

СПИСЪК
на
цитирания в международни издания
(без автоцитирания и автоцитирания на съавтори)
на публикации на
Людмила Петрова Димитрова, доктор,
доцент в секция “Математическа лингвистика”
на Института по математика и информатика, БАН

Публикация (номер 55 в пълен списък на публикациите, номер 24 в списък на публикациите за конкурса)

Dimitrova, L., Koseska-Toszeva, V., Roszko, D., Roszko, R. Application of Multilingual Corpus in Contrastive Studies (on the example of the Bulgarian-Polish-Lithuanian Parallel Corpus. In: International Journal Cognitive Studies/Études Cognitives. Vol. 10, SOW, Warsaw, 2010. 217-240. ISSN 2080-7147.

Цитирана в:

- 1) M. Duškin, J. Satoła-Staškowiak. The Bulgarian-Polish-Russian Parallel Corpus. In: International Journal Cognitive Studies/Études Cognitives. Vol. 11, SOW, Warsaw, 2011. 241-254. ISSN 2080-7147.

Публикация (номер 43 в пълен списък на публикациите, номер 12 в списък на публикациите за конкурса)

Dimitrova, L., Rashkov, P. (2009). A New Version for Bulgarian MULTTEXT-East Morphosyntactic Specifications for Some Verbal Forms. In: Shyrov, Dimitrova (Eds. 2009), Organization and Development of Digital Lexical Resources. Kiev, 2009, 30-37. ISBN 978-966-507-252-2.

Цитирана в:

- 2) T. Erjavec. MULTTEXT-East: Morphosyntactic resources for many languages. In: International Journal Language Resources and Evaluation, 2011. 27 pp. (in print)

Публикация (номер 47 в пълен списък на публикациите, номер 16 в списък на публикациите за конкурса)

Ludmila Dimitrova, Violetta Koseska-Toszeva. Bulgarian-Polish Corpus. In: International Journal Cognitive Studies – Études Cognitives, Vol. 9, SOW, Warszawa 2009, 133–141.

Цитирана в:

- 3) M. Duškin, J. Satoła-Staškowiak. The Bulgarian-Polish-Russian Parallel Corpus. In: International Journal Cognitive Studies/Études Cognitives. Vol. 11, SOW, Warsaw, 2011. 241-254. ISSN 2080-7147.

Публикация (номер 52 в пълен списък на публикациите, номер 21 в списък на публикациите за конкурса)

Dimitrova, L., Koseska, V., Roszko, D., Roszko, R. Bulgarian-Polish-Lithuanian Corpus–Current Development. Association for Computational Linguistics (ACL), available at <http://www.aclweb.org/anthology/W/W09/W09-4001.pdf> (Proceedings of the International Workshop “Multilingual resources, technologies and evaluation for Central and Eastern European languages” in conjunction with International Conference RANPL’2009. Borovec, Bulgaria, 17 September 2009. 1-8. ISBN 978-954-452-008-3.)

Цитирана в:

- 4) M. Duškin, J. Satoła-Staškowiak. The Bulgarian-Polish-Russian Parallel Corpus. In: International Journal Cognitive Studies/Études Cognitives. Vol. 11, SOW, Warsaw, 2011. 241-254. ISSN 2080-7147.

Публикация (номер 53 в пълен списък на публикациите, номер 22 в списък на публикациите за конкурса)

Dimitrova, L., Koseska, V., Roszko, D., Roszko, R. Bulgarian-Polish-Lithuanian Corpus – Problems of Development and Annotation. In: T. Erjavec (Ed.), Research Infrastructure for Digital Lexicography. Ljubljana, 2009, 72-86. ISSN 1581-9973/ISBN 978-961-264-012-5.

Цитирана в:

- 5) M. Duškin, J. Satoła-Staškowiak. The Bulgarian-Polish-Russian Parallel Corpus. In: International Journal Cognitive Studies/Études Cognitives. Vol. 11, SOW, Warsaw, 2011. 241-254. ISSN 2080-7147.

Публикация (номер 48 в пълен списък на публикациите, номер 17 в списък на публикациите за конкурса)

Dimitrova, L., Panova, R., Dutsova, R. Lexical Database of the Experimental Bulgarian-Polish online Dictionary. In: Garabík (Ed.), Metalanguage and Encoding scheme Design for Digital Lexicography. Bratislava, 2009, 36-47. ISBN 978-80-7399-745-8. www.mondilex.org/pro3.pdf

Цитирана в:

- 6) Violetta Koseska and Antoni Mazurkiewicz. Time Flow and Tenses. SOW Publ. House, Warsaw, Poland. 223 pages. 2010. ISBN 978-83-89191-94-6.

Публикация (номер 50 в пълен списък на публикациите, номер 19 в списък на публикациите за конкурса)

L. Dimitrova, V. Koseska–Toszeva, R. Dutsova, R. Panova. Bulgarian-Polish online Dictionary – Design and Development. In: Koseska, Dimitrova, Roszko (Eds. 2009), Representing Semantics in Digital Lexicography. 76-88. SOW Publ. House, Warsaw, Poland. ISBN 978-83-89191-87-8. www.mondilex.org/pro4.pdf

Цитирана в:

- 7) Roszko, Danuta, Roszko, Roman. Morphosyntactic Specifications for Polish and Lithuanian. [Description of Morphosyntactic Markers for Polish and Lithuanian Nouns within MULTEXT-East Morphosyntactic Specifications (Version 3.0 May 10th, 2004)]. In: Koseska, Dimitrova, Roszko (Eds.2009), Representing Semantics in Digital Lexicography. Proceedings of the MONDILEX Fourth Open Workshop 29 June - 1 July, 2009. SOW Publ. House, Warsaw, Poland. 145-158. ISBN 978-83-89191-87. www.mondilex.org/pro4.pdf

Публикация (номер 45 в пълен списък на публикациите, номер 14 в списък на публикациите за конкурса)

Dimitrova, Ludmila, V. Koseska, I. Derzhanski, R. Rozsko. Annotation of Parallel Corpora (on the Example of the Bulgarian-Polish Parallel Corpus). In: Proceedings of the MONDILEX Second Open Workshop, 2-4 February, 2009, Kiev. Dovira Publ. House, Kyiv, Ukraine. 2009. 47-54. ISBN 978-966-507-252-2. www.mondilex.org/pro2.pdf

Цитирана в:

- 8) Adam Przepiókowski. A Comparison of Two Morphosyntactic Tagsets of Polish. In: Koseska, Dimitrova, Roszko (Eds.2009), Representing Semantics in Digital Lexicography. Proceedings of the MONDILEX Fourth Open Workshop, 29 June - 1 July, 2009. SOW Publ. House, Warsaw, Poland. 138-144. ISBN 978-83-89191-87. www.mondilex.org/pro4.pdf

- 9) Alexandr Rosen. 2010. Morphological tags in parallel corpora. In: Čermák, František, Klégr, Aleš, & Corness, Patrick (eds.). 2010. InterCorp: Exploring a Multilingual Corpus. Praha: Nakladatelství Lidové noviny. (in print)

Публикация (номер 44 в пълен списък на публикациите, номер 13 в списък на публикациите за конкурса)

Dimitrova, L., Garabik, R., Majchrakova, D. Comparing Bulgarian and Slovak Multext-East morphology tagset. In: Organisation and Development of Digital Lexical Resources. Proceedings of the MONDILEX Second Open Workshop, Kiev, Ukraine, 2–4 February 2009. 38-46. ISBN 978-966-507-252-2

Цитирана в:

- 10) Christian Chiarcos and Tomaz Erjavec (2011). OWL/DL formalization of the MULTEXT-East morphosyntactic specifications. In: *Proceedings of the 5th Linguistic Annotation Workshop (LAW-V), held in conjunction with the ACL-HLT 2011*, June 2011, Portland, Oregon, USA, p. 11-20. <http://aclweb.org/anthology-new/W/W11/W11-0402.pdf>
- 11) T. Erjavec (2011). MULTEXT-East: Morphosyntactic resources for many languages. In: *International Journal Language Resources and Evaluation*, 2011. 27 pp. (in print)
- 12) Christian Chiarcos (2010). Grounding an Ontology of Linguistic Annotations in the Data Category Registry. In: *Proceedings of the Language Resource and Language Technology Standards – state of the art, emerging needs, and future developments, 18th of May, 2010 within LREC'2010*, Malta. <http://www.lrec-conf.org/proceedings/lrec2010/workshops/W4.pdf>
- 13) Alexandr Rosen (2010). Morphological tags in parallel corpora. In: Čermák, František, Klégr, Aleš, & Corness, Patrick (eds.). 2010. InterCorp: Exploring a Multilingual Corpus. Praha: Nakladatelství Lidové noviny. (in print)
- 14) T. Erjavec (2010). MULTEXT-East Version 4: Multilingual Morphosyntactic Specifications, Lexicons and Corpora. In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation, LREC'10*, 19-21 May 2010, Malta. 2544-2547. Publ. by ELRA, Paris. ISBN 2-9517408-6-7 http://www.lrec-conf.org/proceedings/lrec2010/pdf/138_Paper.pdf

Публикация (номер 30 в пълен списък на публикациите, номер 7 в списък на публикациите за конкурса)

Dimitrova, Ludmila, Tomaz Erjavec, Nancy Ide, Heiki-Jan Kaalep, Vladimír Petkevic, and Dan Tufis. Multext-East: Parallel and Comparable Corpora and Lexicons for Six Central and Eastern European Languages. In *Proceedings of COLING-ACL '98: 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics*, Montréal, Québec, Canada, Ed. by Christian Boitet & Pete Whitelock, 315-319. San Francisco, Calif.: Morgan Kaufmann.

Цитирана в:

- 15) Jernej Vičič, Petr Homola (2010). Speeding up the Implementation Process of a Shallow Transfer Machine Translation System. In *Proceedings of the 14th Annual Conference of the European Association for Machine Translation EAMT'2010*, 27-28 May 2010, St Raphael, France. Also at: <http://www.mt-archive.info/EAMT-2010-Vicic.pdf>
- 16) Homa Baradaran Hashemi, Azaden Shakery, Hesham Faili (2010). Creating a Persian-English Comparable Corpus. In *Proceedings of the International Conference of the Cross-Language Evaluation Forum, CLEF 2010*, Padua, Italy, September 2010. **Multilingual**

- and Multimodal Information Access Evaluation, Lecture Notes in Computer Science, vol. 6360, Springer 2010, 27-39,**
- 17) Emmerich Kelih. (2009) Preliminary Analysis of a Slavic Parallel Corpus. In: Proceedings of the SLOVKO conference. November 25-27, 2009. Slovak AS, Bratislava, 2009, 175-183. ISBN 978-807399-875-2. http://korpus.juls.savba.sk/~slovko/2009/Proceedings_Slovko_2009.pdf
 - 18) Emmerich Kelih. (2009) Slawisches Parellel-Textkorpus: Projektvorstellung von "Kak zakal'jalas' stal' (KZS)". In: E. Kelih, E., Levickij V. V., and Altmann, G. (editors), Methods in text analysis //Методи аналізу тексту: збірник наукових праць. – Чернівці: ЧНУ, 2009. – 350 с.//, pages 106–124, Cernivci. Ruta. 2009. ISBN 978-966-423-043-5
 - 19) Jernej Vičič, Petr Homola and Vladislav Kubon. (2009) A Method to Restrict the Blow-up of Hypotheses of a Non-disambiguated. In: Proceeding of International Conference RANLP'2009, Borovec, Bulgaria, 14-16 September 2009. 460-464. ISSN 1313-8502.
 - 20) Camelia Ignat. (2009) Improving Statistical Alignment and Translation Using Highly Multilingual Corpora. June 2009, **Université de Strasbourg, France.** 215 pp.
 - 21) Ivan Derzhanski, Natalia Kotsyba. (2009) Towards a Consistent Morphological Tagset for Slavic Languages: Extention of the MULTEXT-East Morphosyntactic Specifications for Byelorussian, Polish and Ukrainian. In: Metalanguage and Encoding scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, Bratislava, Slovak Republic, 15–16 April 2009. 36-47. ISBN 978-5-9900813-6-9. www.mondilex.org/pro3.pdf
 - 22) Darja Fišer and Beñot Sagot. (2008) Combining Multiple Resources to Build Reliable Wordnets. In Proceedings of Text Speech and Dialogue Conference, TSD'08, Brno. 61-68. **Springer-Verlag Berlin Heidelberg 2008**, ISSN 0302-9743/ISBN-10 3-540-87390-2 **Springer Berlin Heidelberg NewYork/ISBN-13 978-3-540-87390-7** Springer Berlin Heidelberg NewYork <http://www.springerlink.com/content/j20t5p166654/front-matter.pdf>
 - 23) Darja Fišer (2008): Using Multilingual Resources for Building SloWNet Faster. In: Proceedings of the 4th International Global WordNet Conference, January 22-25, 2008, Szeged, Hungary. 185-193. ISBN 978-963-482-854-9. http://www.inf.u-szeged.hu/projectdirs/gwc2008/GWC2008_Proceedings_Final.pdf
 - 24) Jernej Vičič (2008). Rapid development of data for shallow transfer RBMT translation systems for highly inflective languages. In Proceedings of the Sixth Language Technologies Conference, October 16th-17th, 2008 : proceedings of the 11th International Multiconference Information Society - IS 2008, volume C, 98-103, ISSN 1581-9973, ISBN 978-961-264-006-4
 - 25) L. Felföldi (2008). Classifier Combination Systems and their Application in Human Language Technology. University of Szeged, Doctoral School in Mathematics and Computer Science, Ph.D. Programme in Informatics, May 2008, 118 pages
 - 26) Daniel Jurafsky and James H. Martin. Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition. 1024 pages, **Publisher: Prentice Hall; 2 edition (May 26, 2008), ISBN-10: 0131873210, ISBN-13: 978-0131873216**
 - 27) Ivan Derzhanski, Natalia Kotsyba. (2008) Morphosyntactic Tagging of Slavic Languages. In: Lexicographic Tools and Techniques. Proceedings of the MONDILEX First Open Workshop, 3-4 October 2008, Moscow. 68-79. Russia, Moscow, IITP-RAS Publ. House. ISBN 978-5-9900813-6-9. www.mondilex.org/pro1.pdf
 - 28) Natalia Kotsyba, Adam Radziszewski, Ivan Derzhanski. (2009) Integrating the Polish Language into MULTEXT-East Family: Morphosyntactic Specifications, Converter, Lexicon and Corpus. In: Proceedings of the MONDILEX, Fifth Open Workshop, October

- 14-15, 2009, Ljubljana, Slovene. 37-55. ISBN 978-961-264-012-5, ISSN 1581-9973. www.mondilex.org/pro5.pdf
- 29) TamasVaradi. (2007) NP Modification Structures in Parallel Corpora. In: Incorporating Corpora: The Linguist and the Translator, 2007 - books.google.com Pp. 168-186 Type: eBook Document; English. **Publisher: Clevedon ; Buffalo : Multilingual Matters, ©2008.** ISBN: 9781853599873 1853599875 <http://scholar.google.bg/scholar?hl=en&lr=&cites=13996482698111050937>
- 30) Obradović, I., R. Stanković, M. Utvić. (2007) Integrisano okruženje za pripremu paralelizovanog korpusa. Zbornik radova međunarodnog simpozijuma Razlike između bosanskog/bošnjačkog, hrvatskog i srpskog jezika, Graz, Austria, April 2007, B. Tošović (ur.), 563-578. http://www.rgf.bg.ac.rs/LicnePrezentacije/ivan_obradovic/radovi.html
- 31) Victoria Bobicev. (2007) METODE ȘI ALGORITMI STATISTICI DE PROCESARE A TEXTELOR - Statistica - PhD Thesis 2007. Universitatea Tehnică a Moldovei Cu titlu de manuscris CZU: 004.912 http://www.cnaa.acad.md/files/theses/2007/6478/viorica_bobicev-thesis.pdf.
- 32) Darja Fišer (2007): A Multilingual Approach to Building Slovene Wordnet. In: Proceedings of the International workshop on A Common Natural Language Processing Paradigm for Balkan Languages held within the International Conference Recent Advances in Natural Language Processing RANLP'07. 26 September 2007, Borovets, Bulgaria.
- 33) Julia Trushkina. (2007) Development of a Multilingual Parallel Corpus and a Part-of-Speech Tagger for Afrikaans. In: **IFIP International Federation for Information Processing, Volume 228, Intelligent Information Processing III**, eds. Z. Shi, Shimohara K., Feng D., (Boston: Springer), pp. 453-462. 2007. ISBN 978-0-378-44641-7-47
- 34) Julia Trushkina. (2006) The North-West University Bible corpus: A multilingual parallel corpus for South African languages. In: **Language Matters: Studies in the Languages of Africa, Volume 37, Issue 2, 2006, pages 227 – 245. Publishing House Routledge.** ISSN: 1753-5395 (electronic), ISSN:1022-8195 (paper) <http://www.informaworld.com/10.1080/10228190608566262>
- 35) Anna Feldman. (2006) Portable language technology: a resource-light approach to morpho-syntactic tagging. PhD Thesis, 2006, Ohio State University, <http://etd.ohiolink.edu/> <http://etd.ohiolink.edu/send-pdf.cgi/Feldman%20Anna.pdf?osu1153344391>
- 36) Andras Kornai, P. Halacsy, V. Nagy, Cs. Oravecz, V. Tron, D. Varga. (2006) Web-based frequency dictionaries for medium density languages. In 11th Conference of the European Chapter of the Association for Computational Linguistics, EACL-2006, April 2006, Trento, Italy. A. Kilgariff and M. Baroni (eds) Proceedings of the 2nd International Workshop on Web as Corpus. (EACL 2006 WS01) 1-8. <http://acl.ldc.upenn.edu/W/W06/W06-1701.pdf> http://acl.ldc.upenn.edu/eacl2006/ws01_webcorpus_02.pdf
- 37) Magda Ševčíková Razímová, Zdeněk Žabokrtský. (2006) Systematic Parameterized Description of Pro-forms in the Prague Dependency Treebank 2.0*Institute of Formal and Applied Linguistics, Charles University, Prague. In Proceedings of the Fifth Workshop on Treebanks and Linguistic Theories (TLT), December, 2006, Prague, Czech Republic, pp. 175-186.
- 38) Péter Halácsy, András Kornai, and Dániel Varga (2005). Morfológiai egyértelműsítés maximum entrópia módszerrel. In: III. Magyar Számítógépes Nyelvészeti Konferencia, edited by Zoltán Alexin, Dóra Csendes. Szegedi Tudományegyetem, pages 180–189. http://mokk.bme.hu/archive/halacsy05/doc/data/at_download/http://mokk.bme.hu/archive/halacsy05

- 39) Péter Halácsy, András Kornai, László Németh, Bálint Sas, Dániel Varga, Tamás Váradi, and Attila Vonyó (2005). A Hunglish korpusz és szótár. In: Proceedings III. Magyar Számítógépes Nyelvészeti Konferencia. Szegedi Tudományegyetem, 2005. <http://mokk.bme.hu/archive/hunglish05>.
- 40) Daniel Varga, Peter Halacsy, Andras Kornai, Viktor Nagy, Laszlo Nemeth, Viktor Tron. (2005) Parallel Corpora for Medium Density Languages. Proceedings of the International Conference RANLP-2005. Borovec, Bulgaria, 21-23 September 2005. 590-596. ISBN 954-91743-3-6. (A version of this paper appeared in: Nicolov et al (Eds): **Current Issues in Linguistics Theory, Volume 292, Recent Advances in Natural Language Processing IV**. Selected papers from RANLP 2005. 247-258. **John Benjamins Publ. Company Amsterdam, Philadelphia, 2007, ISBN 978 90 272 4807 7**.)
- 41) Verginica Barbu Mititelu, Radu Ion. (2005) Cross-Language Transfer of Syntactic Relations Using Parallel Corpora. In Diana Inkpen & Carlo Strapparava (eds.), Proceedings of the International Workshop within EuroLAN 2005 Summer School on **Cross-Language Knowledge Induction**, 25 July - 6 August 2005, Cluj-Napoca, Romania, pp. 46-51.
- 42) Radu Jon, Verginica Barbu Mititelu. (2005) Towards ROMANCE FrameNet. The translation task. Proceedings of the 1st ROMANCE FrameNet Workshop, 26-28 July 2005, Cluj-Napoca, Romania.
- 43) Verginica Barbu Mititelu, Radu Ion. (2005) Automatic Import of Verbal Syntactic Relations Using Parallel Corpora. Proceedings of the International Conference RANLP-2005. Borovec, Bulgaria, 21-23 September 2005. 329-332. ISBN 954-91743-3-6.
- 44) Behrang QasemiZadeh, Saeed Rahimi. Persian in MULTEXT-East Framework. In: Advances in Natural Language Processing, Proceedings of the 5th International Conference on NLP, FinTAL 2006, Turku, Finland, August 23-25, 2006. **Lecture Notes in Computer Science, vol. 4139, Springer 2006**, 541-551, DOI: 10.1007/11816508_54. ISBN 3-540-37334-9.
Also at: <http://www.springerlink.com/content/g23370g0w02m00t3>
- 45) Markus Dickinson. (2005) Error detection and correction in annotated corpora. Ph.D. thesis, 2005, Ohio State University. <http://etd.ohiolink.edu> // http://etd.ohiolink.edu/view.cgi?acc_num=osu1123788552
- 46) H. Kornilakis, M. Grigoriadou, E. Galiotou, E. Papakitsos. Building the Greek 1984 Corpus. In: Design and Development of a Multilingual Balkan Wordnet, {WP6: Improvement and Extension of the monolingual WordNets, APRIL 25, 2004}. Balkanet, IST-2000-29388. 75 pp. 2004
- 47) ISO/TC 37 / SC 4 N 080 Language resource management - Morphosyntactic annotation framework, 2003-08-21. 5 pp. 2003
- 48) Roger Evans, Carole Tiberius, Dunstan Brown and Greville Corbett. A large-scale inheritance-based morphological lexicon for Russian. ITRI-03-02, May, 2003, University of Brighton, 9 pp.
- 49) R. Evans, C. Tiberius, D. Brown, G. G. Corbett. A large-scale inheritance-based morphological lexicon for Russian. In Proceedings of the EACL'2003 Workshop on Morphological Processing of Slavic Languages, 9-16. 2003.
- 50) Harry Kornilakis, Maria Grigoriadou, Eleni Galiotou, Evangelos Papakitsos. Aligning, Annotating and Lemmatizing a Corpus for the Validation of BALKAN WORDNETS. In Proceedings of the Workshop on "Balkan Language Resources and Tools" in the framework of the 1st Balkan Conference in Informatics – BCI 2003, Stelios Piperidis, Vangelis Karkaletsis (Eds.), Thessaloniki, Greece, 21 November 2003, ISBN 954 90906-

- 51) András Kuba, András Hócz, and János Csirik. POS Tagging of Hungarian with Combined Statistical and Rule-Based Methods. In Petr Sojka, Ivan Kopeček, and Karel Pala, editors, Text, Speech and Dialogue, Proceedings of the 7th International Conference, TSD 2004, Brno, Czech Republic and LNAI. **Lecture Notes in Computer Science**, vol. 3206, Springer, 2004, 113-120, DOI: 10.1007/b100511 <http://springerlink.metapress.com/openurl.asp?genre=article&issn=0302-9743&volume=3206&spage=113>
- 52) András Kuba, László Felföldi, András Kocsor. POS tagger combinations on Hungarian text. In Proceedings of the Second International Joint Conference on Natural Language Processing, IJCNLP, Jeju Island, Korea, October 11-13, 2005, Companion Volume including Posters/Demos and tutorial abstracts, 191-196, <http://acl.ldc.upenn.edu/I/I05/I05-2033.pdf>, also in **Lecture Notes in Computer Science**, vol. 3651, Springer 2005, ISBN 3-540-29172-5
- 53) F. Mihelič, J. Gros, S. Dobrišek, J. Žibert, N. Pavešić. Spoken Language Resources at LUKS of the University of Ljubljana. In **International Journal of Speech Technology**, Vol. 6, Num. 3, 221–232, 2003, DOI:10.1023/A:1023462002932. **Kluwer Academic Publishers**. Manufactured in The Netherlands.
- 54) Milena Slavcheva. Some aspects of the morphological processing of Bulgarian. In Proceedings of the 10th Conference of the EACL: Workshop on Morphological Processing of Slavic Languages. April 13th, Budapest. 71-78, 2003.

Публикация (номер 31 в Списъка на публикациите)

Ludmila Dimitrova, Tomaž Erjavec, Nancy Ide, Heiki J. Kaalep, Cszaba Oravetz, Vladimir Petkevič, and Dan Tufis. Multext-East: Overview of the project. In Proceedings of the ALLC-ACH '98 Conference, Debrecen, Hungary, July 5-10 1998.

Цитирана в:

- 55) Radu Jon (2007): Metode de dezambiguizare semantica automata. Aplicatii pentru limbile engleza si romana. PhD Thesis. Academia Romana. //Word Sense Disambiguation Methods Applied to English and Romanian. ROMANIAN ACADEMY Research Institute for Artificial Intelligence.// Bucharest, May 2007. 158 pages.

Публикация (номер 7 в пълен списък на публикациите, номер 1 в списък на публикациите за конкурса)

Л. Димитрова, Е. Паскалева, И. Ненова (1981). Пакет програми за автоматично сегментиране на български текст. *Системи и управление*, брой 3, 19-25, 1981, София

Цитирана в:

- 56) Kiril Simov, Galia Angelova (1985). A Program Package for Automatical Segmentation of Bulgarian Wordforms. In Proceedings of the First Conference of Program Designers, July 11-12, 1985, Budapest, Ed. By A. Iványi, 19-23.

Публикация (номер 8 в пълен списък на публикациите, номер 2 в списък на публикациите за конкурса)

Е. Паскалева, Л. Димитрова, И. Ненова (1981). Автоматично сегментиране на български словоформи – описание на алгоритъма. *Автоматизирани системи и управление*, брой 4, 9-16, 1981, София

Цитирана в:

- 57) Kiril Simov, Galia Angelova (1985). A Program Package for Automatical Segmentation of Bulgarian Wordforms. In Proceedings of the First Conference of Program Designers, July 11-12, 1985, Budapest, Ed. By A. Iványi, 19-23.

Публикация (номер 27 в пълен списък на публикациите)

T. Erjavec, M. Monachini (Eds.). Contributors: R. Pavlov, L. Dimitrova, L. Sinapova, K. Simov, V. Petkevic, H.J. Kaalep, L. Tihanyi, D. Tufis, A.M. Barbu, T. Erjavec, P. Holozan. (1997). Specifications and Notation for Lexicon Encoding. COP Project 106 MULTEXT-East: Deliverable Report D1.1Final, Work Package WP1 - Task 1.1, Deliverable D1.1 F Final Report, 17 December 1997. Institute Jozef Stefan, Ljubljana, Slovenia, 167 pages; Versions 1-3 are available at <http://nl.ijs.si/ME/CD/docs/mte-d11f/index.html>.

Цитирана в:

- 58) Csaba Oravecz, Victor Nagy, and Károly Varasdi. Morphological Idiosyncrasy in Hungarian Multiword Expressions. In: Proceedings of the SLOVAKO conference. November 10-12, 2005. Slovak AS, Bratislava, 2005, 147-155. ISBN 80-224-0895-6.
- 59) Zoltán Alexin, János Csirik, Tibor Gyimóthy, Károly Bibok, Csaba Hatvani, Gábor Prószéky, László Tihanyi. (2003) Manually Annotated Hungarian Corpus. In Proceedings of the EACL'2003, 53-56.
- 60) Adam Przepiórkowski (2009). A Comparison of Two Morphosyntactic Tagsets of Polish. In: Koseska, Dimitrova, Roszko (Eds.2009), Representing Semantics in Digital Lexicography. Proceedings of the MONDILEX Fourth Open Workshop, 29 June - 1 July, 2009. SOW Publ. House, Warsaw, Poland. 138-144. ISBN 978-83-89191-87. www.mondilex.org/pro4.pdf
- 61) Markus Dickinson. (2005) Error detection and correction in annotated corpora. Ph.D. thesis, 2005, Ohio State University. <http://etd.ohiolink.edu> // http://etd.ohiolink.edu/view.cgi?acc_num=osu1123788552
- 62) Mirjam Sepesy Maucec and Zdravko Kacic (2007): Statistical Machine Translation from Slovenian to English. In: Journal of Computing and Information Technology - CIT 15, 2007, 1, 47-59. doi:10.2498/cit.1000760
- 63) TTL - Tokenizing, Tagging and Lemmatizing free running texts. (2008) Available at the: <http://www.clarin.eu/tools/ttl-tokenizing-tagging-and-lemmatizing-free-running-texts>
Developers: Radu Ion, Dan Tufis.
- 64) Božo Bekavac, Marko Tadić. (2007) Implementation of Croatian NERC system. In: *Annual Meeting of the ACL 2007. Proceedings of the Workshop on Balto-Slavonic Natural Language Processing: Information Extraction and Enabling Technologies*. SESSION: Information extraction. Prague, Czech Republic. 11-18. Publisher: Association for Computational Linguistics, Morristown, NJ, USA, 2007. <http://ufal.mff.cuni.cz/acl2007/>
- 65) Antoni Oliver, Marko Tadić (2004). Enlarging the Croatian Morphological Lexicon by Automatic Lexical Acquisition from Raw Corpora. In *Proceedings of the LREC'2004*.
- 66) Tadić, M. and S. Fulgosi (2003). Building the croatian morphological lexicon. In Proceedings of the 10th Conference of the EACL: Workshop on Morphological Processing of Slavic Languages. April 13th, Budapest. 41-46, 2003.
- 67) Antoni Oliver, Irene Castellón, Llufs Màrquez (2003). Automatic lexical acquisition from raw corpora: an application to Russian. In Proceedings of the 10th Conference of the EACL: Workshop on Morphological Processing of Slavic Languages. April 13th, Budapest. 17-24, 2003.
- 68) Adam Przepiórkowski, Marcin Woliński (2003). A flexemic tagset for Polish. In Proceedings of the 10th Conference of the EACL: Workshop on Morphological Processing of Slavic Languages. April 13th, Budapest. 33-40, 2003.
- 69) Cvetana Krstev, Duško Vitas (2011). An Aligned English-Serbian Corpus. In: (N. Tomović & J. Vujić, Eds.) ELLSIIR Proceedings (English Language and Literature Studies: Image, Identity, Reality), Volume I, Belgrade, 4-6 December 2009, 495-508, Faculty of Philology, University of Belgrade, ISBN 978-86-6153-005-0, 2011.

- 70) Jernej Vičič, Petr Homola (2010). Speeding up the Implementation Process of a Shallow Transfer Machine Translation System. In Proceedings of the 14th Annual Conference of the European Association for Machine Translation EAMT'2010, 27-28 May 2010, St Raphael, France. Also at: <http://www.mt-archive.info/EAMT-2010-Vicic.pdf>
- 71) Camelia Ignat. (2009) Improving Statistical Alignment and Translation Using Highly Multilingual Corpora. June 2009, Université de Strasbourg, France. 215 pp.
- 72) Behrang QasemiZadeh, Saeed Rahimi. Persian in MULTEXT-East Framework. In: Advances in Natural Language Processing, Proceedings of the 5th International Conference on NLP, FinTAL 2006, Turku, Finland, August 23-25, 2006. **Lecture Notes in Computer Science, 4139 Springer 2006**, 541-551, DOI: 10.1007/11816508_54. ISBN 3-540-37334-9
Also at: <http://www.springerlink.com/content/g23370g0w02m00t3>
- 73) Roszko, Danuta, Roszko, Roman. Morphosyntactic Specifications for Polish and Lithuanian. [Description of Morphosyntactic Markers for Polish and Lithuanian Nouns within MULTEXT-East Morphosyntactic Specifications (Version 3.0 May 10th, 2004)]
.In: Koseska, Dimitrova, Roszko (Eds.2009), Representing Semantics in Digital Lexicography. SOW Publ. House, Warsaw, Poland. 145-158. ISBN 978-83-89191-87. www.mondilex.org/pro4.pdf
- 74) Kotsyba, Natalia, Adam Radziszewski, Ivan Derzhanski, Integrating the Polish Language into MULTEXT-East Family: Morphosyntactic Specifications, Converter, Lexicon and Corpus. In: Proceedings of the MONDILEX, Fifth Open Workshop, October 14-15, 2009, Ljubljana, Slovene. 37-55. ISBN 978-961-264-012-5, ISSN 1581-9973. www.mondilex.org/pro5.pdf
- 75) Derzhanski, Kotsyba, 2009: Towards a Consistent Morphological Tagset for Slavic Languages: Extention of the MULTEXT-East Morphosyntactic Specifications for Byelorussian, Polish and Ukrainian. In: Metalanguage and Encoding scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, Bratislava, Slovak Republic, 15–16 April 2009. 36-47. Tribun EU Publ. House. ISBN 978-5-9900813-6-9. www.mondilex.org/pro3.pdf
- 76) Roszko, Roman. Morphosyntactic Specifications for Polish. Theoretical foundations. Description of morphosyntactic markers for Polish nouns within MULTEXT-East Morphosyntactic Specifications (Version 3.0 May 10th, 2004). In: Metalanguage and Encoding Scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, 15-16 April 2009, Bratislava, Slovak republic. 140–150. Tribun EU Publ. House. ISBN 978-80-7399-745-8. www.mondilex.org/pro3.pdf
- 77) Roszko, Roman. Description of morphosyntactic markers for Polish verbs within MULTEXT-East Morphosyntactic Specifications (Version 3.0 May 10th, 2004). In: Koseska, Dimitrova, Roszko (Eds.2009), Representing Semantics in Digital Lexicography. SOW Publ. House, Warsaw, Poland. 159-165. ISBN 978-83-89191-87-8. www.mondilex.org/pro4.pdf
- 78) Jakub Piskorski, Marcin Sydow, Dawid Weiss (2008): Exploring Linguistic Features for Web Spam Detection: A Preliminary Study. In: Linguistics Features, 2008, In: Proceedings of the 4th international workshop on Adversarial information retrieval on the web, SESSION: Text analysis. Beijing, China. 25-28. AIRWeb, Vol. 295. Publisher: ACM New York, NY, USA. ISBN: 978-1-60558-159-0 7 pages. www.users.pjwstk.edu.pl/~msyd/lingFeat08draft.pdf
- 79) Marco Tadić. (2007) Building the Croatian Dependency Treebank: the initial stages. In: Suvremena lingvistika 63, 2007, 85-92. ISSN 0586-0296
- 80) Kornél Markó, Robert Baud, Pierre Zweigenbaum, Lars Borin, Magnus Merkel, Stefan Schulz (2006): Towards a Multilingual Medical Lexicon. In: AMIA 2006 Symposium Proceedings. 534-538. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839525/>

- 81) Kornél Markó, Robert Baud, Pierre Zweigenbaum, Magnus Merkel, Maria Toporowska-Gronostaj, Dimitrios Kokkinakis, Stefan Schulz. (2006): Cross-Lingual Alignment of Medical Lexicons. In Pierre Zweigenbaum, Stefan Schulz, and Patrick Ruch, editors, LREC 2006 Workshop on Acquiring and Representing Multilingual, Specialized Lexicons: the Case of Biomedicine. Genova, Italy, 2006. ELDA. 5-8. [http://www.math.md/files/csjm/v17-n2/v17-n2-\(pp-160-178\).pdf](http://www.math.md/files/csjm/v17-n2/v17-n2-(pp-160-178).pdf)
- 82) Hristo Tanev, Jens Linge, Jakub Piskorski, Vanni Zavarella, Mijail Kabadjov, Martin Atkinson, Ralf Steinberger (2009): Exploiting Machine Learning Techniques to Build an Event Extraction System for Portuguese and Spanish. In: Linguamatica, Num. 2 - Dezembro 2009, 55-67. ISSN: 1647-0818. www.linguamatica.com/index.php/linguamatica/article/view/37/37
- 83) Alhazov, E. Boian, S. Cojocar, Yu. Rogozhin. Modelling Inflections in Romanian Language by P Systems with String Replication. Preproc. of the Tenth Workshop on Membrane Computing (WMC10), Curtea de Arges, 2009, 116-128.
- 84) Artiom Alhazov, Elena Boian, Svetlana Cojocar, Yurii Rogozhin. Modelling Inflections in Romanian Language by P Systems with String Replication. Computer Science Journal of Moldova, vol.17, no.2 (50), 2009. 160-178. [http://www.math.md/files/csjm/v17-n2/v17-n2-\(pp-160-178\).pdf](http://www.math.md/files/csjm/v17-n2/v17-n2-(pp-160-178).pdf)
- 85) Adam Kilgarriff, Simon Krek (2009): Slovene tagset.//Slovenski nabor oznak.// Adapted from Multext-East website <http://nl.ijs.si/ME/V3/msd/html>. April 2009. <http://www.sketchengine.co.uk/tagsets/slovene.html>
- 86) Daniel Jurafsky and James H. Martin. Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition. 1024 pages, **Publisher: Prentice Hall**; 2 edition (May 26, 2008), **ISBN-10:** 0131873210, **ISBN-13:** 978-0131873216
- 87) Elena Paskaleva (2007): Balkan South-East Corpora Aligned to English. 7 pages.http://www.coli.uni-saarland.de/~tania/block-kurs/web-sofia2007/_BSEC_Elena.pdf
- 88) RH Baud, M Nyström, L Borin, R Evans, S Schulz, P Zweigenbaum (2005). Interchanging Lexical Information for a Multilingual Dictionary. In: AMIA 2005 Symposium Proceedings. 31 – 35. 2005. http://spraakdata.gu.se/svelb/pblctns/Baud_AMIA2005.pdf
- 89) Nenadić, G. Creating digital language resources. In Pregled nacionalnog centra za digitalizaciju. Num. 5, 191-30. Matematički fakultet, Beograd. 2004. <http://eprints.rclis.org/handle/10760/11860>

Цитирания на разработки, публикувани на Интернет-страницата на международния проект COP Project 106 MULTTEXT-East (<http://nl.ijs.si/ME>), в

- 90) Cvetana Krstev, Ranka Stanković, Duško Vitas. (2010) A Description of Morphological Features of Serbian: a Revision using Feature System Declaration. In: (Nicoletta Calcolari et al. Eds. 2010) Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010, Valetta, Malta, May 2010, 816-819, 2010. ISBN 2-9517408-6-7.
- 91) Špela Arhar, Nina Ledinek (2008). Oblikoskladenjske oznake JOS: revizija in nadgradnja nabora oznak za avtomatsko oblikoskladenjsko oznacevanje slovenščine. In Proceedings of the Sixth Language Technologies Conference, October 16th-17th, 2008 within the 11th International Multiconference Information Society - IS 2008, volume C. Ljubljana, Slovene. 54-59. Publ. House Informacijska družba. ISSN 1581-9973, ISBN 978-961-264-006-4.

- 92) Jernej Vičič. (2008) Rapid development of data for shallow transfer RBMT translation systems for highly inflective languages. In: Proceedings of the Sixth Language Technologies Conference, October 16th-17th, 2008 within the 11th International Multiconference Information Society - IS 2008, volume C. Ljubljana, Slovene. 98-103. Publ. House Informacijska družba. ISSN 1581-9973, ISBN 978-961-264-006-4.
- 93) Victoria Bobicev. (2007) METODE ȘI ALGORITMI STATISTICI DE PROCESARE A TEXTELOR - Statistica - PhD Thesis 2007. Universitatea Tehnică a Moldovei Cu titlu de manuscris CZU: 004.912
http://www.cnaa.acad.md/files/theses/2007/6478/viorica_bobicev-thesis.pdf.
- 94) Jaroslava Hlaváčová. (2005) Orwell's 1984 – Playng with Czech and Slovak Versions. In: Proceedings of the SLOVKO conference. November 10-12, 2005. Slovak AS, Bratislava, 2005, 116-123. ISBN 80-224-0895-6.
- 95) В.А Широков и др. (2005): Корпусна лінгвістика. Київ, Видавництво „Довіра”, 2005. ISBN 966-507-189-0.
- 96) A. Rosen. (2005) In Search of the Best Method for Sentence Alignment in Parallel Texts. In: Proceedings of the SLOVKO conference. November 10-12, 2005. Slovak AS, Bratislava, 2005, 174-185. ISBN 80-224-0895-6.
- 97) Karel Pala and Aleš Horák (2009). Multilingual Features of Complex Valency Frames. In Proceedings of the Recent Advances in Intelligent Information Systems Conference, 2009, 41-49. ISBN 978-83-60434-59-8

Цитирания на изданието „East meets West: A Compendium of Multilingual Language Resources (Erjavec, T., Lawson, A., Romary, L., Eds. 1998), Mannheim: TELRI-MULTEXT EAST Edition (CD-ROM)” (двухтомн CD-ROM), в:

- 98) Cvetana Krstev, Duško Vitas (2011). An Aligned English-Serbian Corpus. In: (N. Tomović & J. Vujić, Eds.) ELLSIIR Proceedings (English Language and Literature Studies: Image, Identity, Reality), Volume I, Belgrade, 4-6 December 2009. 495-508, Faculty of Philology, University of Belgrade, 2011, ISBN 978-86-6153-005-0.
- 99) Éric Laporte, Duško Vitas, Cvetana Krstev (2006). Preparation and exploitation of bilingual texts. In: Lux Coreana, No. 1, 110-132, Han-Seine, 2006.
- 100) Krstev, C., Vitas, D., Stanković, R., Obradović, I., Pavlović-Lažetić, G. (2004). Combining Heterogeneous Lexical Resources. In Proceedings of LREC'04, vol. 4, pp. 1103-1106, Lisbon: ARTIPOL.
- 101) Zoltán Alexin, János Csirik, Tibor Gyimóthy, Károly Bibok, Csaba Hatvani, Gábor Prószéky, László Tihanyi. (2003) Manually Annotated Hungarian Corpus. In Proceedings of the EACL'2003, 53-56.
- 102) Tamás Horváth, Zoltán Alexin, Tibor Gyimóthy, Stefan Wrobel (1999). Application of Different Learning Methods to Hungarian Part-of-speech Tagging. In: (Saso Dzeroski, Peter A. Flach Eds.) Proceedings of the 9th International Workshop, Inductive Logic Programming (ILP-99), Bled, Slovenia, June 24-27, 1999. 128-139, **Springer, Lecture Notes in Computer Science** vol. 1634/1999, ISBN 3-540-66109-3, DOI: 10.1007/3-540-48751-4_13, DBLP, <http://dblp.uni-trier.de>