Language, Information, and Computation

言語情報計算

Proceedings of The 16th Pacific Asia Conference

Edited by
Ik-Hwan Lee
Yong-Beom Kim
Key-Sun Choi
Minhaeng Lee

The Korean Society for Language and Information

Language, Information, and Computation

Proceedings of
The 16th Pacific Asia Conference

January 31 - February 2 2002, Jeju, Korea

Edited by
Ik-Hwan Lee, Yong-Beom Kim, Key-Sun Choi, and Minhaeng Lee

The Korean Society for Language and Information

©2002 The Korean Society for Language and Information

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

Published by

The Korean Society for Language and Information (KSLI)

c/o Prof. Minhaeng Lee

Division of European Studies

College of Liberal Arts

Yonsei University

Seoul 120-749

Republic of Korea

The publication of this proceedings was financially supported by Korea Research Foundation.

FOREWORD

We are very pleased to be able to host the 16th Pacific Asia Conference on Language Information and Computation (PACLIC 16) with the help from various scholars, institutions, and universities despite the unfavorable economic and political situations on the global scale. With the turn of the century we are witnessing the quality of our life depends more on the development in the software domains than on the abundance in the material matters. We are proud that many of the scholars who are here are leading this new trend of history in their own country. Now we have gathered together in this beautiful island of Jeju to strengthen our positions in our own field and to pursue and widen our academic curiosity and perspectives into the neighboring fields.

Since the 1994 amalgamation of the two organizations, namely the Asian Conference on Language, Information and Computation (ACLIC) and the Pacific Conference on Formal and Computational Linguistics (PACFoCoL), there have been beneficial interactions between theoretical linguistics and computational fields of research. These seem to be a few of the major contributions that this annual conference has made since its inception.

As the host of this conference, the Korean Society for Language and Information (KSLI) carried out a long process of soliciting and selecting papers. The call for extended abstracts drew 56 responses from various countries including Canada, Germany, Hong Kong, Japan, Korea, Singapore, Taiwan, and United States. The Program Committee of PACLIC 16 conducted a blind review process, and sent out each submission to three reviewers. With the evaluations and comments from the reviewers on hand, the Organizing Committee and the Program Committee of this conference got together and also took on a blind process of deciding how many papers to be presented during the conference. This joint process selected 43 out of 56 papers.

We would like to thank various scholars and institutions for their help. We would like to express our deep gratitude to our distinguished guest speakers, Erhard W. Hinrichs and Kiyong Lee for their keynote lectures. We are also grateful to our co-host Cheju University and for the financial support and organizational matters. We would also like to acknowledge the generous financial support from the Korea Research Foundation. Last, but not the least, we would also like to thank the Research Institute of Humanities and Social Sciences, Kwangwoon University for its financial support.

There are numerous individuals, without whose help the conference would have not been possible. I wish to express my gratitude to Suk-Jin Chang (Conference Chair) and the Advisory Committee members including Jason S. Chang, Chu-Ren Huang, Akira Ikeya, Chungmin Lee, Young-Hern Lee, Kim Teng Lua, Byung-Soo Park and Benjamin K. T'sou, for their advice

and support.

I am also very grateful to Ik-Hwan Lee (Chair of Program Committee), Key-Sun Choi (Co-chair of Program Committee), and the members of the program committee and invited reviewers. I wish to note that Minhaeng Lee has been playing the pivotal role in organizing this conference, and I am also grateful for the help of the three local organizers in Jeju, Chang-Ik Lee, Ki-Suk Lee and Yosong Park whose active roles have been instrumental in running this conference. I am no less grateful to Tom Lai and Olivia Kwong for their friendly support.

Yong-Beom Kim
Chair, Organizing Committee
President, Korean Society for Language and Information

COMMITTEE MEMBERS

Conference Chair

Suk-Jin Chang	Seoul National University, Korea

Advisory Committee

Akira Ikeya, Co-chair.	Toyo Gakuen University, Japan
Jason S. Chang, Co-chair.	Tsing Hua University, Taiwan
Chu-Ren Huang, Co-chair.	Academia Sinica, Taiwan
Chungmin Lee, Co-chair.	Seoul National University, Korea
Young-Hern Lee, Co-chair.	Chosun University, Korea
Kim Teng Lua, Co-chair.	National University of Singapore, Singapore
Byung-Soo Park, Co-chair.	Kyung Hee University, Korea
Benjamin K. T'sou, Co-chair.	City University of Hong Kong, Hong Kong

ORGANIZING COMMITTEE

Yong-Beom Kim, Chair.	Kwangwoon University, Korea
Minhaeng Lee, Secretary General.	Yonsei University, Korea
Zhao-ming Gao	National Taiwan University, Taiwan
Yasunari Harada	Waseda University, Japan
Masahito Kawamori	NTT Research Labs, Japan
Yuji Matsumoto	Nara Advanced Institute of Science and Technology, Japan
Tom b.y Lai	The City University of Hong Kong, Hong Kong
Chang-Ik Lee	Cheju National University, Korea
Ki-Suk Lee	Cheju National University, Korea
Yeo-Sung Park	Cheju National University, Korea
Shu-chuan Tseng	Academia Sinica, Taiwan
Kam-fai Wong	The Chinese University of Hong Kong, Hong Kong

PROGRAM COMMITTEE

Ik-Hwan Lee, Chair.	Yonsei University, Korea
Key-Sun Choi, Co-chair.	Korean Advanced Institute of Science and Technology, Korea
Jae-Woong Choe	Korea University, Korea
Roland Hausser	The University of Erlangen-Nuremberg, Germany
Erhard Hinrichs	The University of Tübingen, Germany
Akira Ishikawa	Sophia University, Japan
Kiyoshi Ishikawa	Hosei University, Japan
Dong Hong Ji	Kent Ridge Digital Labs, Singapore
Xu Jie	National University of Singapore, Singapore
Makoto Kanazawa	Tokyo University, Japan
Beom-Mo Kang	Korea University, Korea
Andreas Kathol	UC Berkeley, USA
Sue-jin Ker	Soochow University, Taiwan
Jong-Bok Kim	Kyung Hee University, Korea
Olivia Kwong	The City University of Hong Kong, Hong Kong
Kiyong Lee	Korea University, Korea.
Robert Luk.	The Hong Kong Polytechnic University, Hong Kong
Yong-Kyoon No	Chungnam National University, Korea
Toshiyuki Ogihara	Kanda Gaigo University, Japan
Byung-Rae Ryu	Chungnam National University, Korea.
Markus Schulze	The University of Erlangen-Nuremberg, Germany
Peter Sells	Stanford University, USA
Lily I-Wen Su	National Taiwan University, Taiwan
Hae-Kyung Wee	Sejong University, Korea.
John Whitman	Cornell University, USA
Dan-Hee Yang	Pyungtaek University, Korea
Young-Eun Yoon	Ewha University, Korea
Kei Yoshimoto	Tohoku University, Japan

ADDITIONAL REVIEWERS

Dale Gerdemann	The University of Tübingen, Germany
Thomas Künneth	The University of Erlangen-Nuremberg, Germany
Frank Richter	The University of Tübingen, Germany
Manfred Sailer	The University of Tübingen, Germany

CONFERENCE PROGRAM

DAY 1 (Thursday, 31 January 2002)

Room 1

08:30-09:00	*Registration*
09:00-09:15	**Session Chair:** *Ik-Hwan Lee*
	Opening Ceremony *Suk-Jin Chang* (Conference Chair, SNU) *Man-Keun Boo* (President of Cheju National University)
09:15-10:15	**Session Chair:** *Yong-Beom Kim*
	Keynote Speech: Robust Syntactic Annotation of Corpora and Memory-Based Parsing *Erhard Hinrichs* (University of Tuebingen)
10:15-11:15	**Session Chair:** *Richard S. Lavin*
	An Event-Based Semantics for Japanese Emphatic Particles *Akira Ishikawa* (Sophia University)
	Toward a Bilingual Legal Term Glossary from Context Profiles *Oi Yee Kwong* (City U of Hong Kong)
11:15-11:30	*Coffee/Tea Break*
11:30-12:10	**Session Chair:** *Byung-Soo Park*
	Morphological Passivization and the Change of Lexical-semantic Structures in Korean *Yoon-shin Kim* (Seoul National U)
	Syntactic Interference in Chinese-English Bilingual Children *Erin Yao-ling Wang* (National Taiwan University)
12:10-14:00	*Lunch*
14:00-15:00	**Session Chair:** *Benjamin K. T'sou*
	The Loom-LAG for Syntax Analysis - Adding a Language-independent Level to LAG *Markus Schulze* (University of Nuremberg-Erlangen)
	Issues in Chinese Prosody: Conceptual Foundations of a Linguistically-motivated Text-to-Speech System for Mandarin *Richard S. Lavin* (Kyushu Tokai U)

15:00-15:40	**Session Chair:** *Erhard Hinrichs*
	Disambiguation of Coordinate Expressions in Japanese by Extracting Mutual Case Relation *Yoshiyuki Kotani* (Tokyo U of Agriculture and Technology)
	A Verb-Pattern Based Korean-Chinese Machine Translation System *Changhyun Kim / Young Kil Kim / Mun Pyo Hong / Young Ae Seo / Seong Il Yang / Seung Kwon Choi* (ETRI)
15:40-16:00	*Coffee/Tea Break*
16:00-17:00	**Session Chair:** *Chungmin Lee*
	Representing Topic-comment Structures in Chinese *Haihua Pan / Jianhua Hu* (City U of Hong Kong)
	A Focus Structure-Based Approach to Scopal Ambiguity in Japanese *Ryoya Okabe* (Sophia University)
17:00-18:00	**Session Chair:** *Sue J. Ker*
	The Interface between Syntax and Morphology: Taiwanese Verbal Complexes *Huei-Ling Lin* (National Chung Cheng U)
	Finite Small Clauses in Japanese and their Theoretical Implications *Masakazu Kuno* (University of Tokyo)
	Small Clauses and Default Case *Youngjun Jang* (Chung-Ang University)
18:00-18:30	*Business Meeting*

DAY 2 (Friday, 1 February 2002)

Room 1

09:00-10:00	**Session Chair:** *Akira Ikeya*
	Keynote Speech: A Simple Syntax for Complex Semantics *Kiyong Lee* (Korea University)
10:00-11:00	**Session Chair:** *Beom-Mo Kang*
	The Structure of Polysemy: A study of Multi-sense Words based on WordNet *Jen-Yi Lin / Chang-Hua Yang / Shu-Chuan Tseng / Chu-Ren Huang* (Academia Sinica)
	A Korean Noun Semantic Hierarchy (WordNet) Construction *Juho Lee / Koaunghi Un / Hee-Sook Bae / Key-Sun Choi* (KAIST)
11:00-11:15	*Coffee/Tea Break*
11:15-12:15	**Session Chair:** *Akira Ishikawa*
	Argumentness and Probabilistic Case Structures *Dan-Hee Yang* (Pyongtaek U) / *Ik-Hwan Lee* (Yonsei U)
	Penn Korean Treebank: Development and Evaluation *Chung-hye Han* (Simon Fraser U) / *Na-Rae Han, Eon-Suk Ko,* *Martha Palmer* (U of Pennsylvania) / *Hee-Jong Yi* (U of Delaware)
12:15-13:45	*Lunch*
13:45-15:15	**Session Chair:** *Key-Sun Choi*
	An Operator Assisted Call Routing System *Chun-Jen Lee / Jason S. Chang* (Chunghwa Telecom Co.)
	Heuristic-based Korean Co-reference Resolution for Information Extraction *Euisok Chung / Soojong Lim / Bo-Hyung Yun* (ETRI)
	Building a Domain-specific French-Korean Lexicon *Aesun Yoon* (Pusan National U)
15:15-15:30	*Coffee/Tea Break*
15:30-16:30	**Session Chair:** *Oi Yee Kwong*
	A Korean Homonym Disambiguation System Based on Statistical Model Using Weights *Jun-Su Kim / Wang-Woo Lee / Cheol-Young Ock* (Ulsan University)
	Concession and Linguistic Inference *Yong-Beom Kim* (Kwangwoon University)

16:30-17:30	**Session Chair:** *Hidekazu Suzuki*
	This Adverbial Accusative: A Corpus-Based Observations and More *Michiyuki Sato* (Sophia University)
	On Negative Imperatives in Korean *Chung-hye Han* (Simon Fraser U) / *Chungmin Lee* (Seoul National U)
17:30-17:45	*Coffee/Tea Break*
17:45-18:45	**Session Chair:** *Dongli Han*
	If a Quantifier is not floated, but moored or even incorporated: Complexity of Presuppositions in Local Domains *Yoshiki Mori* (Tsukuba U) / *Kei Yoshimoto* (Tohoku U)
	An Alignment Based Technique for Text Translation on Traditional Chinese and Simplified Chinese *Sue J. Ker / Chun-Hsien Lin* (Soochow University)
	An important Issue in Data Mining – Data Cleaning *Qi Xiao Yang / Sung Sam Yuan / Lu Chan / Jay Rajasekera* (IHPC)
18:45-19:30	*Break*
19:30-21:00	*Reception*

Room 2 / *DAY 2 (Friday)*

12:15-13:45	*Lunch*
13:45-15:15	**Session Chair:** *Qi Xiao Yang*
	Mismatches in Korean Copula Constructions and Linearization Effects *Chan Chung* (Dongseo U) / *Jong-Bok Kim* (Kyung Hee U)
	Double Subject, Double Nominative Object and Double Accusative Object Constructions in Japanese and Korean *Hiroaki Nakamura* (Japan Coast Guard Academy)
	A Phase-based Approach to ECM across CP in Korean *Youngsun Kim* (Cornell University)
15:15-15:30	*Coffee/Tea Break*
15:30-16:30	**Session Chair:** *Young-Joo Kim*
	Implicit Adjuncts: The Cases of Degree Modifiers in Japanese and English *Akira Ikeya* (Toyo Gakuen U) / *Hisako Ikawa* (Tsurumi University)
	Type Construction of Nouns with the Verb *ha-* 'do' *Seohyun Im / Chungmin Lee* (Seoul National U)
16:30-17: 30	**Session Chair:** *Markus Schulze*
	A Japanese Compound Verb V-*te-iku* and Event Composition *Eri Tanaka* (Osaka U)
	Analysis of Korean Predicative Verb Forms in LAG Framework *Soora Kim* (University of Nuremberg-Erlangen)
17:30-17:45	*Coffee/Tea Break*
17:45-18:45	**Session Chair:** *Minhaeng Lee*
	Grammaticalization and Semantic Typology: Time-relationship Adverbs in Japanese, Korean, English, and German *Tetsuharu Moriya*(Kanazawa University)/*Kaoru Horie*(Tohoku University)
	A Deterministic Method for Structural Analysis of Compound Words in Japanese *Dongli Han / Takeshi Ito / Teiji Furugori* (U of Electro-Communications)
	Edge-Integrity and the Syllable Structure in Korean *Eungyeong Kang* (Cornell University)
18:45-19:30	*Break*
19:30-21:00	*Reception*

DAY 3 (Saturday, 2 February 2002)

Room 1

09:30-11:00	**Session Chair:** *Michiyuki Sato*
	Identification of Chinese Personal Names in Unrestricted Texts *Lawrence Cheung / Benjamin K. Tsou / Maosong Sun* (City U of Hong Kong)
	Do 'transitive adjectives' Really Exist? *Byung-Soo Park* (Kyung Hee U)
	Backward Anaphora *Hyeran Lee* (Kyung Hee U)
11:00-11:20	*Coffee/Tea Break*
11:20-12:00	**Session Chair:** *Aesun Yoon*
	Generating a Category Set of Words Using a Hierarchical Part-of-Speech System and Tagged Corpus *Takeyuki Kojima / Yoshiyuki Kotani* (Tokyo U of Agriculture and Technology)
	Implementation of Long-distance Reflexives in Korean *Yong-hun Lee* (U of Illinois at Urbana-Champaign)
12:00-13:00	*Lunch and Closing*
13:00-13:30	*Break*
13:30-18:00	*City Tour (optional)*

TABLE OF CONTENTS

Foreword i
Committee Members iii
Conference Program vi

SPECIAL LECTURES

Erhard W. Hinrichs
Robust Syntactic Annotation of Corpora and Memory-based Parsing 1

Kiyong Lee
A Simple Syntax for Complex Semantics 2

PAPERS

Lawrence Cheung, Benjamin K. T'sou and **Maosong Sun**
Identification of Chinese Personal Names in Unreistircted Texts 28

Chan Chung and **Jong-Bok Kim**
Mismatches in Korean Copula Constructions and Linearization Effects 36

Euisok Chung, Soojong Lim and **Bo-Hyun Yun**
Heuristic-based Korean Coreference Resolution for Information Extraction 50

Chung-hye Han and **Chung-min Lee**
On Negative Imperatives in Korean 59

Chung-hye Han, Na-Rae Han, Eon-Suk Ko, Martha Palmer and **Heejong Yi**
Penn Korean Treebank: Development and Evaluation 69

Dongli Han, Takeshi Ito and **Teiji Furugori**
A Deterministic Method for Structural Analysis of Compound Words in Japanese 79

Akira Ikeya and **Hisako Ikawa**

Implicit Adjuncts: The Cases of Degree Modifiers in Japanese and English 92

Seohyun Im and **Chungmin Lee**

Type Construction of Nouns with the Verb *ha*-'do' 103

Akira Ishikawa

An Event-Based Semantics for Japanese Emphatic Particles 113

Youngjun Jang

Small Clauses and Default Case 123

Eungyeong Kang

Edge-Integrity and the Syllable Structure in Korean 135

Sue J. Ker and **Chun-Hsien Lin**

An Alignment Based Technique for Text Translation between Traditional Chinese and Simplified Chinese 147

Changhyun Kim, Young Kil Kim, Munpyo Hong, Young Ae Seo, Sung Il Yang and **Sung-Kwon Choi**

Verb Pattern Based Korean-Chinese Machine Translation System 157

Jun-Su Kim, Wang-Woo Lee, Chang-Hwan Kim and **Cheol-young Ock**

A Korean Homonym Disambiguation System Based on Statistical Model Using Weights 166

Soora Kim

Analysis of Korean Predicative Verb Forms in LAG Framework 177

Yong-Beom Kim

Concession and Linguistic Inference 187

Yoon-shin Kim

Morphological Passivization and the Change of Lexical-Semantic Structures in Korean 195

Youngsun Kim

A Phase-based Approach to ECM across CP in Korean 205

Takeyuki Kojima and **Yoshiyuki Kotani**

Generating a Category Set of Words Using a Hierarchical Part-of-Speech System and Tagged Corpus 217

Yoshiyuki Kotani

Disambiguation of Coordinate Expressions in Japanese by Extracting Mutual Case Relation 227

Masakazu Kuno

Finite Small Clauses in Japanese and Their Theoretical Implications 237

Oi Yee Kwong

Toward a Bilingual Legal Terms Glossary from Context Profiles 249

Richard S. Lavin

Issues in Chinese Prosody: Conceptual Foundations of a Linguistically-Motivated Text-to-Speech for Mandarin 259

Chun-Jen Lee and **Jason S. Chang**

An Operator Assisted Call Routing System 271

Hyeran Lee

Backward Anaphora 281

Juho Lee, Koaunghi Un, Hee-Sook Bae and **Key-Sun Choi**

A Korean Noun Semantic Hierarchy (Wordnet) Construction 290

Yong-hun Lee

Implementation of Long-distance Reflexives in Korean 296

Huei-Ling Lin

The Interface between Syntax and Morphology: Taiwanese Verbal Complexes 308

Jen-Yi Lin, Chang-Hua Yang, Shu-Chuan Tseng and **Chu-Ren Huang**
The Structure of Polysemy: A Study of Multi-sense Words Based on WordNet 320

Yoshiki Mori and **Kei Yoshimoto**
If a Quantifier is not floated, but moored or even incorporated:
Complexity of Presuppositions in Local Domain 330

Tetsuharu Moriya and **Kaoru Horie**
Grammaticalization and Semantic Typology: Time-relationship Adverbs in
Japanese, Korean, English and German 348

Hiroaki Nakamura
Double Subject, Double Nominative Object and Double Accusative Object
Constructions in Japanese and Korean 358

Ryoya Okabe
A Focus-based Approach to Scope Ambiguity in Japanese 370

Haihua Pan and **Jianhua Hu**
Representing Topic-Comment Structures in Chinese 382

Byung-Soo Park
Do "Transitive Adjectives" Really Exist? 391

Michiyuki Sato
This Adverbial Accusative: A Corpus-based Observation and More 404

Markus Schulze
The Loom-LAG for Syntax Analysis - Adding a Language-independent Level to
LAG 411

Eri Tanaka
A Japanese Compound Verb V-*te-iku* and Event Composition 421

Erin Yaoling Wang
Syntactic Interference in Chinese-English Bilingual Children 433

Dan-Hee Yang and **Ik-Hwan Lee**

Argumentness and Probabilistic Case Structures 447

Qi Xiao Yang, Sung Sam Yuan, Lu Chun and **Jay Rajasekera**

An Important Issue in Data Mining - Data Cleaning 455

Aesun Yoon

Building a Domain-Specific French-Korean Lexicon 465

AUTHOR INDEX 475