PACLIC18

# Proceedings of The 18th Pacific Asia Conference on Language, Information and Computation

## December 8th -10th, 2004
## Waseda University, Tokyo, Japan

Edited by
Hiroshi Masuichi, Tomoko Ohkuma, Kiyoshi Ishikawa,
Yasunari Harada and Kei Yoshimoto

**DATES:** December 8 (Wednesday) - 10 (Friday), 2004
**VENUE:** International Conference Center, Waseda University, Tokyo

**ORGANIZED BY**
The Logico-Linguistic Society of Japan

**SPONSORED BY**
Waseda University
Media Network Center, Waseda University

**SUPPORTED BY**
Institute for DECODE, Waseda University
Language and Speech Science Research Laboratories, Waseda University
Corporate Research Group, Fuji Xerox Co., Ltd.

**ACKNOWLEDGMENTS**
Preparations for PACLIC18 were financially supported in part by Waseda University Grant for Special Research Projects: International Joint Research No. 2003C-201.
PACLIC18 is financially supported by Waseda University Grant for International Conference Operation.

Copyright © 2004 PACLIC 18 Organizing Committee and PACLIC Steering Committee.
All rights reserved.
Copyright of contributed papers reserved by respective author(s).

# Foreword

On behalf of the Logico-Linguistic Society of Japan, let me express our pleasure to have this opportunity to host PACLIC 18, the 18th Pacific Asia Conference on Language, Information and Computation, at Waseda University in Tokyo.

PACLIC 18 is the nineth conference to convene under the present name "Pacific Asia Conference on Language, Information and Computation," which was inaugurated in Kyoto in 1994, replacing its predecessors "Korean-Japan Joint Conference on Formal Linguistics" and "Pacific Asia Conference on Formal and Computational Linguistics."

In the foreword written on the occasion of PACLIC 14 held in 2000 at Waseda University, I wrote as follows: "The past ten years was the time in which we have witnessed so many wars and so much strife among nations and races. One of the main causes of these tragedies was and still is language differences we employ." Regrettably, this is still true nowadays. Not a single day passes without the news of murder and killing among nations and races. Therefore, we cannot overemphasize the importance of language study since it is through language that we can help to create and achieve peace in the world.

The PACLIC conferences, as is well known, have a long tradition of putting strong focus on the intersection of linguistics and computer science. This focus has a vital importance in the days when we experience a tremendous gap between several branches of language study and the research on computer science.

While the PACLIC conferences have provided a sense of family-like community for the people who attend the conference, we have employed anonymous review process since PACLIC 10, sending out each paper to reviewers for evaluation. With the review results at hand, the Program Committee of this conference selected 18 Oral Presentations and 11 Interactive Presentations out of 53 submissions.

In this conference, we are fortunate to be able to invite the following speakers, Dr. Chu-Ren Huang at Academia Sinica, Dr. Kiyong Lee at Korea University and Dr. Yuji Matsumoto at Nara Institute of Science and Technology.

Our next PACLIC conference is going to be held in Taiwan at Academia Sinica in December 2005. We have also received an important piece of news that China and Thailand have proposed to host the following conferences. With the two additional countries hosting future installments of PACLIC conferences, this series will become "Pacific Asia Conference" in the true sense of the word.

Our appreciation is extended to our committee members. Just to name a few, we would like to thank Dr. Chu-Ren Huang and Dr. Beom-mo Kang for serving as Program Co-Chairs, Dr. Kei Yoshimoto for serving as Acting Program Chair, Prof. Yasunari Harada for serving as Organizing Chair, Dr. Kiyoshi Ishikawa for serving as Organizing Co-Chair and Dr. Hiroshi Masuichi and Ms. Tomoko Ohkuma for serving as key members of PACLIC 18 secretariat.

Last but not least, I want to take this opportunity to thank Waseda University for offering financial support in planning and hosting this conference, Institute for Digital Enhancement of Cognitive Development and Language and Speech Science Research Laboratories of Waseda University and Corporate Research Group of Fuji Xerox Co., Ltd. for providing human and network resources to assist organization and operation of PACLIC 18th.

---

**Akira IKEYA**
Emeritus Professor of Tokyo Gakugei University
Honorary Chair of PACLIC 18 Program Committee
President of the Logico-Linguistic Society of Japan

# Table of Contents

Foreword 3

Invited Talk 7

**Processing and Representing Temporally Sequential Events** 9
Kiyong Lee, Korea University

**Machine Learning based NLP: Experiences and Supporting Tools** 15
Yuji Matsumoto, Nara Institute of Science and Technology

**Text-based Construction and Comparison of Domain Ontology: A Study Based on Classical Poetry** 17
Chu-Ren Huang, Academia Sinica

Oral Session 21

**An Adjacency Constraint on Argument Selection** 23
Kei Takahashi and Kiyoshi Ishikawa

**Towards a Proper Treatment of Adjuncts in Japanese** 35
Kenji Yokota

**An Analysis of the Korean [manyak ... V-telato] Construction: An Indexed Phrase Structure Grammar Approach** 47
Hee-Rahk Chae

**An Analysis of Japanese ta / teiru in a Dynamic Semantics Framework and a Comparison with Korean Temporal Markers a nohta / a twuta** 59
Yoko Mizuta

**Relational Nouns as Anaphors** 71
Hiroaki Nakamura and Yoshiki Mori

**Capturing and Parsing the Mixed Properties of Light Verb Constructions in a Typed Feature Structure Grammar** 81
Jong-Bok Kim, Jaehyung Yang and Incheol Choi

**Japanese Subjects and Information Structure: A Constraint-based Approach** 93
Akira Ohtani and Yuji Matsumoto

**High WSD Accuracy Using Naive Bayesian Classifier with Rich Features** 105
Cuong Anh Le and Akira Shimazu

**Automatic Discovery of Telic and Agentive Roles from Corpus Data** 115
Ichiro Yamada and Timothy Baldwin

**Bilingual Knowledge Extraction Using Chunk Alignment** 127
Young-Sook Hwang, Kyonghee Paik and Yutaka Sasaki

**Pruning False Unknown Words to Improve Chinese Word Segmentation** 139
Chooi-Ling Goh, Masayuki Asahara and Yuji Matsumoto

**Ontology-based Prediction of Compound Relations: A Study Based on SUMO** 151
Jia-Fei Hong, Xiang-Bing Li and Chu-Ren Huang

**Treebank-Based Acquisition of a Chinese Lexical-Functional Grammar** 161
Michael Burke, Olivia Lam, Aoife Cahill, Rowena Chan, Ruth O'Donovan, Adams Bodomo, Josef van Genabith and Andy Way

**Constructing English Reading Courseware** 173
Masao Utiyama, Midori Tanimura and Hitoshi Isahara

**Acquiring Compound Word Translations Both Automatically and Dynamically** 181
Yujie Zhang and Hitoshi Isahara

**Integrated Use of Internal and External Evidence in the Alignment of Multi-Word Named Entities** 187
Takeshi Kutsumi, Takehiko Yoshimi, Katsunori Kotani, Ichiko Sata and Hitoshi Isahara

**Tiny Corpus Applications with Transformation-Based Error-Driven Learning: Evaluations of Automatic Grammar Induction and Partial Parsing of SaiSiyat** 197
Zhemin Lin and Li-May Sung

**A Unified Model of Thai Romanization and Word Segmentation** 205
Wirote Aroonmanakun and Wanchai Rivepiboon

Interactive Session 215

**Scalar Meanings of the Concessive (-to), the Contrastive Topic Marker (-nun) and -man 'only' in Korean (and Japanese)** 217
Chungmin Lee

**On Argument-Adjunct Asymmetry of Sluicing in Mandarin Chinese** 227
Li-Chi Lee Chen

**A Contrastive Study of Function Verbs in English and Japanese: Cut and Kiru** 235
Asako Otomo

**Effects of Mora Phonemes on Japanese Word Accent** 243
Yasuyo Tokuhiro and Shizuo Hiki

**Incorporating Pronunciation Variation into Different Strategies of Term Transliteration** 251
Jin-Shea Kuo and Ying-Kuei Yang

**Extraction of Cognition Results of Travel Routes with a Thesaurus** 259
Kazutaka Takao and Yasuo Asakura

**Adaptive Word Sense Tagging on Chinese Corpus** 267
Sue-Jin Ker and Jen-Nan Chen

**Generating Paired Transliterated-cognates Using Multiple Pronunciation Characteristics from Web Corpora** 275
Jin-Shea Kuo and Ying-Kuei Yang

**Chinese-English Parallel Corpus Construction and its Application** 283
Baobao Chang

**Developing an Automated Test of Spoken Japanese** 291
Yasunari Harada, Masanori Suzuki and Jared Bernstein

**Three English Learner Assistance Systems Using Automatic Paraphrasing Techniques** 299
Masaki Murata and Hitoshi Isahara

Committees 307