

# DIALECTAL ATLAS OF THE ARAB WORLD - BETWEEN INTENTION AND REALITY

Oleg REDKIN

Faculty of Asian and African Studies, Laboratory for Analysis and Modelling of Social Processes,  
St Petersburg State University,  
199034, 11, Universitetskaya nab., St Petersburg, Russia

and

Olga BERNIKOVA

Faculty of Asian and African Studies, Laboratory for Analysis and Modelling of Social Processes,  
St Petersburg State University,  
199034, 11, Universitetskaya nab., St Petersburg, Russia

## ABSTRACT

Arabic dialectology has a long history and achieved significant progress in collecting and analyzing linguistic data and its classification. The present paper analyses modern trends in the linguistic situation in the Arab world and defines the topics essential for the Arabic dialectology, which require an urgent solution. During the last century, several attempts have been undertaken to create dialectal atlases of different regions of the Arab world. Besides this, considerable work also has been done on synchronous descriptions of modern and ancient Arabic dialects.

Meanwhile these researches do not always reflect fully and adequately the current state of the dialectal continuum, which underwent significant transformations as a result of social and economic changes in the region, development of new technologies. Globalization and leveling of dialectal differences, spread of languages of international communication will lead to disappearance of small dialectal groups and vernaculars, among them those that are not sufficiently described or are little known at all, which presupposes their urgent investigation.

In recent years digital methods of research open new horizons for scholars and put on the agenda the task of building of the dialectal atlas of the Arab world basing on the previous experience and data collected as well as application of ICT.

**Keywords:** Arabic, dialectology, technologies, atlas, model, language.

## 1. INTRODUCTION

Arabic dialectology has always been an essential part of Arabic studies. Multinational population, vast geographical range of the Arab world, coexistence of different cultures on its territories contributed to the current linguistic diversity in the region. This diversity finds its reflection in simultaneous use of Modern Standard Arabic (MSA) as a primary mean of official and written communications and dialects, which provide everyday oral communication. MSA and dialects differ typologically and correspond to different linguistic systems. At present time these systems are characterized by dynamic development and need to be monitored on a constant basis. The present research defines

some solutions which can be effective in dialectal studies. They rely upon implementation of digital technologies in research and successive description of the linguistic picture of the Arab world. Today the use of ICT in humanities is urgently required, because “meta-research efforts should be done in order to relate research results in an adequate and more useful way” [1].

## 2. METHOD AND LITERATURE REVIEW

The study of the Arabic dialects has a long history. The pioneers of researches of this kind were traditional Arabic grammarians - representatives of grammar schools in Baṣra and Kūfa, such as Sibawayhi, al-Farāhīdī, al-Kisā'ī, Al-Fārisī and others. Thanks to their studies, and other secondary sources we have information related to the peculiarities of tribal dialects during the first centuries of the Hijra. Some features of the dialect of the Quraysh tribe may be also found in the text of the Qur'ān. Linguistic work of medieval philologists was based on a huge empirical data collected in the linguistic environment of nomadic tribes. They used a variety of methods for collecting and selecting language material from different informants [2]. In its turn, the European Arabic dialectology dates back to the 19th century. Since that time, it has achieved significant progress, and data related to various dialectal groups and local dialects has been collected and thoroughly studied.

This process has been facilitated by the fact that previously inaccessible areas in the Arab world have become more open, making it possible to carry on field linguistic researches there, and as a result, many lacunae in the dialectal picture of individual regions of the Arab world were filled up.

More than that some issues that previously were left out of the attention of researchers (such as gender or age) have become subjects of linguistic investigation.

It has broadened scholars' awareness about the problems related to the Arabic linguistics in general and Arabic dialectology in particular and the publications dedicated to it may be found in the reference bibliography compiled by M. Bakalla [3].

Later on, Arabic Linguistics Society developed “Bibliography of Arabic Linguistics 1979-1994” [4].

As for the example of ICT implementation in the Arabic dialectology in the last decade several e-resources were developed. One of them is Vienna Corpus of Arabic varieties -

an international project aiming at the collection of digital language resources documenting varieties of spoken Arabic. It provides a wide range of materials such as language profiles, dictionaries, annotated texts, bibliographies and more [5]. For the last years Natural Language Processing for Arabic Dialects has grown widely. One of the latest research in this regard was dedicated to the methods of natural language processing for Maghrebi Arabic Dialects [6]. Previously this dialectal group received less attention.

Thus, different tools for the Arabic language processing were developed by the present time, which could facilitate the investigation of the dialectal atlas of the Arab world.

### 3. MULTYDISCIPLINARY AND NEW ERA IN ARABIC DIALECTOLOGY

In recent years, there has been widespread use of digital technologies in the field of humanitarian research, which has been called Digital Humanities [7]. Development of new technologies paved way for new types of researches which has been implemented in such areas as the computer processing of text of the Qur'ān, digital analysis of Arabic manuscripts, automatic recognition of the Arabic speech [8], analysis of mass media [9] and development of computer models that reflect the history of the Arab caliphate.

Effectiveness of the dialectal researches also have risen up due to the implementation of digital technologies of linguistic data processing which range from accurate sound recording to computer research methods, such as data mining, spectral analysis, etc.

These opportunities allow to consider new problems and to carry on analysis basing on significant amounts of linguistic material. New technologies also allow to reach new practical results and implementations of linguistic researches, for example, to determine the type of dialect in an automated mode [10].

As a result, modern scholars may combine ICT and traditional ways of dialectal studies which rely on results of linguistic expeditions from various parts of the Arab world and desk researches as well.

Thus, during the last decades, much of the work has been done in the description of the ethno-linguistic situation in various parts of the Arab world. At the same time, a number of issues related to modern dialectal continuum should be studied more detailed due to the changes in the linguistic situation. For example, the influence of Modern Standard Arabic on the local dialects as a result of the development of education or the interference of foreign languages, English or French in particular.

The problem related to the impact of the official policy of the state on the *status quo* of linguistic landscape also has not been finally solved. Besides that, the consequences of the influence of mass media and the Internet on the local dialects has not been properly studied as well. It is also timely to develop further such kind of digital methods with respect to the classification of data with regard to the stratigraphy of Arabic dialects.

Some of these lacunae may be explained by the objective difficulties the dialectologists face with. First of all, the dialects are primarily a mean of verbal communication and the number of written texts is limited to advertising slogans, humor in newspapers and magazines (cartoons, anecdotes, etc.) and only occasionally occur in the literature.

Secondly, as a rule the area of territorial distribution of a dialect - is the homeland of its speakers and hence is associated with a certain territory (but the dialects of nomadic tribes), which

presupposes the data investigation within the dialectal environment.

The only exception seems to be the Egyptian dialect which, thanks to 'soap operas' and Egyptian-produced feature films and programs of Egyptian television entered every house in the Arab world. The spread of this dialect is facilitated by the fact that a large number of Egyptians, most of them are skilled labor force - doctors, engineers, teachers work in other Arab countries, mostly in the Gulf, are the actors of the dialectal influence.

### 4. ATLASES FROM VARIOUS REGIONS OF THE ARAB WORLD

During the last century, several attempts have been undertaken to create dialectal atlases of different regions of the Arab world. In one way or another most of the works on Arabic dialectology are supplied with graphics and maps and special attention in this respect was paid to the dialects of Yemen. The current state of the dialectal studies is regarded and scrupulously examined in series of publications [11].

In spite of some progress in Arabic dialectology there are a lot of lacunas in Arabic dialectal geography. Thus, P. Behnshteds notes that the "...Arabic dialectology has not contributed to the general theory of dialect geography... existing Arabic dialect atlases do not have such dense nets of research points as European ones" [12].

In fact, the information presented in the existing atlases is fragmentary and geographically heterogeneous. The following table summarizes the main dialectal atlases of various regions of the Arab world, presented in [12].

**Table 1.** Dialectal atlases of the Arab world

Author	Year	Region	Number of localities
Bergsträsser G.	1915	Syria Palestine	67
Cantineau J.	1940-1946	the Hōrān	60
Behnstedt P. and M. Woidich	1985	Egypt	560
Behnstedt P.	1985	North Yemen	165
Arnold W., Behnstedt P.	1993	Syria	50
Behnstedt P.	1997	Syria	500
Arnold W.	1998	Hatay (Turkey)	70
Mejri S.	2000	Tunisia	250

Even a fragment analysis of the data mentioned above demonstrates how much research remains undone with respect to the Arabic dialectal geography. There are no atlases to present which include the entire linguistic picture of the Arabic-speaking world. Besides that, the Arab world has undergone through rapid and dramatic developments, which affected Arabic dialects as well. Hence new field investigations of local Arabic vernaculars are urgently needed which must be carried out not only within

the natural language environment using advantages of modern linguistics but also with the implementation of computer tools [13].

Up today considerable work also has been done on synchronous and diachronic descriptions of modern and ancient Arabic dialects. It allows their detailed classification both in terms of geography and functional distribution, analysis of their morphological system, as well as vocabulary and syntax. All of these make it possible to create a linguistic dialectal atlas of the Arab world. Most of the publications on the Arabic dialectology include maps with isoglosses that could facilitate the future work.

However, in spite of existing theoretical base and modern technologies the task is much more complicated since the current linguistic situation is changing rapidly.

Due to the migrations of population the dialectal isoglosses often overlap each other. As a result, the peculiarities of local dialects, which were preserved to a large extent due to their isolation from one another, are gradually fading away. Political changes, the growth of living standards, spread of literacy and education affected all spheres of the life of the society, including the language (dialect).

For example, development of the system of education brought along the spread of Modern Standard Arabic (MSA) and the English language as well. The latter often plays the role of a language bridge between foreign labor force and local citizens and, to some extent, it has become the language of science and education in the Arab world. Invention of digital mass media and ICT brought into being super-dialectal areas of communication such as Twitter, chats, social networks.

Globalization and leveling of dialectal differences, spread of languages of international communication will lead to the disappearance of small dialectal groups and dialects, and among them those that are not sufficiently described or are little known at all, which presupposes their urgent investigation [14], [15].

## 5. THEORETICAL AND PRACTICAL IMPLEMENTATION OF THE ATLAS

Development of atlas of the Arabic dialects has practical as well as theoretical application. In the realm of science, it would help to solve a number of theoretical problems of Arabic and Semitic as well as comparative and historical linguistics. For instance, it would make it possible to trace the diachrony of distribution of the "bedouinized" variant of the phoneme /g/ vs. /q/ and hence to follow the ways of nomadic tribes in the past.

The atlas may be also useful in regard of historical studies, for example, in the study of the Arab conquests and formation of the caliphate during the first centuries of the Hijra, when Moslem armies from both nomads and sedentary tribes of the Arabian Peninsula - speakers of individual tribal dialects, migrated beyond the borders of their traditional homeland, and finally settled in Northern Africa, Spain, Mesopotamia and Palestine.

Today monitoring of the dialectal features in individual regions of the Arab world will foster our awareness of the consequences of migration processes and their global perspectives, cultural and ideological impacts of external linguistic influences.

Besides that, knowledge of the local dialects on the territories mentioned above may give hints about the traces of ancient migrations of the Arab tribes and their relations with autochthonous population, for instance in Maghreb and Malta.

This kind of atlas may be also used for pure educational purposes such as in the process of foreign language acquisition and fostering dialectal awareness and it would shed light on global perspectives of MSA and historical linguistics.

Practical implementation of the data presented in the atlas may be taken into consideration while developing speech recognition technologies, localization of software, development and optimization of search engines on the web etc. Besides, this kind of linguistic data may help improve cultural and educational policies of the local governments and administrations.

Since the regular language of everyday communication in one way or another, affects the entire mindset and the choice of behavior models of its speakers, the dialectal peculiarities such as vocabulary variations, gender defined varieties of the speech, etc., should be taken into account when building advertising and marketing strategies targeting consumers in certain regions of the Arab world. In this regard, the atlas would be one of the most helpful reference sources.

## 6. CONCLUSIONS

The study of modern Arabic dialects presupposes field investigations 'on the research site', i.e. in the dialectal environment while analysis of the collected data should be carried on mostly 'on the desk'.

Although modern technologies have expanded the sphere of dialects which today play a role as a mean of communication in the Internet, for example, in social networks, this does not exclude the need of carrying on field researches directly among of the dialect speakers.

The atlas may be presented as a traditional edition on paper periodically updated and improved, but above all it must be a computer model showing both the evolution of idioms and the real state of the dialects during certain period of time.

As for the necessary resources in recent years significant dialectal material has been collected and research methodologies for field and desk studies were developed in addition to the apparatus that allow to process scholar data with the help of digital technologies. The proposed project does not require significant material resources and lies in the area of scholar interests of researchers engaged in the Arabic dialectology.

In this regard, it is important that since 1993, the international school for the study of the Arabic dialects - International Association of Arabic Dialectology (AIDA) [16] has become a panel for exchange of ideas for international scholars, which may be considered the prototype of a research team working on this project.

To a large extent, it will represent the implementation and synthesis of their recent achievements and results of researches. It will also sum up the work of the dialectological community which has been carried on for over the past 150 years. The proposed project should be based on principles of international cooperation, interdisciplinary studies as well as usage of the cutting-edge technologies, and methods of linguistic research. Implementation of these principles and methodologies will multiply collective efforts of the project team members and combine skills of a large pool of scholars who may carry on common researches.

The proposed project of the dialectal atlas of the Arab world requires implementation of tools and attitudes similar to the Rosetta Project<sup>1</sup>, bringing together international scholars and using digital technologies and online resources, combining

<sup>1</sup><https://rosettaproject.org/>

results of field investigations and desk researches, along with the real-time monitoring of the current state of the linguistic picture of the Arab world.

In the future, it is necessary to make promptly upgrades of the map of the distribution of modern Arabic dialects and to check and if necessary modify existing isoglosses of phonetic, morphological and lexical peculiarities of these vernaculars.

## 7. ACKNOWLEDGMENTS

The funding no. 15.61.2219.2013 from Saint Petersburg State University (Russia) to support this research project is gratefully acknowledged.

## 8. REFERENCES

[1] N. Callaos, "Interdisciplinary Communication", **Systemics, Cybernetics and Informatics**, Vol. 11, No 9, 2013, p.23.

[2] E. Mishkurov, **Theoretical Course of the Arabic Language (in Russian)**, Moscow, 2004, p.479-480.

[3] M. Bakalla, **Arabic Linguistics: An Introduction and Bibliography**, London, Mansell, 1983.

[4] Arabic Linguistics Society, **Bibliography of Arabic Linguistics 1979-1994**,  
<https://community.dur.ac.uk/daniel.newman/Arabib.pdf>, retrieved on August, 2017.

[5] Vienna Corpus of Arabic varieties,  
<https://vicav.acdh.oeaw.ac.at/>, consulted online on August, 2017.

[6] S. Harrat, K. Meftouh, K. Smaïli, **Maghrebi Arabic dialect processing: an overview**. ICNLSSP 2017 - International Conference on Natural Language, Signal and Speech Processing, Dec 2017, Casablanca, Morocco. 2017.

[7] Humboldt Kosmos. Digital Humanities, Alexander von Homboldt Stiftung Foundation No. 102/2014.

[8] Z. Y. Mohammed, A. S. M. Khidhir. "Real-Time Arabic Speech Recognition", **International Journal of Computer Applications**, Vol. 81, No.4, November 2013, pp. 43-45.

[9] Zeev Volkovich, O. Granichin, O. Redkin, O. Bernikova, "Modeling and Visualization of Arabic Media", **Journal of Informetrics**, Vol. 10, Issue 2, pp. 439-453.  
<https://doi.org/10.1016/j.joi.2016.02.008>.

[10] O.F. Zaidan, Ch. Callison-Burch, "Arabic Dialect Identification", **Computational Linguistics**, Vol. 1, N. 1, 2012, pp. 1-35.

[11] P. Behnstedt, M. Woidich, "Arabische Dialektgeographie. Eine Einführung", Leiden-Boston, 2005.

[12] P. Behnstedt, "Dialect Geography", in: **Encyclopedia of Arabic Language and Linguistics**, Managing Editors Online

Edition: Lutz Edzard, Rudolf de Jong. Consulted online on 16 September 2017.

[13] O. Redkin, O. Bernikova, "The Arabic Diglossia: What is Next?" **3rd International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM 2016**, [www.sgemsocial.org](http://www.sgemsocial.org), SGEM2016 Conference Proceedings, ISBN 978-619-7105-70-4 / ISSN 2367-5659, 24 - 31 August, 2016, Book 1 Vol. 1, 1075-1080 pp, DOI:10.5593/SGEMSOCIAL2016/B11/S03.137.

[14] I. Vinnikov, "**Language and folklore of Bukharian Arabs**" (in Russian), Moscow, 1969.

[15] K. Versteegh, "**The Arabic Language**", Edinburgh, 2014, p. 214.

[16] International Association of Arabic Dialectology, <http://aidabucharest2015.lls.unibuc.ro/> Consulted online on August, 2017.