



CESAR

Central and South-East European Resources
Project no. 271022

Deliverable D5.4

Dissemination material – M18

Version No. 1.2
31/07/2011

Document Information

Deliverable number:	5.4b
Deliverable title:	Dissemination material
Due date of deliverable:	31/07/2012
Actual submission date of deliverable:	31/07/2012
Main Author(s):	Marko Tadic (FFZG)
Participants:	
Internal reviewer:	Tibor Pintér (HASRIL)
Workpackage:	WP5
Workpackage title:	Outreach, awareness and sustainability
Workpackage leader:	FFZG
Dissemination Level:	Public
Version:	1.2
Keywords:	dissemination

History of Versions

Version	Date	Status	Name of the Author (Partner)	Contributions	Description/ Approval Level
0.2	31/07/2012		Tibor Pintér (HASRIL)	proofreading	
0.1	28/07/2012		Marko Tadic (FFZG)	preparation	

EXECUTIVE SUMMARY

The deliverable describes the overall activity on dissemination efforts made from M7 to M18.

Table of Contents

1. Types of dissemination materials.....	4
2. Dissemination materials.....	5
2.1 Webpage	6
2.2 Flyers	7
2.3 Poster.....	9
2.4 T-shirts.....	11
2.5 Cups, bags, pens, USB-sticks, calendars, roll-up.....	12
2.6 Presentations at conferences and events.....	16
2.7 Publications.....	20
2.8 Language White Papers	22

1. Types of dissemination materials

Within the second and third six-month periods dissemination activities in CESAR project were done through several channels on different occasions.

The channels of dissemination covered mainly traditional dissemination means (flyers, posters, public web-page, scientific / professional conferences / events presences / lectures / presentations / demonstrations, press releases

The channels used for dissemination purposes were:

- webpage – external webpage redesigned
- flyer – used for presenting the general information about the project
- coordinated META-NET flyer in all CESAR languages
- posters – posters presented on conferences and road-show events
- presentations – presentations presented on conferences and meetings
- video recording of CESAR partner’s presentations at different events
- coordinated META-NET Language Whitepapers for CESAR languages
- publications by CESAR authors at conferences, proceedings and journals

2. Dissemination materials

Following the dissemination strategy set down in D5.1, a general visual identity of the CESAR project was produced. Beside the general visual identity elements (usage of colours, typefaces, etc.), it included:

- official logo: used in all dissemination materials
 - <http://cesar.nytud.hu/documents/archive/logo>
- official PPT template: used by project partners for presentations at conferences and other events
 - <http://cesar.nytud.hu/documents/archive/presentation-templates>
 - as PPT, PPTX and Beamer presentation template
- QR code: for making simple and efficient shortcut to the URL of the CESAR project website
 - <http://cesar.nytud.hu/documents/archive/qr-code-for-cesar-url>
- deliverables template: used for producing deliverables
 - <http://cesar.nytud.hu/documents/archive/deliverables-template>



Figure 1: Official CESAR logo



Figure 2: CESAR QR code

2.1 Webpage

The external webpage serves for the presentation of the project and it is made not just for LT community, but also for the wider public giving insight into the details about the project, list of the project partners, current news, list of related events, overall workplan, list of deliverables and specialised information packages for media and researchers. The address www.cesar-project.net is not pointing to URL <http://www.meta-net.eu/projects/cesar>, but it points to a stand-alone CESAR project web site. However, the webpages show the clear linkage of CESAR to the META-NET community since the webpages underwent the major redesign and they now follow the joint META-NET visual identity elements adding some of CESAR's own features.



The screenshot shows the CESAR external webpage. The header features the CESAR logo (CENTRAL AND SOUTH-EAST EUROPEAN RESOURCES) and the META logo (A Network of Excellence forging the Multilingual Europe Technology Alliance). Navigation links include HOME, ABOUT, EVENTS, WORKPLAN, DELIVERABLES, LINKS, PARTNERS, and INFORMATION FOR. A search bar and a 'Log in' button are also present. The main content area is titled 'Project' and 'Latest news'. The featured news item is 'CESAR Road show in Poland, September 27–28th, 2012'. The text of the news item reads: 'I am very pleased to invite you to join the international event "Human Language Technology Days 2012" to be held on September 27-28, 2012 in Warsaw. The event is organised by the Institute of Computer Science, Polish Academy of Sciences and the University of Łódź in the frame of ICT-PSP project CESAR (Central and South-East European Resources, part of META-NET) and is collocated with ICT Proposers' Day 2012. Language Technology Days 2012 intends to promote knowledge about language technology, its potential, but also possible threats accompanied with its development, by gathering experts who will try to answer important questions on the future of language and language processing in a globalized digital information society. How does massive digitization of information, knowledge and everyday communication affect our language? Will our language change or even disappear? Will Internet always be divided by the languages of its users? Another important topic we intend to cover is a Road Show of European language technology, aiming at presenting state-of-the-art, directions and visions of development of language resources and tools for the common scientific and commercial market. Particular attention will be drawn to language technology for Polish by presenting new advances in the field and their application in administration and business as well as involvement of open source and research community in the process of development of language resources. I hope that you will find our proposal worth of notice and you will attend the conference and the follow up discussion (please see the preliminary programme of the event in general outline). I wish you a pleasant stay in Warsaw.'

At the bottom left, there is a logo for 'cip' (Competitiveness and Innovation Programme) and the European Union flag, with the text: 'The project has received funding from the Competitiveness and Innovation Framework Programme under Grant Agreement n° 271022.'

Figure 3: CESAR external webpage

2.2 Flyers

In the coordination with META-NET, the localisation of general META-NET flyer was performed before the Budapest META-FORUM 2011. The localisation was done for all CESAR languages: Bulgarian, Croatian, Hungarian, Polish, Serbian and Slovak following the prescribed procedure.

However, the separate CESAR-specific flyer was also prepared and produced presenting the general information about the project.

CESAR CENTRAL AND SOUTH-EAST EUROPEAN RESOURCES

CESAR

Project partners

- HASRIL:** Research Institute for Linguistics, Hungarian Academy of Sciences
- BME-TMIT:** Department of Telecommunications and Media Informatics, Budapest University for Technology and Economics
- FFZG:** University of Zagreb, Faculty of Humanities and Social Sciences
- IPIPAN:** Institute of Computer Science, Polish Academy of Sciences
- ULODZ:** University of Łódź
- UBG:** Faculty of Mathematics, University of Belgrade
- IPUP:** Institute Mihajlo Pupin, Belgrade
- IBL:** Institute for Bulgarian Language Prof. Lyubomir Andreychin, Bulgarian Academy of Sciences
- LSIL:** Ľudovít Štúr Institute of Linguistics, Slovak Academy of Sciences

Contact
 Tamás Váradi (project coordinator)
 Research Institute for Linguistics
 Hungarian Academy of Sciences
 Benczur u 33
 1068 Budapest
 Hungary
 Tel: +36 1 342 9372 ext. 6010
 Fax: +36 1 322 9297
 E-mail: varadi@nytud.hu

The CESAR project has received funding from the ICT Policy Support Programme (ICT PSP), grant agreement no 271022.

META-NET

CESAR

CENTRAL AND SOUTH-EAST EUROPEAN RESOURCES

Project duration
 2011-02-01 – 2013-01-31

www.cesar-project.net

Figure 4: CESAR flyer external pages



CESAR
CENTRAL AND SOUTH-EAST
EUROPEAN RESOURCES

Main goal

CESAR project is oriented towards the stimulation of ICT-based cross-lingual communication, collaboration and participation and thereby contribute to the creation of a pan-European digital single market. ICT research has started in Central- and Southeast-Europe with a lag behind Western European countries and has experienced rather short-term and fortuitous cooperation on the level of national research. Language technology has emerged in the respective participating countries autonomously, i.e., with national support both in the academic and in the private sector. As a result, the resources developed often reflect the isolated circumstances of their creation, and still often lack standardisation.



European dimension

The CESAR project is fully committed to providing synergy with relevant strategies and activities at the European, regional and

www.cesar-project.net

national level. The partners are key players in the language technology scene for their respective languages and have built extensive links with LRT centres across the whole of Europe but especially in the geographical region that the consortium covers. At the same time, they have served and will continue to serve as catalytic force in the language technology scene of their countries. Typically, they have led national consortia involving the other major partners in the countries and have an established track record of their capability to mobilise their national partners.

The European and global dimension of the CESAR project results will be visible through the META-NET Multilingual European Technology Alliance and its sharing facility META-SHARE. CESAR project resources and tools will also be accessible through other types of linguistic infrastructures as well as partners' individual sites.

Target audience

Target user community of language resources and services embraces all stakeholders at the modern digital market: everyday end-users, professional end-users (business, administration, media, education, libraries, etc.) and expertise holders (researchers, industrialists, policy makers, etc). Our concern is a careful investigation of the needs of various types of users – from individual users to large multinational organisations – from the perspective of the current status as well as from the near future prospects.

CESAR is a part of www.meta-net.eu



Languages and technology involved

The project will make available a comprehensive set of language resources and tools covering the following languages:

- Bulgarian
- Croatian
- Hungarian
- Polish
- Serbian
- Slovak.

Resources will include interoperable mono- and multilingual:

- speech databases
- corpora
- dictionaries
- wordnets.

Relevant language technology processing tools will include:

- tokenisers
- lemmatisers
- taggers
- chunkers and parsers.

Depending on local capacities and efficiency considerations, the resources and tools may be either made available as web services through partners' sites or could be contributed through the META-SHARE digital exchange facility.

META-SHARE

Figure 5: CESAR flyer internal pages

Beside the CESAR overall flyer, an introductory META-NET flyer production was coordinated in all CESAR languages.

2.3 Poster

The overall CESAR introductory poster was prepared, produced and distributed to partners.

The poster features a red header with the text 'META-NET' and the CESAR logo below it. The main content is organized into several text boxes and a map of Europe. The 'Main goal' section describes the project's focus on stimulating ICT-based cross-lingual communication. The 'European dimension' section highlights the project's commitment to providing synergy with relevant strategies at the European, regional, and national levels. The 'Languages and technology involved' section lists the languages covered (Bulgarian, Croatian, Hungarian, Polish, Serbian, Slovak) and the types of resources and tools provided. The 'Individual goals' section lists seven specific objectives, including describing the national landscape, contributing to a pan-European digital resource exchange facility, and bridging the technological gap. The 'Target audience' section identifies the user community of language resources and services. The 'Consortium partners' section lists the participating institutions from Hungary, Poland, and Slovakia. The 'Project duration' section provides the dates 2014-02-01 to 2015-01-31. A QR code and the website 'www.meta-net.eu' are also present.

Meta-NET

CESAR
CENTRAL AND SOUTH-EAST EUROPEAN RESOURCES

Main goal
CESAR project is oriented towards the stimulation of ICT-based cross-lingual communication, collaboration and participation and thereby contribute to the creation of a pan-European digital single market. ICT research has started in Central- and Southeast-Europe with a lag behind Western European countries and has experienced neither short-term and fortuitous cooperation on the level of national research. Language technology has emerged in the respective participating countries autonomously, i.e., with national support both in the academic and in the private sector. As a result, the resources developed often reflect the isolated circumstances of their creation, and still often lack standardisation.

European dimension
The CESAR project is fully committed to providing synergy with relevant strategies and activities at the European, regional and national level. The partners are key players in the language technology scene for their respective languages and have built extensive links with LRT centres across the whole of Europe but especially in the geographical region that the consortium covers. At the same time, they have served and will continue to serve as catalytic force in the language technology scene of their countries. Typically, they have led national consortia involving the other major partners in the countries and have an established track record of their capability to mobilise their national partners. The European and global dimension of the CESAR project results will be visible through the META-NET Multilingual European Technology Alliance and its sharing facility META-SHARE. CESAR project resources and tools will also be accessible through other types of linguistic infrastructures as well as partners' individual sites.

Languages and technology involved
The project will make available a comprehensive set of language resources and tools covering the following languages:
• Bulgarian • Polish
• Croatian • Serbian
• Hungarian • Slovak.
Resources will include interoperable mono- and multilingual:
• speech databases
• corpora
• dictionaries and wordnets.
Relevant language technology processing tools will include:
• tokenisers • taggers
• lemmatisers • chunkers and parsers.
Depending on local capacities and efficiency considerations, the resources may be either made available as web services at partners' site or contributed to the META-SHARE digital exchange facility.

Individual goals

1. Description of the national landscape in the terms of:
 - a. language use;
 - b. language-savvy products and services;
 - c. language technologies and resources;
 - d. main actors (research, industry, government, society);
 - e. public policies and programmes;
 - f. prevailing standards and practices;
 - g. current level of development (main drivers and roadblocks);
 synthesized this in a simple, clear, standardized format.
2. Contribute to a pan-European digital resource exchange facility by
 - a. collecting resources;
 - b. documenting, linking and upgrading.
3. Collaborate with other partner projects from META-NET as well as FlaReNET and CLARIN, to ensure consistent approaches, practices and standards aimed at enabling accessibility of, easier access to and reuse of quality language resources.
4. Help build and operate broad, non-commercial, community-driven, inter-connected repositories, exchange facilities etc. that can be used by language researchers, developers and professionals.
5. Mobilise national and regional actors, public bodies and funding agencies by
 - a. raising awareness;
 - b. organizing meetings and focused events.
6. Reinvalidate cooperation between key technology partners in the region, building on previous collaboration in TELRI, MULTITEXT-EAST and other projects.
7. Bridge the technological gap between this region and the other parts of Europe by filling obvious and important gaps in language resources and tools infrastructure.

Target audience
Target user community of language resources and services embraces all stakeholders at the modern digital market: everyday end-users, professional end-users (business, administration, media, education, libraries, etc.) and expertise holders (researchers, industrialists, policy makers, etc). Our concern is a careful investigation of the needs of various types of users – from individual users to large multinational organisations – from the perspective of the current status as well as from the near future prospects.

Consortium partners

- ILTEK: Research Institute for Linguistics, Hungarian Academy of Sciences
- IAIK: Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics
- ILP: Faculty of Humanities and Social Sciences, University of Zagreb
- ICIS: Institute of Computer Science, Polish Academy of Sciences
- UL: University of Lodz
- FM: Faculty of Mathematics, University of Belgrade
- ILP: Institute for Bulgarian Language Prof. Lyubomir Andreychik, Bulgarian Academy of Sciences
- ILU: Institute for Linguistics, Slovak Academy of Sciences

Project duration
2014-02-01 – 2015-01-31

CESAR is a part of
www.meta-net.eu

www.cesar-project.net

Contact
Tünde Vrárdi (co-ordinator) • Research Institute for Linguistics, Hungarian Academy of Sciences • E: +36 1214630 • F: +36 13225297 • E: vrad@nyelvi.hu • W: http://www.cesar-project.net

The research within the project CESAR leading to these results has received funding from the ICT Policy Support Programme (ICT PSP), grant agreement no 271022.

Figure 6: CESAR overall poster in English

Българска академия на науките, Институт за български език
участва в

CESAR

CENTRAL AND SOUTH-EAST EUROPEAN RESOURCES

Основна цел
Проектът CESAR се стреми да подобри многоезиковата комуникация с помощта на информационните технологии, както и да подкрепи сътрудничеството и участието в тази сфера, за да допринесе за развитието на Паневропейския дигитален пазар.
Централни и Източни Европейски гранични езиково-лингвистични центри за информационните и комуникационните технологии в сътрудничество с партньорите от Западна Европа, Езиковите ресурси и технологии в страните – участници в проекта, се реализират самостоятелно с поддръжка и академична, и в частния сектор, но предимно на национално равнище. Запознати ползватели от тези езикови ресурси често отразяват ограниченията на използването, при които да създадени както и липсата на стандартизация.

Конкретни цели
1. Характеристика на националните езиково-лингвистични ресурси с оглед на:
а. употребата на езика;
б. специализираните езикови продукти и услуги;
в. езиковите технологии и ресурси;
г. основните участници в сектора (научноизследователски организации, индустрия, правителство, обществени организации);
д. държавните политики и практики;
е. най-често използваните стандарти и практики;
ж. актуалното развитие на сектора (основни стимули, индикатори и трудности).
Резултатите от изследването ще бъдат обобщени в достъпен и стандартизиран формат.
2. Пренос към Паневропейския езиково-лингвистичен ресурс чрез:
а. проучване, събиране, създаване и разпространение на съществуващите езикови ресурси и технологии;
б. създаване на документация, съхраняване и нагледване на ресурсите, за да отговарят на действащи стандарти и ръководства.
3. Сътрудничество с партньорски проекти, като META-NET – Multilingual European Technology Alliance, F4U4NET и CLARIN, за да се осигурят поддръжка, практики и стандарти с цел по-лесен достъп до и употреба на качествени езикови ресурси и технологии.
4. Изградена на достъпен, некоммерчески, обществено полезен и свързани помежду си връзки на езикови ресурси и технологии, които ще бъдат използвани от изследователите на езикови аспекти в областта на езиковите технологии.
5. Мобилизиране на националните и регионалните участници в сектора, обществените институции и финансиращите агенции чрез:
а. по-добра информираност за спецификата на изследванията в областта;
б. организиране на срещи и други информационни събития.

Езикови ресурси и технологии
В рамките на проекта ще бъдат направени достъпни множество езикови ресурси и програми за текста обработка, които обхващат следните езика:
• Български
• Полски
• Словаки
• Сръбски
• Унгарски
• Хърватски.
Езиковите ресурси ще включват едно- и многоезикови:
• бази данни от реч
• корпузи
• речници и лексикално-семантични мрежи (wordnets)
Програмите за езикова обработка ще включват, но не само да се ограничават до:
• токанематори (разделяне на текста на единички)
• лексикематори (определяне на основна форма)
• тагери (определяне на граматични характеристики)
• членери (определяне на кратки форми)
• парсери (определяне на фраза структура).
Езиковите ресурси ще бъдат достъпни или като уеб страници, или като част от услугите за дигитален обмен на META-SHARE.

Бъдещи потребители
Потребители на езиковите ресурси и услуги по проекта са всички групи от свързаните дигитален пазар хората, които използват езикови справочници и програми за обработка на езика в академичното си и в професионалните си занимания (бизнес, администрация, медиите, образованието, библиотечните системи и други), както и агентите (изследователи, мениджъри, политими и други). Подробно е идентифицирано съществуващите на различни групи от потребители – от индивидуалните до големите многонационални организации – с оглед както на актуалния им статус и потребности, така и на бъдещите им планове и перспективи.

Партньори по проекта
IMMAGI Research Institute for Linguistics, Hungarian Academy of Science
IME-IMM Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics
IKM Faculty of Humanities and Social Sciences, University of Zagreb
IPRIM Institute of Computer Science, Polish Academy of Sciences
IUBS University of Ljubljana
IMM Faculty of Mathematics, University of Belgrade
IMM Institute IMbaj Papp
IMM Институт за български език "Проф. Любомир Андрейков", Българска академия на науките
IMM Institut für Linguistik, Slovak Academy of Science

Срок на проекта
2019-02-01 – 2023-01-31

За контакти
Тамара Варади (координатор) • Институт за лингвистика, Унгарска академия на науките • тел: +36 1 321 4530 • факс: +36 1 322 8267 • e-mail: tamara@nyud.hu • интернет адрес: <http://www.cesarproject.net>

ICTPSP Проектът CESAR е получил финансиране от ICT Policy Support Programme (ICT PSP) договор N 271022

Figure 7: CESAR overall poster in Bulgarian

Beside that overall poster, the partners also produced a number of posters that were covering different issues or tasks that had to be covered for CESAR languages. These posters are accessible on the internal webpage at the link <http://cesar.nyud.hu/documents/archive/posters>.

2.4 T-shirts

In a number of other projects a special, project-related t-shirts have proven to be a very efficient dissemination material since they often transfer message outside of the narrow LT community. For that reasons, we have produced CESAR t-shirts in two background colours (black, white) and four sizes (S, M, L, XL). They were successfully distributed at the major events in the field starting from the November 2011.

In the front of t-shirt CESAR logo is presented, while on the back a META-NET is applied. The URL of project webpages is printed on the edge of the left sleeve.

The first run of t-shirts was produced in 100 copies in each background colour and most of them were distributed until M17. This is pointing us to the need to produce more t-shirts for the coming events.



Figure 8: CESAR black background t-shirts

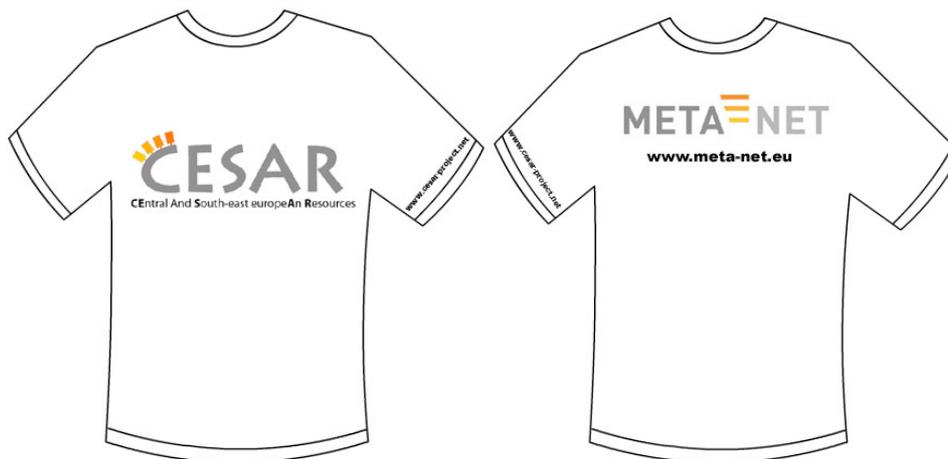


Figure 9: CESAR white background t-shirts

2.5 Cups, bags, pens, USB-sticks, calendars, roll-up...

For successful and continuous dissemination about project CESAR, we produced a number of different other dissemination materials all bearing CESAR logo at least, but also URL of the project website.



Figure 10: CESAR small cup



Figure 11: CESAR tall cup



Figure 12: CESAR bag



Figure 13: CESAR ball-point pens



Figure 14: CESAR 4 Gb USB-sticks

2.6 Presentations at conferences and events

Between M7 and M18, the CESAR project was presented in a number of occasions using the official template for presentation of CESAR project activities. It is available on the internal website: <http://cesar.nytud.hu/documents/archive/presentation-templates>

Also, the CESAR partners that are participants in different events store their presentations on the internal webpage at the link <http://cesar.nytud.hu/events>.

For tracking the dissemination activities, an on-line form was produced for collecting details about each and every dissemination activity. It is available at the link <https://docs.google.com/spreadsheets/ccc?key=0ApVXUpftZrIsdDBVR0Z0eUJkQk1FeXZQVkM1OERaYnc#gid=0>.

Participations in conferences were conducted in several ways:

- CESAR partners presenting project and its results
 - papers with oral or poster presentations
- CESAR project being an organiser of the workshops/tutorials/conference
 - e.g. LREC2012 NooJ Tutorial (incl. presentation of the CESAR project)
 - e.g. LREC2012 workshop @META (Advanced Treebanking as a part of META-RESERCH activity), organised jointly with META-NET, METANET4U and META-NORD
- CESAR appearing at project exhibitions
 - e.g. EAMT2012 EU-projects poster exhibition, LREC2012 EU-projects village, etc.

The detailed list of public appearance is given at D5.2, p13-14 while here we will just list the events where CESAR dissemination materials were presented:

- **EuroLan2011**, Iași, Romania, 24 August-4 September 2011
 - HASRIL held a tutorial on NooJ
 - flyer and poster presented
- **SlaviCorp2011**, Dubrovnik, Croatia, 12-14 September 2011
 - FFZG with two papers, IIPAN one paper, UBG two papers, LSIL one paper
 - flyer and poster, lectures video recorded
- **META-NET** meeting, Berlin, Germany, 20-21 October 2011
 - presentation on CESAR batches
 - flyer and poster presented
- **InFuture**, Zagreb, Croatia, 9-11 November 2011
 - CESAR project presented in a plenary session
 - flyer presented
- **LTC2011**, Poznań, Poland, 25-27 November 2011
 - joint CESAR paper on gaps in LRT
 - joint CESAR demo
 - UBG a paper, ULODZ a paper
 - IIPAN a poster on Polish CESAR resources

- flyer, poster, roll-up, t-shirts, cups presented
- lectures video recorded
- **LREC2012**, Istanbul, Turkey, 21-27 May 2012
 - eight papers presented
 - CESAR project presented at the EU projects special session
 - CESAR project presented at the EU projects village
 - poster, flyer, t-shirts, cups, pens, bags, USB-sticks presented
- **EAMT2012**, Trento, Italy, 28-30 May 2012
 - poster and flyer presented
- **META-FORUM2012**, Bruxelles, Belgium, 20-21 June 2012
 - poster and flyer presented
- **NooJ2012**, Paris, France, 14-16 June 2012



Figure 15: CESAR project's results presentation at SlaviCorp2011.



Figure 16: Collective META-NET picture in Berlin, 2011-10



Figure 17: CESAR poster presented at LTC2011 conference in Poznań, 2011-11



Figure 18: CESAR at EU projects special session, LREC2012 conference, Istanbul, 2012-05



Figure 19: CESAR poster presented at EAMT2012 conference, Trento, 2012-05

2.7 Publications

CESAR partners published a number of papers in journals and conference proceedings. The list until M12 is given in D5.2, so here we will add only the papers appearing after M12:

- Abari, K., Olaszy, G. (2012) A formánsmenetek rendszere CVC kapcsolatok magánhangzóiban a C képzési helyének függvényében (Formant trajectory types in vowels of CVC sequences as the function of the articulatory place of Hungarian consonants), *Beszédkutatás*.
- Bekavac, B., Šojat, K. (2012) Syntactic Patterns of Verb Definitions in Croatian WordNet. In: Vučković, K., Bekavac, B., Silberztein, M. *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 International Conference*, Cambridge Scholars Publishing, pp. 109-117, Isbn: 978-1-4438-3711-8 [pdf](#)
- Koeva, S., Stoyanova, I. and Dekova, R. (2012) Applications of Bulgarian-English Parallel Corpus for Exploring Translational Asymmetries, In: Vučković, K., Bekavac, B., Silberztein (eds.), M. *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 International Conference*, Cambridge Scholars Publishing, pp. 227-240 (Isbn: 978-1-4438-3711-8) [pdf](#)
- Maciej Ogrodniczuk, Radovan Garabík, Svetla Koeva, Cvetana Krstev, Piotr Pezik, Tibor Pintér, Adam Przepiórkowski, György Szaszák, Marko Tadić, Tamás Váradi, Duško Vitas (2012) Central and South-European language resources in META-SHARE In: *Infotheca*, 12(1), 2012
- Merkler, D., Berović, D., Agić, Ž. (2012) Disambiguation of Homographic Adjective and Adverb Forms in Croatian Texts. In: Vučković, K., Bekavac, B., Silberztein, M. *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 International Conference*, Cambridge Scholars Publishing, pp. 109-117, (Isbn: 978-1-4438-3711-8) [pdf](#)
- Pezik P., Ogrodniczuk M., Przepiórkowski A. (2011) Towards an open repository of Polish language resources. In: Vetulani, Z. (ed.) *Proceedings of the 5th Language & Technology Conference*, LTC2011, Poznań, pp. 511-515.
- Gellért Sárosi, Tibor Fegyó, Péter Mihajlik, Balázs Tarján, Judit Pancza, and Zoltán Hans (2012) LVCSR-based speech analytics of a Hungarian language call-center. In: *Workshop on Innovation and Applications in Speech Technology*, Dublin, Ireland, March 9-10. [pdf](#)
- Stanković, R., Utvić, M., Vitas, D., Krstev, C., Obradović, I. (2012) On the Compatibility of Lexical Resources for NooJ, In: Vučković, K., Bekavac, B., Silberztein (eds.), M. *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 International Conference*, Cambridge Scholars Publishing, pp. 95-107 (Isbn: 978-1-4438-3711-8) [pdf](#)
- Szaszák, György. (2011) Beszédatbázis érzelmek felismeréséhez (A speech database for automatic emotion recognition) In: Festival of the Science, Hungary, 17 November, 2011. [pdf](#)
- Tóth B., Németh G., Olaszy G. (2012) Beszédkorpusz tervezése magyar nyelvű, rejtett Markov-modell alapú szövegfelolvasóhoz (Speech corpus design for Hungarian hidden Markov model based speech synthesis), Gósy M.: *Beszédkutatás*

- Tóth, B., Németh G. (in press) Optimizing HMM Speech Synthesis for Low Resource Devices, *Journal of Advanced Computational Intelligence & Intelligent Informatics*, Vol. 16, No. 2., 2012

Particularly strong appearance CESAR partners had at the major conference in Language Technology field, LREC2012 and its workshops

- IPIPAN, ULODZ
 - Ogrodniczuk, Pęzik, Przepiórkowski *Towards a comprehensive open repository of Polish language resources*
- UBG
 - Stanković, Krstev, Obradović, Trtovac, Utvić *A tool for enhanced search of multilingual digital libraries of e-journals*
 - Tanasijević, Sikimić, Pavlović-Lažetić *Multimedia database of the cultural heritage of the Balkans*
- IBL
 - Koeva, Stoyanova, Dekova, Rizov, Genov *Bulgarian X-language Parallel Corpus*
 - Iliev, Genov *Expanding Parallel Resources for Medium-Density Languages for Free*
- FFZG
 - Šojat, Mikelić Preradović, Tadić *Generation of Verbal Stems in Derivationally Rich Language*
 - Berović, Agić, Tadić *Croatian Dependency Treebank: Recent Development and Initial Experiments*
- HASRIL
 - Héja, Takács *Automatically Generated Online Dictionaries*

2.8 Language White Papers

The coordinated effort with META-NET produced the Language White Papers series of books describing 30 different European languages and the state of their Language Technologies. For CESAR languages all LWPs were produced and they should be published with the Springer publishers soon.

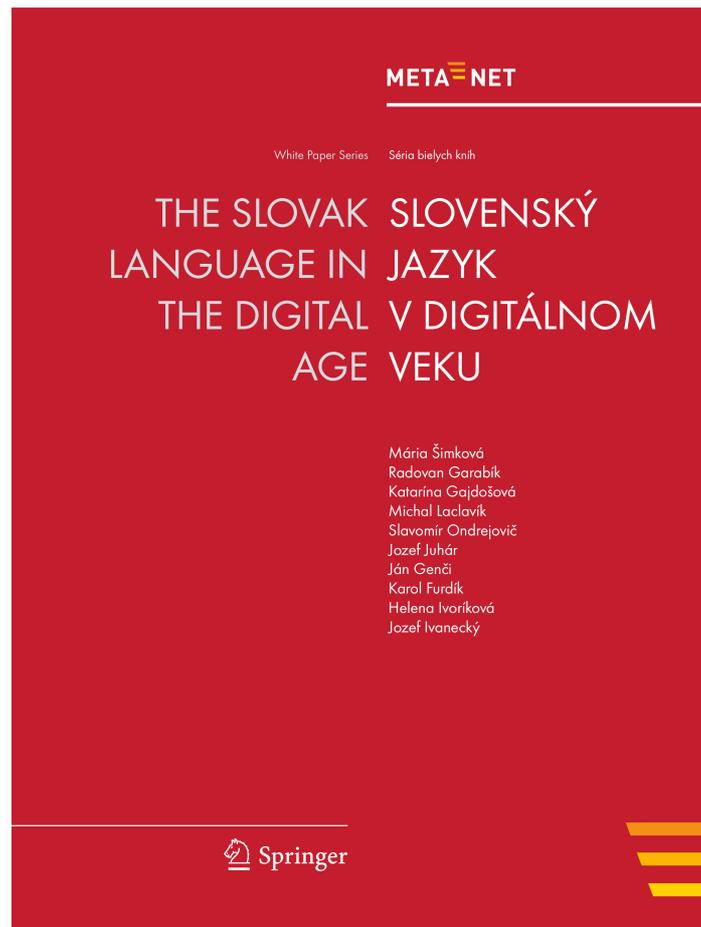


Figure 20: Front page of the Slovakian Language White Paper.