

Irina Podtergera

eHumanities: Benefits for Historical Philologies / eHumanities: Nutzen für die historischen Philologien (Freiburg im Breisgau, 8 October – 10 October 2015): An Analytical Conference Report*

1. *Introduction*

From 8 to 10 October 2015, the Department of Slavonic Studies at the University of Freiburg, hosted the international interdisciplinary conference titled ‘eHumanities: Nutzen für die historischen Philologien’ (‘eHumanities: Benefits for Historical Philologies’). The conference launched the final phase of the project ‘slavaComp – COMPUTER-aided research on VARIABILITY in Church SLAVONIC’ (‘slavaComp – COMPUTERGESTÜTZTE UNTERSUCHUNG VON VARIABILITÄT IM KIRCHENSLAVISCHEN’, cf. <<http://www.slavaComp.uni-freiburg.de/>>), sponsored by the Federal Ministry of Education and Research (BMBF). The project is being carried out in cooperation between the Department of Slavonic Studies and the University IT Services. The aim of the conference was to discuss in a broad interdisciplinary exchange recent findings and current research approaches in the realm of Digital Humanities. It gathered specialists in different philologies, Linguistic and Literary Computing, Ancient, Classical and Mediaeval Studies as well as computer scientists.

Our concern was twofold. On the one hand, we intend to contribute to the establishment of computer-based research methods in Historical Philologies. On the other hand, we attach great importance to interdisciplinary networking between historically-oriented scholars (mediaevalists, historians, philologists) who have made use of digital methods for their research work and/or cooperate closely with computer scientists. We are convinced that all this can help find new ways of strengthening the position of Historical Philologies and Ancient, Classical and Mediaeval Studies within the traditional Humanities disciplines.

2. *October 8th*

The conference started on the evening of October 8 with the welcoming words offered by Juliane Besters-Dilger (Vice-President for Academic Affairs of the University of Freiburg) and Daniel Jacob (Dean of the Faculty of Philology). The keynote speaker, Manfred Thaller (Professor Emeritus at the Department of Computer Science for the Hu-

* This publication appears as a part of the project ‘slavaComp – COMPUTER-aided research on VARIABILITY in Church SLAVONIC’, sponsored by the German Federal Ministry of Education and Research (BMBF), funding code 01UG1251, as well as of the habilitation project of Dr. Irina Podtergera in the framework of the Margarete von Wrangell Habilitation Programme supported by the Ministry of Science, Research and the Arts (MWK Baden-Württemberg).

manities, University of Cologne) spoke on the application of computational technologies to philological studies, discussing the modern concepts of ‘Digital Humanities’, ‘Big Data’ and ‘big’ and ‘small’ academic disciplines (‘große’ und ‘kleine’ Fächer) and ‘big’ and ‘small’ philologies respectively. He assumes that the concept ‘eHumanities’ originates from the concept ‘eScience’. It involves academic disciplines in the research environment, in which data collection, analysis of these data as well as publication of results are supported equally well by the distributed information technology. Nevertheless, the ‘eHumanities’ are not the same as ‘eScience’ particularly with regard to data collection and analysis methods. He therefore offers his own definition of ‘Digital Humanities’. He regards it as an intellectual agenda that seeks to achieve substantive results which are unavailable or unverifiable otherwise, for example because the data are so diverse that they are not otherwise analysable than by formal means. The special feature of this definition is that the emphasis is placed on the analytical value of the digital methods in the Humanities and not only on the sharing of results or on the challenges of ‘big data’. The notion of big data is associated in the Humanities with the possibility to apply one method of analysis to the whole written heritage of a certain language. However, that is precisely why the boundary between the so-called ‘big’ and ‘small’ philologies is blurred. This distinction refers exclusively to the university infrastructure but not to the academic disciplines themselves. Thus the well-known database Perseus that includes the classic Latin and Greek texts contains 68,925,971 words, while the database of the American texts from the 19th century contains 58,332,095 words. However, as we know, English studies is a ‘big’ academic discipline while Classical Philology is not. Discussing the concept ‘Big Data’ on the basis of Perseus and other data bases, Manfred Thaller formulated criteria for two paradigms: one of the ‘big’ and the other of the ‘small’ Philologies and confronted us with the question which of these criteria our slavaComp-Database satisfies.

3. October 9th

The next day of the conference, 9 October, began with two opening talks. The first one was held by Gaja Di Luzio, a representative of the DLR Project Management Agency in Bonn that supports the Federal Ministry of Education and Research (BMBF) in the implementation of project funding. She highlighted BMBF’s principal activities and funding initiatives, which should make the humanities more effective in solving the long-term challenges of our time. She gave a brief overview of the BMBF current framework programme *Humanities, Cultural and Social Science* as well as specific priority topics from last years that were designed to promote new forms of collaboration among scholars and to develop innovative research approaches and thus new prospects for research. One of the funding measure was the call for proposal *Guidelines for Funding Research and Development Projects in the eHumanities*, published in 2011. Within the framework of this funding programme, the project slavaComp was launched in October 2012. After this overview, the Project Director Juliane Besters-Dilger allowed a glimpse into the slavaComp. She outlined the background of the project, the process of application and organization of the work and

explained the project contents and objectives as well as some problems and difficulties that we have encountered. The goals of the project are to establish an extensive bilingual (Church Slavonic-Greek resp. Greek-Church Slavonic) glossary and to create a lemmatizer to return the respective lemma of any valid Church Slavonic word regardless of its specific graphic features. What lies behind this is the desire to make analyzable the lexical and graphic variation of the Church Slavonic written heritage in its regional and chronological development until the 16th century. These objectives met the criteria which Manfred Thaller had laid out in his keynote speech for a ‘big’ philology (see § 2 above).

3.1. After these introductory presentations, the conference took off with a dense programme of eighteen papers organised in six interconnected sessions. The two morning sessions were chaired by Christian Mair (Professor for English Linguistics, University of Freiburg). The first session brought together experts in Slavonic Studies, Information Science and German Linguistics.

3.1.1. Ralph Cleminson (Winchester), formerly Professor of Slavonic Studies at the University of Portsmouth, University of Otago and the Central European University, provided some theoretical and practical reflections on the topic of *Encoding Text and Encoding Texts*. He took stock of what had been achieved so far in the field of Digital Palaeoslavonic Studies and considered what should be done next. Based on his extensive experience in codicology, palaeography, textual criticism, and other areas of mediaeval Slavonic studies and on his activities in developing techniques and standards for the digital representation and description of Slavonic manuscripts and early-printed books, he built new bridges between possibilities of computer technology and theory and methodology of textual criticism. He emphasized that an encoded text as a digital edition must accommodate the cultural function of the text to be encoded – both the ‘ideal’ function and its particular realisations in manuscripts.

3.1.2. Alexander Mehler, Professor for Computational Humanities and Text Technology and Head of the Text Technology Lab at the Goethe University Frankfurt, introduced *Wikidition*, a new text technology that allows automatic lexiconization, i.e. lemmatization and grammatical analysis of each syntactical word, and cross-linking of text corpora. From the technical point of view, Wikidition incorporates a large variety of text mining tools that recognise lexical, morphological, and syntactic information and links automatically lexical, sentential, and textual units within the whole corpus. This technical approach is of crucial importance for it opens mediaevalists, historians and philologists new perspectives for research in the field of texts in contact and text reuse. Finally, Wikidition was exemplified by means of a corpus of Mediaeval Latin texts (cf. <http://capitwiki.hucompute.org/>).

3.1.3. Stefan Engelberg, Professor for German Linguistics (University of Mannheim) and Head of the Department of Lexical Studies at the Institute of German Language (IDS, Mannheim) discussed the topic *Internet Lexicography and the Lexicon Dynamics*. He focused on a gap between everyday language usage and our current state of knowledge in linguistics: because of the strong dynamics of the lexicon, traditional lexicography finds

it difficult to record word usage entirely. Corpus-based studies show that only 1% of the contemporary inventory of lexemes is documented in paper dictionaries. The reasons for this include constant fluctuations due to loan words, word creation, and meaning shifts, idiosyncrasy in the formation of compound words, blurred boundaries between lexical and grammatical phenomena in syntagmatic patterns and conventionalized multiword expressions, as well as lexical variation within idioms. It is not possible to reflect this continuous dynamics in paper dictionaries. The speaker demonstrated some attempts to solve this problem in internet lexicography by means of the new features in the access structures of electronic dictionaries, data networking, and visualisation.

3.2. The second morning session was dedicated to new tasks facing the historical Slavonic lexicography.

3.2.1. Professor Lora Taseva (Bulgarian Academy of Science, Institute of Balkan Studies in Sofia) spoke on *Multiple Translations as a Research Object of Philological Mediaeval Studies and Challenge for Computational Linguistics*. As we know, multiple translations are a significant feature of the mediaeval Slavonic culture. Many of them have not been located or/and dated until now or their origins are controversial. Lexical factors play a key role for the dating and localisation of translated texts as well as for the description of translation techniques. It helps to determine the scriptorium in which the text has been translated. But because of constant text reuse and text contact there are many lexical overlaps, so that a clear distinguishing between related mediaeval translations is anything but trivial. Texts with several traditions and translations of the same text, which are prepared in different scriptoria within a short time, make this task even more difficult. Furthermore, polysemy of the original word, text type-specific word usage and distinctive language characteristics of the source text must be always taken into account for those factors which impacted on the choice of words. An accurate dating and exact location have to be tackled only by means of statistical analyses of 'big data'. This contribution revealed how extremely useful such text technology as Wikidition (see § 3.1.2 above) would be for slavists. Linkification and lexiconization of text corpora are counted among the major tasks of the historical eSlavistics. Meanwhile, that is exactly the case, when the substantive results cannot be got or verified otherwise than by formal means (see § 2 above).

3.2.2. Development of a tool for automatic lexiconization of Church Slavonic texts irrespective of their origins and graphic and orthographic peculiarities, in other words, creation of a lemmatizer is a slavaComp's core task. Irina Podtergera and Susanne Mocken reported about how it can be accomplished. In order to develop lemmatization algorithms we generated a vast MetaGlossary with hyperlemmata, headwords, syntactical words respectively word forms, and multiword expressions including their grammatical characteristics, both in Church Slavonic and Greek. For this purpose we brought together under one roof eighteen bi- and monolingual glossaries and word form indices.

3.2.2.1. Irina Podtergera (Department of Slavonic Studies, University of Freiburg) concentrated in her contribution entitled *From Historical Paper-Lexicography to Histori-*

cal *E-Lexicography* on philological and linguistic aspects of the issue. She highlighted the macro- and microstructure of the glossaries in question, underlining how differently the same kind of information may be presented. At the same time, she brought out the formal and substantive advantages of an electronic dictionary. Its format is much more flexible and thus enables us to compensate any inconsistencies of the existing paper dictionaries. From a philological point of view, the greatest benefit of the Church Slavonic electronic dictionary is that it facilitates significant investigations of history of lexemes and concepts and helps distinguish more exactly the mediaeval schools of translation according to lexical properties of the translated texts. Thereby, she picked up the thread of the discussion of the preceding contribution (see §. 3.2.1 above).

3.2.2.2. Susanne Mocken (IT-Services, University of Freiburg) dealt with issues of markup. She answered the question of her title *How Can Diversity Be Unified*. The main focus of her presentation lay on the technical aspects of our project. The first module in developing an electronic Church Slavonic-Greek dictionary was an automatic conversion of all the glossaries, which had been prepared in non-Unicode capable systems, into the Unicode standard. Thereafter, the converted documents had to be encoded in TEI/XML, which allowed us to unify the heterogeneous structures of the glossaries and compensate their substantial inconsistencies. The speaker gave a description of the XML-structure of all encoded glossaries and showed the participants how the preliminary version of our Church Slavonic-Greek MetaGlossary works.

3.3. The two afternoon sessions were chaired by Georg Lausen, professor of Databases and Information Systems at the University of Freiburg. The intention of six afternoon lectures was to discuss Slavonic and related approaches with a strong digital component. The focal point was on current developments in the field of corpus-linguistic treatment of language change, linguistic and literary computing, and graphic visualisation of language data.

3.3.1. Roland Meyer, professor of West Slavonic Linguistics at the Humboldt University of Berlin, illustrated the application of specific computational linguistic methods to the study of Slavonic languages. He evaluated the *Data Driving Identification of Registers in the Historical Texts* by a synchronous and diachronic comparison of the relative pronouns in Polish, Czech, and Russian. Using statistical methods, he provided additional arguments concerning the origins of the Russian relative pronoun *kotoryi*. On the one hand, based on Alexander Isachenko's statement that in 17th century Russian, there were relative clauses only in the high variety, and on the other hand, following Achim Rabus' paper about language contact between Ruthenian and Russian, he considered that this pronoun was a result of the contact-induced changes in 17th and 18th century and appeared in Russian under the influence of Ruthenian, which itself had been influenced by West Slavonic, and in particular Polish. In addition, he employed stylometric analysis of texts from the 18th and 19th century to argue that the usage of relative clauses is a predominantly characteristic of scientific texts.

3.3.2. David J. Birnbaum, professor at the Department of Slavic Languages and Literatures of the University of Pittsburgh and the Chair of the Department, reported

about his collaborative work with Hanne M. Eckhoff (Department of Language and Linguistics, University of Tromsø) on the digital edition of the Codex Suprasliensis. His paper was devoted to the *Machine-Assisted Normalization* of the encoded Old Church Slavonic manuscript text. Despite the wide range of different types of editions – viz. diplomatic, normalized with reading-view, text-critical or comparative, annotated and interpreted facsimile – in the final analysis, there is always only one physical artefact. In this sense, the multi-layered reality of the text in question must always be taken into account for editing, regardless of the respective research focus. The edited text must be accessible to all scholars. This thesis was exemplified by the above mentioned text. At the present time, the electronic edition of the Codex Suprasliensis is supplied with diplomatic transcriptions of all Slavonic texts, parallel Greek correspondences, and high-quality facsimile of the manuscript (cf. <<http://suprasliensis.obdurodon.org/>>). But we still lack a normalized reading view of it. On the basis of the experience gained within the historical Slavonic data bases PROIEL (i.e. Pragmatic Resources in Old Indo-European Languages) and TOROT (i.e. Tromsø Old Russian and Old Church Slavonic Treebank), David Birnbaum and Hanne M. Eckhoff have developed a machine-assisted method to convert a diplomatic edition of the manuscript into normalized canonic Old Church Slavonic. Special attention in the paper was given to the difficulties with which they were confronted during this work.

3.3.3. The third lecture in this session was that of Aleksandr Moldovan, head of the Vinogradov's Institute of Russian Language (Russian Academy of Sciences, Moscow). He spoke on *Essentials of Language Documentation* by focusing upon the old Cyrillic written heritage. He dwelt on two problems areas. At first, he discussed complications in the encoding of old Cyrillic texts emphasising linguistic relevance of graphic and orthographic distinctive features as well as of regional and historical variations of writing and grammatical systems. He recalled the absolute necessity to link numerous linguistic tools (see § 3.1.2 above) in order to automatically recognise the morphological forms in Mediaeval Cyrillic texts. The best result achieved so far in this matter is a semi-automatic guessing tool, which is integrated into the Russian National Corpus and offers suggestions for further manual processing. Secondly, he addressed the deficiencies of metalinguistic text encoding, particularly in regard to dating and taxonomic determination of encoded texts. The last point is of particular importance for an objective linguistic interpretation of the language usage in the text in question. Following Nikita I. Tolstoj, he explained that it seems more plausible to classify the texts according to their functional domains rather than to their types. This is implemented for Old and Middle Russian in the beta-version of the Russian National Corpus.

3.3.4. The second afternoon session began with a contribution of Achim Rabus, Aleksander Brückner professor for Slavonic Linguistics at the Friedrich Schiller University in Jena. He presented a talk on *Multiple Use of Data and Code* focusing on two recent Slavonic di- and synchronic corpus-linguistics projects, in which he was involved or which he initiated. He took the Freiburg diachronic VMČ corpus as a starting point in order to

argue that graphical user interfaces, data, and codes can be recycled and subsequently performed. This has been realised in his synchronic project on the subject *Rusyn Language as a Minority Language across National Boundaries: Dynamic Processes*, in which framework a corpus of spoken Rusyn is being developed. The faceted navigation enables researchers to browse the information space by playing audio files, surveying the location where it was recorded, differentiation by speakers' sex and age, etc. and results in new geo-temporal and interpretive contexts. The ensuing presentation showed the search options of both corpora by browsing selected linguistic features.

3.3.5. From the technical point of view, the faceted search was the focus of interest in the next lecture: Thomas Efer from the Department for Natural Language Processing at the University of Leipzig described the *Use of Graph Databases in the Analysis of historical corpora*. Setting the tone by pointing to the limitation of the text processing with XML as a simple hierarchy of elements, he brought to attention the benefits of graph databases for text technology by using the example of the Leipzig historical project *exchange*. Unlike XML, which mirrors only one hierarchy, graph databases, thanks to their flexibility, cover many parallel hierarchies. They permit a direct connection of various index card catalogues or charts, i.e. various relational databases. Thus, it is possible to interlink geo-temporal and interpretive contexts, so that all information about the document, time, place, type, author, group of author, etc. are interconnected. Moreover, it is possible to embed each word in different hierarchies. This text technology is particularly suitable for conceptualisation of language resources because it allows an optional segmentation of the text for the search. Finally, the speaker demonstrated faceted browsing in Leipzig historical corpora by searching for co-occurrences of word *ισονομία* in the works of Galen of Pergamon, and all tokens of *φάρμακον* and *μανία* surrounded by seven words, three of which are nouns in the medical works written by Athens authors in the 2nd century BC. Furthermore, a special advantage of graph databases exists in the possibility of various visualisations of the search results. This contribution showed us what should be done next in ePalaeoslavistics.

3.3.6. The final lecture on this day was on the *Annotation of Zeros*, held by Christine Grillborzer (Department of Slavonic Studies, University of Freiburg). She signalled the difficulties faced by linguists by searching for clauses with zero dative subject in the Russian National Corpus (RNC) and comparing them to the clauses with a nominative subject. The ratio between the dative zero subject hits and the nominative subject hits is 1 : 115 (dat.-sub. : nom.-sub.), which doesn't correspond to Russian grammatical realities because it is an inherent property of Russian syntax that the dative subject can be realised covertly. The divergence in the RNC arises from the fact that only overt arguments have been annotated. However, it would be highly preferable to annotate zero arguments of the verb too. This would be of particular importance for investigations of syntactic change. After analysing four overt and covert constructions with dative subject, the lecturer discussed possibilities for an automatic, machine-assisted annotation of zero dative subject.

4. *October 10th*

The last conference day, 10 October, consisted of two morning sessions and featured six papers from Slavonic Studies, Classical Philology and Computer Science. Both sessions were chaired by Gerhard Schneider, head of the University IT-Services and co-director of the project *slavaComp*. The thematic focus of the first session was on data aggregation and enrichment, whereas the second session broached the issue of machine-assisted solutions for specific philological questions.

4.1. Anissava Miltenova, chair of the Department of Old Bulgarian Literature at the Institute of Literature (Bulgarian Academy of Science in Sofia) presented a talk on *Rethinking Old Church Slavonic Digital Library by Ontologies* giving insights into the project *Scripta Bulgarica*. This innovative project is intended not only for scholars but also for students and can be used for research and educational purposes. It involves an electronic thesaurus and a library of Old Church Slavonic Texts and pursues a double aim, namely, to collect data concerning mediaeval Bulgarian written heritage and provide models and samples for the presentation of metadata, terminological articles, and articles on Byzantine writers, etc. The integrated thesaurus contains terms and concepts in Palaeoslavistics in eight languages. The text resources and metadata are extracted from already existing databases and corpora, for instance from the *Repertorium of Old Bulgarian Literature and Letters* (<cf. <http://repertorium.obdurodon.org/>>), the digital edition of the Codex Suprasliensis (both designed by David J. Birnbaum), electronic collection of Bulgarian manuscripts, etc. Furthermore, there are articles of encyclopaedic type on topics relating to mediaeval Slavonic Studies. In order to achieve this compilation, an internet aggregator was conceived. It is based on multi-layer data modelling and multi-standard framework and collects information from extant digital libraries and various resources for knowledge exchange.

4.2. Toma Tasovac, head of the Belgrad Center of Digital Humanities, entitled his lecture *The Devil is in the Detail: From Data Modelling to Data Enrichment in Legacy Dictionaries*. He began it by referring to a New York Times article *Justices Turning More Frequently to Dictionary, and Not Just for Big Words* (appeared on June 13, 2011, cf. <<http://www.nytimes.com/2011/06/14/us/14bar.html>> [28.11.2015]) to underline the social significance of dictionaries and their role in the knowledge taxonomy. However, the main emphasis of his lecture was on historical dictionaries. They function nowadays not as reference works for the contemporary language usage but as research objects. Taking for instance the definition of *coffée* (“Coffa or Cauphe”) in Blounts *Glossographia* of 1661, the speaker reflected on the thesis that historic dictionaries are a source for reconstruction of social and culture-dependent lexical knowledge of specific periods of time. Therefore, the mere digitisation of such dictionaries in order to make them available for all is of little value for studies in the field of conceptual history. The top challenge for eLexicography must be to incorporate the available electronic editions into an efficient research environment for the exploration of historical semantics. This approach has been realised by the creation of the cutting-edge *Plattform*

for the *Transcription and Digital Editions of the Serbian Manuscript* (<<http://prepis.org>>). Toma Tasovac shortly reported on his experience with encoding of Vuk Karadžić's *Lexicon Serbico-Germanico-Latinum* (1818, 1852), which is rich on historical, cultural and (socio) linguistic information. In order to offer all this information for further studies, he modelled structural, semantic, encyclopaedic and serendipitous access paths to dictionary content and integrated the dictionary into the platform the WordNet, a semantic navigation tool, which provides access to the relations between lexical concepts within a dictionary. Another major project of which he was in charge was digitising some 23,000 lexicographic paper slips compiled by Serbian amateur lexicographer Dimitrije Čemerikić (1882-1960). This hand-picked slips collection that was never published contains about 16,000 lemmas with definitions and examples testifying to the now endangered Serbian dialect from the historic city of Prizren and documenting manifold aspects of Serbian popular culture and urban life in the middle of the twentieth century. As in the case of Karadžić's dictionary, the most important goal of the project was to model search options based on multiple access paths that go beyond the alphabetic macrostructure. The paper slips were not transcribed but scanned and combined into entries with headwords, which were marked up with their standardised orthographic counterparts. In order to implement reliable faceted navigation and targeted search capabilities, the data were enriched by including additional, annotated information, namely, by providing lexical paradigmatic equivalents and assigning semantic fields. The result was an increased use value of the scanned paper slips without transcription or structural modelling of their content.

4.3. Jürgen Fuchsbauer (Institute of Slavonic Languages and Literature, University of Regensburg) dealt with *Paralleling Different Versions of Slavic Texts*. Using the example of Church Slavonic and Balkan Slavic Lives of Paraskeva of Epibatai (Petka Tärnovska), he raised the question of how several versions of one text should be aligned within one digital edition and what preliminary work would be necessary for this. He gave an overview of the whole corpus, from the original Church Slavonic text situated in the Middle Bulgarian 'Miscellany of German' (*Germanov Sbornik*, 1358/59), through shortened and extended Church Slavonic redactions of the text, which had been composed by Patriarch Euthymius of Tärnovo between 1376 and 1382, up to Russian Church Slavonic redactions and Bulgarian vernacular versions from 17th, 18th, and 19th century. All these redactions and versions differ from each other in respect of their language, text structure and contents. Together, as a corpus, they attract great interest both for the history of language and historical text-linguistics: they enable the visualisation of the language change between the Church Slavonic regional and chronological varieties and from Church Slavonic to Middle and Early Modern Bulgarian as well as changes in the text structure. In order to achieve comparability between all versions within the *Paraskeva of Epibatai Corpus*, thematic and text units must be linked to each other, possibly including the predication level. However, this is anything but trivial. The lecturer gave an account of the preparatory philological work which is needed for designing and building up an efficient corpus of parallel texts with highly complicated tradition.

4.4. The second morning session, which concluded the conference, was opened with Simon Skilevic's report *Dealings with the graphic variety*. As a student assistant in the slavaComp project, he has developed a tool for converting non-Unicode files into Unicode format. Until recently, palaeoslavists edited old Slavonic manuscripts with non-Unicode fonts because the Unicode standard was incomplete. The result is that today we have dozens of non-Unicode texts, the reuse of which is not possible offhand. Most of the bilingual – Church Slavonic-Greek and Greek-Church Slavonic – glossaries that are made available to us are also prepared with non-Unicode fonts. Hence, developing a conversion tool for Church Slavonic was a priority: without conversion, we could not begin processing data (s. § 3.2.2.2 above). Simon Skilevic reported on technical aspects of programme development and illustrated how it works. The software is called *slavaComp-converter* and is one of the SlaVaComp principal outcomes. The default profile has a tabular form whereby a table can always be extended not only for Cyrillic script. Consequently, it permits the conversion of texts in all Indo-European languages from non-Unicode into Unicode fonts and can thus be used by all historically oriented scholars and not only by palaeoslavists.

4.5. Stylianos Chronopoulos (Department for Greek and Latin Philology, University of Freiburg) presented his ongoing research project on *Pollux' WordNet*, concerning a digital edition of a famous Greek thesaurus from the 2nd century AD. The thesaurus consists of ten books and contains ca. 120,000 words which are pooled in hierarchically-structured semantic fields organised according to subject-matter. What is special about this dictionary is that lists of words are embedded in a continuous text, so that the microstructure of the semantic field depends on syntax of this text. The aim of the proposed digital edition is to represent the microstructure of the semantic fields. This should be accomplished by means of WordNet ontology. On the basis of five examples, the lecturer discussed specific problems requiring solution. These included dealings with gaps, inconsistencies, and descriptions of word formation, as well as appropriate definition of category, region, and usage pointers which should correspond with thesaurus' ontology.

4.6. The conference was closed by the lecture of Evgenii Filimonov (Department of Slavonic Studies, University of Freiburg) *Greek-Slavonic Asymmetries in Syntax and Lexis*. As a representative of the slavaComp project, he talked about an urgent problem, namely, dealing with discrepancies between Church Slavonic translation and the Greek original. He focused primarily on two types of lexical asymmetries: free or vague translation of the original term and multiword expressions for one-word equivalents and vice versa in the source and target language. For the latter, we use *collocation* as a terminus technicus. There are different types of collocations in the glossaries we work on. The lecturer analysed principles of collocation encoding and showed how the multi-word expressions can be searched automatically. The second part of his lecture was dedicated to the encoding of discrepancies in the syntactic structure between Greek original and Church Slavonic translation.

5. *Conclusions*

To sum up, this conference stressed how Slavistics and especially Palaeoslavistics benefit from Digital Humanities. Not only are computer-assisted methods of great importance because they offer new perspectives for analysing written heritage, but also the eHumanities *per se* because they stimulate interdisciplinary networking and the exchange of knowledge between representatives of different disciplines and different scientific cultures. One of the results of this networking is the unification of research instruments and tools which leads to the elimination of the boundaries between 'big' and 'small' philologies.

Conference abstracts can be downloaded from <<http://www.slavacomp.uni-freiburg.de/konferenz.html>>.

Abstract

Irina Podtergera

eHumanities: Benefits for Historical Philologies. An Analytical Conference Report

This contribution is an analytical report on the eHumanities conference held in Freiburg on 8-10 October 2015. The aim of the conference was to discuss recent findings and current research approaches in the field of Digital Humanities in a broad interdisciplinary exchange. It gathered specialists in different philologies, Linguistic and Literary Computing, Ancient, Classical and Mediaeval Studies as well as computer scientists. This report summarizes their presentations.

Keywords

Digital Humanities; Historical Philologies; Computer-Based Research Methods.