

ESF Exploratory Workshop on

Methods of Digital Philology for the Study of Early and Classical Arabic Language

Berlin (Germany), 19-21 October 2011

Scientific Report

Convened by:
Kirill Dmitriev^①, and **Michael Marx**^②

① School of Modern Languages, University of St Andrews, Scotland

② Corpus Coranicum, Berlin-Brandenburgische Akademie der Wissenschaften, Potsdam, Germany

Executive Summary

The ESF SCH Exploratory Workshop “Methods of Digital Philology for the Study of Early and Classical Arabic Language” was held over three days on 19-21 October 2011 in the Brugsch-Pascha Hall of the National Museums in Berlin, Germany. With 27 participants from ten European countries and the USA the workshop brought together leading scholars of digital Arabic philology as well as related fields of research including Egyptology and Roman philology. The workshop was dedicated to the exploration of new methods of lexicographical research in the field of the history of the Arabic language. The language of the workshop was English, one paper was presented in German.

The scientific agenda of the workshop was structured upon thematic working sessions on Arabic Lexicography, Inscriptions and Papyri, Qur’an, Digital Approaches in Comparative Perspective and Planning of Follow-up Research Activities and Collaborative Actions. In the centre of the ESF-workshop was the discussion of relevant materials for the study of the Early and Classical Arabic language. To date no etymological or historical dictionary of the Arabic language exists. Especially epigraphical material from the Ancient-North-Arabian and Ancient South-Arabian Languages has never been systematically referred to, in order to study the genesis of the Arabic Language. Especially the varieties of Ancient North-Arabian (600 BCE-400CE) in its Safaitic, Hismaitic or the yet unclassified so called Thamudic variations) of which several tenthousand inscriptions exist, belong like Arabic to the Central Semitic Group and can offer promising material for the understanding of the emergence of the Arabic language. This is also true for the vast corpus of Ancient South-Arabian languages (800BCE-600CE), of which e.g. Sabaic offers some morphological and lexicographical features very close to Arabic. Both the inscriptions of the North and the South offer through dated inscriptions material indispensable for the historical linguistics. Apart from inscriptions on stone and bronze, also the corpus of Arabic papyri represents a neglected source for dated references in the field of the early Arabic language.

The main goals of the workshop were twofold: on the one hand, the meeting gave the opportunity to discuss the current state and the future of research on the Arabic language and vocabulary from the perspectives of different study fields, focusing in particular on the epigraphy, papyrology, Qur’an and poetry. And on the other hand, the workshop represented a forum for exchange of working experience with different digital databases, which already exist or are being created. Sharing of this research experience with new digital methods was carried out in the context of the new databased project – the Analytical Database of Early and Classical Arabic poetry proposed by the convenors.

This database aims to complement existing digital resources for the study of the early and classical Arabic language. The new database will make the significant lexicographic material provided by the works of Arabic poetry available for research on the origins of the Arabic language and the semantic development of its vocabulary. The database will implement comprehensive analytical tools. Along with the lexical entries in the format of an electronic dictionary, it will include all other relevant data such as information about the authors, their tribal affiliation, chronological periodization and geographical classification. Such analytical tools will make the database useful not just as a convenient lexical source, but also as a reference work for wider research on Arabic literature, history and culture. An innovative feature of the database comes from the

possibility to include in it cross-references to the parallel linguistic material provided by other sources (inscriptions, papyri and the Quran), which have not been studied in relation to each other yet. The participants of the workshop discussed the perspectives of linking the different lexical databases together as well as the scientific, methodological and technical issues required for the establishment of a closer cooperation in the field. Various questions related to the architecture of the existing databases and the new project have been discussed such as the digital output formats, statistical functions, representation and analysis of graphical sources, implementation of a compatible root matrix for Arabic vocabulary and others. Once fully developed, the database will be available online for researchers, who will be able to access it and also actively contribute to it. Thus, the database will have a long-term impact on promoting collaborative international research in the field of Arabic studies. The ESF SCH Exploratory Workshop offered unique format to explore concrete steps to achieve an effective and sustainable teamwork between European scholars, who actively implement digital tools and methods in their research on the Arabic language.

Scientific content of the event

The workshop focused on the new research platform for the Arabic language including the analytical database of the Early and Classical Arabic poetry proposed by the convenors. The prototype of this database, which was developed in a pilot project, was presented to the participants by the convenors. Arabic poetry, especially the corpus of its early works (6th-8th centuries A.D.), embodies an important repository of the Arabic language and its vocabulary. As such, it has attracted close attention from the earliest Arab philologists, and remained a central source of linguistic reference throughout the classical period of the Arabic philological tradition and beyond. In their written form, the works of early Arabic poetry have been transmitted not only in separate poetic collections of poetry (dīwāns), but most widely in the form of poetic quotations in all kinds of literary production in Arabic. Accordingly, such rich textual evidence of Arabic poetry provides us with direct insights into the cultural and intellectual context in which this poetry has been transmitted and which reflects the development of language consciousness among Arabs. This corpus of texts carries multiple and layered imprints of the history of the Arabic language and philology, a feature of it that has not yet been acknowledged; whereas all other sources of our knowledge about the origins of the Arabic language – inscriptions, papyri and the Qur’ān – have already become the subject of European research and digitisation initiatives: the Digital Archive for the Study of pre-Islamic Arabian Inscriptions (ERC Advanced Grant 269774-DASI, Università di Pisa, Italy), the Safaitic Database (University of Oxford, UK), the Arabic Papyrology Database (Universität Zürich, CH/LMU Munich, Germany), the Corpus Coranicum Project (BBAW, Berlin, Germany). Prof. Albert Arazi (Hebrew University Jerusalem) published in 1999 a concordance of preislamic poetry and had informed the organizers of the workshop that he has transformed parts of his paper concordance into a digitalized version and signalled that he is willing to discuss further developments, plannings and possible cooperation with the organizers. Apart from Prof. Arazi’s attempt, the resources of early Arabic poetry have never been subjected to a systematic analysis due to their vast but fragmented distribution. In order to meet this complex, yet very promising,

challenge, the workshop explored the ways to implement new methods of investigation through digital technologies.

The workshop gave the participants an opportunity to discuss the architecture of the new Analytical Database of Arabic Poetry in order to design it as an accessible and useful research platform for the scholars working in various domains of the Arabic studies. Dr Kathrin Müller (BAW, Munich) gave an insightful presentation on the history, the current state and the perspectives of the research on the lexicography of Arabic conducted at the Bavarian Academy of Sciences and Humanities. As one of the leading research groups in this field the Commission for the Semitic Philology at the Bavarian Academy of Sciences and Humanities has valuable experience and profound expertise in Arabic philology, which is highly relevant for the new Analytical Database of Arabic Poetry. Dr Kathrin Müller shared with the participants the information on the resources of the Anton Spitaler Archive on Arabic poetry preserved in the Bavarian Academy of Sciences and Humanities. The future development of the new research platform in form of the proposed Analytical Database of Arabic Poetry will have to rely on this important archive. Thus, the workshop allowed us to make the first and very important step forward in establishing a long-term cooperation with the Commission for the Semitic Philology at the Bavarian Academy of Sciences and Humanities. Several presentations at the workshop including the talks by Michael Macdonald on the Sabaic Database project, by Dr Johannes Thoman on the Arabic Papyrology Database, by Dr Alessia Prioretta on the Corpus of South Arabian Inscriptions and by Dr Anne Multhoff on the digital resources on South Arabian in the University of Jena helped the participants to exchange information on the state-of-the-art digital research on early Arabic inscriptions.

Another highly relevant source of our knowledge of the origins of the Arabic language is represented by the Qur'an. Dr Daniel Birnstiel, Dr Orhan Elmaz, David Kiltz and Dr Catherine Pennacchio shared with the participants their expertise on the current research in this field. Especially the contribution by Dr Birnstiel and Dr Elmaz on technical issues and methods of semantic analysis provided important information for the future progress of the project. Primarily, resources on the processes of the transmission, usage and interpretation of Arabic poetry have to be made accessible in the form of a reliable research platform, which will be used for a comprehensive study of the history of the Arabic philological tradition and its cultural implications. This goal can be achieved through an Analytical Database of Arabic Poetry. This database will include the complete vocabulary from all texts of early Arabic poetry, as well as all relevant information on their attestation in written sources. The database will also provide references and links to the lexical resources of other relevant databases. Thus, it will create a unique and integrative tool for research on the semantic changes in the Arabic language, its historical development and the hermeneutical context of its interpretation. With the help of the analytical tools of the database, each lexeme of this entire lexical corpus will be assessed in relation to the literary framework of its attestation including information on the genre of the relevant poetic text and on the tribal, chronological and geographical background of its author. Moreover, the database will record in detail the data of textual transmission of the works of early Arabic poetry in the context

of Arab-Muslim scholarship of the 8th to 10th centuries. This comprehensive collection of data and its analytical classification will for the first time allow systematic investigation into the process of semantic change in the Arabic language and the development of a philological approach to the language.

Valuable comparative perspectives and interdisciplinary approaches were shared during sessions 4 and 5 of the workshop. Presentations on relevant projects such as the Glossarium graeco-arabicum, Arabic and Latin Glossary as well as contributions from colleagues in the field of Roman philology and Egyptology widened significantly the horizons of projects and research initiatives which have to be evaluated and taken into account for the further development of the Analytical Database of Arabic Poetry. Colleagues working on the digital edition of the Egyptian dictionary for example possess over twenty years of experience in developing digital philological resources including tools for visual representation of inscriptions and full-search options for scanned manuscripts. Bringing together experts from various relevant areas of digital philology, the workshop created a unique platform for information exchange and fruitful research cooperation in the future.

Assessment of the results, contribution to the future direction of the field, outcome

The workshop explored new research methods in lexicography of the Arabic language and helped to promote European collaboration in the field of the digital Arabic philology. The workshop provided important feedback and help to design the structure and the interface of the new Analytical Database of Arabic Poetry in an appropriate way, making it accessible and useful for the scholars working on different Arabic primary sources. This important outcome of the workshop provides strong foundation for scientific exchange and future collaboration in the field. As the first tangible outcome of the workshop the “Arabic digital philology”-WIKI platform for the exchange of information and relevant resources in the field of digital Arabic philology has been launched:

http://arabicdigitalphilology.wikia.com/wiki/Arabic_digital_philology_Wiki.

This platform will be further developed and maintained in the future.

Besides the database of Ancient poetry during the workshop the idea of a web-portal-database was discussed, with the help of which lexicographical evidence coming from Ancient North-Arabian, Ancient South-Arabian and Arabic papyri would enrich our view of the emergence of the Arabic language manifested in pre-Islamic Arabic poetry and the first parchments of the Qur'anic text. The web portal will require further discussion and cooperation especially regarding the information technique standards. At the next meeting of the group scheduled for 2013 this issue will be discussed again.

The workshop has to be regarded as an important step towards the establishment of a European research network of researchers on the history of the Arabic language.

Final programme

Wednesday, 19 October 2011

Morning	<i>Arrival</i>
14:00	<i>Get-together, reception</i> (Brugsch-Pascha-Hall, Geschwister-Scholl-Str. 2-8, fifth floor)
15.00-15.20	Welcome by Convenors Kirill Dmitriev (FU Berlin, Germany / University of St Andrews, Scotland) Michael Marx (BBAW, Potsdam, Germany)
15.20-15.40	Presentation of the European Science Foundation (ESF) Hanne Ruus (ESF Standing Committee for the Humanities (SCH))
15.40-18.20	Session 1: Arabic Lexicography
15.40-16.00	Klassisch-arabische Philologie in der Kommission für Semitische Philologie der Bayerischen Akademie der Wissenschaften (BAdW), München Kathrin Müller (BAdW, Munich, Germany)
16.00-16.20	Discussion
16.20-16.40	<i>Coffee / tea break</i>
16.40-17.00	Analytical Database of Early and Classical Arabic Poetry Kirill Dmitriev (FU Berlin, Germany / University of St Andrews, Scotland)
17.00-17.20	Database for the Early Arabic Language: Inscriptions, Papyri, Qur'an and Poetry Michael Marx (BBAW, Potsdam, Germany)
17.20-18.20	Discussion
19.00	<i>Dinner</i>

Thursday, 20 October 2011

09.20-11.40	Session 2: Inscriptions and Papyri
09.20-09.40	The Safaitic Database Online for the 21st Century Michael Macdonald (University of Oxford, UK) Daniel Burt (University of Oxford, UK)
09.40-10.00	From APD to EpiDoc: Standards and Procedures for Editions with Critical Apparatus in Arabic Papyrology Johannes Thomann (Universität Zürich, Switzerland)
10.00-10.20	Discussion
10.20-10.40	<i>Coffee / tea break</i>
10.40-11.00	The Corpus of South Arabian Inscriptions: the CSAI Project Alessia Prioletta (Università di Pisa, Italy)
11.00-11.20	South Arabian Lexicographical Data of the University of Jena Anne Multhoff (Universität Jena, Germany)
11.20-11.40	Discussion
11.40-13.00	<i>Lunch</i>
13.00-15:20	Session 3: Qur'an
13.00-13.20	First Steps towards a Text Grammar of the Qur'an: Some Reflections and Examples Daniel Birnstiel (University of Cambridge, UK)

13.20-13.40	Computational Approches to Word and Root Statistics of the Qur'anic vocabulary Orhan Elmaz (Universität Wien, Austria)
13.40-14.00	Discussion
14.00-14.20	<i>Coffee / tea break</i>
14.20-14.40	Some Remarks on Foreign Lexems in the Qur'an and the Work of Arthur Jeffery' Foreign Vocabulary David Kiltz (BBAW, Potsdam, Germany)
14.40-15.00	Presentation of the Database Project on the Qur'anic Vocabulary Catherine Pennacchio (Paris, France)
15.00-15.20	Discussion
15.20-15.40	<i>Coffee / tea break</i>
15.40-17:00	Session 4: Digital approaches in comparative perspective (1)
15.40-16.00	Greek into Arabic: the Glossarium graeco-arabicum, Pinakes Text Greek and Arabic, and the pseudo-Theology of Aristotle Cristina D'Ancona (Università di Pisa, Italy) Andrea Bozzi (ILC/CNR, Pisa, Italy)
16.00-16.20	Presentation of the Glossarium graeco-arabicum Yury Arzhanov (Ruhr-Universität Bochum, Germany) Torsten Roeder (BBAW, Germany)
16.20-16.40	The "Arabic and Latin Glossary" – A Research Tool for Medieval Translations from Arabic into Latin Barbara Jockers (Universität Würzburg, Germany)
16.40-17.00	Discussion
17.00	<i>End of Session 4</i>

Friday, 21 October 2011

09.00-11:00	Session 5: Digital approaches in comparative perspective (2)
09.00-09.20	Medico-botanical Synonym Lists in Hebrew Letters from the Iberian Peninsula. A Contribution to Arabic Lexicography Guido Mensching (Freie Universität Berlin, Germany) Dorothea Köhler (Freie Universität Berlin, Germany)
09.20-09.40	The Ancient Egyptian Dictionary Project at the Berlin-Brandenburg Academy of Sciences and Humanities Frank Feder (BBAW, Berlin, Germany)
09.40-10.00	Discussion
10.00-10.20	<i>Coffee / tea break</i>
10:20-10.40	Technical Approches and Database Technologies used for the Corpus Curanicum Markus Schnöpf (BBAW, Berlin, Germany) Michael Marx (BBAW, Berlin, Germany)
10.40-11.00	Discussion
11.00-11.30	<i>Coffee / tea break</i>
11.30-13:00	Session 6: Planning of Follow-up Research Activities and Collaborative Actions
13.00	<i>End of Workshop and departure</i>

Final list of participants

Convenor:

1. **Kirill DMITRIEV**
University of St Andrews
School of Modern Languages
Buchanan Building, Union Street
St Andrews, Fife KY16 9PH
Scotland, UK
Email: dmitriev@gmx.net

Co-Convenor:

2. **Michael MARX**
Corpus Coranicum
Berlin-Brandenburgische Akademie der
Wissenschaften
Am Neuen Markt 8
14467 Potsdam
Germany
Email: marx@bbaw.de

ESF Representative:

3. **Hanne RUUS**
Institut for Nordiske Studier og Sprogvidenskab
Københavns Universitet
Njalsgade 120
23000 København
Denmark
Email: haru@hum.ku.dk

Observer:

4. **Verena LEPPER**
Egyptian Museum and Papyrus Collection
National Museums Berlin
Bodestr. 1-3
10178 Berlin
Germany
Email: v.lepper@smb.spk-berlin.de

Participants:

5. **Yury ARZHANOV**
Seminar für Orientalistik
Fakultät für Philologie
Ruhr-Universität Bochum
Gebäude GB 2/34
Universitätsstr. 150
44780 Bochum
Germany
Email: Yury.Arzhanov@ruhr-uni-bochum.de

6. **Daniel BIRNSTIEL**
Faculty of Asian and Middle Eastern Studies
Cambridge University
Sidgwick Avenue
Cambridge CB3 9DA
United Kingdom
Email: db437@cam.ac.uk

7. **Andrea BOZZI**
CNR Istituto di Linguistica Computazionale
(ILC)
Via G. Moruzzi 1
56124 Pisa
Italy
Email: Andrea.Bozzi@ilc.cnr.it

8. **Daniel BURT**
The AALC Projekt
The Khalili Research Centre
3 St John Street
Oxford OX1 2LG
United Kingdom
Email: daniel.burt@orinst.ox.ac.uk

9. **Orhan ELMAZ**
Institut für Orientalistik
Philologisch-Kulturwissenschaftliche Fakultät
Universität Wien
Spitalgasse 2-4
1090 Wien
Austria
Email: orhan.elmaz@univie.ac.at

10. **Alba FEDELI**
Facoltà di Lettere e Filosofia
Università degli Studi di Milano
via Festa del Perdono, 7
20122 Milano
Italy
Email: Alba.Fedeli@unimi.it

11. **Frank FEDER**
Berlin-Brandenburgische Akademie der
Wissenschaften
Jägerstraße 22/23
10117 Berlin
Germany
Email: feder@bbaw.de

12. **Beatrice GRÜNDLER**
Department of Near Eastern Languages and
Civilizations
Yale University
P.O. Box 208236
New Haven, CT 06520-8236
USA
Email: Beatrice.Gruendler@yale.edu

13. **Stephan GUTH**
Department of Culture Studies and Oriental
Languages
The Faculty of Humanities
Universitetet i Oslo
Postboks 1010 Blindern
0315 Oslo
Norway
Email: stephan.guth@ikos.uio.no
14. **Robert HOYLAND**
Oriental Institute
Faculty of Oriental Studies
University of Oxford
Pusey Lane
Oxford, OX1 2LE
United Kingdom
Email: robert.hoyland@orinst.ox.ac.uk
15. **Barbara JOCKERS**
Institut für Philosophie
Universität Würzburg
Residenz - Südflügel
97070 Würzburg
Germany
Email: barbara.jockers@uni-wuerzburg.de
16. **David KILTZ**
Corpus Coranicum
Berlin-Brandenburgische Akademie der
Wissenschaften
Am Neuen Markt 8
14467 Potsdam
Germany
Email: kiltz@bbaw.de
17. **Dorothea KÖHLER**
Institut für Romanische Philologie
Freie Universität Berlin
Habelschwerdter Allee 45
Raum JK 30/111
14195 Berlin
Germany
Email: koehler@lingrom.fu-berlin.de
18. **Michael MACDONALD**
The Oriental Institute
University of Oxford
Pusey Lane
Oxford OX1 2LE
United Kingdom
Email: michael.macdonald@orinst.ox.ac.uk
19. **Guido MENSCHING**
Institut für Romanische Philologie
Freie Universität Berlin
Habelschwerdter Allee 45
Raum JK 30/111
14195 Berlin
Germany
Email: mensch@lingrom.fu-berlin.de
20. **Kathrin MÜLLER**
Kommission für Semitische Philologie
Bayerische Akademie der Wissenschaften
Alfons-Goppel-Strasse 11
80539 München
Germany
Email: kathrin.mueller@semphil.badw.de
21. **Anne MULTHOFF**
Lehrstuhl für Semitische Philologie und
Islamwissenschaft
Universität Jena
Löbdergraben 24a,
07737 Jena
Germany
Email: anne.multhoff@uni-jena.de
22. **Catherine PENNACCHIO**
CERMOM INALCO
7 rue Galliéni
92100 Boulogne Billancourt
France
Email: leah@orange.fr
23. **Alessia PRIOLETTA**
Dipartimento del Scienze Storiche di Mondo
Antico
Università degli studi di Pisa
Via Galvani1
56126 Pisa
Italy
Email: a.prioletta@sta.unipi.it
24. **Torsten ROEDER**
Berlin-Brandenburgische Akademie der
Wissenschaften
Jägerstraße 22/23
10117 Berlin
Germany
Email: roeder@bbaw.de
25. **Markus SCHNÖPF**
Berlin-Brandenburgische Akademie der
Wissenschaften
Jägerstraße 22/23
10117 Berlin
Germany
Email: schnoepf@bbaw.de

26. Peter STEIN

Theologische Fakultät
Universität Jena
Fürstengraben 6
07737 Jena
Germany
Email: peter.stein@uni-jena.de

27. Johannes THOMANN

Universität Zürich
Orientalisches Seminar
Wiesenstr.9
8008 Zürich
Switzerland
Email: johannes.thomann@access.uzh.ch

Statistical information on participants (age bracket, countries of origin, etc.)

Country representation

The workshop had 27 participants. 26 of them came from eight European countries and one from the USA.

Austria	1	Denmark	1 (ESF rapporteur)
France	1	Germany	13
Italy	3	Norway	1
Switzerland	1	United Kingdom	5
United States	1		

Gender repartition

One third of the participants were female.

Male: 17 Female (incl ESF rapporteur): 10

More than a half of the participants belong to the young generation of researchers with the age under 35 years.

