

# **27th Pacific Asia Conference on Language, Information and Computation 2013**

**(PACLIC 27)**

**Taipei, Taiwan  
21 - 24 November 2013**

ISBN: 978-1-63266-288-0

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2013) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

### Main Conference Keynote Speeches

1. Words and Rules Revisited: Reassessing the Role of Construction and Memory in Language..... 1  
*Crez'O arantz*
2. A Principle-Based Approach for Natural Language Processing: Eliminating the Shortcomings of Rule-Based Systems with Statistics ..... 2  
*Wen/Nian J sw*
3. Extracting “Critical Features” for the CEFR Levels Using Corpora of EFL Learners’ Written Essays..... 4  
*[ wno'Vqno*
4. It’s about Time: More and More Sophisticated Statistical Methods in Corpus Linguistics..... 8  
*Uelan'Vj 0I ries*
5. Linking Text with Data and Knowledge Bases..... 9  
*Lwiej i'Vswlii*

### Workshop Keynote Speeches

1. Mining Language Learners’ Production Data for Understanding of L2 Learning Systems..... 11  
*[ wno Vono*
2. Introducing Linggle: From Concordance to Linguistic Search Engine ..... 12  
*Lason'UEj ani*

### Main Conference Invited Talks

1. Statistical Machine Translation Based on Predicate-Argument Structure ..... 15  
*Ej gpi s kpi \ qpi*
2. Using Hierarchically Structured Lexicon as Key Clues Solving Data Sparseness Problems in Word Sense Disambiguation: A Case for Korean and Its Applications to English and Chinese..... 17  
*Ceswn'I qon.'O kpj q'Mko . "cpf "J { wmlEj wlmMy qp*
3. Effects of Constituent Orders on Grammaticalization Patterns of the Serial Verbs for ‘Give’ in Thai and Mandarin Chinese ..... 18  
*Mkpi nctp'Vj grncplcpc"cpf "Ucxyuj k'Wgj ctc*

## PACLIC 27 Papers

### Oral Presentation Session 1A: Grammar and Syntax

1. Global Approach to Scalar Implicatures in Dynamic Semantics.....27  
*Lae/m'f gqo*
2. Unification of Numeral Classifiers and Plural Markers: Empirical Facts and Implications.....37  
*Qne/Uqqp'J gt 'cpf 'f wp/Tw'Ej gp*
3. Head-Internal Relatives in Japanese as Rich Context-Setters.....47  
*Voj rw Ugtcmw*

### Oral Presentation Session 1B: Language Generation

1. An Abstract Generation System for Social Scientific Papers .....57  
*Okej kq'Mcpqmq'cpf 'F qpi rk'J cp*
2. Automatic Utterance Generation in Consideration of Nominatives and Emoticon Annotation .....66  
*[ wawmg'Pkij kq.'Ok c'k'O km c.'cpf 'F qpi rk'J cp*
3. Ensemble Approach for Fine-Grained Question Classification in Bengali ..... P IC

### Oral Presentation Session 2A: Discourse and Pragmatics

1. Prosodic Convergence, Divergence, and Feedback: Coherence and Meaning in Conversation .....85  
*Nkej kqpi 'f cpi*
2. A Quantitative Comparative Study of Prosodic and Discourse Units, the Case of French and Taiwan Mandarin .....92  
*Ncwt gpv'Rt gxqv.'Uj w'Ej wcp'Vugpi . 'Cnkp'Ej gpi /J ukgp'Ej gp.'cpf 'Mko 'Rguj nqx*
3. Corpus-Based Research on Tense Analysis and Rhetorical Structure in Journal Article Abstracts.....102  
*Rkp/pkpi 'Vw'cpf 'Uj kj /Rkpi 'Y cpi*

## Oral Presentation Session 2B: Lexical Knowledge Learning

1. A Novel Schema-Oriented Approach for Chinese New Word Identification 108  
*\j cq'Nw '\j kkp' [ cp.'cpf 'Lwp/j qpi 'T w*
2. A Study of the Effectiveness of Suffixes for Chinese Word Segmentation ... 118  
*Zkqs kpi 'Nk'Ej gpi skpi '\ qpi .'cpf 'Mgj/[ kj 'Uw*
3. Learning Fine-Grained Selectional Restrictions .....P IC

## Oral Presentation Session 3A: Speech Perception

1. Towards a Revised Motor Theory of L2 Speech Perception ..... 136  
*[ kj qw'Ncp*
2. Difficulties in Perception and Pronunciation of Mandarin Chinese Disyllabic Word Tone Acquisition: A Study of Some Japanese University Students ..... 143  
*[ wkpi 'F qpi .' [ ciwuj k'Vwdqvc.'cpf 'Ocucwng'F cpwmlk*

## Oral Presentation Session 3B: Language Learning

1. Exploring the Chinese Mental Lexicon with Word Association Norms ..... 153  
*Qk' [ gg'My qp*
2. Towards Automatic Error Type Classification of Japanese Language Learners' Writings..... 163  
*J k qo kQ{co c.'Oco qt w'Mqo cej k'cpf '[ wk'O ciwwo qvq*

## Oral Presentation Session 4A: Language Acquisition

1. The Development of Coherence in Narratives: Causal Relations ..... 173  
*Wen/hwk'Ucj*
2. Clausal-Packaging of Path of Motion in Second Language Acquisition of Russian and Spanish ..... 181  
*Mcy ck'Ej wk'J ukpi /rkp' [ gj .'Y gp/Ej wp'Ncp.'cpf '[ wJ cp'Ej gpi*
3. Age Related Differences in Language Usage and Reading between English Monolinguals and Bilinguals ..... 190  
*Dylan Marshall anf Hamid Gomari*

## Oral Presentation Session 4B: Corpus Compilation and Analysis

1. Compiling a Corpus of Taiwanese Students' Spoken English ..... 199  
*Ncp/lgp"J wpi*
2. BCCWJ-TimeBank: Temporal and Event Information Annotation on Japanese Text ..... 206  
*Ocuc{wnk'Cucj ctc."Uej k" cuf c."J knt k'Mqpkuj k'O k'wj q"Kó cf c."cpf "Mmwq'O cgnvy c*
3. A Corpus-Based Approach to Linguistic Function ..... 215  
*J gpi dkp"J cp"cpf "Lqpcyj cp"Y gdwgt*

## Oral Presentation Session 5A: Grammar and Syntax

1. A Case Study of a Free Word Order ..... 222  
*Xrcf kuxc" Mwdqp."O ctngxc "Nqr cwnqxc."cpf "Lk'k'O kt qxunf*
2. Some Formal Properties of Higher Order Anaphors ..... 232  
*T0\ wdg*
3. ChinGram: A TRALE Implementation of an HPSG Fragment of Mandarin Chinese ..... 240  
*Ugkcp"O wngt "cpf "Lcpc "Nk' gpnqxc*

## Oral Presentation Session 5B: Machine Translation

1. Vietnamese to Chinese Machine Translation via Chinese Character as Pivot ..... 250  
*J ck\ j cq."Vlcpkq"J kp."cpf "Lpi {k\ j cpi*
2. Transliteration Extraction from Classical Chinese Buddhist Literature Using Conditional Random Fields ..... 260  
*[wEj wp"Y cpi "cpf "Tkej ctf "V] qpi /J cp"Vuck*
3. Effects of Parsing Errors on Pre-Reordering Performance for Chinese-to-Japanese SMT ..... 267  
*Fcp"J cp."Rcucwcn"O ct wpg/I qo gl."J wmwng"O k'cq."Mcwuj kq"Uwf qj."cpf "*  
*OcuccnkPci cv*

## Oral Presentation Session 6A: Morphology and Syntax

1. Reduplication across Categories in Cantonese .....277  
*Charles Lam*
2. Yet Another Piece of Evidence for the Common Base Approach to Japanese  
Causative/Inchoative Alternations .....287  
*Tomokazu Takehisa*

## Oral Presentation Session 6B: Phonetics

1. Are Mandarin Sandhi Tone 3 and Tone 2 the Same or Different? The Results  
of Functional Data Analysis.....296  
*Chierh Cheng, Jenn-Yeu Chen, and Michele Gubian*
2. An Application of Comparative Corpora of Interactional Data—Toward the  
Sound Profiles of Sites of Initiation in French and Mandarin Recycling Repair  
.....302  
*Helen Kai-yun Chen*

## Oral Presentation Session 7A: Semantics

1. *Of*-Constructions in the Predicate of *Demonstrate* and *Show* in Academic  
Discourse.....312  
*Liyin Chen and Siaw-Fong Chung*
2. Spatial Particles in English: A Quantitative Corpus-Based Approach to the  
Conceptualization of Symmetry in Bodily Orientation .....322  
*Alvin Cheng-Hsien Chen*
3. Typological Stage of Counterfactuals in Mandarin .....329  
*Qian Yong*

## Oral Presentation Session 7B: Sentiment Analysis

1. Event Sequence Model for Semantic Analysis of Time and Location in  
Dialogue System .....339  
*Yasuhiro Noguchi, Satoru Kogure, Makoto Kondo, Hideki Asoh, Ichiro  
Kobayashi, Akira Takagi, Tatsuhiko Konishi, and Yukihiro Itoh*
2. #Irony or #Sarcasm—A Quantitative and Qualitative Study Based on Twitter  
.....349  
*Po-Ya Angela Wang*

3. Collective Sentiment Classification Based on User Leniency and Product Popularity .....357  
*Wenliang Gao, Naoki Yoshinaga, Nobuhiro Kaji, and Masaru Kitsuregawa*
4. KOSAC: A Full-Fledged Korean Sentiment Analysis Corpus .....366  
*Hayeon Jang, Munhyong Kim, and Hyopil Shin*

### Poster Session 1

1. Cross-Lingual Link Discovery between Chinese and English Wiki Knowledge Bases .....374  
*Qingliang Miao, Huayu Lu, Shu Zhang, and Yao Meng*
2. Locative Postpositions and Conceptual Structure in Japanese .....382  
*Akira Ohtani*
3. Transliteration Systems across Indian Languages Using Parallel Corpora ...390  
*Rishabh Srivastava and Riyaz Ahmad Bhat*
4. Exploiting Parallel Corpus for Handling Out-of-Vocabulary Words .....399  
*Juan Luo, John Tinsley, and Yves Lepage*
5. Classifying Questions in Question Answering System Using Finite State Machines with a Simple Learning Approach.....409  
*Mohammad Moinul Hoque, Teresa Goncalves, and Paulo Quaresma*
6. Use of Combined Topic Models in Unsupervised Domain Adaptation for Word Sense Disambiguation.....415  
*Shinya Kunii and Hiroyuki Shinnou*
7. Vietnamese Text Accent Restoration with Statistical Machine Translation...423  
*Luan-Nghia Pham, Viet-Hong Tran, and Vinh-Van Nguyen*
8. A Compact FP-Tree for Fast Frequent Pattern Retrieval .....430  
*Tri Thanh Nguyen*
9. ML-Tuned Constraint Grammars .....440  
*Eckhard Bick*

### Poster Session 2

1. Comparative Analyses of Textual Contents and Styles of Five Major Japanese Newspapers .....450  
*Takafumi Suzuki, Erina Kanou, and Yui Arakawa*
2. The Island Effect in Postverbal Constructions in Japanese .....459  
*Kohji Kamada*
3. Evaluation of Corpus Assisted Spanish Learning .....467



<i>Hui-Chuan Lu and Yu-Hsin Chu</i>	
4. Augmented Parsing of Unknown Word by Graph-Based Semi-Supervised Learning .....	474
<i>Qiuping Huang, Derek F. Wong, Lidia S. Chao, Xiaodong Zeng, and Liangye He</i>	
5. Tonal Patterns in the 15th Century: A Corpus-Based Approach.....	483
<i>Chihkai Lin</i>	
6. Basic Principles for Segmenting Thai EDUs.....	491
<i>Nalinee Intasaw and Wirote Aroonmanakun</i>	
7. Automatic Clause Boundary Annotation in the Hindi Treebank .....	499
<i>Rahul Sharma, Soma Paul, Riyaz Ahmad Bhat, and Sambhav Jain</i>	
8. Myths in Korean Morphology and Their Computational Implications .....	505
<i>Hee-Rahk Chae</i>	

### **Workshop Session 1**

1. Creative Language Learning Projects with Digital Media .....	512
<i>Chin-chi Chao</i>	
2. iPad Reading: An Innovative Approach to New Literacies .....	520
<i>Hsin-Chou Huang</i>	
3. A Generic Cognitively Motivated Web-Environment to Help People to Become Quickly Fluent in a New Language .....	526
<i>Michael Zock, Guy Lapalme, and Lih-Juang Fang</i>	

### **Workshop Session 2**

1. Evaluation on Second Language Collocational Congruency with Computational Semantic Similarity .....	534
<i>Ching-Ying Lee and Chih-cheng Lin</i>	
2. A Corpus-Based Tool for Exploring Domain-Specific Collocations in English .....	542
<i>Ping-Yu Huang, Chien-Ming Chen, Nai-Lung Tsao, and David Wible</i>	
3. Automatic Identification of English Collocation Errors Based on Dependency Relations .....	550
<i>Zhao-Ming Gao</i>	

### Workshop Session 3

1. A Japanese Learning Support System Matching Individual Abilities.....556  
*Takahiro Ohno, Zyunitiro Edani, Ayato Inoue, and Dongli Han*
2. PADS Restoration and Its Importance in Reading Comprehension and  
Meaning Representation .....563  
*Shian-jung Chen*