine Translation System for IWSLT-2004&quot;</td>
</tr>
<tr>
<td></td>
<td>Nicola BERTOLDI, Roldano CATTONI, Mauro CETTOLO and Marcello FEDERICO</td>
</tr>
<tr>
<td>11:40</td>
<td>&quot;Alignment Templates: the RWTH SMT System&quot;</td>
</tr>
<tr>
<td></td>
<td>Oliver BENDER, Richard ZENS, Evgeny MATUSOV and Hermann NEY</td>
</tr>
<tr>
<td>12:05</td>
<td>&quot;TALP: Xgram-based Spoken Language Translation System&quot;</td>
</tr>
<tr>
<td></td>
<td>Adria DE GISPERT and Jose B. MARINO</td>
</tr>
<tr>
<td></td>
<td>Lunch break</td>
</tr>
</tbody>
</table>

#### Invited Talk (chair: Gianni LAZZARI)

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:30</td>
<td>The Statistical Approach to Spoken Language Translation</td>
</tr>
<tr>
<td></td>
<td>Hermann Ney</td>
</tr>
<tr>
<td>14:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
</tr>
<tr>
<td>14:45</td>
<td>Evaluation</td>
</tr>
<tr>
<td>15:10</td>
<td>Campaign</td>
</tr>
<tr>
<td>15:10</td>
<td>Campaign</td>
</tr>
<tr>
<td>15:35</td>
<td>&quot;Example-based and Rule-based MT&quot; (chair: Stephan VOGEL)</td>
</tr>
</tbody>
</table>

**Banquet**

- **16:40** 17:30 bus transfer from the workshop venue to the banquet venue
- **17:30** 19:30 restaurant "NOU" (located at the 1F of the Nara-ken New Public Hall)

**October 1, 2004**

**Technical Paper : Session 1 (chair: Noriko KANDO)**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00</td>
<td>Technical Paper</td>
<td>&quot;Multi-Lingual Speech Recognition System for Speech-To-Speech Translation&quot;</td>
<td>Satoshi NAKAMURA, Konstantin MARKOV, Takatoshi JITSUHIRO, Jin-Song ZHANG, Hirofumi YAMAMOTO and Genichiro KIKUI</td>
</tr>
<tr>
<td>09:25</td>
<td>Technical Paper</td>
<td>&quot;Minimum Error Training of Log-Linear Translation Models&quot;</td>
<td>Mauro CETTOLO and Marcello FEDERICO</td>
</tr>
<tr>
<td>09:50</td>
<td>Technical Paper</td>
<td>&quot;On Feature Selection in Maximum Entropy Approach to Statistical Concept-based Speech-to-Speech Translation&quot;</td>
<td>Liang GU and Yuqing GAO</td>
</tr>
<tr>
<td>10:15</td>
<td>Technical Paper</td>
<td>&quot;Statistical Machine Translation of Spontaneous Speech with Scarce Resources&quot;</td>
<td>Evgeny MATUSOV, Maja POPOVIC, Richard ZENS and Hermann NEY</td>
</tr>
</tbody>
</table>

**Evaluation Campaign : "Hybrid MT" (chair: Youngjik LEE)**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:55</td>
<td>Evaluation</td>
<td>&quot;EBMT, SMT, Hybrid and More: ATR Spoken Language Translation System&quot;</td>
<td>Eiichiro SUMITA, Yasuhiro AKIBA, Takao DOI, Andrew FINCH, Kenji IMAMURA, Hideo OKUMA, Michael PAUL, Mitsuo SHIMOHATA and Taro WATANABE</td>
</tr>
<tr>
<td>11:20</td>
<td>&quot;Multi-Engine Based Chinese-to-English Translation System&quot;</td>
<td>Yuncun ZUO, Yu ZHOU and Chengqing ZONG</td>
<td></td>
</tr>
<tr>
<td>11:45</td>
<td>&quot;Experimenting with Phrase-Based Statistical Translation within the IWSLT 2004 Chinese-to-English Shared Translation Task&quot;</td>
<td>Philippe LANGLAIS, Michael CARL and Oliver STREITER</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:10</td>
<td>&quot;The ISL EDTRL System&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Juergen REICHERT and Alex WAIBEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lunch break</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13:30</td>
<td>Invited Talk (chair: Herve BLANCHON)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>How long will we be able to ignore linguistic knowledge and their formalisms?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jun'ichi TSUJII</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14:45</td>
<td>Technical Paper : Session 2 (chair: Satoshi NAKAMURA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;Phrase-based alignment combining corpus cooccurrences and linguistic knowledge&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adria DE GISPERT, Jose B. MARINO and Josep M. CREGO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:10</td>
<td>&quot;Toward Named Entity Extraction and Translation in Spoken Language Translation&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fei HUANG, Stephan VOGEL and Alex WAIBEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:35</td>
<td>&quot;Spoken Dialogue Translation Systems Evaluation: Results, New Trends, Problems and Proposals&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Herve BLANCHON, Christian BOITET and Laurent BESACIER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:00</td>
<td>&quot;PolyphraZ : a tool for the quantitative and subjective evaluation of parallel corpora&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Najeh HAJLAOUI and Christian BOITET</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:40</td>
<td>Open Discussion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;Toward the Evaluation of Speech Translation&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marcello FEDERICO (ITC-irst) [moderator]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Young-suk LEE (IBM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hermann NEY (RWTH)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stephan VOGEL (CMU)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17:40</td>
<td>Workshop Closing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Closing Remarks</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Copyright(c) 2004 ATR All rights reserved.