

الأطلس اللساني الرقمي: الواقع والمأمول. Digital Linguistic Atlas: State and Perspectives.

مریم صدیقی *

جامعة أبو القاسم سعد الله- الجزائر 2 (الجزائر)

Meriem Seddiki

Abu Alkacem Saad Allah - Algiers 2(Algeria)

meriem.seddiki@univ-alger2.dz

Received date: 02/05/2021	Acceptance date: 15/09/2021	Publication date: 30/09/2021
---------------------------	-----------------------------	------------------------------

ملخص

تحاول هذه الورقة تسليط الضوء على الاهتمام المتزايد بالأطالس، والمقاربات الأولى التي أدت إلى إنشاء خرائط ضخمة للظواهر اللسانية، وذلك في ظل ظهور اللسانيات الجغرافية، وهو التخصص الذي يقوم بتحديد المعارف التي يتم تنظيمها عبر الخرائط والأطالس في إطار مجال جديد من الأبحاث يتمثل في الأطلس اللساني الرقمي. كما تروم هذه الورقة تقديم نظرة عامة حول الأطالس اللسانية الأولى في العالم الغربي، وأهم تطلعاتها المستقبلية، ومن ثم إيجاد إجابة عن بعض الأسئلة، ومن أهمها: ما هو النطاق العام لمشاريع الأطالس اللسانية الحديثة في الغرب؟ وكيف يمكن للعالم العربي الاستفادة من هذه المشاريع لإنشاء الأطالس اللسانية الرقمية الخاصة بهم؟

كلمات مفتاحية: أطلس لساني رقمي؛ مشاريع؛ حوسبة؛ رسم خرائط لسانية؛ لسانيات جغرافية.

Abstract

This paper tends to highlight the growing interest in atlases, and the very earliest approaches led to a creation of comprehensive maps for linguistic phenomena, in the time of the emergence of geographical linguistics, a new discipline that identify the knowledge organized via maps and atlases, under the umbrella of a new field of research that is digital linguistic atlas. This paper tries, accordingly, to find an answer to some questions: What is the general scope of the recent western linguistic atlas projects? And how can Arab world benefit from these projects in this field to create their own digital linguistic atlases? Besides, it attempts to provide an overview of the earliest linguistic atlases in the Western world and their main perspectives.

Key Words: digital linguistic atlas; projects; computerization; linguistic cartography; geographical linguistics.

* Corresponding author: Meriem Seddiki

E-mail: meriem.seddiki@univ-alger2.dz

1. Introduction

During the last century, several attempts have been carried out to create linguistic atlases of different areas of the world including Arab regions. In recent years, however, there has been widespread use of digital technologies in the field of humanitarian research. The development of new technologies paved way for new types of research which has been realized in many areas of the globe. Thus, the computational use of maps and atlases are investigated and seen as a current focus of attention. The effectiveness of the linguistic and 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. (Bernikova & Redkin, 2017, p. 45). Besides, Western languages dialectology has achieved a significant progress in collecting and analyzing linguistic data and its classification. Arabic linguistic atlas, in its turn, witnessed an important but less recognized effort in the field. Accordingly, this paper attempts to spot the light on the growing interest in linguistic atlases mainly digital ones and provide a theoretical overview of some western as well as Arab world linguistic atlases and their main perspectives.

2. Definition of terms

First and for most, it is vital to know and identify the major key terms of the topic in question by giving a short but an accurate definition for each:

2.1. Atlas

An atlas is a coherent and systematic collection of geographical data in analogue or digital form, representing a particular area and/or one or more geographical themes, together with tools for information retrieval, analysis and presentation (Koop, 1993).

2.2. Digital atlas

In comparison with a GIS (a *Geographic Information System* that is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data), a digital atlas is a computerized geographic information system related to a certain area or theme in connection with a given purpose – with an additional narrative faculty in which maps play a dominant role (Van Elzakker, 1993); i.e, there would be some extra narrative virtue in a digital atlas that can be used, like explanation.

2.3. Linguistic atlas

Linguistic atlas, also dialect atlas, is a book of maps which show the distribution of language features over a chosen area. The maps show, with conventional signs such as dots, circles, and triangles, the locations of features as used by native speakers, such as sounds, words, or syntactic features (Oxford dictionary, 1998). Ideally, the speakers are directly interviewed in their home communities and their responses immediately noted, but the data are sometimes gathered by postal enquiry.

*A **linguistic map** is a thematic map showing the geographic distribution of the speakers of a language, or isoglosses of a dialect continuum of the same language. A collection of linguistic maps is a linguistic atlas (Martijn & Nerbonne, 2007, p. 572).

2.4. Digital linguistic atlas

A digital linguistic atlas may be defined, then, as a collection of a computerized thematic maps showing the geographic distribution of the speakers of a language, or *isoglosses* (*isogloss* is a line drawn on a map around the area in which a linguistic feature is to be found, such as a particular pronunciation of a given word) of a dialect continuum of the same language (Martijn & Nerbonne, 2007, p. 581).

2.5. Geographical linguistic

It is the branch of linguistics that involves the study of regional variations of speech forms. It is a language mapping for particular linguistic phenomena such as dialect distribution and isoglosses. Also called dialect geography or "geolinguistics" (Moseley, 2010).

2.6. Linguistic cartography

Linguistic cartography is a map survey documenting the cartographic characteristics of an existing language maps. The survey not only consistently categorizes language map symbology, it also captures unique strategies observed for handling locations with linguistic plurality as well (Oxford dictionary, 1998).

3. General aspects of linguistic atlases

Broadly speaking, when producing paper atlases in the past, the number and size of the map pages and the use of black and white or color maps as well as

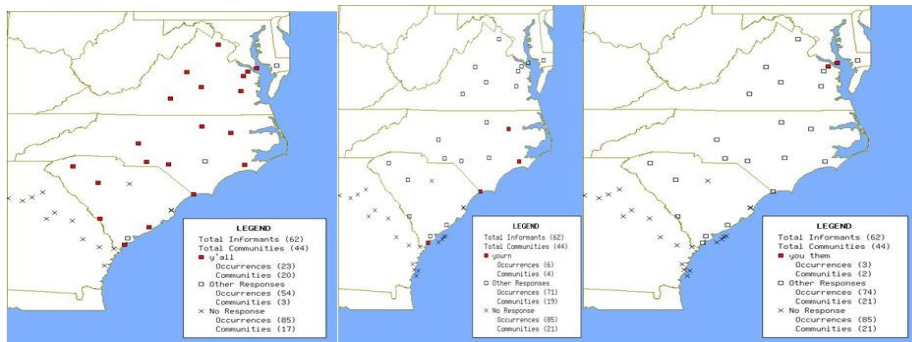
the print run would set the cost price for reproducing the atlas. Nowadays, however, this aspect can be ignored for digital atlases, apart from the production of DVD's or other data carriers (Ormeling & Elzakker, 2009, p. 3). Sometimes, it would be decided beforehand that the information can be available for free. Yet, the reason why atlases still provide an essential contribution is because the information in them has been processed in a systematic way.

Accordingly, it is worth mentioning that linguistic atlases have other general aspects which are highly prominent. Hotzenköcherle (1962) summarizes them as follows:

1. A linguistic atlas should present a cross-section of sounds, forms, words, etc. over space and time.
2. Building on the spatial structure that it results from, the goal should be a "language internal" analysis leading to a spatial typology.
3. Finally, the goal should be to relate the linguistic findings (whether of individual, structural or typological nature) to extra-linguistic phenomena in space.

In fact, the 1st stated aspect is an expression of the documentation principle, while the 2nd and the 3rd one refer to application, and thus to the analytical potential of an atlas.

Figure1: An example of a linguistic atlas



Source: Retrieved from a website: (<https://cutt.us/fXCvy>)

Besides, the use of atlases covers all the aspects of the use of individual, single maps, but has additional dimensions. By showing different maps in a specific sequence, specific points can be made since the main purposes of atlas production is to allow the comparison of maps, by which datasets are linked to one another, and relationships between their resulting pattern may be established throughout this comparison (Ormeling & Elzakker, Op.Cit., p. 2). (See figure 1).

4. Digitalizing linguistic Atlas projects

4.1. Historical overview

Historically speaking, the earliest atlas was the "*Sprachatlas des Deutschen Reiches*" of Georg Wenker and Ferdinand Wrede, published in the beginning of 1888, followed by the "*Atlas Linguistique de la France*", of Jules Gilliéron between 1902 and 1910, and the "*AIS - Sprach- und Sachatlas Italiens und der Südschweiz*" of Karl Jaberg and Jakob Jud, published 1928–1940 while the first linguistic atlas of the US was published by Hans Kurath. (Chambers & Trudgill, 1998, pp. 15-18).

In the last two decades of the twentieth century, language maps covering whole continents and regions came into view, but it was not until the *Atlas of the World's Languages* (Moseley and Asher, 1994) was published that a truly global coverage became possible. The world's languages cataloguing was initiated by the Summer Institute of Linguistics (now, SIL International) in their regularly revised volume, the *Ethnologue: Languages of the World* (Lewis, 2009); in its printed and online versions, this indeed include maps, but the maps are additional to the textual identification and description of a language (Moseley, 2010).

The following table may summarize the above earliest atlases' appearance:

Table1: The earliest linguistic atlases

Names	Date	Atlases
GeorgWenker Ferdinand Wrede	beginning of 1888	" <i>Sprachatlas des Deutschen Reiches</i> "
Jules Gilliéron	1902 -1910	" <i>Atlas Linguistique de la France</i> "
Karl Jaberg and Jakob Jud	1928–1940	" <i>AIS - Sprach- und Sachatlas Italiens und der Südschweiz</i> "
Hans Kurath	USA	The first linguistic atlas
Moseley & Asher	1994	" <i>Atlas of the World's Languages</i> " (by it, a truly global coverage became possible)
Lewis	2009	The <i>Ethnologue: Languages of the World</i> (a printed and online versions)

Source: the author (2019)

Whereas, during the 1970s, the first projects to use computer methods in linguistic cartography, especially in the planning of atlases, emerged, like the *Computer Developed Linguistic Atlas of England (CLAE)*, Viereck and Ramisch 1991_1997), the *Atlas Linguarum Europae (ALE)*, Alinei, Viereck and Weijnen

1983) and the *Kleiner Deutscher Sprachatlas* (KDSA, Veith, Putschke and Hummel 1984_1999) (Martijn & Nerbonne, Op.Cit., p. 585).

More recently, however, almost every atlas is realized using computers, and the most common procedure is to digitize the collected data and then analyze it statistically. Even though computer-based methods yield an additional benefit, the data continues to be explored, encoded and mostly classified using traditional techniques. Many atlas projects are currently renovating their (sometimes historical) data by providing digital editions with enhanced functions. Occasionally such work is conducted within larger organizational frames, as with the North American atlases (e.g., the *Linguistic Atlas of the Middle and South Atlantic States* [LAMSAS], cf. Kretzschmar and Schneider 1996) or the Bavarian atlases (e.g., the *Bavarian Database of Dialects* [BAYDAT], cf. Zimmermann 2007) (Ibid., p. 586).

4.2. Advantages and challenges

Although the idea of linking data is not new and it is not necessarily bound to the computerization of linguistics, still the digitization of data makes it possible to cross-connect projects around the world.

4.2.1. Advantages

Actually, a general consideration for producing atlases would be their ease of handling (just compare handling topographic map sheets or paper road maps with handling topographic atlases or road atlases or visualizing their contents on the screen of one's laptop) (Ormeling & Elzakker, Op.Cit., p.3). More importantly is the general concern that through atlases one is able to select and compare different views on geographical information: different topics for the same area, different snapshots in time for the same area and topic, or even different areas for the same topic and time frame.

Furthermore, it should be pointed out that thanks to GIS technology GIS (a *Geographic Information System* that is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data), cartographic quality in linguistics has improved in recent years. Whereas, in the 1990s and the beginning of the 21st century, linguists often still designed their atlases on their own. Yet, both the computational basis and collaborations within and across philologies seem set to become more relevant than ever via web mapping and other more techniques. (Martijn & Nerbonne, Op.Cit., p. 587)

One of the most ambitious undertakings in this regard, for instance, is the launched *regionalsprache.de* project (REDE; cf.

<<http://www.regionalsprache.de>>) at the Marburg research center. Among other things (e.g., the collection of fresh linguistic data), REDE aims to digitize a wide range of German language atlases and bring them together via the internet along with numerous sound files and bibliographic information (cf. in this regard the design of *DiWA* as well). (Ibid., p. 586)

As per the Arab world, the atlas can be 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. (Bernikova, & Redkin, Op.Cit., p.46)

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. (Idem)

In the sphere 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. 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 Modern Standard Arabic (MSA) and historical linguistics. (Idem)

4.2.2. Challenges

Despite the existing theoretical bases 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 with 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 (Ibid., p. 47)

^[22]. Nevertheless, such intentions are overwhelmed with both fundamental and more specific problems. One technical and organizational issue is the problem of how to handle cross-connectivity. Hence, as a future challenge, Geolinguists (and others) have the difficult task of ensuring that students and scholars are adequately educated.

For instance, 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 (Vinnikov, 1969). Invention of digital mass media and ICT "Information and Communication Technologies" 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 (Versteegh, 2014, p. 214).

5. The Arab world's Linguistic Atlas

Arabic language and its dialectology has always been an essential part of Arabic studies. Many factors such as 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 present everyday oral communication (Bernikova & Redkin, Op.Cit., p. 44).^[25]

Besides, with the arrival of Islam, Arabic gained a foothold in North Africa too, by which Arabic became the language of scholarship and religion. But, with the withdrawal of the Muslims from Spain in 1492, North Africa experienced a relative cultural decline. Classical Arabic fell into disuse, while colloquial Arabic began to spread throughout the region at the expense of Tamazight, and this movement continued with the growth of urban centers in the colonial era. (El Kirat *et al*, 2010, p. 28)

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. Yet, 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 (Bernikova,

& Redkin, Op.Cit., p. 45). By the implementation of ICT 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.

As a result, modern scholars can 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. Moreover, due to the changes in the linguistic situation, a number of issues related to modern dialectal continuum ought to be studied with more details. For instance, the impact 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. (Bernikova & Redkin, Loc.Cit.)

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 (Ibid., p. 46) ^[30]. Most of the publications on the Arabic dialectology include maps with isoglosses that could facilitate the future work.

6. Perspectives of the Arab world's Linguistic Atlas

Development of atlas of the Arabic dialects has practical as well as theoretical application. 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. Thus, Arab world can benefit from the latest projects in this field to create their own digital linguistic atlases.

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. The international school for the study of the Arabic dialects - International Association of Arabic Dialectology (AIDA) 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, linguistic atlases must represent the implementation and synthesis of their recent achievements and results of researches to sum up the work of the dialectological community which has been carried on for over the past 150 years. Any proposed project, then, should be based on principles of international cooperation, interdisciplinary studies as well as usage of the cutting-edge technologies, and methods of linguistic research. Thus, 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. (Op.Cit., p. 47)

However, the existed projects of the dialectal atlas of the Arab world require implementation of tools and attitudes by bringing together international scholars and using digital technologies and online resources, combining 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.

7. Conclusion

All in all, although modern technologies have expanded the sphere of linguistic atlases 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. With the help of digital technologies, significant dialectal materials as well as the necessary resources can be collected and research methodologies for field and desk studies are developed in recent years together with the apparatus that allow to process scholar data, the fact which give helpful hints to improve the current digital linguistic atlases in the world.

References

Books

- Chambers, J.K.; Trudgill, P. (1998). *Dialectology* (2nd ed.). Cambridge University Press. pp. 15–18
- Concise Oxford Companion to the English Language. (1998), originally published by Oxford University Press.

- El Kirat, El Allame, Y. and Mezhoud, S. (2010). "North Africa and the Middle East". In: Moseley, C.(ed.), Atlas of the World's Languages in Danger. 3rd edition, UNESCO Publishing. P. 28
- Hotzenköcherle, R. (1962). Einführung in den Sprachatlas der deutschen Schweiz. A: Zur Methodologie der Kleinraumatlanten. (Introduction to the Language Atlas of German Switzerland. A: The methodology of the small-space atlases). Berne: Francke. In: Martijn, W. and Nerbonne, J. (2007). Linguistic atlases: Traditional and modern. Vol. 32, p. 581
- Martijn, W. and Nerbonne, J. (2007). Linguistic atlases: Traditional and modern. Dialect pronunciation comparison and spoken word recognition. In: Osenova, Hinrichs and Nerbonne (eds.), p. 572
- Moseley, C. (2010). Atlas of the World's Languages in Danger. 3rd edition, UNESCO Publishing.
- Vinnikov, I. (1969)."Language and folklore of Bukharian Arabs" (in Russian), Moscow.
- Versteegh, K. (2014)."The Arabic Language", Edinburgh, p. 214

Articles

- Bernikova, O. & Redkin, O. (2017). "Dialectal Atlas of the Arab World - between Intention and Reality". In: Systemics, Cybernetics and Informatics Journal. Vol. 15, N. 5, p. 45
- Francke. In: Martijn, W. and Nerbonne, J. (2007). Linguistic atlases: Traditional and modern. Vol. 32, p. 581
- Freudenberg, R. (1965). Die Entwicklungsgeschichte der dialektgeographischen Methode. *Zeitschrift für Mundartforschung* (The development history of the dialectgeographical method. Journal for dialect research). In: Martijn, W. and Nerbonne, J. (2007). Linguistic atlases: Traditional and modern. Vol. 32.
- Koop, R.O. (April, 1993), Tools for the Electronic Production of Atlases. In: Klinghammer, I.; L.Zentai & F.J.Ormeling (eds.). *Proceedings of the Seminar on electronic atlases* held at Eötvös Lorand University, Visegrad, Hungary, Budapest: ICA.
- Ormeling, F., & van Elzakker, C. P. J. M. (November, 2009). Digital atlas user requirements and use scenarios. In *ICC 2009 : Proceedings of the 24th international cartographic conference ICC : The world's geo - spatial solutions, Santiago, Chile*. International Cartographic Association. P. 3

- Van Elzaker, C.P.J.M. (April, 1993). The use of electronic atlases. In: Klinghammer, I.; L.Zentai & F.J.Ormeling (eds.) *Proceedings of the Seminar on electronic atlases* held at Eötvös Lorand University, Visegrad, Hungary, Budapest: ICA, pp. 145-155.

Websites

- Siiri Lehtinen.(2018). "Dialectology & Language Variation", retrieved from: <https://cutt.us/fXCvy>
 - Vienna Corpus of Arabic varieties, retrieved from: <https://vicav.acdh.oeaw.ac.at/>, consulted online on August, 2019.
 - International Association of Arabic Dialectology, retrieved from: <http://aidabucharest2015.ils.unibuc.ro/> Consulted online on August, 2019.
-