

Church

Second Annual Workshop on Very Large Corpora (WVLC2)

Thursday, 4 August 1994

Kyoto International Community House

Kyoto, Japan

**Second Annual Workshop on
Very Large Corpora
(WVLC2)**

Program and Proceedings

Thursday, 4 August 1994

Kyoto International Community House

Kyoto, Japan

Workshop Program

Corpora

09:30 McKelvie, David; Thompson, Henry S.;

TEI Conformant Structural Markup of a Trilingual Corpus in the ECI Multilingual Corpus 1

09:55 Yarowsky, David;

A Comparison of Corpus-based Techniques for Restoring Accents in Spanish and French Text

10:20 Uramoto, Naohiko;

Extracting Disambiguated Thesaurus from Parallel Dictionary Definitions

10:45 Kita, Kenji; Omoto, Takashi; Yano, Yoneo; Kato, Yasuhiko;

Application of Corpora in Second Language Learning - The Problem of Collocational Knowledge Acquisition

11:10 break

Alignment

11:35 Grishman, Ralph

Iterative Alignment of Syntactic Structures for a Bilingual Corpus

12:00 Fung, Pascale; Wu, Dekai;

Statistical Augmentation of a Chinese Machine-Readable Dictionary

12:25 break for lunch

14:00 Invited Speaker: Martin Kay

15:00 break

Information Retrieval and Matching

15:25 Fujii, Hideo; Croft, Bruce;

Comparing the Retrieval Performance of English and Japanese Text Databases

15:50 Merkel, Magnus; Nilsson, Bernt; Ahrenberg, Lars;

A Phrase-Retrieval System Based on Recurrence

16:15 Sekine, Satoshi;

Automatic Sublanguage Identification for a New Text

16:40 Umemura, Kyoji;

String Comparison Based on Substring Equations

Table of Contents

<i>TEI Conformant Structural Markup of a Trilingual Corpus in the ECI Multilingual Corpus 1</i> Mackelvie, David; Thompson, Henry S. (University of Edinburgh) -----	7
<i>A Comparison of Corpus-based Techniques for Restoring Accents in Spanish and French Text</i> Yarowsky, David (Univ. of Pennsylvania) -----	19
<i>Extracting Disambiguated Thesaurus from Parallel Dictionary Definitions</i> Uramoto, Naohiko (IBM Tokyo Research Laboratory) -----	33
<i>Application of Corpora in Second Language Learning - The Problem of Collocational Knowledge Acquisition</i> Kita, Kenji*; Omoto, Takashi*; Yano, Yoneo*; and Kato, Yasuhiko** (*Tokushima Univ., **The National Language Research Institute) -----	43
<i>Iterative Alignment of Syntactic Structures for a Bilingual Corpus</i> Grishman, Ralph (New York University) -----	57
<i>Statistical Augmentation of a Chinese Machine-Readable Dictionary</i> Fung, Pascale*; Wu, Dekai** (*Columbia University, **HKUST) -----	69
<i>Comparing the Retrieval Performance of English and Japanese Text Databases</i> Fujii, Hideo; Croft, Bruce (University of Massachusetts) -----	87
<i>A Phrase-Retrieval System Based on Recurrence</i> Merkel, Magnus; Nilsson, Bernt; Ahrenberg, Lars (Linköping University) -	99
<i>Automatic Sublanguage Identification for a New Text</i> Sekine, Satoshi (New York University) -----	109
<i>String Comparison Based on Substring Equations</i> Umemura, Kyoji (NTT Basic Research Laboratory) -----	121
<i>Bilingual Alignment and Tense</i> Santos, Diana (INESC)-----	129
<i>Comparative Discourse Analysis of Parallel Texts</i> Pim van der Eijk (Digital Equipment Corp.) -----	143

Authors' Index

Ahrenberg, Lars	99
Croft, Bruce	87
Fujii, Hideo	87
Fung, Pascale	69
Grishman, Ralph	57
Kato, Yasuhiko	43
Kita, Kenji	43
McKelvie, David	7
Merkel, Magnus	99
Nilsson, Bernt	99
Omoto, Takashi	43
Pim van der Eijk	143
Santos, Diana	129
Sekine, Satoshi	109
Thompson, Henry S.	7
Umemura, Kyoji	121
Uramoto, Naohiko	33
Yano, Yoneo	43
Yarowsky, David	19
Wu, Dekai	69