

Lucrările conferinței
*Resurse lingvistice și instrumente pentru
prelucrarea limbii române*
București, 6–7 mai 2010

Volum apărut cu sprijinul Ministerului Educației, Cercetării și
Tineretului, prin Centrul Național de Management Programe
(CNMP), proiectul eDTLR – Dicționarul Tezaur al Limbii
Române în format electronic (contract 91–013/18.09.2007)

ISSN 1843-911X

Lucrările conferinței
*Resurse lingvistice și instrumente pentru
prelucrarea limbii române*
București, 6-7 mai 2010

Editori:
Adrian Iftene
Horia-Nicolai Teodorescu
Dan Cristea
Dan Tufiș

Organizatori:
Facultatea de Informatică,
Universitatea „Alexandru Ioan Cuza” Iași

Institutul de Cercetări pentru Inteligență Artificială
Academia Română, București

Institutul de Informatică Teoretică
Academia Română, Filiala Iași

Muzeul Național al Literaturii Române
Filiala București

Intelligentics, Cluj-Napoca

Editura Universității „Alexandru Ioan Cuza” Iași

COMITETUL DE PROGRAM:

Vasile Apopei, Institutul de Informatică Teoretică, A.R., Iași
Verginica Barbu Mititelu, Institutul de Cercetări în Inteligență Artificială, A.R., București
Corneliu Burileanu, Facultatea de Electronică, Universitatea Politehnică București și Institutul de Cercetări în Inteligență Artificială, A.R., București
Monica Busuioc, Institutul de Lingvistică „Iorgu Iordan - Al. Rosetti”, A.R., București
Alexandru Ceaușu, Institutul de Cercetări în Inteligență Artificială, A.R., București
Lucian Chișu, Muzeul Național al Literaturii Române, București
Constantin Ciubotaru, Institutul de Matematică și Informatică, Academia de Științe a Moldovei, Chișinău
Svetlana Cojocaru, Institutul de Matematică și Informatică, Academia de Științe a Moldovei, Chișinău
Alexandra Cornilescu, Facultatea de Litere, Universitatea București
Dan Cristea, Facultatea de Informatică, Universitatea „Al. I. Cuza” și Institutul de Informatică Teoretică, A.R., Iași
Nicolae Curteanu, Institutul de Informatică Teoretică, A.R., Iași
Gabriela Czibula, Universitatea „Babeș-Bolyai”, Cluj-Napoca
Cristina Florescu, Institutul de Filologie Română „Al. Philippide”, A.R., Iași
Corina Forăscu, Facultatea de Informatică, Universitatea „Al. I. Cuza”, Iași și ICIA, A.R., București
Maria Georgescu, ISSCO / TIM, ETI, Université de Genève
Gabriela Haja, Institutul de Filologie Română „Al. Philippide”, A.R., Iași
Catalina Hallett, Open University, London
Sanda Harabagiu, Department of Computer Science University of Texas at Dallas
Maria Husarciuc, Centrul de Studii Biblico-Filologice „Monumenta linguae Dacoromanorum”, Universitatea „Al. I. Cuza”, Iași
Adrian Iftene, Facultatea de Informatica, Universitatea „Al. I. Cuza” Iași
Diana Inkpen, Université d’Ottawa
Radu Ion, Institutul de Cercetări în Inteligență Artificială, A.R., București
Elena Irimia, Institutul de Cercetări în Inteligență Artificială, A.R., București
Doina Jitcă, Institutul de Informatică Teoretică, A.R., Iași
Daniel Marcu, University of Southern California
Rodica Marian, Institutul de Lingvistică și Istorie Literară „Sextil Pușcariu”, A.R., Cluj-Napoca
Dan Moldovan, University of Texas at Dallas
Vivi Năstase, EML Research, Heidelberg
Constantin Orasan, University of Wolverhampton
Oana Postolache, ISI - University of Southern California, Los Angeles
Irina Prodanoff, ILC-Pisa și Università di Pavia
Marius Pașca, Google Research
Georgiana Pușcașu, University of Wolverhampton
Vasile Rus, University of Memphis
Vitalie Scurtu, Intelligentic, Cluj-Napoca
Violeta Seretan, Departament de Linguistique, Université de Geneve
Mihai Surdeanu, University of Stanford
Valentin Tablan, University of Sheffield
Amalia Todirașcu, Université Marc Bloch, Strasbourg
Horia-Nicolai Teodorescu, Institutul de Informatică Teoretică, A.R. și Universitatea Tehnică, Iași
Dan Tufiș, Institutul de Cercetări în Inteligență Artificială, A.R., București
Ioana Vintilă-Rădulescu, Institutul de Lingvistică „Iorgu Iordan - Al. Rosetti”, A.R., București
Adriana Vlad, Facultatea de Electronică, Universitatea Politehnică București și Institutul de Cercetări în Inteligență Artificială, A.R., București

COMITETUL DE ORGANIZARE:

Vlad Alexa, FII-UAIC (vlad.alex@infoiasi.ro)

Lucian Chișu, MNLR (lucianchisu@ gmail.com)

Ioana Corăci, (ioana.coraci@yahoo.com)

Marius Corăci, Intelligentics (marius@ intelligentics.ro)

Dan Cristea, FII-UAIC și IIT-AR (dcristea@ info.uaic.ro)

Florin-Tudor Cristea, FII-UAIC (florin.cristea@ infoiasi.ro)

Lucian Gădioi, FII-UAIC (lucian.gadioi@info.uaic.ro)

Adrian Iftene, FII-UAIC (adiftene@info.uaic.ro)

Angela Ioniță, ICIA-AR (aionita@racai.ro)

Petru-Adrian Istrimschi, MediaEC-UAIC (petru.istrimschi@infoiasi.ro)

Mihai-Alex Moruz, FII-UAIC și IIT-AR (mmoruz@info.uaic.ro)

Marius Răschip, FII-UAIC (mraschip@info.uaic.ro)

Horia-Nicolai Teodorescu, IIT-AR și Universitatea Tehnică, Iași (hteodor@etc.tuiasi.ro)

Diana-Maria Trandabăț, FII-UAIC și IIT-AR (dtrandabat@ info.uaic.ro)

Dan Tufiș, ICIA-AR (tufis@racai.ro)

Eugenia Țărălungă, MNLR (eugenia12@gmail.com)

Cuprins

Cuvânt înainte	9
----------------------	---

Capitol I – Corpusuri vocale și prelucrarea vorbirii

Metodologie pentru constituirea și analiza unui corpus adnotat de semnale vocale - cazul SROL	13
<i>Horia-Nicolai Teodorescu</i>	
Tehnici de identificare a zonelor vocalice în secvențe rostite în limba română	23
<i>Marius Zbancioc, Horia-Nicolai Teodorescu și Monica Feraru</i>	
Aspecte metodologice de organizare a datelor și de analiză statistică a vocilor emoționale	35
<i>Horia-Nicolai Teodorescu, Ioan Păvăloi și Monica Feraru</i>	
Program de editare de contururi intonaționale în limba română bazat pe ierarhii de forme prosodice funcționale	45
<i>Vasile Apopei, Doina Jitcă și Otilia Păduraru</i>	
Corpus pentru gnatofonie: protocol, metodologie, adnotare	51
<i>Horia-Nicolai Teodorescu și Alina Untu</i>	
Dispozitiv electronic de preprocesare în regim paralel a spectrului vocal	61
<i>Mircea Hulea și Alina Untu</i>	

Capitol II – Platforme, dicționare și corpusuri pentru prelucrarea textelor

Resurse lingvistice românești în flux continuu	73
<i>Dan Cristea</i>	
Comunicarea electronică și problemele noastre ortografice – fără soluții?	81
<i>Lucian Chișu</i>	
Când migrăm la diacriticele corecte?	89
<i>Bogdan Stăncescu</i>	
Construcția automată de corpusuri multilinguale	103
<i>Tiberiu Boroș, Dan Tufiș și Alexandru Ceaușu</i>	
Parsarea comparativă a dicționarelor-tezaur românești, franceze și germane	113
<i>Neculai Curteanu, Alex Moruz și Diana Trandabăț</i>	
Realizarea unui treebank românesc	123
<i>Cenel-Augusto Perez</i>	
Monitorizarea presei în cadrul proiectului Neorom	131
<i>Ana-Maria Barbu</i>	
Emoții în cuvinte: elaborarea unei resurse multilingve	141
<i>Victoria Bobicev, Victoria Maxim, Tatiana Prodan, Natalia Burciu și Victoria Angheluș</i>	

Capitol III – Aplicații ale tehnologiilor lingvistice textuale

Sistem întrebare-răspuns antrenabil pentru limba română	153
<i>Dan Ștefănescu, Radu Ion, Alexandru Ceaușu, Dan Tufiș, Elena Irimia și Verginica Barbu-Mititelu</i>	
Un sistem de traducere automată română-franceză	165
<i>Mirabela Navlea și Amalia Todirașcu</i>	

Servicii Web interoperabile și multilinguale	175
<i>Radu Ion, Alexandru Ceaușu, Dan Ștefănescu și Dan Tufiș</i>	
Tipare hiponimice pentru limba română	185
<i>Verginica Barbu-Mititelu</i>	
Mecanismele generative ale morfologiei derivaționale	195
<i>Mircea Petic</i>	
Dezvoltarea unui parser de roluri semantice pentru limba română	203
<i>Diana Trandabăț și Dan Cristea</i>	
Supravegherea pe Internet: părerea consumatorilor despre anumite produse sau despre anumite evenimente	213
<i>Adrian Iftene, Alina-Elena Mihaila, George-Alexandru Vlad și Geta Stancu</i>	
Folosirea verbelor pentru determinarea inferențelor textuale	223
<i>Mihai Alex Moruz</i>	
Summaries / Abstracts	233
Index de autori	244