Workshop on Multiword Expressions: from Theory to Applications

(MWE 2010)

At the 23rd International Conference on Computational Linguistics 2010

Beijing, China
28 August 2010

Table of Contents

Being Theoretical is Being Practical: Multiword Units and Terminological Structure Revitalised  
Kyo Kageura ................................................................. 1

Computational Lexicography of Multi-Word Units. How Efficient Can It Be?  
Filip Gralinski, Agata Savary, Monika Czerepowicka and Filip Makowiecki ............... 2

Construction of Chinese Idiom Knowledge-base and Its Applications  
Lei Wang and Shiwen Yu ................................................... 11

Automatic Extraction of Arabic Multiword Expressions  
Mohammed Attia, Antonio Toral, Lamia Tounsi, Pavel Pecina and Josef van Genabith ........ 19

Sentence Analysis and Collocation Identification  
Eric Wehrli, Violeta Seretan and Luka Nerima ........................................ 28

Automatic Extraction of Complex Predicates in Bengali  
Dipankar Das, Santanu Pal, Tapabrata Mondal, Tanmoy Chakraborty and Sivaji Bandyopadhyay 37

Handling Named Entities and Compound Verbs in Phrase-Based Statistical Machine Translation  
Santanu Pal, Sudip Kumar Naskar, Pavel Pecina, Sivaji Bandyopadhyay and Andy Way ...... 46

Application of the Tightness Continuum Measure to Chinese Information Retrieval  
Ying Xu, Randy Goebel, Christoph Ringlstetter and Grzegorz Kondrak ....................... 55

Standardizing Complex Functional Expressions in Japanese Predicates: Applying Theoretically-Based Paraphrasing Rules  
Tomoko Izumi, Kenji Imamura, Genichiro Kikui and Satoshi Sato .......................... 64

Identification of Reduplication in Bengali Corpus and their Semantic Analysis: A Rule Based Approach  
Tanmoy Chakraborty and Sivaji Bandyopadhyay ........................................ 73

Contrastive Filtering of Domain-Specific Multi-Word Terms from Different Types of Corpora  
Francesca Bonin, Felice Dell’Orletta, Giulia Venturi and Simonetta Montemagni ....... 77

A Hybrid Approach for Functional Expression Identification in a Japanese Reading Assistant  
Gregory Hazelbeck and Hiroaki Saito ........................................... 81

An Efficient, Generic Approach to Extracting Multi-Word Expressions from Dependency Trees  
Scott Martens and Vincent Vandeghinste ............................................ 85

Multiword Expressions as Discourse Relation Markers (DRMs)  
Aravind Joshi ................................................................. 89