LREC 2008

http://www.lrec-conf.org/lrec2008

List of accepted papers

- Conference Programme -

Hélène Mazo

Published : mercredi 2 avril 2008
Modified : mercredi 16 avril 2008
Created : mardi 21 mai 2013

ID	Authors	Title
7	Dennis Spohr	A General Methodology for Mapping EuroWordNets to the Suggested Upper Merged Ontology
8	Thomas Mandl, Frederic Gey, Giorgio Di Nunzio, Nicola Ferro, Mark Sanderson, Diana Santos, Christa Womser-Hacker	An Evaluation Resource for Geographic Information Retrieval
10	M. Attia, M. Rashwan, A. Ragheb, M. Al-Badrashiny, H. Al-Basoumy	A Compact Arabic Lexical Semantics Language Resource Based on the Theory of Semantic Fields
11	Diana Maynard	Benchmarking Textual Annotation Tools for the Semantic Web
13	Knight, D. and Tennent, P.	Introducing DRS (The Digital Replay System) : A tool for the future of Corpus Linguistic research and analysis
15	Dimitrios Kokkinakis	A Semantically Annotated Swedish Medical Corpus
17	Volha Petukhova and Harry Bunt	LIRICS Semantic Role Annotation : Design and Evaluation of a Set of Data Categories
20	Estelle Delpech, Patrick Saint-Dizier	Investigating the Structure of Procedural Texts for Answering How-to Questions
21	Satoshi Sekine	Extended Named Entity Ontology with Attribute Information
22	Jorge Civera and Alfons Juan-Cscar	Bilingual Text Classification using the IBM 1 Translation Model
23	Michaela Atterer, Hinrich Schtze	An Inverted Index for Storing and Retrieving Grammatical Dependencies
24	Florian Khler , Hinrich Schtze, Michaela Atterer	A Question Answering System for German. Experiments with Morphological Linguistic Resources
25	David Bamman, Marco Passarotti, Roberto Busa, Gregory Crane	The annotation guidelines of the Latin Dependency Treebank and Index Thomisticus Treebank. The treatment of some specific syntactic constructions in Latin
26	Jchym KolY, Jan Švec	Structural Metadata Annotation of Speech Corpora : Comparing Broadcast News and Broadcast Conversations
28	Shuichi Itahashi, Chiu-yu Tseng	The 2008 Oriental COCOSDA Book Project - In Commemoration of the First Decade of Sustained Activities in Asia
30	Mikko Lounela	Process Model for Composing High-quality Text Corpora
32	H. Hge, Z. Kacic, B. Kotnik, M. Rojc, N. Moreau, H.-U. Hain	Evaluation of Modules and Tools for Speech Synthesis - The ECESS Framework -
35	Mariona Taul, M. Antnia Mart, Marta Recasens	AnCora : Multilevel Annotated Corpora for Catalan and Spanish
37	Wei Liu, Ben Allison, Louise Guthrie	Professor or screaming beast ? Detecting Words Misuse in Chinese
38	Hiroyuki Shinnou, Minoru Sasaki	Ping-pong Document Clustering using NMF and Linkage-Based Refinement
39	Hiroki Yamazaki, Keisuke Kitamura, Takashi Harada and Seiichi Yamamoto	Creation of Learner Corpus and its Application to Speech Recognition
42	Nria Bel, Sergio Espeja, Montserrat Marimon, Marta Villegas	COLDIC, a Lexicographic Platform for LMF Compliant Lexica
43	Christophe Jouis, Julien Bourdaillet	Representation of Atypical Entities in Ontologies

44	Stphanie Weiser, Philippe Laublet, Jean-Luc Minel	Automatic Identification of Temporal Information in Tourism Web Pages
45	Muath Alzghool and Diana Inkpen	Combining Multiple Models for Speech Information Retrieval
46	Shisanu Tongchim, Randolf Altmeyer, Virach Sornlertlamvanich, Hitoshi Isahara	A Dependency Parser for Thai
48	Henk van den Heuvel, Jean-Pierre Martens, Bart D’hoore, Kristof D’hanens, Nanneke Konings	The AUTONOMATA Spoken Names Corpus
49	Iris Hendrickx, Gosse Bouma, Frederik Coppens, Walter Daelemans, Veronique Hoste, Geert Kloosterman, Anne-Marie Mineur, Joeri Van Der Vloet, Jean-Luc Verschelde	A Coreference Corpus and Resolution System for Dutch
50	Xavier Tannier, Philippe Muller	Evaluation Metrics for Automatic Temporal Annotation of Texts
51	Lothar Lemnitzer, Holger Wunsch and Piklu Gupta	Enriching GermaNet with verb-noun relations - a case study of lexical acquisition
52	Jens Nilsson, Joakim Nivre	MaltEval : An Evaluation and Visualization Tool for Dependency Parsing
54	Dan Tufi_, Alexandru Ceau_u	DIAC+ : A Professional Diacritics Recovering System
55	Briony Williams, Rhys James Jones	Acquiring Pronunciation Data for a Placenames Lexicon in a Less-Resourced Language
56	Dan Tufi_, Elena Irimia, Radu Ion, Alexandru Ceau_u	Unsupervised Lexical Acquisition for Part of Speech Tagging
57	Krzysztof Marasek, Ryszard Gubrynowicz	Design and Data Collection for Spoken Polish Dialogs Database
60	Zhuli Xie, Barbara Di Eugenio, Peter C. Nelson	From Extracting to Abstracting : Generating Quasi-abstractive Summaries
61	Chun-Yuan Teng and Hsin-Hsi Chen	Event Detection and Summarization in Weblogs with Temporal Collocations
62	Hiroyuki Shinnou, Minoru Sasaki	Spectral Clustering for a Large Data Set by Reducing the Similarity Matrix Size
64	Danica Damljanovic, Valentin Tablan, Kalina Bontcheva	A Text-based Query Interface to OWL Ontologies
66	Daniel Zeman	Reusable Tagset Conversion Using Tagset Drivers
67	Cvetana Krstev, Ranka Stankovi , Duško Vitas, Ivan Obradovi	The Usage of Various Lexical Resources and Tools to Improve the Performance of Web Search Engines
68	Lynette Melnar, Chen Liu	Borrowing Language Resources for Development of Automatic Speech Recognition for Low- and Middle-Density Languages
69	Nick Campbell	Tools & Resources for Visualising Coversational-Speech Interaction
70	Marco Guerini, Carlo Strapparava, Oliviero Stock	Resources for Persuasion
71	Marco Guerini, Carlo Strapparava, Oliviero Stock	Valentino : A tool for Valence Shifting of Natural Language Texts
72	Yuhang Yang, Qin Lu, Tiejun Zhao	Chinese Term Extraction Based on Delimiters
73	Diana Santos, Rosrio Silva, Susana Incio	What's in a Colour ? Studying and Contrasting Colours with COMPARA
75	Erhard Hinrichs and Monica L. Lau	In contrast - A Complex Discourse Connective
76	Einav Itamar, Alon Itai	Using Movie Subtitles for Creating a Large-Scale Bilingual Corpora
77	Beata Trawinski and Jan-Philipp Soehn	A Multilingual Database of Polarity Items

78	Serge Sharoff, Mikhail Kopotev, Toma~ Erjavec, Anna Feldman, Dagmar Divjak	Designing and Evaluating a Russian Tagset
79	Jonas Sjbergh, Kenji Araki	What is Poorly Said is a Little Funny
83	Nasser Abouzakhar, Ben Allison, Louise Guthrie	Unsupervised Learning-based Anomalous Arabic Text Detection
84	Adrian Popescu, Gregory Grefenstette	A Conceptual Approach to Web Image Retrieval
86	Alain Joubert, Mathieu Lafourcade	Evolutionary basic notions for a thematic representation of general knowledge
87	Eduardo Blanco, Nuria Castell and Dan Moldovan	Causal Relation Extraction
88	Ekaterina Lapshinova-Koltunski, Ulrich Heid	Head or Non-head ? Semi-automatic procedures for extracting and classifying subcategorisation properties of compounds
89	Toma~ Erjavec, Simon Krek	The JOS morphosyntactically tagged corpus of Slovene
90	Dan Tufi_, Radu Ion, Alexandru Ceau_u and Dan ^tef nescu	RACAI's Linguistic Web Services
91	Denis Maurel	Prolexbase : A multilingual relational lexical database of proper names
92	Hanno Biber, Evelyn Breiteneder, Karlheinz Mrth	Words in Contexts. Digital Editions of Literary Journals in the "AAC - Austrian Academy Corpus"
93	Harry Bunt, Chwhynny Overbeeke	Towards formal interpretation of semantic annotation
94	Georg Rehm, Marina Santini, Alexander Mehler, Pavel Braslavski, Rdiger Gleim, Andrea Stubbe, Svetlana Symonenko, Mirko Tavosanis and Vedrana Vidulin	Towards a Reference Corpus of Web Genres for the Evaluation of Genre Identification Systems
95	Thomas Schmidt, Susan Duncan, Oliver Ehmer, Jeffrey Hoyt, Michael Kipp, Dan Loehr,Magnus Magnusson, Travis Rose, Han Sloetjes	An exchange format for multimodal annotations
96	Han Ren, Donghong Ji, Lei Han	A Research on Automatic Chinese Catchword Extraction
97	Georg Rehm, Oliver Schonefeld, Andreas Witt, Timm Lehmberg, Christian Chiarcos, Hanan Bechara, Florian Eishold, Kilian Evang, Magdalena Leshtanska, Aleksandar Savkov, Matthias Stark	The Metadata-Database of a Next Generation Sustainability Web-Platform for Language Resources
99	Brian Weiss, Craig Schlenoff, Greg Sanders, Michelle Steves, Sherri Condon, Jon Phillips and Dan Parvaz	Performance Evaluation of Speech Translation Systems
10	Annie Zaenen, Danny Bobrow, Cleo Condoravdi	The Encoding of lexical implications in VerbNet Predicates of change of locations
10	Nicols Morales, Javier Tejedor, Javier Garrido, Jos Cols, Doroteo T. Toledano	STC-TIMIT : Generation of a Single-channel Telephone Corpus
2 10	Janne Bondi Johannessen, Torbjrn Nordgrd, Lars Nygaard	Evaluation of Linguistics-Based Translation
10	Stephen Purpura, John Wilkerson and Dustin Hillard	The U.S. Policy Agenda Legislation Corpus Volume 1 - a Language Resource from 1947 - 1998
5 10	Inderjeet Mani, Janet Hitzeman, Justin Richer, Dave Harris, Rob Quimby, Ben Wellner	SpatialML : Annotation Scheme, Corpora, and Tools
10 7	Qing Ma, Nakao Koichi, Masaki Murata, Hitoshi Isahara	Selection of Japanese-English Equivalents by Integrating High-quality Corpora and Huge Amounts of Web Data
10	Ernesto William De Luca, Birte Lnneker-Rodman	Integrating Metaphor Information into RDF/OWL EuroWordNet
10	Jack Halpern	Exploiting Lexical Resources for Disambiguating CJK and Arabic Orthographic Variants
9 11		

L
is
t
o
f
a
c
c
e
p
t
e
d
p
a
p
e
r
s

4

JiY Mrovsk	Does Netgraph Fit Prague Dependency Treebank ?

11	Jahna Otterbacher, Dragomir R. Radev	Modeling Document Dynamics : An Evolutionary Approach
5 11 6	Vincent Vandeghinste, Peter Dirix, Ineke Schuurman, Stella Markantonatou, Sokratis Sofianopoulos, Marina Vassiliou, Olga Yannoutsou, Toni Badia, Maite Melero, Gemma Boleda, Michael Carl, Paul Schmidt	Evaluation of a Machine Translation System for Low Resource Languages : METIS-II
11	Sebastian Gottwald, Matthias Richter, Gerhard Heyer and Gerik Scheuermann	Tapping Huge Temporally Indexed Textual Resources with WCTAnalyze
7 11	Carlos D. Martnez-Hinarejos, Vicent Tamarit	Evaluation of different segmentation techniques for dialogue turns
12	Barbara Plank, Khalil Sima'an	Subdomain Sensitive Statistical Parsing using Raw Corpora
0 12	Beata Megyesi, Bengt Dahlqvist, Eva Pettersson Eva, Joakim Nivre	Swedish-Turkish Parallel Treebank
12 3	Paulo C F de Oliveira, Edson Wilson Torrens, Alexandre Cidral, Sidney Schossland, Evandro Bittencourt	Evaluating Summaries Automatically - a system proposal
12	Richard Johansson, Pierre Nugues	Comparing Dependency and Constituent Syntax for Frame-semantic Analysis
12	Julia Trushkina, Lieve Macken, Hans Paulussen	Sentence Alignment in DPC : Maximizing Precision, Minimizing Human Effort
6 12	Dimitrios Kokkinakis	MeSH - From a Controlled Vocabulary to a Processable Resource
13	Yves Bestgen	Building Affective Lexicons from Specific Corpora for Automatic Sentiment Analysis
0 13	R.J.J.H. van Son, Wieneke Wesseling, Eric Sanders, Henk van den Heuvel	The IFADV corpus : A free dialog video corpus
13	Jonas Sjbergh, Kenji Araki	A Multi-Lingual Dictionary of Dirty Words
3 13	Kate Forbes-Riley, Diane Litman, Scott Silliman, Amruta Purandare	Uncertainty Corpus : Resource to Study User Affect in Complex Spoken Dialogue Systems
13	Jin-Ji Li, Dong-Il Kim and Jong-Hyeok Lee	Annotation Guidelines for Chinese-Korean Word Alignment
7 13	Georg Rehm, Richard Eckart, Christian Chiarcos, Johannes Dellert	Ontology-Based XQuery'ing of XML-Encoded Language Resources on Multiple Annotation Layers
14	Cdric Messiant, Anna Korhonen, Thierry Poibeau	LexSchem : A Large Subcategorization Lexicon for French Verbs
2 14	Aljoscha Burchardt, Marco Pennacchiotti	FATE : a FrameNet-Annotated corpus for Textual Entailment
14	Thierry Poibeau and Cdric Messiant	Do we still Need Gold Standards for Evaluation ?
4 14	Neil Newbold and Lee Gillam	Automatic Document Quality Control
14	Jeremy Bensley and Andrew Hickl	Unsupervised Resource Creation for Textual Inference Applications
6 14	K. L. Kermanidis, A. Thanopoulos, M. Maragoudakis, N. Fakotakis	Eksairesis : A Domain-adaptable System for Ontology Building from Unstructured Text
14	Milan Gnjatovi , Dietmar Rsner	On the Role of the NIMITEK Corpus in Developing an Emotion Adaptive Spoken Dialogue System
9 15	Oana Frunza	A Trainable Tokenizer, solution for multilingual texts and compound expression tokenization
15	Veronique Hoste, Els Lefever, Klaar Vanopstal, Isabelle Delaere	Learning-based Detection of Scientific Terms in Patient Information
4 15 5	Gwnol Lecorv, Guillaume Gravier, Pascale Sbillot	On the use of Web resources and natural language processing techniques to improve automatic speech recognition systems
15		

List of accepted papers

6

Kjell Elenius, Eva Forsbom, Beta Megyesi	Language Resources and Tools for Swedish : A Survey

15	Markus Dickinson, Chong Min Lee	Detecting Errors in Semantic Annotation
15	Mark Greenwood, Jos Iria, Fabio Ciravegna	Saxon : An Extensible Multimedia Annotator
8 15	Lars Nygaard, Joel Priestley, Anders Nklestad, Janne Bondi Johannessen	Glossa : A multilingual, multimodal, configurable user interface
16	Muriel Amar, Sophie David, Rachel Panckhurst, Lisa Whistlecroft	Classification procedures for software evaluation
0 16	Marilisa Amoia, Claire Gardent	A Test Suite for Inference Involving Adjectives
16	Marco Baroni, Francis Chantree, Adam Kilgarriff, Serge Sharoff	CleanEval : a competition for cleaning webpages
2 16	Piroska Lendvai, Steve Hunt	From Field Notes Towards a Knowledge Base
16	Laura Stoia, Darla Magdalene Shockley, Donna K. Byron and Eric Fosler-Lussier	SCARE : A Situated Corpus with Annotated Referring Expressions
4 16	Satoshi Sato, Suguru Matsuyoshi, Yohsuke Kondoh	Automatic Assessment of Japanese Text Readability Based on a Textbook Corpus
16	Isaac G. Councill, C. Lee Giles, Min-Yen Kan	ParsCit : An open-source CRF reference string parsing package
6 16	Emilie Chtelat-Pel and Annelies Braffort	Sign Language Corpus Annotation : Toward A New Methodology
16	Shunsuke Kozawa, Hitomi Tohyama, Kiyotaka Uchimoto and Shigeki Matsubara	Automatic Acquisition of Usage Information for Language Resources
9 17 2	Jean-Yves Antoine, Abdenour Mokrane, Nathalie Friburger	Automatic Rich Annotation of Large Corpus of Conversational transcribed speech : the Chunking Task of the EPAC Project
17	Vincent Claveau	Automatic Translation of Biomedical Terms by Supervised Machine Learning
3 17	Stphane Huet, Guillaume Gravier, Pascale Sbillot	Morphosyntactic resources for automatic speech recognition
17	Hiroyuki Kaji, Shin'ichi Tamamura, Dashtseren Erdenebat	Automatic Construction of a Japanese-Chinese Dictionary via English
5 17	Tomas By	The Kalashnikov 691 dependency bank
17	Michael Wiegand, Jochen L. Leidner, Dietrich Klakow	Cost-Sensitive Learning in Answer Extraction
7 18	Isa Maks, Piek Vossen, Roxane Segers, Hennie van der Vliet	Adjectives in the Dutch semantic lexical database CORNETTO
18 5	I. Leturia, A. Gurrutxaga, N. Areta, E. Pociello	Analysis and performance of morphological query expansion and language-filtering words on Basque web searching
18	Svetla Koeva, Borislav Rizov, Svetlozara Leseva	Chooser - A Multi-Task Annotation Tool
18	Sebastian Mller, Florian Gdde and Maria Wolters	Corpus Analysis of Spoken Smart-Home Interactions with Older Users
7 18	Antonio Toral, Rafael Muoz, Monica Monachini	Named Entity WordNet
18	Anna Joan, Jorge Vivaldi, Merc Lorente	Turning a term extractor into a new domain : first experiences
9 19	Marie-Jeanne Derouin, Andr Le Meur	Presentation of the new ISO-Standard for the representation of entries in dictionaries : ISO 1951
19	Hiroaki SATO	New Functions of FrameSQL for Multilingual FrameNets
1 19	Adrian Iftene, Alexandra Balahur-Dobrescu	Named Entity Relation Mining Using Wikipedia
19	Stanislas Oger, Georges Linars, Frdric Bchet	Local methods for on-demand out-of-vocabulary word retrieval

3

19	David Griol, Llus F. Hurtado, Encarna Segarra, Emilio Sanchis	Acquisition and Evaluation of a Dialog Corpus through WOz and Dialog Simulation Techniques
19	Toni Badia, Maite Melero and Oriol Valentn	Rapid deployment of a new METIS language pair : Catalan-English
8 20	Laura Kallmeyer, Timm Lichte, Wolfgang Maier, Yannick Parmentier, Johannes Dellert	Developing a TT-MCTAG for German with an RCG-based Parser
20	Marco Pennacchiotti, Diego De Cao, Paolo Marocco, Roberto Basili	Towards a Vector Space Model for FrameNet-like Resources
2 20	Juan Aparicio, Mariona Taul, Ma. Antnia Mart	AnCora-Verb : A Lexical Resource for the Semantic Annotation of Corpora
20	M. Kemps-Snijders, A. Klassmann, C. Zinn, P. Berck, A. Russel, P. Wittenburg	Exploring and Enriching a Language Resource Archive via the Web
5 20	Sylwia Ozdowska	Cross-corpus evaluation of word alignment
20	Han Sloetjes, Peter Wittenburg	Annotation by category - ELAN and ISO DCR
8 21	Markus Dickinson and Charles Jochim	A Simple Method for Tagset Comparision
21	Adam Przepirkowski, RafaB L. Grski, Barbara Lewandowska-Tomaszczyk, Marek AaziDski	Towards the National Corpus of Polish
1 21	Aleksander BuczyDski, Adam Przepirkowski	• Demo : An Open Source Tool for Partial Parsing and Morphosyntactic Disambiguation
21 3	Aukasz Degrski, MichaB Marcinczuk, Adam Przepirkowski	Definition extraction using a sequential combination of baseline grammars and machine learning classifiers
21 4	Thepchai Supnithi, Suchinder Singh, Taneth Ruangrajitpakorn, Prachya Boonkwan, Monthika Boriboon	Open CCG Workbench and Visualization Tool
21	Ekaterina Buyko, Christian Chiarcos, Antonio Pareja-Lora	Ontology-Based Interface Specifications for an NLP Pipeline Architecture
5 21	Marc Finthammer, Irene Cramer	Exploring and Navigating : Tools for GermaNet
21	Mriam Lujn, Carlos D. Martnez, Vicente Alabau	Evaluation of several Maximum Likelihood Linear Regression variants for language adaptation.
7 21	. Andersen, J. Nioche, E.J. Briscoe and J. Carroll	The BNC Parsed with RASP4UIMA
21	Mari Carmen Surez-Figueroa, Asuncin Gmez-Prez	Towards a Glossary of Activities in the Ontology Engineering Field
9 22	Matthieu Hermet, Alain Dsilets, Stan Szpakowicz	Using the Web as a Linguistic Resource to Automatically Correct Lexico-Syntactic Errors
22	Francisco Alvarez, Antonio Vaquero, Fernando Senz	Conceptual Modeling of Ontology-based Linguistic Resources with a Focus on Semantic Relations
1 22	Marc Kemps-Snijders, Menzo Windhouwer, Peter Wittenburg, Sue Ellen Wright	ISOcat : Corralling Data Categories in the Wild
22 6	Doroteo T. Toledano, D. Hernandez-Lopez, C. Esteve-Elizalde, J. Fierrez, J. Ortega-Garcia, D. Ramos, J. Gonzalez-Rodriguez	BioSec Multimodal Biometric Database in Text-Dependent Speaker Recognition
22	Davide Fossati, Barbara Di Eugenio	I saw TREE trees in the park : How to correct real-word spelling mistakes
22	Dafydd Gibbon, Jolanta Bachan	An automatic close copy speech synthesis tool for large-scale speech corpus evaluation
8 22	Steven Bethard, William Corvey, Sara Klingenstein, James H. Martin	Building a Corpus of Temporal-Causal Structure
23	Yoshihiko Hayashi, Chiharu Narawa, Monica Monachini, Claudia Soria, Nicoletta Calzolari	Ontologizing Lexicon Access Functions based on an LMF-based Lexicon Taxonomy
1 23	Silke Scheible	Annotating Superlatives
23		

List of accepted papers

5

Berenike Loos and Lasse Schwarten	A Semantic Memory for Incremental Ontology Population

23	Paula Estrella, Andrei Popescu-Belis, Maghi King	Improving Quality Models for MT Evaluation Based on Evaluators' Feedback
237	Kallirroi Georgila, Maria Wolters, Vasilis Karaiskos, Melissa Kronenthal, Robert Logie, Neil Mayo, Johanna D. Moore, Matt Watson	A Fully Annotated Corpus for Studying the Effect of Cognitive Ageing on Users' Interactions with Spoken Dialogue Systems
238	Shigeyoshi Kitazawa, Shinya Kiriyama, Tomohiko Kasami, Shogo Ishikawa, Naofumi Otani, Hiroaki Horiuchi and Yoichi Takebayashi	A Multimodal Infant Behavior Annotation for Developmental Analysis of Demonstrative Expressions
23	Jette Viethen, Robert Dale, Emiel Krahmer, Mariet Theune, Pascal Touset	Controlling Redundancy in Referring Expressions
924	Stefan Scherer, Petra-Maria Strau	A Flexible Wizard of Oz Environment for Rapid Prototyping
241	Yannick Versley, Simone Paolo Ponzetto, Massimo Poesio, Vladimir Eidelman, Alan Jern, Jason Smith, Xiaofeng Yang, Alessandro Moschitti	BART : A Modular Toolkit for Coreference Resolution
24	Laurianne Sitbon, Patrice Bellot, Philippe Blache	Evaluating Robustness Of A QA System Through A Corpus Of Real-Life Questions
24	Laurianne Sitbon, Patrice Bellot, Philippe Blache	Evaluation Of Lexical Resources And Semantic Networks On A Corpus Of Mental Associations
6247	Bruno Cartoni	Lexical Resources for Automatic Translation of Constructed Neologisms : the Case Study of Relational Adjectives
24	Lena Grothe, Ernesto William De Luca and Andreas Nrnberger	A Comparative Study on Language Identification Methods
925	Joao Graa, Joana Paulo Pardal, Lusa Coheur and Diamantino Caseiro	Building a golden collection of parallel Multi-Language Word Alignment
25	Steliana Ivanova, Sandra Kbler	POS Tagging for German : How Important is the Right Context ?
325	Leen Cleuren, Jacques Duchateau, Pol Ghesquire, Hugo Van hamme	Children's Oral Reading Corpus (CHOREC) : Description and Assessment of Annotator Agreement
25	Piek Vossen, Isa Maks, Roxane Segers, Hennie VanderVliet	Integrating lexical units, synsets and ontology in the Cornetto Database
525	Marie-Jean Meurs, Frdric Duvert, Frdric Bchet, Fabrice Lefvre, Renato de Mori	Semantic Frame Annotation on the French MEDIA corpus
25	Yasuharu Den, Junpei Nakamura, Toshinobu Ogiso, Hideki Ogura	A proper approach to Japanese morphological analysis : Dictionary, model, and evaluation
825	Yujie Zhang, Zhulong Wang, Kiyotaka Uchimoto, Qing Ma, Hitoshi Isahara	Word Alignment Annotation in a Japanese-Chinese Parallel Corpus
26	Paul Buitelaar, Thomas Eigner	Ontology Search with the OntoSelect Ontology Library
226	Pitel Guillaume, Grefenstette Gregory	Semi-automatic Building Method for a Multidimensional Affect Dictionary for a New Language
27	Cssia Trojahn, Paulo Quaresma, Renata Vieira	A Framework for Multilingual Ontology Mapping
027	Heike Bieler, Stefanie Dipper	Measures for Term and Sentence Relevances : an Evaluation for German
27	Diana Maynard, Wim Peters, Yaoyong Li	Evaluating Evaluation Metrics for Ontology-Based Applications : Infinite Reflection
327	Iaki Alegria, Klara Ceberio, Nerea Ezeiza, Aitor Soroa, Gregorio Hernandez	Spelling Correction : from two-level morphology to open source
27	Diana McCarthy	Lexical Substitution as a Framework for Multiword Evaluation
527	Ann Devitt, Khurshid Ahmad	Sentiment Analysis and the Use of Extrinsic Datasets in Evaluation
27	Thierry Bazillon, Yannick Estve, Daniel Luzzati	Manual vs assisted transcription of prepared and spontaneous speech
7278	Eric Sanders, Asuncion Moreno, Herbert Tropf, Lynette Melnar, Nurit Dekel, Breanna Gillies, Niklas Paulsson	LILA : Cellular Telephone Speech Databases from Asia

27	Jeroen Geertzen, Volha Petukhova, Harry Bunt	Evaluating Dialogue Act Tagging with Naive and Expert Annotators
9 28 0	Cyril Grouin	Certification and cleaning up of a text corpus : towards an evaluation of the "grammatical" quality of a corpus
28	Špela Vintar, Darja Fišer	Harvesting Multi-Word Expressions from Parallel Corpora
1 28 3	Antonio Moreno Sandoval, Doroteo T. Toledano, Ral de la Torre, Marta Garrote, Jos M. Guirao	Developing a Phonemic and Syllabic Frequency Inventory for Spontaneous Spoken Castilian Spanish and their Comparison to Text-Based Inventories
28	Francesca Fallucchi, Fabio Massimo Zanzotto	Yet Another Platform for Extracting Knowledge from Corpora
4 28	Christian Hnig, Stefan Bordag, Uwe Quasthoff	UnsuParse : Unsupervised Parsing with unsupervised Part of Speech tagging
28	Philippe Dreuw, Carol Neidle, Vassilis Athitsos, Stan Sclaroff and Hermann Ney	Benchmark Databases for Video-Based Automatic Sign Language Recognition
7 28	Tony Veale, Yanfen Hao	Acquiring Naturalistic Concept Descriptions from the Web
28 9	Susan Robinson, David Traum, Midhun Ittycheriah, Joe Henderer	What would you ask a Conversational Agent ? Observations of Human-Agent Dialogues in a Museum Setting
29	Drahomra "johanka" Spoustov, Pavel Pecina, Jan Haji, Miroslav Spousta	Validating the Quality of Full Morphological Annotation
29	Mark D. Arehart, Keith J. Miller	A Ground Truth Dataset for Matching Culturally Diverse Romanized Person Names
1 29	Michael Carl	Using Log-linear Models for Tuning Machine Translation Output
29	Alexandre Patry, Philippe Langlais	MISTRAL : A Statistical Machine Translation Decoder for Speech Recognition Lattices
3 29	Margarita Alonso Ramos, Owen Rambow and Leo Wanner	Using Semantically Annotated Corpora to Build Collocation Resources
29 5	Kirk Baker, Chris Brew	Statistical Identification of English Loanwords in Korean Using Automatically Generated Training Data
29	Rogelio Nazar, Jorge Vivaldi, Teresa Cabr	A Suite to Compile and Analyze an LSP Corpus
29	Massimo Poesio, Ron Artstein	Anaphoric Annotation in the ARRAU Corpus
7 29	Mark Przybocki, Kay Peterson, Sbastien Bronsart	Translation Adequacy and Preference Evaluation Tool (TAP-ET)
30 0	Dominic Widdows, Kathleen Ferraro	Semantic Vectors : A Scalable Open Source Package and Online Technology Management Application
30 1	Hiroyuki Shinnou, Minoru Sasaki	Division of Example Sentences Based on the Meaning of a Target Word Using Semi-supervised Clustering
30	Cristina Mota, Ralph Grishman	Is this NE tagger getting old ?
3 30	OndYej Bojar, Miroslav Janek, Zden k }abokrtsk, Pavel eška, Peter BeHa	CzEng 0.7 : Parallel Corpus with Community-Supplied Translations
30	Peter Anick, Vijay Murthi, Shaji Sebastian	Similar Term Discovery using Web Search
6 30 8	Jonathan H. Clark, Robert Frederking, Lori Levin	Toward Active Learning in Data Selection : Automatic Discovery of Language Features During Elicitation
30	Joydeep Nath, Monojit Choudhury, Animesh Mukherjee, Christian Biemann, Niloy Ganguly	Unsupervised Parts-of-Speech Induction for Bengali
9 31		

L
is
t
o
f
a
c
c
e
p
t
e
d
p
a
p
e
r
s

3

Michael Mohler, Rada Mihalcea	Babylon Parallel Text Builder : Gathering Parallel Texts for Low-Density Languages

31	Alessandra Zarcone, Alessandro Lenci	Computational Models for Event Type Classification in Context
5 31 6	Jana Trojanov, Marek Hrz, Pavel Campr, Milos Zelezny	Design and Recording of Czech Audio-Visual Database with Impaired Conditions for Continuous Speech Recognition
31	Tams Vradi, Steven Krauwer, Peter Wittenburg, Martin Wynne, Kimmo Koskenniemi	CLARIN : Common Language Resources and Technology Infrastructure
7 31	Markpong Jongtaveesataporn, Chai Wutiwiwatchai, Koji Iwano, Sadaoki Furui	Thai Broadcast News Corpus Construction and Evaluation
32	Davide Buscaldi, Paolo Rosso	Geo-WordNet : Automatic Georeferencing of WordNet
1 32	L. Devillers, J.-C. Martin	Coding Emotional Events in Audiovisual Corpora
32	Ulrich Heid, Marion Weller	Tools for collocation extraction : preferences for active vs. passive
3 32 4	Pavlina Fragkou, Georgios Petasis, Aris Theodorakos, Vangelis Karkaletsis, Constantine D. Spyropoulos	BOEMIE ontology-based text annotation tool
32 6	Grazyna Demenko, Stefan Grocholewski, Katarzyna Klessa, Jerzy Ogrkiewicz, Agnieszka Wagner, Marek Lange, Daniel ZledziDski, Natalia Cylwik	Jurisdic - Polish Speech Database for taking dictation of legal texts
32	Stavros Ntalampiras, Ilyas Potamitis, Todor Ganchev, Nikos Fakotakis	Audio Database in Support of Potential Threat and Crisis Situation Management
32	Prokopis Prokopidis, Vassia Karra, Aggeliki Papagianopoulou, Stelios Piperidis	Condensing sentences for subtitle generation
8 32	JindYich Matoušek, Daniel Tihelka, Jan Romportl	Building of a Speech Corpus Optimised for Unit Selection TTS Synthesis
33 0	Hennie Brugman, Vronique Malais, Laura Hollink	A Common Multimedia Annotation Framework for Cross Linking Cultural Heritage Digital Collections
33	Tiago Freitas, Fabola Santos	CORP-ORAL : Spontaneous Speech Corpus for European Portuguese
33	Nria Bel, Sergio Espeja, Montserrat Marimon	Automatic acquisition for low frequency lexical items
4 33	Katrin Tomanek and Udo Hahn	Approximating Learning Curves for Active-Learning-Driven Annotation
33 6	Stefan Scherer, Hansjrg Hofmann, Malte Lampmann, Martin Pfeil, Steffen Rhinow, Friedhelm Schwenker, Gnther Palm	Emotion Recognition from Speech : Stress Experiment
33 7	Sankaran Baskaran, Kalika Bali, Tanmoy Bhattacharya, Pushpak Bhattacharyya, Monojit Choudhury, Girish Nath Jha, Rajendran S., Saravanan K., Sobha L., KVS Subbarao	A Common Parts-of-Speech Tagset Framework for Indian Languages
34	Bogdan Babych, Serge Sharoff, Anthony Hartley	Generalising lexical translation strategies for MT using comparable corpora
0 34	Claire Grover, Sharon Givon, Richard Tobin, Julian Ball	Named Entity Recognition for Digitised Historical Texts
34	Ghazi Abuhakema, Reem Faraj, Anna Feldman, Eileen Fitzpatrick	Annotating an Arabic Learner Corpus for Error
3 34 4	Tomas Dekens, Yorgos Patsis, Werner Verhelst, Frdric Beaugendre and Franois Capman	A Multi-Sensor Speech Database with Applications towards Robust Speech Processing in Hostile Environments
34	G. Aguado de Cea, J. Puche Alosete, J.A. Ramos	Tagging Spanish Texts : the Problem of ‘se’
5 34	Andrea Agili, Marco Fabbri, Alessandro Panunzi, Manuel Zini	Integration of a Multilingual Keyword Extractor in a Document Management System
34	Milena Yankova, Horacio Saggion, Hamish Cunningham	A Framework for Identity Resolution and Merging for Multi-source Information Extraction
7 34		

L
is
t
o
f
a
c
c
e
p
t
e
d
p
a
p
e
r
s

8

Martin Emms	Tree-distance and some other variants of evalb

34	Peter Adolphs, Stephan Oepen, Ulrich Callmeier, Berthold Crysmann, Dan Flickinger, Bernd Kiefer	Some Fine Points of Hybrid Natural Language Parsing
935	Eneko Agirre, Aitor Soroa	Using the Multilingual Central Repository for Graph-Based Word Sense Disambiguation
35	Simon Mille, Leo Wanner	Making Text Resources Accessible to the Reader : The Case of Patent Claims
235	Daiga Deksne, Raivis SkadiFš, Inguna SkadiFa	Dictionary of Multiword Expressions for Translation into Highly Inflected Languages
35	Marianne Santaholma, Nikos Chatzichrisafis	A knowledge-modeling approach for multilingual Regulus lexica
535	Ute Ziegenhain, Hanne Ferse, Henk van den Heuvel, Asuncion Moreno	LC-STAR II : starring more lexica
35	Isabel Trancoso, Rui Martins, Helena Moniz, Ana Isabel Mata, M. Cu Viana	The LECTRA Corpus - Classroom Lecture Transcriptions in European Portuguese
936	Reut Tsarfaty, Yoav Goldberg	Word-Based or Morpheme-Based ? Annotation Strategies for Modern Hebrew Clitics
36	Fredric Gey, David Kirk Evans, Noriko Kando	A Japanese-English Technical Lexicon for Translation and Language Research
3364	Daan Broeder, Thierry Declerck, Erhard Hinrichs, Stelios Piperidis, Laurent Romary, Nicoletta Calzolari and Peter Wittenburg	Foundation of a Component-based Flexible Registry for Language Resources and Technology
365	N. Oostdijk, M. Reynaert, P. Monachesi, G. van Noord, R. Ordelman, I. Schuurman, V. Vandeghinste	From D-Coi to SoNaR : A reference corpus for Dutch
366	C. Cucchiarini, J. Driesen, H. Van hamme, E. Sanders	Recording Speech of Children, Non-Natives and Elderly People for HLT Applications : the JASMIN-CGN Corpus.
36	Daniela Goecke, Maik Sthrenberg, Andreas Witt	Influence of Text Type and Text Length on Anaphoric Annotation
8370	Daan Broeder, David Nathan, Sven Strmqvist, Remco van Veenendaal	Building a federation of Language Resource Repositories ; the DAM-LR project and its continuation within CLARIN
373	Piek Vossen, Eneko Agirre, Nicoletta Calzolari, Christiane Fellbaum, Shu-kai Hsieh, Chu-Ren Huang, Hitoshi Isahara, Kyoko Kanzaki, Andrea Marchetti, Monica Monachini, Federico Neri, Remo Raffaelli, German Rigau, Maurizio Tescon, Joop VanGent	KYOTO : A System for Mining, Structuring, and Distributing Knowledge Across Languages and Cultures
37	Mario Crespo Miguel, Paul Buitelaar	Domain-Specific English-To-Spanish Translation of FrameNet
37	Paul Trilsbeek, Daan Broeder, Tobias van Valkenhoef, Peter Wittenburg	A Grid of Regional Language Archives
637	Hagen Frstenau	Enriching Frame Semantic Resources with Dependency Graphs
37	Siham Boulaknadel, Beatrice Daille, Driss Aboutajdine	A multi-word term extraction program for Arabic language
8380	Dieter van Uytvanck, Alex Dukers, Jacquelijn Ringersma, Paul Trilsbeek	Language-Sites : Accessing and Presenting Language Resources via Geographic Information Systems
38	A. Cneyd Tantug, Kemal Oflazer, Ilknur D. El-Kahlout	BLEU+ : A Tool for Fine-Grained BLEU Computation
2383	Jussi Karlgren, Hercules Dalianis, Bart Jongejan	Experiments to investigate the connection between case distribution and topical relevance of search terms in an information retrieval setting
38	Kirk Roberts and Andrew Hickl	Scaling Answer Type Detection to Large Hierarchies
4385	P. Spyns, E. D'Halleweyn, C. Cucchiarini	The Dutch-Flemish comprehensive approach to HLT stimulation and innovation : STEVIN, HLT Agency and beyond
38		

List of accepted papers

6

38	Borislav Rizov	Hydra : A Modal Logic Tool for Wordnet Development, Validation and Exploration
390	Javier lvez, Jordi Atserias, Jordi Carrera, Salvador Climent, Egoitz Laparra, Antoni Oliver and German Rigau	Complete and Consistent Annotation of WordNet using the Top Concept Ontology
391	Bento Carlos Dias-da-Silva, Ariani Di Felippo, Maria das Graas Volpe Nunes	The Automatic Mapping of Princeton WordNet Lexical-Conceptual Relations onto the Brazilian Portuguese WordNet Database
39	Matthias Eck, Stephan Vogel, Alex Waibel	Communicating Unknown Words in Machine Translation
239	Roser Morante	Semantic role labeling tools trained on the Cast3LB-CoNNL-SemRol corpus
39	Dan Cristea, Corina For scu, Marius R schip, Michael Zock	How to Evaluate and Raise the Quality in a Collaborative Lexicographic Approach
739	Grazyna Vetulani, Zygmunt Vetulani, Tomasz Obr bski	Verb-Noun Collocation SyntLex Dictionary - Corpus-Based Approach
399	Gregory A. Sanders, Sbastien Bronsart, Sherri Condon, Craig Schlenoff	Odds of Successful Transfer of Low-level Concepts : A Key Metric for Bidirectional Speech-to-speech Machine Translation
400	Tomoyosi Akiba, Kiyoaki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou	Test Collections for Spoken Document Retrieval from Lecture Audio Data
402	Susana de Azeredo, Silvia Maria W. Moraes, Vera Lcia Strube de Lima	Keywords, k-NN and Neural Networks : A Support for Hierarchical Categorization of Texts in Brazilian Portuguese
403	Masaki Itagaki, Takako Aikawa	Post-MT Term Swapper : Supplementing a Statistical Machine Translation System with a User Dictionary
404	Junko Kubo, Keita Tsuji, Shigeo Sugimoto	Temporal Aspects of Terminology for Automatic Term Recognition : Case Study on Women's Studies Terms
40	Jun Okamoto, Kiyoko Uchiyama and Shun Ishizaki	A Contextual Dynamic Network Model for WSD Using Associative Concept Dictionary
40	Hossam Ibrahim, Kareem Darwish, Abdel-Rahim Abdel-sabor	Automatic Extraction of Textual Elements from News Web Pages
740	C. Ray Graham, Deryle Lonsdale, Casey Kennington, Aaron Johnson and Jeremiah McGhee	Elicited Imitation as an Oral Proficiency Measure with ASR Scoring
41	Eiko Yamamoto, Hitoshi Isahara, Akira Terada, Yasunori Abe	Extraction of Informative Expressions from Domain-specific Documents
041	Francis Bond, Hitoshi Isahara, Kyoko Kanzaki, Kiyotaka Uchimoto	Boot-strapping a WordNet using multiple existing WordNets
41	Ruifeng Xu, Yunqing Xia, Kam-Fai Wong, Wenjie Li	Opinion Annotation in On-line Chinese Product Reviews
541 8	David Manzano-Macho, Asuncin Gmez-Prez, Daniel Borrajo	Unsupervised and Domain Independent Ontology Learning. Combining Heterogeneous Sources of Evidence
41	Florian Schiel, Christian Heinrich, Sabine Barfsser, Thomas Gilg	ALC - Alcohol Language Corpus
942	Torsten Zesch, Christof Mller and Iryna Gurevych	Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary
421	Petr Pollk, Jan Voln, Radek Skarnitzl	Phone Segmentation Tool with Integrated Large Pronunciation Lexicon and Czech Phonetically Labelled Reference Database
422	Tokunaga Takenobu, Dain Kaplan, Chu-Ren Huang, Shu-Kai Hsieh, Nicoletta Calzolari, Monica Monachini, Claudia Soria, Shirai Kiyoaki, Virach Sornlertlamvanich, Thatsanee Charoenporn, Xia YingJu	Adapting International Standard for Asian Language Technologies
42		

L
is
t
o
f
a
c
c
e
p
t
e
d
p
a
p
e
r
s

3

Takashi Tsunakawa, Naoaki Okazaki, Jun'ichi Tsujii	Building Bilingual Lexicons Uusing Lexical Translation Probabilities via Pivot Languages

424	Michael Kluck, Axel Huckstorf	The European Thesaurus on International Relations and Area Studies - A Multilingual Resource for Indexing, Retrieval, and Translation
42	Kathrin Eichler, Holmer Hemsen and Gnter Neumann	Unsupervised Relation Extraction From Web Documents
542	Evi Marzelou, Maria Zourari, Voula Giouli, Stelios Piperidis	Building a Greek corpus for Textual Entailment
42	Hitomi Tohyama, Shunsuke Kozawa, Kiyotaka Uchimoto, Shigeki Matsubara and Hitoshi Isahara	Construction of a Metadata Database for Efficient Development and Use of Language Resources
842	YaMin Chou, Chu-Ren Huang, Jia-Fei Hong	The Extended Architecture of Hantology for Kanji
43	Kyoko Kanzaki, Francis Bond, Noriko Tomuro, Hitoshi Isahara	Extraction of Attribute Concepts from Japanese Adjectives
343	Horacio Rodrguez, David Farwell, Javi Farreres, Manuel Bertran, Musa Alkhalifa, M. Antonia Mart	Arabic WordNet : Semi-automatic Extensions using Bayesian Inference
435	Bolette S. Pedersen, Anna Braasch, Lina Henriksen, Sussi Olsen, Claus Povlsen	Merging a Syntactic Resource with a WordNet - a Feasibility Study of a Merge between STO and DanNet
437	Iaki Sainz, Ibon Saratxaga, Eva Navas, Inmaculada Hernez, Jon Sanchez, Iker Luengo, Igor Odriozola	Subjective evaluation of an emotional speech database for Basque
438	Magnus Rosell, Sumithra Velupillai	Revealing Relations between Open and Closed Answers in Questionnaires through Text Clustering Evaluation
43	Isa Maks, Carole Tiberius, Remco van Veenendaal	Standardising bilingual lexical resources according to the Lexicon Markup Framework
44	Reiko Kaji, Hajime Mochizuki	Constructing a Database of Non-Japanese Pronunciations of Different Japanese Romanizations
0442	Rune Stre, Brian Kemper, Kanae Oda, Naoaki Okazaki, Yukiko Matsuoka, Norihiro Kikuchi, Hiroaki Kitano, Sophia Ananiadou, Jun'ichi Tsujii	Connecting Text Mining and Pathways using the PathText Resource
443	Pierrette Bouillon, Sonia Halimi, Yukie Nakao, Kyoko Kanzaki, Hitoshi Isahara, Nikos Tsourakis, Marianne Starlander, Beth Ann Hockey, Manny Rayner	Developing Non-European Translation Pairs in a Medium-Vocabulary Medical Speech Translation System
444	Marta R. Costa-juss, Jos A. R. Fonollosa and Enric Monte	Using Reordering in Statistical Machine Translation based on Recursive Alignment Block Classification
445	Steven Bird, Robert Dale, Bonnie J. Dorr, Bryan Gibson, Mark T. Joseph, Min-Yen Kan, Dongwon Lee, Brett Powley, Dragomir R. Radev, Yee Fan Tan	The ACL Anthology Reference Corpus : A Reference Dataset for Bibliographic Research in Computational Linguistics
44	Chris Biemann, Uwe Quasthoff, Gerhard Heyer, Florian Holz	ASV Toolbox – A Modular Collection of Language Exploration Tools
448	Michelina Savino, Laura Scivetti, Mario Refice	Integrating Audio and Visual Information for Modelling Communicative Behaviours Perceived as Different
44	Philippe Blache, Roxane Bertrand and Galle Ferr	Creating and exploiting multimodal annotated corpora
45	Eli Pociello, Antton Gurrutxaga, Eneko Agirre, Izaskun Aldezabal and German Rigau	WNTERM : Combining the Basque WordNet and a Terminological Dictionary
1454	Rubn Fernndez, Luis A. Hernndez, Eduardo Lpez, Jos Alczar, Guillermo Portillo, Doroteo T. Toledano	Design of a multimodal database for research on automatic detection of severe apnoea cases
45	A. Laurent, T. Merlin, S. Meignier, Y. Estve, P. Delglise	Combined systems for automatic phonetic transcription of proper nouns
545	Atsushi Fujii, Masao Utiyama, Mikio Yamamoto, Takehito Utsuro	Producing a Test Collection for Patent Machine Translation in the Seventh NTCIR Workshop
45		

List of accepted papers

9

Bartosz Broda, Magdalena Derwojedowa, Maciej Piasecki, Stanislaw Szpakowicz	Corpus-based Semantic Relatedness for the Construction of Polish WordNet

46	Xinglong Wang, Claire Grover	Learning the Species of Biomedical Named Entities from Annotated Corpora
46	Iker Luengo, Eva Navas, Iaki Sainz, Ibon Saratxaga, Jon Sanchez, Igor Odriozola, Inma Hernaez	Text independent speaker identification in multilingual environments
1		
46	Le An Ha, Gabriela Fernandez, Ruslan Mitkov, Gloria Corpas	Mutual bilingual terminology extraction
46	Germn Sanchis-Trilles, Joan Andreu Snchez	Using Parsed Corpora for Estimating Stochastic Inversion Transduction Grammars
5 46 8	Takanobu Nishiura, Masato Nakayama, Yuki Denda, Norihide Kitaoka, Kazumasa Yamamoto, Takeshi Yamada, Satoru Tsuge, Chiyomi Miyajima, Masakiyo Fujimoto, Tetsuya Takiguchi, Satoshi Tamura, Shingo Kuroiwa, Kazuya Takeda, Satoshi Nakamura	Evaluation Framework for Distant-talking Speech Recognition under Reverberant Environments — Newest Part of the CENSREC Series—
46	Hiroaki Saito, Shunta Kuboya, Takaaki Sone, Hayato Tagami, Kyoko Ohara	The Japanese FrameNet Software Tools
9 47	Adriana Roventini, Nilda Ruimy	Mapping Events and Abstract Entities from PAROLE-SIMPLE-CLIPS to ItalWordNet
47 2	Akira Ozaki, Sunao Hara, Takashi Kusakawa, Chiyomi Miyajima, Takanori Nishino, Norihide Kitaoka, Katunobu Itou, Kazuya Takeda	In-car Speech Data Collection along with Various Multimodal Signals
47 3	Michael Roth, Sabine Schulte im Walde	Corpus Co-Occurrence, Dictionary and Wikipedia Entries as Resources for Semantic Relatedness Information
47	Kathrin Spreyer, Jonas Kuhn, Bettina Schrader	Identification of Comparable Argument-Head Relations in Parallel Corpora
5 47 7	Martin Reynaert	All, and only, the errors : more complete and consistent spelling and OCR-error correction evaluation
47	Petya Osenova, Kiril Simov, Eelco Mossel	Language Resources for Semantic Document Annotation and Crosslingual Retrieval
8 47	Svitlana Kurella, Serge Sharoff, Anthony Hartley	Corpus-based Tools for Computer-assisted Acquisition of Reading Abilities in Cognate Languages
48	Jan Pomiklek, Pavel Rychl	Detecting co-derivative documents in large text collections
1 48	Yi Zhang, Valia Kordoni	Robust Parsing with a Large HPSG Grammar
48	Jess Gimnez, Llus Mrquez	Towards Heterogeneous Automatic MT Error Analysis
3 48	Jrg Tiedemann	Synchronizing Translated Movie Subtitles
48	Ingunn Amdal, Ole Morten Strand, Jrn Almberg and Torbjrn Svendsen	RUNDKAST : An Annotated Norwegian Broadcast News Speech Corpus
6 48	Olga Uryupina	Error Analysis for Learning-based Coreference Resolution
49	Sara Tonelli, Rodolfo Delmonte, Antonella Bristot	Enriching the Venice Italian Treebank with dependency and grammatical relations
0 49	Rafiya Begum, Samar Husain, Lakshmi Bai, Dipti Misra Sharma	Developing Verb Frames for Hindi
49 3	INH Quang Th¯ng , L HÓng Ph°¡ng, NGUYÄN ThË Minh HuyÁn, NGUYÄN C©m T, Mathias ROSSIGNOL, Vh Xun L°¡ng	Word segmentation of Vietnamese texts : a comparison of approaches
49	Ana-Maria Barbu	Romanian Lexical Data Bases : Inflected and Syllabic Forms Dictionaries
49	Yuka Tateisi, Yusuke Miyao, Kenji Sagae, Jun'ichi Tsujii	GENIA-GR : a Grammatical Relation Corpus for Parser Evaluation in the Biomedical Domain
6 49	Christian Raymond, Kepa Joseba Rodriguez, Giuseppe Riccardi	Active Annotation in the LUNA Italian Corpus of Spontaneous Dialogues

9

500	Amalia Todira_cu, Dan Tufi_, Ulrich Heid, Christopher Gledhill, Dan ^tefanescu, Marion Weller, Franois Rousselot	A Hybrid Approach to Extracting and Classifying Verb+Noun Constructions
501	Siaw-Fong Chung, Laurent Prvot, Mingwei Xu, Kathleen Ahrens, Shu-Kai Hsieh, Chu-Ren Huang	Extracting Concrete Senses of Lexicon through Measurement of Conceptual Similarity in Ontologies
50	Nick Webb, Ting Liu, Mark Hepple, Yorick Wilks	Cross-Domain Dialogue Act Tagging
2507	Ineke Schuurman	Spatiotemporal annotation using MiniSTEx : How to deal with alternative, foreign, vague and/or obsolete names ?
509	Mart Quixal, Toni Badia, Francesc Benavent, Jose R. Boullosa, Judith Domingo, Bernat Grau, Guillem Mass, Oriol Valentn	User-centred design of language checking tools
51	Florian Schiel, Hannes Mgele	Talking and Looking : the SmartWeb Multimodal Interaction Corpus
51	L. Lamel, S. Rosset, C. Ayache, D. Mostefa, J. Turmo, P. Comas	Question Answering on Speech Transcriptions : the QAST evaluation in CLEF
1512	Takeshi Abekawa and Kyo Kageura	Constructing a corpus that indicates patterns of modification between draft and final translations by human translators
51	Mehdi Embarek and Olivier Ferret	Learning patterns for building resources about semantic relations in the medical domain
451	Scott Piao, John McNaught and Sophia Ananiadou	Clustering Related Terms with Definitions
51	Barry Haddow and Beatrice Alex	Exploiting Multiply Annotated Corpora in Biomedical Information Extraction Tasks
651	Tiit Hennoste, Olga Gerassimenko, Riina Kasterpalu, Mare Koit, Andriela Rbis, Krista Strandson	From Human Communication to Intelligent User Interfaces : Corpora of Spoken Estonian
51	Atsushi Fujii	Producing an Encyclopedic Dictionary Using Patent Documents
952	W. Heeren, F. de Jong, L. van der Werff, M. Huijbregts and R. Ordelman	Evaluation of Spoken Document Retrieval for Historic Speech Collections
52	Michael Kipp	Spatiotemporal Coding in ANVIL
152	Costanza Navarretta and Sussi Olsen	Annotating abstract pronominal anaphora in the DAD project
52	Lothar Lemnitzer and Paola Monachesi	Extraction and evaluation of keywords from Learning Objects – a multilingual approach
352	Masatoshi Tsuchiya, Satoru Kogure, Hiromitsu Nishizaki, Kengo Ohta, Seiichi Nakagawa	Developing Corpus of Japanese Classroom Lecture Speech Contents
52	C. Bosco, A. Mazzei, V. Lombardo, G. Attardi, A. Corazza, A. Lavelli, L. Lesmo, G. Satta, M. Simi	Comparing Italian parsers on a common treebank : the Evalita experience
853	Maria Teresa Pazienza, Armando Stellato, Alexandra Tudorache	JMWNL : An extensible multilingual library for accessing wordnets in different languages
53	Emiliano Giovannetti, Simone Marchi, Simonetta Montemagni, Roberto Bartolini	Ontology Learning and Semantic Annotation : a Necessary Symbiosis
4535	Sherri Condon, Jon Phillips, Christy Doran, John Aberdeen, Dan Parvaz, Beatrice Oshika, Greg Sanders and Craig Schlenoff	Applying Automated Metrics to Speech Translation Dialogs
53	Dino Ienco, Serena Villata, Cristina Bosco	Automatic extraction of subcategorization frames for Italian
653	Kremena Ivanova, Ulrich Heid, Sabine Schulte im Walde, Adam Kilgarriff, Jan Pomikalek	Evaluating a German Sketch Grammar : A Case Study on Noun Phrase Case
53	Ziqi Zhang, Jos Iria, Christopher Brewster and Fabio Ciravegna	A Comparative Evaluation of Term Recognition Algorithms
853	Constantin Or san, Oana Andreea Chiorean	Evaluation of a Cross-lingual Romanian-English Multi-document Summariser

9

54	Peng Zhang, Wenjie Li, Furu Wei, Qin Lu, Yuexian Hou	Exploiting the Role of Position Feature in Chinese Relation Extraction
0 54	Gaoying Cui, Qin Lu, Wenjie Li, Yirong Chen	Corpus Exploitation from Wikipedia for Ontology Construction
54 2	Bogdan Babych, Anthony Hartley	Sensitivity of automated MT evaluation metrics on higher quality MT output : BLEU vs task-based evaluation methods
54	Julia Ritz, Stefanie Dipper, Michael Gtze	Annotation of Information Structure : An Evaluation Across Different Types of Texts
54	Laura Kassner, Vivi Nastase, Michael Strube	Acquiring a Taxonomy from the German Wikipedia
4 54	Konrad Hofbauer, Stefan Petrik, Horst Hering	The ATCOSIM Corpus of Non-Prompted Clean Air Traffic Control Speech
54 6	Mara Teresa Vicente-Dez, Doaa Samy and Paloma Martnez	An empirical approach to a preliminary successful identification and resolution of temporal expressions in Spanish news corpora
54 7	Paul Thompson, Philip Cotter, Sophia Ananiadou, John McNaught, Simonetta Montemagni, Andrea Trabucco, Giulia Venturi	Building a Bio-Event Annotated Corpus for the Acquisition of Semantic Frames from Biomedical Corpora
54	Saša Hasan and Hermann Ney	A multi-genre SMT system for Arabic to French
9 55	Yorick Wilks, David Benyon, Christopher Brewster, Pavel Ircing and Oli Mival	Dialogue, Speech and Images : The Companions Project Data Set
55	Zhiyi Song, Stephanie Strassel	Entity Translation and the Alignment in the ACE-07 ET Task
1 55	Ben Allison & Louise Guthrie	Authorship Attribution of E-Mail : Comparing Classifiers Over a New Corpus for Evaluation
55	Simon Krek, Vojko Gorjanc, Špela Arhar	Slovene Terminology Web Portal and the TBX-Compatible Simplified DTD/schema
3 55 4	Rena Nemoto, Ioana Vasilescu, Martine Adda-Decker	Speech errors on frequently observed homophones in French : Perceptual evaluation vs automatic classification
55 6	C.J. Rupp, Ann Copestake, Peter Corbett, Peter Murray-Rust, Advaith Siddharthan, Simone Teufel, Benjamin Waldron	Language Resources and Chemical Informatics
55 7	Thomas Winkler, Theodoros Kostoulas, Richard Adderley, Christian Bonkowski, Todor Ganchev, Joachim Khler, Nikos Fakotakis	The MoveOn Motorcycle Speech Corpus
55	Hiromi itoh Ozaku, Akinori Abe, Kaoru Sagara and Kiyoshi Kogure	Relationships between nursing conversations and activities
8 55	Folkert de Vriend, Jan Pieter Kunst, Louis ten Bosch, Charlotte Giesbers, Roeland van Hout	Evaluating the Relationship between Linguistic and Geographic Distances using a 3D Visualization
56	Sanaz Jabbari, Ben Allison, Louise Guthrie	Using A Probabilistic Model Of Context To DetectWord Obfuscation
0 56	Shiyan Ou, Viktor Pekar, Constantin Or san, Christian Spurk, Matteo Negri	Development and Alignment of a Domain-Specific Ontology for Question Answering
56	Maria Teresa Pazienza, Marco Pennacchiotti, Armando Stellato	A Web Browser Extension for growing-up Ontological Knowledge from Traditional Web Content
2 56	Keiji Shinzato, Daisuke Kawahara, Chikara Hashimoto, Sadao Kurohashi	A Large-Scale Web Data Collection as a Natural Language Processing Infrastructure
56 5	Michael Kaisser, John B. Lowe	Creating a Research Collection of Question Answer Sentence Pairs with Amazon’s Mechanical Turk
56	Andrea Esuli, Fabrizio Sebastiani and Ilaria C. Urciuoli	Annotating Expressions of Opinion and Emotion in the Italian Content Annotation Bank
56	Sara Tonelli, Emanuele Pianta	Frame information transfer from English to Italian

7

568	U. Hahn, E. Beisswanger, E. Buyko, M. Poprat, K. Tomanek, J. Wermter	Semantic Annotations for Biology — A Corpus Development Initiative at the Jena University Language & Information Engineering (JULIE) Lab
56	Manuel Kountz, Ulrich Heid, Kerstin Eckart	A LAF/GrAF based Encoding Scheme for underspecified Representations of syntactic Annotations
957	Laurent Blin, Olivier Boeffard and Vincent Barreaud	WEB-based listening test system for speech synthesis and speech conversion evaluation
57	Christoph Draxler, Florian Schiel, Tania Ellbogen	F0 Of Adolescent Speakers – First Results for the German Ph@ttSessionz Database
4575	Yoshiko Arimoto, Sumio Ohno, Hitoshi Iida	Automatic Emotional Degree Labeling for Speakers' Anger Utterance during Natural Japanese Dialog
57	Valeria Quochi, Monica Monachini, Riccardo Del Gratta, Nicoletta Calzolari	A Lexicon for Biology and Bioinformatics : The BOOTStrep Experience
657	Maria Teresa Pazienza, Armando Stellato, Alexandra Tudorache	A Bottom-up Comparative Study of EuroWordNet and WordNet 3.0 Lexical and Semantic Relations
57	Viktor Bielick, Otakar Smr~	Building the Valency Lexicon of Arabic Verbs
858	Andrei Popescu-Belis, Mike Flynn, Pierre Wellner, Philippe Baudrion	Task-based evaluation of meeting browsers : from task elicitation to user behavior analysis
58	Jordi Atserias, Hugo Zaragoza, Massimiliano Ciaramita, Giuseppe Attardi	Semantically Annotated Snapshot of the English Wikipedia
158	Pedro Concejero, Daniel Tapias, Juan Jos Rodrguez, Juan Carlos Luengo, Sebastin Snchez	Methodology for Evaluating the Usability of User Interfaces in Mobile Services
58	Sandra Kbler, Wolfgang Maier, Ines Rehbein, Yannick Versley	How to Compare Treebanks
358	A. Brutti, L. Cristoforetti, W. Kellermann, L. Marquardt, M. Omologo	WOZ Acoustic Data Collection for Interactive TV
58	Linus Sellberg, Arne Jnsson	Using Random Indexing to improve Singular Value Decomposition for Latent Semantic Analysis
658	Jerid Francom and Mans Hulden	Parallel multi-theory annotations of syntactic structure
58	Chen Yirong, Lu Qin, Li Wenjie, Cui Gaoying	Chinese Core Ontology Construction from a Bilingual Term Bank
859	Massimo Poesio, Udo Kruschwitz, Jon Chamberlain	ANAWIKI : Creating Anaphorically Annotated resources Through Web Cooperation
592	Verena Rieser and Oliver Lemon	Automatic Learning and Evaluation of User-Centered Objective Functions for Dialogue System Optimisation
59	Grzegorz Chrupala, Georgiana Dinu, Josef van Genabith	Learning Morphology with Morfette
59	Alessandra Potrich, Emanuele Pianta	L-ISA : Learning Domain Specific Isa-Relations from the Web
559	Laure Charonnat, Galle Vidal and Olivier Boeffard	Automatic phone segmentation of expressive speech
59	Riccardo Del Gratta, Nilda Ruimy, Antonio Toral	Simple-Clips ongoing research : more information with less data by implementing inheritance
859	Davide Picca, Alfio Massimiliano Gliozzo, Massimiliano Ciaramita	Supersense Tagger for Italian
600	Wei-Hao Lin and Alexander Hauptmann	Vox Populi Annotation : Measuring Intensity of Ideological Perspectives by Aggregating Group Judgments
604	Jordi Carrera, Irene Castelln, Salvador Climent, Marta Coll-Florit	Towards Spanish Verbs’ Selectional Preferences Automatic Acquisition. Semantic Annotation of the SenSem Corpus
60	Pierre Lanchantin, Andrew C. Morris, Xavier Rodet, Christophe Veaux	Automatic Phoneme Segmentation with Relaxed Textual Constraints
660		

List of accepted papers

7

Ngan Nguyen, Jin-Dong Kim, Junichi Tsujii	Challenges in Pronoun Resolution System for Biomedical Text

60	Davide Picca, Alfio Massimiliano Gliozzo, Aldo Gangemi	LMM : an OWL-DL MetaModel to Represent Heterogeneous Lexical Knowledge
860	Hitoshi Isahara, Francis Bond, Kiyotaka Uchimoto, Masao Utiyama, Kyoko Kanzaki	Development of the Japanese WordNet
61	Tommaso Caselli, Nancy Ide, Roberto Bartolini	A Bilingual Corpus of Inter-linked Events
061	Neil Newbold, Bogdan Vrusias, Lee Gillam	Lexical Ontology Extraction Using Terminology Analysis : Automating Video Annotation
61	Victoria Bobicev and Tatiana Zidra_co	Estimating Word Phonosemantics
2613	Petra-Maria Strau, Holger Hoffmann, Wolfgang Minker, Heiko Neumann, Gnther Palm, Stefan Scherer, Harald C. Traue and Ulrich Weidenbacher	The PIT Corpus Of German Multi-Party Dialogues
61	Feiyu Xu, Hans Uszkoreit, Hong Li, Niko Felger	Adaptation of Relation Extraction Rules to New Domains
461	Edouard Geoffrois	An Economic View on Human Language Technology Evaluation
61	Nancy Ide, Collin Baker, Christiane Fellbaum, Charles Fillmore, Rebecca Passonneau	MASC : The Manually Annotated Sub-Corpus of American English
7618	Asuka Sumida, Naoki Yoshinaga, Kentaro Torisawa	Boosting Precision and Recall of Hyponymy Relation Acquisition from Hierarchical Layouts in Wikipedia
61	Rajat Kumar Mohanty, Pushpak Bhattacharyya	Lexical Resources for Semantics Extraction
962	Nikos Tsourakis, Maria Georgescul, Pierrette Bouillon and Manny Rayner	Building Mobile Spoken Dialogue Applications Using Regulus
62	P. Paroubek, I. Robba, A. Vilnat, C. Ayache	EASY, Evaluation of Parsers of French : what are the results ?
162	Nadine Perera, Michael Pitz, Manfred Pinkal	CLIoS : Cross-lingual Induction of Speech Recognition Grammars
62	Rita Marinelli, Melissa Tiberi, Remo Bindi	Encoding Terms from a Scientific Domain in a Terminological Database : Methodology and Criteria
3624	Mukda Suktarachan, Dussadee Thamvijit, Sachit Rajbhandari, Daoyos Noikongka, Puwarat P. Na Mahasarakram, Panita Yongyuth, Asanee Kawtrakul, Margherita Sini	Workbench with Authoring Tools for Collaborative Multi-lingual Ontological Knowledge Construction and Maintenance
62	Benjamin Farber, Dayne Freitag, Nizar Habash, Owen Rambow	Improving NER in Arabic Using a Morphological Tagger
5626	Violaine Prince, Jacques Chauch	Building a Bilingual Representation of the Roget Thesaurus for French to English Machine Translation
62	Paula Cristina Vaz, David Martins de Matos, Nuno J. Mamede	Using Lexical Acquisition to Enrich a Predicate Argument Reusable Database
7628	Elena Cabrio, Milen Kouylekov, Bernardo Magnini, Matteo Negri, Laura Hasler, Constantin Or san, David Toms, Jos L. Vicedo, Guenter Neumann, Corinna Weber	The QALL-ME Benchmark : a Multilingual Resource of Annotated Spoken Requests for Question Answering
629	Maria Teresa Pazienza, Armando Stellato	Clustering of terms from translation dictionaries and synonyms lists to automatically build more structured Linguistic Resources
630	B. Magnini, A. Cappelli, F. Tamburini, C. Bosco, A. Mazzei, V. Lombardo, F. Bertagna, N. Calzolari, A. Toral, V. Bartalesi Lenzi, R. Sprugnoli, and M. Speranza	Evaluation of Natural Language Tools for Italian : EVALITA 2007
63	Kristina Vukovi , Marko Tadi , Zdravko Dovedan	Rule Based Chunker for Croatian
163	N. Moreau, D. Mostefa, R. Stiefelhagen, S. Burger, K. Choukri	Data Collection for the CHIL CLEAR 2007 Evaluation Campaign
63	Stephan Busemann, Yajing Zhang	Identifying Foreign Person Names in Chinese Text

5

638	Lucie Mladov, Šrka Ziknov, Eva Hajiov	From Sentence to Discourse : Building an Annotation Scheme for Discourse Based on Prague Dependency Treebank
639	Cong-Phap Huynh, Christian Boitet, Herv Blanchon	SECTra_w.1 : an Online Collaborative System for Evaluating, Post-editing and Presenting MT Translation Corpora
64	Antnio Branco, Francisco Costa, Pedro Martins, Filipe Nunes, Joo Silva, Sara Silveira	LXService : Web Services of Language Technology for Portuguese
64	Luka Nerima, Eric Wehrli	Generating Bilingual Dictionaries by Transitivity
164	Sonja Bosch, Laurette Pretorius, Kholisa Podile, Axel Fleisch	Experimental fast-tracking of morphological analysers for Nguni languages
64	Valeria Quochi, Basilio Calderone	Learning properties of Noun Phrases : from data to functions
464	Emanuele Pianta, Christian Girardi, Roberto Zanoli	The TextPro tool suite
64	Mark D. Arehart, Chris Wolf, Keith J. Miller	Adjudicator Agreement and System Rankings for Person Name Search
764	Chris Reed, Raquel Mochales Palau, Glenn Rowe, Marie-Francine Moens	Language Resources for Studying Argument
65	Piotr BaDski, RadosBaw MoszczyDski	Enhancing an English-Polish electronic dictionary for multiword expression research
1652	Martine Garnier-Rizet, Gilles Adda, Frederik Cailliau, Jean-Luc Gauvain, Sylvie Guillemin-Lanne, Lori Lamel, Stephan Vanni, ClaireWaast-Richard	CallSurf - Automatic transcription, indexing and structuration of call center conversational speech for knowledge extraction and query by content
65	Mehrnoush Shamsfard	Towards Semi Automatic Construction of a Lexical Ontology for Persian
4656	Riccardo Del Gratta, Roberto Bartolini, Tommaso Caselli, Monica Monachini, Claudia Soria and Nicoletta Calzolari	UFRA : A UIMA-based Approach to Federated Language Resource Architecture
657	Simon Scerri, Myriam Mencke, Brian Davis, Siegfried Handschuh	Evaluating the Ontology underlying sMail - the Conceptual Framework for Semantic Email Communication
658	D. Llorens, F. Prat, A. Marzal, J. M. Vilar, M. J. Castro, J. C. Amengual, S. Barrachina, A. Castellanos, S. Espaa, J. A. Gmez, J. Gorbe, A. Gordo, V. Palazn, G. Peris, R. Ramos-Garijo, F. Zamora	The UJIpenchars Database : A Pen-Based Database of Isolated Handwritten Characters
66	Margot Mieskes, Michael Strube	Parameters for Topic Boundary Detection in Multi Party Dialogues
066	Sopheap Seng, Sethserey Sam, Laurent Besacier, Brigitte Bigi, Eric Castelli	First Broadcast News Transcription System for Khmer Language
66	Weiruo Qu, Christoph Ringlstetter, Randy Goebel	Targeting Chinese Nominal Compounds in Corpora
266	Eva Banik and Alan Lee	A Study of Parentheticals in Discourse Corpora— Implications for NLG Systems
66	Theodoros Kostoulas, Todor Ganchev, Iosif Mporas, Nikos Fakotakis	A Real-World Emotional Speech Corpus for Modern Greek
466	Margot Mieskes, Michael Strube	A Three-stage Disfluency Classifier for Multi Party Dialogues
67	Renata Dividino, Massimo Romanelli and Daniel Sonntag	Semiotic-based Ontology Evaluation Tool S-OntoEval
1674	Beatrice Alex	Comparing Corpus-based toWeb-based Lookup Techniques for Automatic English Inclusion Detection
67	Laura Hasler	Centering Theory for Evaluation of Coherence in Computer-Aided Summaries
667		

List of accepted papers

7

67	Xiwen Cheng, Feiyu Xu	Fine-grained Opinion Topic and Polarity Identification
8 67	Jean Tavernier, Rosa Cowan, Michelle Vanni	Holy Moses ! Leveraging Existing Tools and Resources for Entity Translation
68	Sandra Williams, Richard Power	Deriving rhetorical complexity data from the RST-DT Corpus
5 68 6	Chu-Ren Huang, Lung-Hao Lee, Wei-guang Qu, Jia-Fei Hong, Shiwen Yu	Quality Assurance of Automatic Annotation of Very Large Corpora : a Study based on Heterogeneous Tagging Systems
69	Mark McConville, Myroslava O. Dzikovska	Evaluating Complement-Modifier Distinctions in a Semantically Annotated Corpus
1 69	Theresa Wilson	Annotating Subjective Content in Meetings
69	Vclav Novk, Keith Hall	Inter-sentential Coreferences in Semantic Networks : An Evaluation of Manual Annotation
5 69	Eugenio Picchi, Eva Sassolini, Sebastiana Cucurullo, Francesca Bertagna, Paola Baroni	Semantic Press
69	Claire Cardie, Cynthia Farina, Matt Rawding, Adil Aijaz	An eRulemaking Corpus : Identifying Substantive Issues in Public Comments
9 70	Carmen Banea, Rada Mihalcea, Janyce Wiebe	A Bootstrapping Method for Building Subjectivity Lexicons for Languages with Scarce Resources
70	Lei Xia, Jos Iria	An Approach to Modeling Heterogeneous Resources for Information Extraction
2 70	Mohamed Maamouri, Seth Kulick, Ann Bies	Diacritic Annotation in the Arabic Treebank and Its Impact on Parser Evaluation
70	Viktor Trn	On the durational reduction of repeated mentions : recency and speaker effects
7 70	Josef Ruppenhofer, Swapna Somasundaran, Janyce Wiebe	Finding the Sources and Targets of Subjective Expressions
71	Chantal Enguehard, Harouna Naroua	Evaluation of Virtual Keyboards for West-African Languages
0 71	Georgiana Pu_ca_u and Verginica Barbu Mititelu	Annotation of WordNet Verbs with TimeML Event Classes
71	Constantin Or san, Dan Cristea, Ruslan Mitkov, Antnio Branco	Anaphora Resolution Exercise : An overview
3 71	Diana Trandab c and Maria Husarciuc	Romanian Semantic Roles Resource
71	Nikola Ljubeši , Tomislava Lauc, Damir Boras	Generating a Morphological Lexicon of Organization Entity Names
6 71	Jonathan Butters, Fabio Ciravegna	Using Similarity Metrics For Terminology Recognition
71	Ralf Krestel, Sabine Bergler and Ren Witte	Minding the Source : Automatic Tagging of Reported Speech in Newspaper Articles
8 72	Mark-Christoph Mller, Margot Mieskes, Michael Strube	Knowledge Sources for Bridging Resolution in Multi-Party Dialog
72	Johan Bos	Let's not argue about semantics
1 72 2	Rebecca J. Passonneau, Tom Lippincott, Tae Yano and Judith Klavans	Relation between Agreement Measures on Human Labeling and Machine Learning Performance : Results from an Art History Domain
72	Claire Brierley, Eric Atwell	ProPOSEL : A Prosody and POS English Lexicon for Language Engineering
4 72 5	Christian Monson, Ariadna Font Llitjs, Vamshi Ambati, Lori Levin, Alon Lavie, Alison Alvarez, Roberto Aranovich, Jaime Carbonell, Robert Frederking, Erik Peterson, Katharina Probst	Linguistic Structure and Bilingual Informants Help Induce Machine Translation of Lesser-Resourced Languages
72	Fabio Rinaldi, Gerold Schneider, Kaarel Kaljurand, Michael Hess	Dependency-Based Relation Mining for Biomedical Literature
8 72		

L
is
t
o
f
a
c
c
e
p
t
e
d
p
a
p
e
r
s

9

Joachim Gasch, Caren Brinckmann, Sylvia Dickgieer	memasysco : XML schema based metadata management system for speech corpora

73	Anca D. Dinu	On classifying coherent/incoherent Romanian short texts
2 73	Yu Chen, Andreas Eisele, Martin Kay	Improving Statistical Machine Translation Efficiency by Triangulation
73	Cosmin Bejan and Sanda Harabagiu	A Linguistic Resource for Discovering Event Structures and Resolving Event Coreference
4 73 5	Udhyakumar Nallasamy, Alan W Black, Tanja Schultz and Robert Frederking	NineOneOne : Recognizing and Classifying Speech for Handling Minority Language Emergency Calls
73	Kugatsu Sadamitsu, Satoshi Sekine, Mikio Yamamoto	Sentiment analysis based on probabilistic models using inter-sentence information
6 73	Catalina Hallett, David Hardcastle	Automatic rewriting of patient record narratives
73	Jalal Maleki, Lars Ahrenberg	Converting Romanized Persian to the Arabic Writing System
8 73	Natalie Schluter, Josef van Genabith	Treebank-Based Acquisition of LFG Parsing Resources for French
74	Carlos Gmez Gallo, T. Florian Jaeger, James Allen, Mary Swift	Production In A Multimodal Corpus : How Speakers Communicate Complex Actions
0 74	Lus C. Oliveira, Srgio Paulo, Lus Figueira, Carlos Mendes, Ana Nunes, Joaquim Godinho	Methodologies for Designing and Recording Speech Databases for Corpus Based Synthesis
74	Stephan Walter	Linguistic Description and Automatic Extraction of Definitions from German Court Decisions
2 74	Lorraine Goeuriot, Natalia Grabar, Batrice Daille	Characterization of scientific and popular science discourse in French, Japanese and Russian
74	Karel Pala, Lukš Svoboda, Pavel Šmerk	Czech MWE Database
6 74 8	Jan Bungeroth, Daniel Stein, Philippe Dreuw, Hermann Ney, Sara Morrissey, Andy Way, Lynette van Zijl	The ATIS Sign Language Corpus
74	Stefan Hahn, Patrick Lehnen, Christian Raymond, Hermann Ney	A Comparison of Various Methods for Concept Tagging for Spoken Language Understanding
9 75	Jonathan Chevelu, Nelly Barbot, Olivier Boeffard and Arnaud Delhay	Comparing set-covering strategies for optimal corpus design
75 1	Meghan Lammie Glenn, Stephanie Strassel, Lauren Friedman, Haejoong Lee	Management of Large Annotation Projects Involving Multiple Human Judges : a case study of GALE Machine Translation Post-editing
75 4	Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, Bonnie Webber	The Penn Discourse TreeBank 2.0.
75	Marian Reed, Denise DiPersio and Christopher Cieri	The Linguistic Data Consortium Member Survey : Purpose, Execution and Results
5 75	Angus Roberts, Robert Gaizasukas, Mark Hepple, Yikun Guo	Combining terminology resources and statistical methods for entity recognition : an evaluation
75	Kim Luyckx, Walter Daelemans	Personae : a corpus for author and personality prediction from text
9 76	Diana Santos, Alberto Simes	Portuguese-English word alignment : some experiments
76	David Day, Janet Hitzeman, Michael Wick, Keith Crouch, Massimo Poesio	A Corpus for Cross-Document Co-reference
2 76	Alessandro Lenci, Barbara McGillivray, Simonetta Montemagni, Vito Pirrelli	Unsupervised acquisition of verb subcategorization frames from shallow-parsed corpora
76	Karin Kipper-Schuler, Vinod Kaggal, James Masanz, Philip Ogren and Guergana Savova	System evaluation on a named entity corpus from clinical notes
4 76	Mrton Mihltz	Knowledge-based Coreference Resolution for Hungarian
76	Branimir Boguraev and Mary Neff	Navigating through Dense Annotation Spaces

9

77	Thierry Declerck	A Framework for Standardized Syntactic Annotation
77	Jade Goldstein-Stewart, Kerri A. Goodwin, Roberta E. Sabin, Ransom K. Winder	Creating and Using a Correlated Corpora to Glean Communicative Commonalities
1 77	Stephanie Strassel, Lauren Friedman, Safa Ismael, Linda Brandschain	New Resources for Document Classification, Analysis and Translation Technologies
77	Ulrich Schfer, Hans Uszkoreit, Christian Federmann, Torsten Marek, Yajing Zhang	Extracting and Querying Relations in Scientific Papers on Language Technology
3 77	Jennifer Foster and Josef van Genabith	Parser Evaluation and the BNC : Evaluating 4 constituency parsers with 3 metrics
77 5	Kazuaki Maeda, Haejoong Lee, Shawn Medero, Julie Medero, Robert Parker, Stephanie Strassel	Annotation Tool Development for Large-Scale Corpus Creation Projects at the Linguistic Data Consortium
77	Mark Fishel, Heiki-Jaan Kaalep	Experiments on Processing Overlapping Parallel Corpora
77 7	Eric Atwell, Bayan Abu Shawar	An AI-Inspired Intelligent Agent/Student Architecture to Combine Language Resources Research and Teaching
77 9	Kazuaki Maeda, Xiaoyi Ma, Stephanie Strassel	Creating Sentence-Aligned Parallel Text Corpora from a Large Archive of Potential Parallel Text using BITS and Champollion
78	Hitoshi Isahara, Masao Utiyama, Eiko Yamamoto, Akira Terada, Yasunori Abe	Application of Resource-based Machine Translation to Real Business Scenes
0 78	Eline Westerhout, Paola Monachesi	Creating glossaries using pattern-based and machine learning techniques
78	Yves Peirsman, Simon De Deyne, Kris Heylen, Dirk Geeraerts	The Construction and Evaluation of Word Space Models
4 78	Arne Mauser, Saša Hasan and Hermann Ney	Automatic Evaluation Measures for Statistical Machine Translation System Optimization
78 6	Rudolf Muhr	The Pronouncing Dictionary of Austrian German (AGPD) and the Austrian Phonetic Database (ADABA) – Report on a large Phonetic Resources Database of the three Major Varieties of German
78	Einar Meister, Jaak Vilo	Strengthening the Estonian Language Technology
78	M. Adda-Decker, C. Barras, G. Adda, P. Paroubek, P. Boula de Mareil, B. Habert	Annotation and analysis of overlapping speech in political interviews
8 78	Stephen A. Boxwell and Michael White	Projecting Propbank Roles onto the CCGbank
79 0	Kiyotaka Uchimoto and Yasuharu Den	Word-level Dependency-structure Annotation to Corpus of Spontaneous Japanese and Its Application
79	Caroline Lavecchia, David Langlois and Kamel Smali	Phrase-Based Machine Translation based on Simulated Annealing
79 3	Christopher Cieri, Stephanie Strassel, Meghan Glenn, Reva Schwartz, Wade Shen, Joseph Campbell	Bridging the Gap between Linguists and Technology Developers : Large-Scale, Sociolinguistic Annotation for Dialect and Speaker Recognition
79	Jeremy Nicholson, Valia Kordoni, Yi Zhang, Timothy Baldwin, Rebecca Dridan	Evaluating and Extending the Coverage of HPSG Grammars : A Case Study for German
79	Malvina Nissim, Sara Perboni	The Italian Particle ne : Corpus Construction and Analysis
5 79	Philip V. Ogren, Guergana K. Savova, Christopher G. Chute	Constructing Evaluation Corpora for Automated Clinical Named Entity Recognition
79	David Hardcastle, Donia Scott	Can We Evaluate the Quality of Generated Text ?
7 79	Tejaswini Deoskar and Mats Rooth	Induction of Treebank-Aligned Lexical Resources
80		

List of accepted papers

2

Meni Adler, Yael Netzer, Yoav Goldberg, David Gabay and Michael Elhadad	Tagging a Hebrew Corpus : The Case of Participles

80	Pavel Campr, Marek Hrz, Jana Trojanov	Collection and Preprocessing of Czech Sign Language Corpus for Sign Language Recognition
805	Keith J. Miller, Mark Arehart, Catherine Ball, John Polk, Alan Rubenstein, Ken Samuel, Elizabeth Schroeder, Eva Vecchi, Chris Wol	An Infrastructure, Tools and Methodology for Evaluation of Multicultural Name Matching Systems
80	Caren Brinckmann, Stefan Kleiner, Ralf Knbl, Nina Berend	German Today : an areally extensive corpus of spoken Standard German
80	Mohamed Maamouri, Ann Bies, Seth Kulick	Enhancing the Arabic Treebank : A Collaborative Effort toward New Annotation Guidelines
781 1	Arno Hartholt, Thomas Russ, David Traum, Eduard Hovy, Susan Robinson	A Common Ground for Virtual Humans : Using an Ontology in a Natural Language Oriented Virtual Human Architecture
812	Thorsten Trippel, Michael Maxwell, Greville Corbett, Cambell Prince, Christopher Manning, Stephen Grimes and Steve Moran	Lexicon Schemas and Related Data Models : when Standards Meet Users
81	Veselin Stoyanov, Claire Cardie	Annotating Topics of Opinions
81	Majid Razmara and Leila Kosseim	Answering List Questions using Co-occurrence and Clustering
481 5	Susanne Burger, Kornel Laskowski and Matthias Wlfel	A Comparative Cross-Domain Study of the Occurrence of Laughter in Meeting and Seminar Corpora
816	Wolodja Wentland, Johannes Knopp, Carina Silberer, Matthias Hartung	Building a Multilingual Lexical Resource for Named Entity Disambiguation,Translation and Transliteration
81	Corina For scu	GMT to +2 or How Can TimeML Be Used in Romanian
81	Kris Heylen, Yves Peirsman, Dirk Geeraerts, Dirk Speelman	Modelling Word Similarity. An Evaluation of Automatic Synonymy Extraction Algorithms
881	R. Catizone, A. Dingli, H. Pinto, Y. Wilks	Information Extraction Tools and Methods for Understanding Dialogue in a Companion
82	Youssef Drissi, Branimir Boguraev, David Ferrucci, Paul T. Keyser and Anthony Levas	A Development Environment for Configurable Meta-Annotators in a Pipelined NLP Architecture
082	Kyoko Hirose Ohara	Lexicon, Grammar, and Multilinguality in the Japanese FrameNet
82	Marianna Apidianaki	Translation-oriented Word Sense Induction Based on Parallel Corpora
282	Lynne Cahill	Using similarity measures to extend the LinGO lexicon
82	Dave Toney, Sophie Rosset, Aurlien Max, Olivier Galilbert, Eric Bilinski	An evaluation of spoken and textual interaction in the RITEL interactive question answering system
582	Todor Arnaudov, Ruslan Mitkov	Smarty – Extendable Framework for Bilingual and Multilingual Comprehension Assistants
82	Karine Megerdoomian and Dan Parvaz	Low-Density Language Bootstrapping : The Case of Tajiki Persian
782 8	Doaa Samy and Ana Gonzlez-Ledesma	Pragmatic Annotation of Discourse Markers in a Multilingual Parallel Corpus (Arabic-Spanish-English)
83	Gerard de Melo, Gerhard Weikum	Mapping Roget's Thesaurus and WordNet to French
083	Daisuke Kawahara, Kiyotaka Uchimoto	A Method for Automatically Constructing Case Frames for English
832	Eric Ringger, Marc Carmen, Robbie Haertel, Kevin Seppi, Deryle Lonsdale, Peter, McClanahan, James Carroll, Noel Ellison	Assessing the Costs of Machine-Assisted Corpus Annotation through a User Study
83	Antonio Bonafonte, Jordi Adell, Ignasi Esquerra, Silvia Gallego, Asuncin Moreno, Javier Prez	Corpus and Voices for Catalan Speech Synthesis
84		

List of accepted papers

3

Nizar Habash and Ryan Roth	Identification of Naturally Occurring Numerical Expressions in Arabic

844	Reginald Hobbs, Jamal Laoudi, Clare R. Voss	MTriage : Web-enabled Software for the Creation, Machine Translation, and Annotation of Smart Documents
847	Veronika Vincze, Gyrgy Szarvas, Attila Almsi, Dra Szauter, Rbert Ormndi, Richrd Farkas, Csaba Hatvani and Jnos Csirik	Hungarian word sense disambiguated corpus
84	Harald Hammarstm, Christina Thornell, Malin Petzell, Torbjrn Westerlund	Bootstrapping Language Description : The case of Mpiemo (Bantu A, Central African Republic)
84	Olga N. Lashevskaja, Olga Yu. Shemanaeva	Semantic Annotation Layer in Russian National Corpus : Lexical Classes of Nouns and Adjectives
9851	Olga Pustylnikov, Alexander Mehler, Rdiger Gleim	A Unified Database of Dependency Treebanks. Integrating, Quantifying & Evaluating Dependency Data.
85	M. Adda-Decker, T. Pellegrini, E. Bilinski, G. Adda	Developments of Ltzebuergesch resources for automatic speech processing and linguistic studies
585	A. Allauzen, H. Bonneau-Maynard	Training and evaluation of POS taggers on the French MULTITAG corpus
85	Pter Halcsy, Andrs Kornai, Pter Nmeth, Dniel Varga	Parallel creation of gigaword corpora for medium density languages – an interim report
886	Takahiro Ono, Hitomi Tohyama, Shigeki Matsubara	Construction and Analysis of Word-level Time-aligned Simultaneous Interpretation Corpus
86	Christopher Cieri, Mark Liberman	15 Years of Language Resource Creation and Sharing : A Progress Report on LDC Activities
186	Liviu P. Dinu, Marius Popescu, Anca Dinu	Authorship Identification of Romanian Texts with Controversial Paternity
86	Christophe Veaux, Gregory Beller, Xavier Rodet	IrcamCorpusTools : an extensible platform for speech corpora exploitation
386	Bodil Nistrup Madsen, Hanne Erdman Thomsen	A Taxonomy of Lexical Metadata Categories
86	David Guthrie, Louise Guthrie, Yorick Wilks	An Unsupervised Approach for the Detection of Outliers in Corpora
686	Peter Adolphs	Acquiring a Poor Man's Inflectional Lexicon for German
86	George Demetriou, Robert Gaizauskas, Haotian Sun and Angus Roberts	ANNALIST — ANNotation ALIgnment and Scoring Tool
987	Patcharee Varasai, Chaveevan Pechsiri, Thana Sukvari, Vee Satayamas and Asanee Kawtrakul	Building an Annotated Corpus for Text Summarization and Question Answering
87	Michael Rosner	ODL : An Object Description Language for Lexical Information
187	Rodney D. Nielsen, Wayne Ward, James H. Martin and Martha Palmer	Annotating Students’ Understanding of Science Concepts
87	Erin Fitzgerald, Frederick Jelinek	Linguistic Resources for Reconstructing Spontaneous Speech Text
487	Mehrnoush Shamsfard, Hakimeh Fadaee	A Hybrid Morphology-Based POS Tagger for Persian
87	Nianwen Xue, Hua Zhong and Kai-Yun Chen	Annotating "tense" in a tense-less language
7880	Martha Palmer, Olga Babko-Malaya, Ann Bies, Mona Diab, Mohammed Maamouri, Aous Mansouri, Wajdi Zaghouani	A Pilot Arabic Propbank
88	Maarten Janssen, Tiago Freitas	Spock - a Spoken Corpus Client
1882	Yoji Kiyota, Noriyuki Tamura, Satoshi Sakai, Hiroshi Nakagawa, Hidetaka Masuda	Automated Subject Induction from Query Keywords through Wikipedia Categories and Subject Headings
88	Mrk Fk, Nicolas Audibert, Jnos Szab, Albert Rilliard, Gza Nmeth, Vronique Auberg	Multimodal Spontaneous Expressive Speech Corpus for Hungarian
488		

List of accepted papers

5

Stefan Evert	A lightweight and efficient tool for cleaning Web pages

886	Marius Pa_ca	Low-Complexity Heuristics for Deriving Fine-Grained Classes of Named Entities from Web Textual Data
887	Clare R. Voss, Jamal Laoudi, Jeffrey Micher	Exploitation of an Arabic Language Resource for MT Evaluation : Using Buckwalter-based Lookup Tool to Augment CMU Alignment Algorithm
89	L.M. Spracklin, D.Z. Inkpen, A. Nayak	Using the Complexity of the Distribution of Lexical Elements as a Feature in
289	Anil Kumar Singh, Kiran Pala and Harshit Surana	Authorship Attribution
89	Marine Carpuat, Dekai Wu	Evaluation of Context-dependent Phrasal Translation Lexicons for Statistical Machine Translation
689	Victoria Arranz, Franck Gandcher, Valrie Mapelli and Khalid Choukri	A Guide for the Production of Reusable Language Resources
89	Valrie Mapelli, Victoria Arranz, Hlne Mazo, Khalid Choukri	Latest Developments in ELRA's Services
990	Olivier Hamon and Djamel Mostefa	An Experimental Methodology for an End-to-End Evaluation in Speech-to-Speech Translation
90	Djamel Mostefa, Arnaud Vallee	New telephone speech databases for French : a children database and an optimized adult corpus
190	Linda Brandschain, Christopher Cieri, David Graff, Abby Neely, Kevin Walker	Speaker Recognition : Building the Mixer 4 and 5 Corpora
90	Nilda Ruimy, Antonio Toral	More Semantic Links in the SIMPLE-CLIPS Database
390	Romaric Besanon, Stphane Chaudiron, Djamel Mostefa, Ismal Timimi, Khalid Choukri	The InFile project : a crosslingual filtering systems evaluation campaign
908	ric Villemonte de la Clergerie, Olivier Hamon, Djamel Mostefa, Christelle Ayache, Patrick Paroubek and Anne Vilnat	PASSAGE : from French Parser Evaluation to Large Sized Treebank
90	Olga Babko-Malaya	Annotation of Nuggets and Relevance in GALE Distillation Evaluation
91	Brian Davis, Siegfried Handschuh, Alexander Troussov, John Judge, Mikhail Sogrin	Linguistically Light Lexical Extensions for Ontologies
1912	Aicha Bouhjar	Amazigh Language Terminology in Morocco or Management of a ‘Multidimensional’ Variation
91	Carol Peters, Martin Braschler, Giorgio Di Nunzio, Nicola Ferro, Julio Gonzalo, Mark Sanderson	From Research to Application in Multilingual Information Access : the Contribution of Evaluation
391	J.V. White, D. Hunter, J.D. Goldstein	Statistical Evaluation of Information Distillation Systems
91	Bente Maegaard, M. Atiyya, K. Choukri, S. Krauwer, C. Mokbel, M. Yaseen	MEDAR – collaboration between European and Mediterranean Arabic partners

6